Location Effects or Sorting? Evidence from Firm Relocation*

Pauline Carry	Benny Kleinman	Elio Nimier-David
Princeton University	Stanford University	Cornell University

June 9, 2025

Abstract

Why are wages in New York or Paris higher than in other cities? This paper uses establishment mobility to separate the role of "location effects" (e.g., local geography, infrastructure, and agglomeration) from the spatial sorting of workers and firms. Using French administrative records and U.S. commercial data, we document that 4% of establishments relocate annually. Establishments retain their main activity and structure as they move, but adjust their workforce and wages. Combining establishment and worker mobility, we decompose wage disparities across French commuting zones. We find that spatial wage differences are largely driven by the sorting and co-location of workers and firms: location effects account for only 2–4% of disparities, while differences in the composition of workers and establishments account for 30% and 17%, respectively. The remaining half is accounted for by the co-location of high-wage workers and establishments, 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 that do not control for establishment composition.

*We are grateful to Hadar Avivi, Michèle Belot, Marianne Bertrand, Zach Bleemer, Janet Currie, Ellora Derenoncourt, Jonathan Dingel, Cecile Gaubert, Matthew Gentzkow, Joshua Gottlieb, John Grigsby, Rick Hornbeck, Anders Humlum, Simon Jäger, Philipp Kircher, Francis Kramarz, Ilyana Kuziemko, Thibault Lamadon, Jeremy Lise, Jack Mountjoy, Matthew Notowidigdo, Ezra Oberfield, Evan Riehl, Roman Rivera, Richard Rogerson, Evan Rose, Esteban Rossi-Hansberg, Karthik Sastry, Benjamin Schoefer, Gabriel Swagel, Chad Syverson, Owen Zidar and Eric Zwick for useful feedback. We would like to thank conference and seminar participants at Cowles, the University of Chicago, Cornell University, Princeton University, HEC Paris, Houston University and Boston University for their comments. We thank Michael Amior for discussing our paper at the Northeast Labor Symposium for Early Career Economists. We thank Cedric Christensen, Silvia Pedersoli, Noah Simon, and Alex Weinberg for outstanding research assistance. This work is supported by a public grant overseen by the French National Research Agency (ANR) as part of the "Investissements d'Avenir" program (reference: ANR-10-EQPX-17 - Centre d'accès sécurisé aux données – CASD).

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 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; Slattery and Zidar, 2020) under the assumption that bringing in "high-paying firms" can raise local prosperity. This distinction 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 provide the first empirical decomposition of spatial wage disparities that separately identifies the contributions of locations, establishments, and workers. We leverage establishment mobility across space to disentangle "location effects" from spatial sorting. To this end, we first establish a set of stylized facts about the understudied phenomenon of establishment mobility. Using administrative data from France and commercial data from the United States, we examine how prevalent relocation is, what it entails for an establishment to move, and which establishments 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—across both establishments and space—to decompose local wage disparities in France. Our main finding is that location effects account for only 2-4% of wage differences between commuting zones (CZs), while the spatial sorting of workers and firms accounts for most regional wage disparities. Using these estimated location effects, we revisit the city-size wage premium and provide new estimates of the elasticity of wages to local density that control for worker and firm composition.

Studying relocation decisions presents a key challenge: in many administrative data sources, an establishment's identifier changes following a move. 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 unique to these data, we observe the establishment's identity, its old location, and its

[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*" by the French statistical office. To obtain a valid business license, entrepreneurs are required to submit proof of both closure and reopening to the local court within three months of the move.

new location on the same form, 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 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, we find that headquarters are more likely to relocate.

We begin by presenting key facts about establishment 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 relocations involve a change in commuting zone. Second, we explore the geography of relocation. Similar to worker migration, establishment mobility is characterized by sizable bidirectional flows between all pairs of locations, suggesting little role for shared location factors like market size or local taxation in driving destination choices. Also similar to worker flows, establishment relocations are well approximated by a gravity equation, with distance playing a strong role in movement rates between locations. Third, we examine heterogeneity in establishments' location decisions and find evidence of spatial sorting: conditional on origin location and firm age, larger and higher-wage establishments are more likely to choose larger and higher-wage destinations.

We next examine how establishments evolve following a relocation to a different commuting zone. We find that relocation results in minimal changes to an establishment's core activities or production structure, confirming that we observe the same establishment before and after the move. 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 experience no 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. Half of former employees are separated from the establishment during the relocation. The average wage of an establishment also 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 wage decomposition, most of the change reflects a different composition of workers.

We develop a simple spatial equilibrium framework to rationalize the descriptive evidence on relocation decisions and to further motivate our wage decomposition exercise. The framework incorporates both firm and worker mobility, providing a theoretical interpretation of our estimated worker, establishment, and location effects. In the model, firms with heterogeneous productivity select locations to maximize profits, subject to idiosyncratic location preferences, akin to the migration literature. Workers with heterogeneous abilities choose locations and match with firms in frictional local labor markets. The model predicts that higher (lower) productivity firms tend to relocate to areas with higher (lower) regional productivity or more (less) productive workers relative to their origin, due to complementarities in the production technology. These drivers of mobility align with our main findings on relocation decisions, including gravity-like relocation patterns and sorting of large and high-pay establishments to large and high-pay cities. Notably, the model also delivers a simple log-linear wage equation featuring distinct location, firm, and worker components, consistent with our wage decomposition exercise and its underlying assumptions.

In our main decomposition exercise, we use 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 establishment mobility connects locations. Identification relies on the assumption that, conditional on worker, establishment, and location types, establishment and worker mobility is as good as random. We present four tests supporting this assumption. First, we consider event studies for establishment and worker wages as in Card et al. (2013) and Badinski et al. (2023). 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 relocating establishments did not experience shocks that directly affect wages. Fourth, we replicate our analysis using a subset of relocations that are likely exogenous, driven by entrepreneurs' personal hometown preferences.

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%–4%

of the total variance in log wages across CZs. 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 establishment composition contribute an additional 17%. 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 establishments and high-type workers in the same areas. Finally, the sorting of high-type establishments and workers into CZs with strong location effects accounts for an additional 15% of the total variance. Notably, we find little heterogeneity in this decomposition when estimating the model across broad sectors, showing that the contribution of location remains small even when allowing for industry-specific effects.

These findings suggest that spatial sorting and segregation play a far greater role in explaining regional wage disparities than local geography, physical infrastructure, or regional agglomeration effects. A key question, then, is whether standard evidence for agglomeration effects holds once we account for the composition of both workers and establishments. 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 establishment composition is also accounted for, the estimate drops further to a precise and significant 0.007. Although the wage-density relationship remains positive, it weakens substantially once spatial sorting is accounted for, indicating that existing estimates largely reflect differences in local establishment composition.

We highlight two key implications of our results. First, location effects account for only a small share of spatial wage disparities, while the sorting of workers and firms explains the majority. These results suggest a relatively small role for agglomeration spillovers that manifest themselves through higher regional productivity, such as those typically considered in spatial equilibrium models. They also offer guidance on how to build and discipline spatial models of cities and regions. Second, our results 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 the 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; Moretti, 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 (and the co-location with workers) 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; Baum-Snow and Pavan, 2012; Moretti, 2012; 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 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 that examines firm sorting as a key mechanism behind spatial disparities in economic activity, including Combes et al. (2012), Behrens et al. (2014), Suárez Serrato and Zidar (2016), Gaubert (2018), Lindenlaub et al. (2022), Bilal (2023), Hong (2023), Mann (2023), Oh (2023), Kleinman (2024), and Lhuillier (2025).[4] Most of these papers incorporate heterogeneity in two of the three dimensions—location, workers, and firms—and none estimates the separate contributions of each along with their co-location patterns. Our results offer a set of empirical moments that can help discipline the importance of sorting relative to location fundamentals and agglomeration forces in spatial equilibrium models. Although our focus is empirical, we also present a spatial equilibrium framework with three-way heterogeneity and endogenous mobility of both workers and firms, consistent with our empirical findings.

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 system-

[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 27% 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, and to Bilal (2023) who studies dispersion in job loss rates across space.

[4]See also Diamond and Gaubert (2022) on worker sorting and Diamond and Suárez Serrato (2025) for a recent survey.

atic 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. This paper is also 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 consider spatial wage dispersion driven by both single-establishment firms—that constitute most firms—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 U.S. 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. We then introduce the data used to proxy relocation rates for the U.S.

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 business (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 has no incentive to declare a relocated establishment as a new one, as this would involve more bureaucratic procedures than reporting a change of address. Second, new establishments are unlikely mistaken as relocations, because 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 the 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, work, and birth) 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 location data are highly granular, reflecting France's division into approximately 35,000 municipalities ($\approx 40\%$ of the total in Europe). We leverage this granularity to distinguish mobility within and across CZs. We focus on companies that operate in the 304 CZs for metropolitan France as defined by INSEE in 2010.

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 2002–2016.[5] 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 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.[6]

For a subset of results, we also use firms' balance sheet data (Ficus-Fare), which allow us to compute the stock of capital. As balance-sheet data are reported at the firm level, 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

[5] We end our period of consideration in 2016 because the data collection changed in 2017, creating a temporary break in the series.

[6] For a detailed description of the evolution of spatial disparities in France, see Kramarz et al. (2022).

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 of their establishments.

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 identity and operations. 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 five-digit 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.[7]

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.

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 2021.[8] For each year, we compute the ratio of relocating establishments to the total number of operating establishments 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 single-unit firms and headquarters of multi-unit firms 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

[7]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.

[8]For the US, we compute several measures of the relocation rate using two datasets. In the D&B data, we identify relocations based on changes in zip codes (see Section 2.3). In the Compustat data, we track relocations using changes in both zip codes and street addresses. In both datasets, a relocation refers to a change in location for single-establishment firms or a change in headquarters location for multi-establishment firms.

between 5.0% and 6.0% over the period (consistent with Duranton and Puga (2001) who focus on this industry).[9] 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%. Zooming in on moves between CZs, we find that movers and non-movers have a similar industry composition by major sector (Panel (b)). Finally, Figure B.2 shows that, in each year, establishments that relocate account for approximately 2.5% of the French workforce, and that relocations occur across the entire size distribution.

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), indicating that a substantial share of moves correspond to a change in the local labor market.[10] In line with patterns observed in worker migration, we also document a strong gravity structure for relocation rates. To this end, we estimate the following model:

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

where the dependent variable is the log of the number of establishments relocating from province i to province j (or CZs in the case of the US). 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 role for spatial frictions in shaping relocation decisions.[11]

Symmetry of flows. A key characteristic of worker mobility across space is the prevalence of bidirectional flows, even between very different local markets. This pattern often motivates models of location choice to allow for substantial dispersion in idiosyncratic

[9] The business service industry includes telecommunications, real estate, IT, R&D, and services provided to companies such as consulting or marketing.

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

[11] In France, the average and median relocation distance for an establishment is 151 and 42 kilometers respectively. It is 178 and 69 kilometers for workers (see Appendix Table B.1). Appendix Figure B.5 plots the distribution of relocation distances for establishments, highlighting the existence of relocations spanning several hundred kilometers.

preferences. We find a similar pattern for firm relocations. 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, such as market size or taxes.

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 to highlight sorting patterns that are not driven by life-cycle dynamics. We also control for the commuting zone of origin, η_i, to avoid capturing systematic differences in firms' original location. 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 CZs, with elasticities of 0.086 (standard error: 0.008) for France and 0.056 (standard error: 0.001) for the US. Second, higher-paying establishments also relocate to CZs where wages are higher on average. Using the French data, we estimate an elasticity of 0.022 (standard error: 0.002). These correlations suggest that establishment relocation contributes to the sorting of higher-paying establishments into higher-paying locations. In Section 6.1, we reproduce this analysis using estimated fixed effects instead of characteristics and find that the sorting is driven by higher fixed effects establishments moving to locations with higher average worker fixed effects.

3.3 What Happens to an Establishment that Relocates?

We now analyze 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.[12] The last column of Table 1 indicates that the top manager remains the same in 67% of relocating establishments, compared to 76% 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 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. We match

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

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, 55% of entrepreneurs report no major changes since firm creation, and 38% 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%), whether they serve local, regional, national or international clients (9%), the number of clients (31%), and the type 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 (6%) or competition (28% 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 following relocation, we now examine whether they adjust their labor input. Specifically, we investigate whether they begin employing workers from the new location and whether compensation adjusts in response to the local wage level.

First, we find that a significant portion of an establishment's workforce is replaced by workers from the new location following a move. About half of the establishment's workers from the former location cease to be employees, and this share increases over time as establishments continue to separate from their original workforce. 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)). Employment in the new location begins prior to the relocation, jumping from around 20%-25% to around 50% upon a move, and continues to grow thereafter.

Given that both location and workforce composition change with relocation, we now examine how wages evolve within establishments. Specifically, we test whether establishments 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 CZs. 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 calendar year fixed effects, respectively. Δ_J represents the difference in average log hourly wages between the destination and origin CZs, 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 CZs. 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 our results are not driven by composition effects. 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 CZs.[13]

Figure 6 presents the results for establishment relocations (Panel (a)) and worker relocations (Panel (b)). In both cases, wages follow similar trends 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 establishments, relocating to a commuting zone with 1% higher wages is associated with a wage increase of 0.14-0.17% (the corresponding elasticity for worker relocations is higher, at 0.23-0.27). The fact that this elasticity is below one highlights the significant role of establishments in wage setting. Meanwhile, the coefficient being significantly different from zero indicates that the location matters as well—either through local labor market conditions or shifts in workforce composition following relocation. In Section 5, we further disentangle the contribution of the local labor force from that of the location itself.

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. By observing the same establishment—with the same activity and input composition—operating in different locations, we can directly assess how wages vary

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

across space. Combined with worker mobility data, this approach allows us to disentangle the respective 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 featuring heterogeneous workers, firms, and locations, along with 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:

$$\max_{n \in \{1...N\}} w_n(a) - \kappa_{on}^w + \epsilon_n^w,$$

where κ_{on}^w represents the bilateral migration friction between locations o and n (equal to zero if $n = o$), 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). Expected profits depend on the local composition of workers. A firm owner who starts their business in location o chooses the location n that maximizes:

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

where κ^f_{on} is a bilateral relocation friction between locations o and n, and ϵ^f_n is an idiosyncratic preference shock for location n. These shocks are drawn from a Type-1 extreme value distribution, with shape parameter ξ^f. 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.

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, $\Phi_n = \bar{\Phi}_n L_n^{\alpha}$.[14]

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 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 $p_n = \vartheta_n^{1-\eta}$.

[14] 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. Each firm is atomistic and does not internalize its role on the location size.

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 we assume that both parties' outside options after the sunk cost c has been paid are zero.[15] 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.[16]

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 $s^f_{on}(z)$ 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:*

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

(b) ***The model features gross bilateral relocation flows:*** *if κ^f_{on} and κ^f_{no} are finite, then $s^f_{on}(z) > 0$ and $s^f_{no}(z) > 0$ for all o, n, z. Other things equal, an increase in the dispersion of idiosyncratic firm shocks (higher ξ^f) leads to a lower ratio of net-flows to gross-flows between locations, $\frac{|M_{on}(z)-M_{no}(z)|}{M_{on}(z)+M_{no}(z)}$.*

(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:*

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

[16] 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.

$H(v) = \frac{v^{1+\delta}}{1+\delta}$. *Then,* $s_{on}^{f}(z)$ *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, 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*,* $s_{on}^{w}(a)$*, 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 *who 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-paying firms into higher-paying locations (see Figure 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) and complementarities in the production function.

5 Empirical Model and Identification

In this section, we introduce the empirical model we use to decompose 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, establishment, and location heterogeneity to spatial wage disparities. Our main analysis is based on the following model:

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

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 a move. This implies, for instance, that establishments do not use relocation as an opportunity to modify their type of activity in a way that would affect wages. This assumption is supported by the findings in Section 3.3, which show that establishments maintain the same type of activity, legal status, headquarter status, and input composition. Moreover, only changes in ϕ that are systematically correlated with the location effect would bias our results. Specifically, we would overestimate the contribution of location effects if firms that become more productive (or adopt higher pay policies) tend to relocate to areas with higher fixed effects.[17] However, since our estimates indicate a small role for location effects, this is unlikely to be substantial.

In our main specification, each establishment has its own fixed effect. As an alternative, we consider a specification that clusters establishments together, following Bonhomme et al. (2019), to account for the potentially limited mobility of workers between establishments and to reduce dimensionality.

The location fixed effect, ψ_c, captures compensation differences between locations while controlling for worker and establishment heterogeneity. 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. We define locations at the level of the commuting zone in our main analysis and present a robustness at the province level.[18]

[17] Conversely, we would underestimate the role of location if establishments experiencing negative shocks tend to relocate to higher-paying areas.

[18] Commuting zones provide a relevant definition of local labor markets, because they are determined by commuting flows between place of residence and place of work. 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 worker's age and year fixed effects.

In Section 6.2, we present evidence supporting our specification choices and 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 double-mover 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 effect (ψ). 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.[19] 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 relocating establishments to other firms in both the origin and destination locations. In the context of Figure 7, this means that some workers transition to or from Firm B while 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 long as Firm B remains connected to the rest of the sample through worker mobility in both locations. In our main estimation, we exploit these two sources of identification. As a robustness check, we re-estimate the model excluding workers who move with their establishments to rule out potential biases from wage rigidities. This alternative specification produces comparable results. We now turn to a discussion of the main identification assumptions.

[19] The concern that wage rigidity may bias our estimates is mitigated by the legal requirement mandating that all employment contracts be renegotiated when a relocation exceeds 10 kilometers.

Additive separability. Following a large body of literature, we assume the additive separability of log wages, meaning that they can be expressed as a sum of fixed effects. In Section 4, we show that this assumption is supported by a simple search and matching model with worker and firm mobility in which the production function features complementarities. Empirically, the high adjusted R^2 of 0.86 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 different types of establishments to move to different types of locations, and allows for different types of workers to sort into different establishments and locations (e.g., assortative matching). For example, our theoretical model predicts that higher fixed-effect establishments will relocate to more productive areas and to locations with higher average worker effects, due to the super-modularity of the production function. This assumption is also consistent with moves driven by idiosyncratic preference shocks (e.g., home bias) that are orthogonal to time-varying unobservables (ϵ_{ict}). In practice, it rules out two main scenarios: (i) mobility induced by shocks or trends 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 match. We implement four 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. 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 quartiles of estimated fixed effects (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 are 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 in absolute value. This symmetry is consistent with our log additive model and supports the limited role of match effects in driving relocation or job changes.[20] Third, while wage changes for workers switching establishments are large—reaching up to thirty log points—the corresponding wage changes for relocating establishments is much smaller, suggesting greater heterogeneity in establishment effects than in location effects.

Second, our framework allows for a novel placebo test: we examine the 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 that pass-through to wages or if the process itself impacts wages. Panel (d) 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 SINE 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 37% of both groups reporting no difficulties. Appendix Figure B.6, panel (a), further shows that both groups report comparable changes in local competition, with 58% of movers indicating no change and 55% of non-movers. Panel (b) confirms that movers and non-movers faced similar obstacles when opening their establishments. In sum, these results suggest that establishment relocations are not primarily driven by shocks that affect wages or binding constraints.

Fourth, we show that our main results are robust to re-estimating model (5) using only the subset of establishments that relocate to the hometown region of their top manager.

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

We interpret these moves as being primarily driven by managers' personal preferences, rather than economic motives also affecting wages. We detail our approach in Section 6.2.

Dynamics. We do not model explicitly the possibility that the effects of changing firm or location may take time to fully materialize, for instance because of learning (De la Roca and Puga, 2016). Instead, our estimates reflect the average wage change observed in the years following a job change or a relocation, as we include all subsequent observations. Nevertheless, our evidence suggests that dynamic effects are limited. Event studies estimates based on Equation (4) show that, for both relocating establishments and workers, most wage adjustments materialize immediately in the year of the move, with only limited subsequent dynamics (see Figure 6). Specifically, the wage coefficient increases from 0.23 to 0.27 for relocating workers and from 0.15 to 0.18 for relocating establishments between the year of the move and four years later, suggesting that dynamic effects exist but are relatively modest compared to the overall impact. Similarly, Appendix Figure B.8 shows that wage changes for workers switching firms—whether within or across commuting zones—also occur largely in the year of the transition.

5.3 Sample and Connected Set

We estimate Equation (5) on the universe of companies, operating in the private sector, with at least one full-time equivalent employee, over the period 2002–2016. Table 2 presents summary statistics for establishments and workers by mover status. Although movers and stayers are not identical, they exhibit substantial overlap in observable characteristics. On average, establishments that relocate have 14.5 employees, compared to 14.3 for non-movers. They employ a slightly higher share of skilled workers. Finally, establishment relocation is observed in all major industries, including manufacturing, services, retail and construction.

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 average of the number of unique workers and establishments per CZ, as well as the number of movers.

Existing studies using two-way fixed effect models emphasize that 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 concern is mitigated in our case for two reasons. First, we observe a substantial volume of worker and establishment moves. Second, we decompose differences in wages *across* a small number of commuting zones. As explained in Section 5.4, we aggregate estimated fixed effects at the CZ level before computing the variance, over a large number of workers and plants. With a sufficient number of observations per area, the noise should be close to zero (as confirmed by the estimated variance of residuals across CZs which is approximately null).

To further verify that our results are not affected by limited mobility bias, we implement two additional strategies. First, following Bonhomme et al. (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 population-weighted commuting zone-level averages of log hourly wages, the fixed effects estimated from Equation (5), the time-varying controls, and the residuals. Each average is weighted by the number of worker-year observations. Our goal is to assess the contribution of these components to the dispersion in wages across commuting zones. We rely on the following variance decomposition:

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

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 co-locating and sorting between locations. A positive $cov(\psi_c, \overline{\phi_c})$ (respectively, $cov(\psi_c, \overline{\alpha_c})$) suggests that high-paying establishments tend to locate in high-paying areas (respectively, high-ability workers sort into high-paying locations). Finally, $cov(\overline{\alpha_c}, \overline{\phi_c})$ informs us on the degree of co-location of higher-paying establishments and high-ability workers, irrespective of the productivity of the location.

In the following analysis, 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 results from the baseline model, where each establishment has its own fixed effect. Column (2) presents the specification with clustered establishments, and column (3) reports the variance-corrected version. The table provides the standard deviation of average log hourly wages between CZs, along with the standard deviations of worker, plant, and location effects, time-varying controls, and two times the covariance terms, following Equation (6).

First, the central finding is that location effects alone account for only 4.2% of the total variance in log wages across CZs (2.4% in the clustered version). This indicates that location-specific characteristics contribute relatively little to spatial wage disparities. Instead, most of the variation arises from heterogeneity in workers and establishments, and from their co-location patterns over space.

Second, a substantial share of variation is explained by differences in worker and establishment effects. Worker effects ($\overline{\alpha_c}$) explain approximately 30% of the variance between CZs (31% in the clustered version), while establishment effects ($\overline{\phi_{J(c)}}$) explain around 17% (13% in the clustered version). Appendix Figure D.2 maps average wages and estimated fixed effects across France. The spatial distribution of worker and establishment effects closely mirrors the distribution of gross hourly wages, with geographically concentrated pockets of high productivity, particularly in the Paris region and around other large cities such as Lyon, Grenoble and Toulouse. Location effects are higher in the Paris area, but otherwise feature low dispersion.

Third, the sorting of workers and establishments across locations accounts for the remaining half of the variance. Approximately one-third of spatial wage disparities is due to the co-location of more productive workers and higher-paying establishments—independent of location effects (40% in the clustered version). Moreover, both productive workers and establishments disproportionately sort into high-type locations, further reinforcing this mechanism. The covariance between worker and area effects explains about 11% of the variance (7.9% in the clustered version), while the covariance between establishment and location effects account for another 4.9% (7.2% 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, although variation in location effects alone explains only a modest share of spatial wage disparities, the sorting of workers and firms across heterogeneous locations amplifies the aggregate impact of these differences.

In Section 3, we document that, on average, larger and higher-paying establishments tend to relocate to larger and higher-paying CZs. To further disentangle the drivers of this sorting, we leverage our estimated fixed effects. Appendix Figure D.3 plots the correlation between the fixed effects of relocating establishments and the difference between destination and origin in (i) location effects (panel (a)), (ii) the average worker effect (panel (b)), and (iii) the average establishment effect (panel (c)). The results indicate that higher fixed-effect establishments primarily relocate to areas with a higher concentration of high-ability workers and high-fixed effect plants. This suggests that human capital plays a role in shaping relocation decisions.

6.2 Robustness and Heterogeneity Analysis

Alternative specifications. We consider several alternative versions of our model and sample. The corresponding variance decompositions are presented in Table 4, with Column (1) reproducing our baseline estimates for reference.

First, we include additional time-varying controls in our specification to account for establishment's age (see columns (2) and (3)). Specifically, we include a polynomial of the establishment's age to capture potential changes in firm pay-setting practices over the firm lifecycle (Brown and Medoff, 2003; Babina et al., 2019).[21] The inclusion of these controls has negligible impact on our results.

Second, we re-estimate our model excluding workers who remain with their establishment when it relocates (column (4)). As discussed in Section 5.1, our parameters of interest remain identified through establishment mobility between locations and worker mobility between establishments in the origin and destination locations. This analysis serves to ensure that our findings are not influenced by potential wage rigidity for incumbent workers or increased compensation due to longer commute (Verdugo and Kandoussi

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

(2024)). When using only new hires for identification, we obtain very similar results, with location effects account for 2.7% of wage disparities in a clustered model.

Third, we examine alternative methods of clustering establishments and varying the number of clusters. Column (5) reports results from clustering on means and standard deviation rather than percentiles. In column (6), we increase the number of clusters from 10 to 100. These changes have minimal impact on the decomposition: while adding more clusters slightly increases the contribution of establishments and reduces that of locations, the overall pattern remains stable.

Fourth, while the baseline model estimates all parameters jointly, we now adopt a two-step procedure. In the first step, we estimate worker and establishment × location effects using a standard AKM framework. In the second step, we decompose the establishment × location effect by leveraging establishment mobility across locations.[22] As shown in column (7), this approach produces results nearly identical to those of the baseline specification.

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 study period. As a robustness check, we re-estimate our model separately for the first and second halves of the sample. The contributions of the different components to spatial wage disparities remain almost identical in both periods (see Appendix Figure D.4).

Subset of exogenous relocations. As an additional robustness check, we re-estimate Equation (5) using a subset of establishment relocations that are plausibly driven by exogenous, non-economic factors. Specifically, we focus on moves to the top manager's province of birth. These relocations are likely motivated by personal ties to the hometown rather than economic considerations, and hence are less likely to reflect complementarities between the establishment and the location. The analysis is conducted at the province level—96 provinces in total—as birthplace information is only observed at this level. We continue to use 10 clusters of establishments. For comparison, we also re-estimate the baseline model using all province-level moves. Appendix Table D.3 presents the results. We find that (i) the province-level estimates closely match those obtained at the CZ level, and (ii) the subset of plausibly exogenous moves yields variance decompositions nearly identical to those from the full sample. In particular, the contribution of location effects to spatial wage disparities is 2.2% when using all province-level moves and 2.3% using hometown-driven relocations—versus 2.4% in the between-CZ decomposition using clustered establishments.

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

Location effects by industry. We examine whether our results vary across major industries: manufacturing, lower-skill services (e.g., retail, accommodation, services to households), and higher-skill services (e.g., consulting, marketing, accounting). Appendix Table D.2 reports the results from separate estimations by industry, where establishments are clustered. While the contribution of location effects slightly varies across industries, it remains consistently below 5%. Interestingly, sorting patterns do vary by industry. In manufacturing, we observe stronger co-location of high-ability workers and high-paying establishments, but little evidence that these establishments and workers systematically sort into high-paying locations. In contrast, in services industries, both productive workers and higher-paying establishments are more likely to locate in high-fixed-effect areas. Overall, these results show that location effects, even when allowing them to be industry-specific, account for very little of wage disparities.

6.3 Comparison with Alternative Approaches

In this section, we compare our results to alternative approaches from the literature. Table 5 presents the estimates, with column (1) reporting our baseline results.

We first replicate the seminal approach of Glaeser and Maré (2001) and Combes et al. (2008) using our main sample. These papers use a two-way fixed effects model to decompose spatial wage disparities into worker heterogeneity and other location-specific factors (including establishment composition). This method relies on worker mobility across locations for identification.[23] As shown in column (2), location-specific factors explain 14% of the variance—approximately three times higher than in our baseline model–while worker composition accounts for 48%, and sorting across locations for the remaining 40%.

However, as emphasized by Card et al. (2025), worker mobility across locations often coincides with changes in establishment *hierarchy* (i.e., it's rank in the local job ladder). This change in hierarchy introduces a downward bias in the estimated location effects.[24] To address this bias, Card et al. (2025) estimate a standard AKM model with worker and establishment fixed effects, identified through worker mobility between establishments. They then aggregate the establishment effects at the CZ level to recover the location effects. These location effects reflect the typical change in wage experienced by a worker moving from the average establishment in their origin location to the average establishment in their destination location.

Column (3) shows their estimates for the US; column (4) applies the same method to our French data. The results for France and for the US are remarkably similar despite strong differences in labor markets and geography: location effects explain close to 30%

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

[24] In particular, workers moving to higher paying locations were working on average in relatively higher paying establishments in their initial location and move to relatively lower paying establishments in their new location.

of wage disparities in both countries—twice as large as in column (2), consistent with the "hierarchy" bias. However, the interpretation of the role of location differs from that in our main results, as it is the combination of the pure location effect, the establishment composition, and the co-location. In this approach, location and establishment effects are perfectly collinear.

As a final benchmark, we estimate an AKM model with establishment and location fixed effects, but omit 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 in location c at time t. This approach mirrors that of Glaeser and Maré (2001); Combes et al. (2008), but identifies effects using establishment rather than worker mobility. It allows us to separately estimate the contributions of establishments and locations, as well as their sorting patterns. Column (5) shows that establishments explain 43% of the variance, locations 24%, and their sorting 38%. However, this model abstracts from worker composition, which influences both establishment and location effects.

Despite differences in methods and interpretations, all approaches in columns (2)-(5) highlight the substantial role of location effects. Our approach makes progress in two directions. First, it isolates the *pure* location effect, enabling us to empirically assess the contribution of geography, infrastructure, and agglomeration to spatial disparities. Second, it allows us to study simultaneously the sorting of plants and workers over space, as well as their co-location patterns. By explicitly accounting for workers, establishments, and locations, our model reveals that the direct contribution of location effect to wage disparity is relatively small. Much of the spatial variation captured in earlier approaches reflects establishment composition and the spatial sorting of productive workers and establishments.

6.4 Revisiting the Location Size Premium

We now illustrate another application of our method by revisiting the well-established relationship between city size and wages, often referred to as the "urban wage premium". The literature has established that larger locations tend to offer higher wages, even after controlling for worker characteristics and ability (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., workers in larger cities may work in employers intrinsically more productive (irrespective of their location). The model estimated from Equation (5) allows us to assess the relationship between city size and wages while accounting for the spatial sorting of both workers and establishments.

We begin by estimating the correlation between the average local gross hourly wage and local population density in our sample. We find an elasticity equal to 0.058 (standard error: 0.0024—see Appendix Table D.4). A doubling of city size implies wages higher by 5.8%. We then decompose this elasticity between the respective elasticities of the location effect, the average worker effect and the average establishment effect to local density.

The elasticity of the estimated location effect from Equation (5) is much smaller: equal to 0.007 (standard error: 0.0006). We depict the correlation on Panel (a) of Figure 9.[25] This implies that city size positively affects wages, even after controlling for the type of establishments where workers are employed. However, our estimate is two to three times smaller than in methods that do not account for establishment composition. Table D.4 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 0.031 and 0.021. Panel (b) of Figure 9 further compares our estimated elasticity to results from the literature and to results obtained using the alternative approaches discussed in the previous Section. First, elasticities from the literature tend to be much higher, ranging from 3-5%. Second, estimates using the method from Card et al. (2025) on our French data, yield an elasticity to density of 0.028, while a two-way fixed effects model similar to Glaeser and Maré (2001) yields an elasticity of 0.021. These elasticities have distinct interpretations: 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. This distinction is meaningful, as the urban wage premium has long been viewed as evidence of agglomeration effects. Our analysis suggests, however, that much of the wage advantage in larger and denser cities reflects the sorting of high-wage firms and workers. Once we account for the composition of establishments, density appears to play a more limited role in determining wages.

In addition, Appendix Table D.5 shows that our elasticity estimates stay unchanged when we instrument location size using historical population data on municipalities, following the approach of Ciccone and Hall (1996) and Combes et al. (2010). We rely on the earliest census year available, 1789.

Finally, we examine the characteristics of workers and firms *within* locations. Appendix Figure D.5 reproduces the distribution of worker and establishment fixed effects, separately for the largest cities and the rest of France. In larger cities, worker effects exhibit greater dispersion, driven by a thicker right tail. The establishment distribution features both a higher mode and greater dispersion. These patterns imply that larger cities tend to host more productive plants on average, and a more diverse workforce, particularly with a higher concentration of highly productive workers.

[25]Panel A of Appendix Table D.4 demonstrates that this estimate remains robust across different definitions of CZ size, including employment and population per kilometer square.

7 Conclusion

Similar to worker migration, firm relocation is a widespread phenomenon and a key mechanism allocating economic activity across space. In this paper, we combine data from France and the U.S. to provide a detailed characterization of firm relocation decisions and use it as a tool to decompose spatial disparities. Building on 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—rather than the sorting of workers and firms across locations? Our findings indicate that location effects account for only a small fraction of spatial wage disparities: the majority of the city-size premium is driven by the local composition of workers and firms.

These findings have several implications for both economists and policymakers. First, our results suggest that agglomeration spillovers play a relatively modest role in explaining wage differences across space, compared to the sorting of workers and firms. Identifying whether these disparities originate from workers, firms, or locations is essential for modeling cities, and our estimates provide informative moments for calibrating and disciplining spatial equilibrium models. Second, the prominent role of establishment composition in driving spatial wage disparities aligns with the view of local policymakers, who often regard the location decisions of high-paying firms as key determinants of local earnings. Third, observed relocation patterns underscore the importance of local human capital in shaping firms' location choices. Finally, from a national policy perspective, our findings raise concerns about inter-locality competition, particularly the use of potentially counterproductive "beggar-thy-neighbor" policies.

This paper opens several avenues 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, local geography or specific infrastructures.

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.

Baum-Snow, Nathaniel and Ronni Pavan, "Understanding the City Size Wage Gap," *The Review of Economic Studies*, 2012, *79* (1), 88–127.

Behrens, Kristian, Gilles Duranton, and Frédéric Robert-Nicoud, "Productive Cities: Sorting, Selection, and Agglomeration," *Journal of Political Economy*, 2014, *122* (3), 507–553.

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.

Bilal, Adrien, "The geography of unemployment," *The Quarterly Journal of Economics*, 2023, *138* (3), 1507–1576.

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.

Ciccone, Antonio and Robert E. Hall, "Productivity and the Density of Economic Activity," *The American Economic Review*, 1996, *86* (1), 54–70.

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.

— , — , — , **Diego Puga, and Sébastien Roux**, "The productivity advantages of large cities: Distinguishing agglomeration from firm selection," *Econometrica*, 2012, *80* (6), 2543–2594.

Diamond, Rebecca and Cecile Gaubert, "Spatial sorting and inequality," *Annual Review of Economics*, 2022, *14* (1), 795–819.

— **and Juan Carlos Suárez Serrato**, "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980–2023," 2025.

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.

Enamorado, Ted, Benjamin Fifield, and Kosuke Imai, "Using a Probabilistic Model to Assist Merging of Large-Scale Administrative Records," *American Political Science Review*, 2019, *113* (2), 353–371.

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.

Kleinman, Benny, "Wage inequality and the spatial expansion of firms," Technical Report, Tech. rep., mimeo 2024.

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

Kramarz, Francis, Elio Nimier-David, and Thomas Delemotte, "Inequality and earnings dynamics in France: National policies and local consequences," *Quantitative Economics*, 2022, *13* (4), 1527–1591.

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," 2025.

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

Mann, Lukas, "Spatial sorting and the rise of geographic inequality," Technical Report, Technical report 2023.

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

_ , *The new geography of jobs*, Houghton Mifflin Harcourt, 2012.

_ **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.

Slattery, Cailin and Owen Zidar, "Evaluating state and local business incentives," *Journal of Economic Perspectives*, 2020, *34* (2), 90–118.

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.

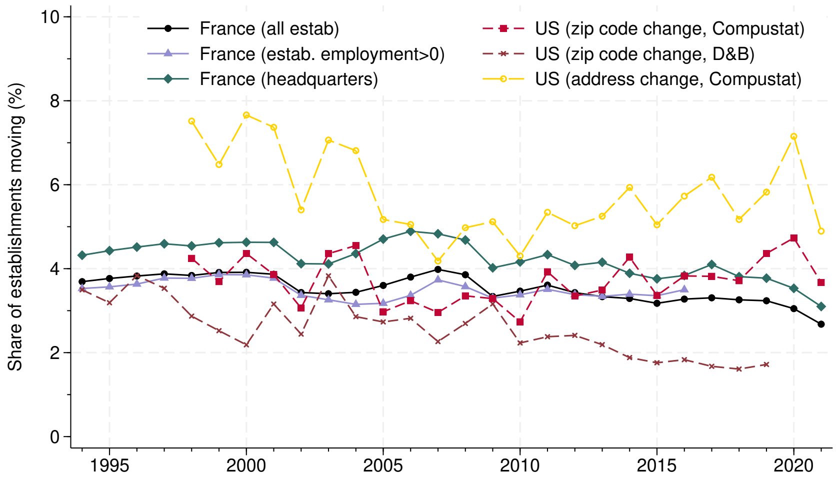
Verdugo, Gregory and Malak Kandoussi, "Will You Follow me to the Suburbs?," 2024.

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

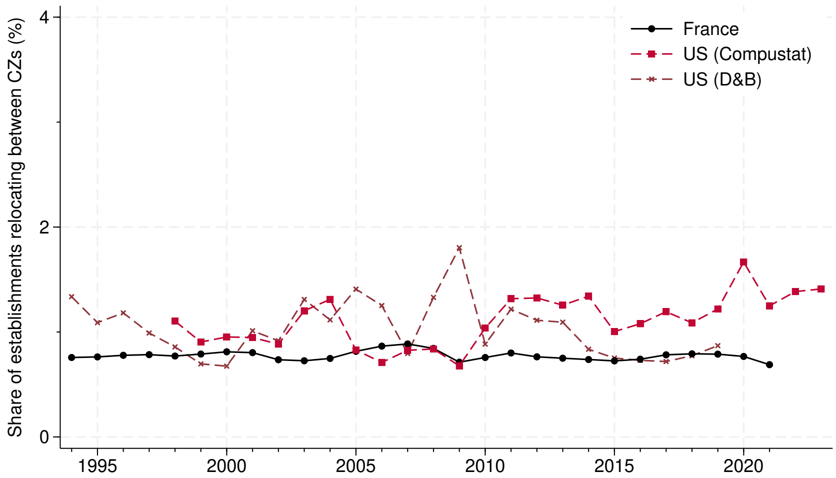
Figures

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

(a) All Moves



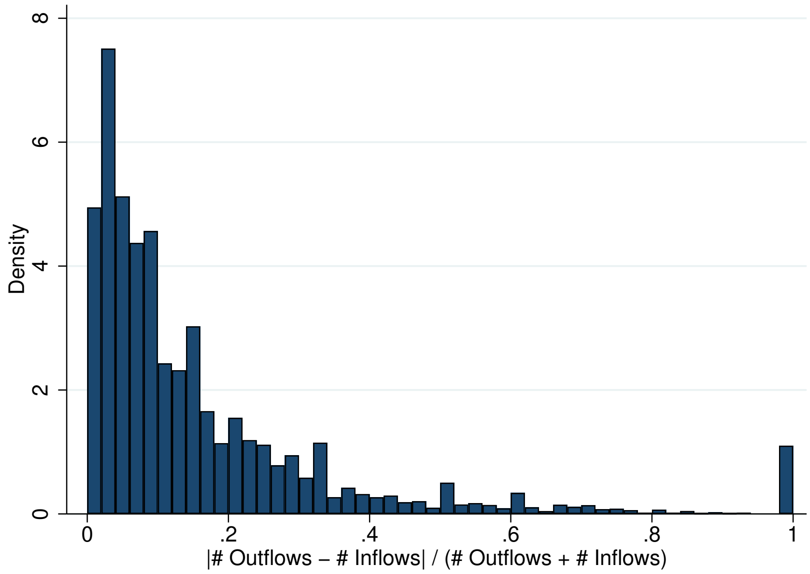
(b) Moves Between Commuting Zones



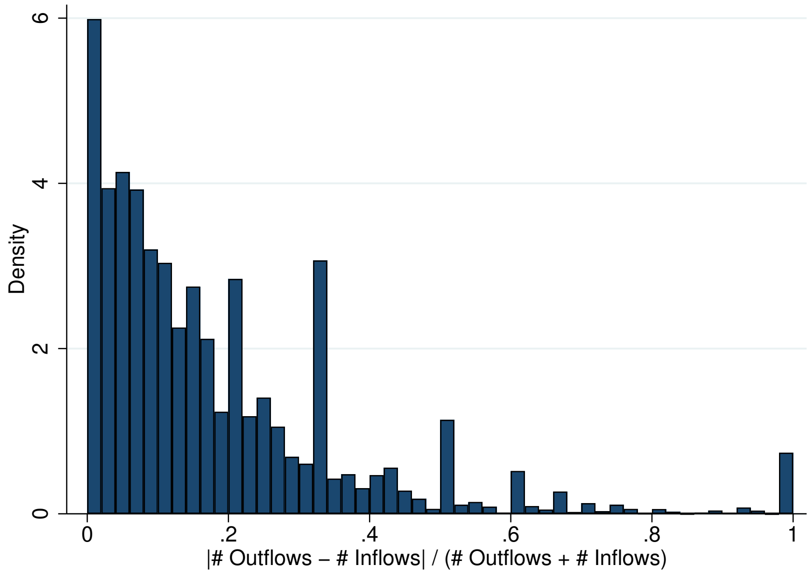
Notes: Figure 1 plots, for each year between 1994 and 2021, the share of establishments relocating as a fraction of the total number of establishments for France (solid lines) and for the United States (dashed lines). The solid dark line with circular markers includes all types of establishments. The 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 moves 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



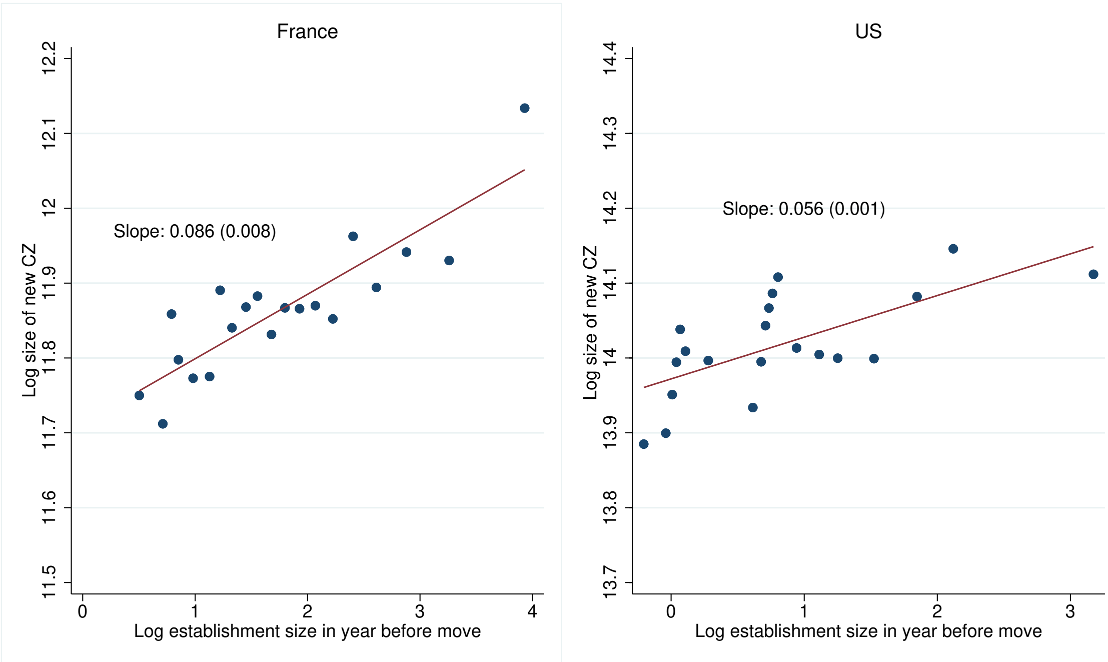
(b) US



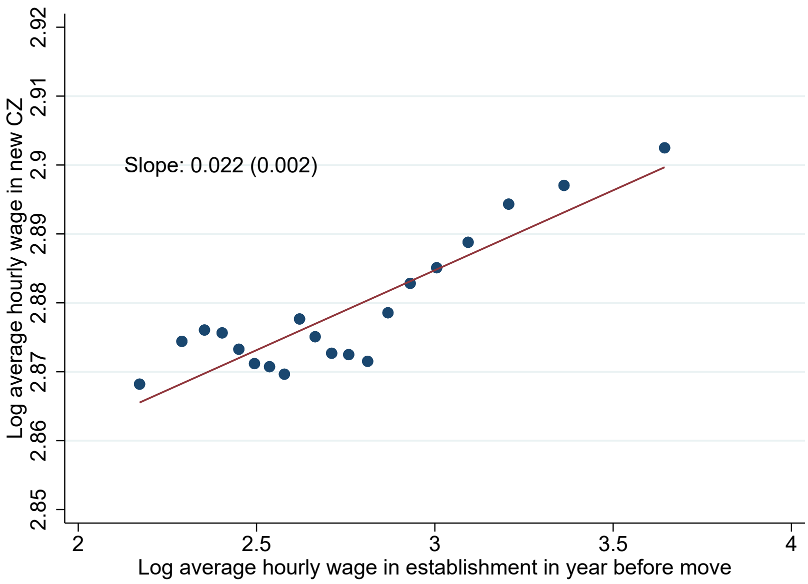
Notes: Figure 2 plots the distribution of the ratio between net and gross flows, for every pair of provinces (France) or commuting zones (US). This index takes values between 0 and 1. A value of 1 implies unidirectional flows: establishments are either going from location A to B, or from location B to 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 sub-periods 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 and Dun and Bradstreet. 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)



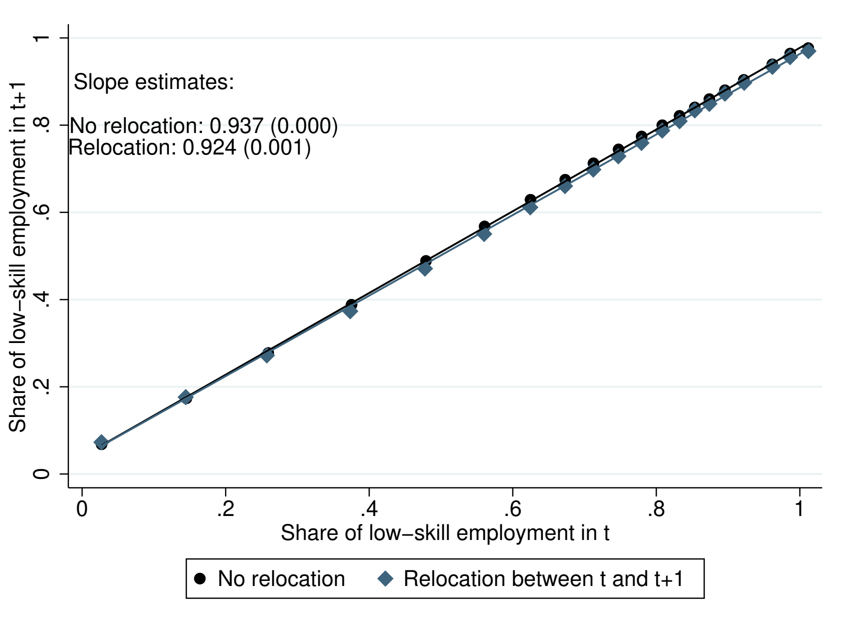
(b) Establishment Wage and Destination Wage (France)



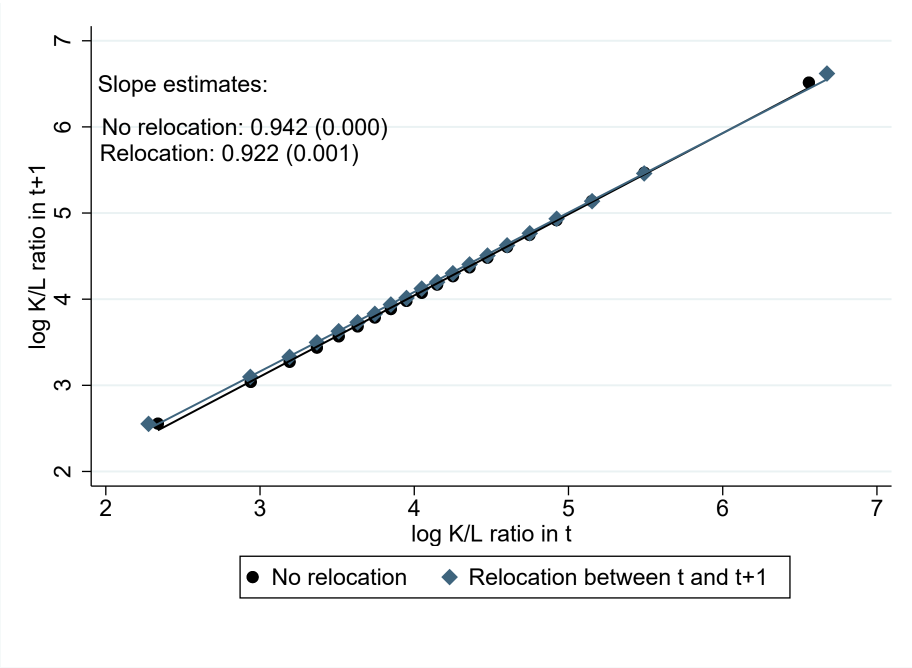
Notes: Figure 3 shows the correlation between characteristics of relocating establishments and characteristics of destination commuting zones. The fitted lines and slopes report estimates of γ_1 in Equation (3), for $x = \{\text{log size, log average hourly wage}\}$. The regression controls for establishment 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, relocation register, and Dun and Bradsreet. Go back to main text

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

(a) Skill Composition



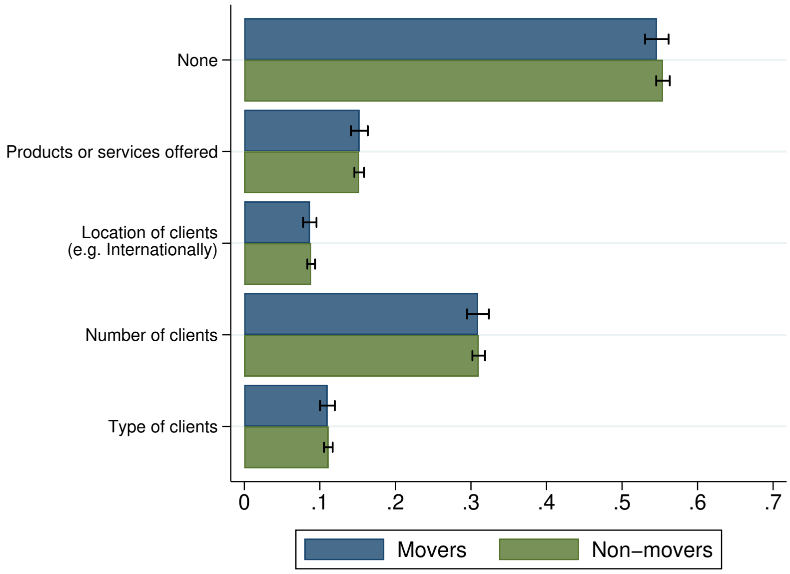
(b) Capital-to-Labor Ratio



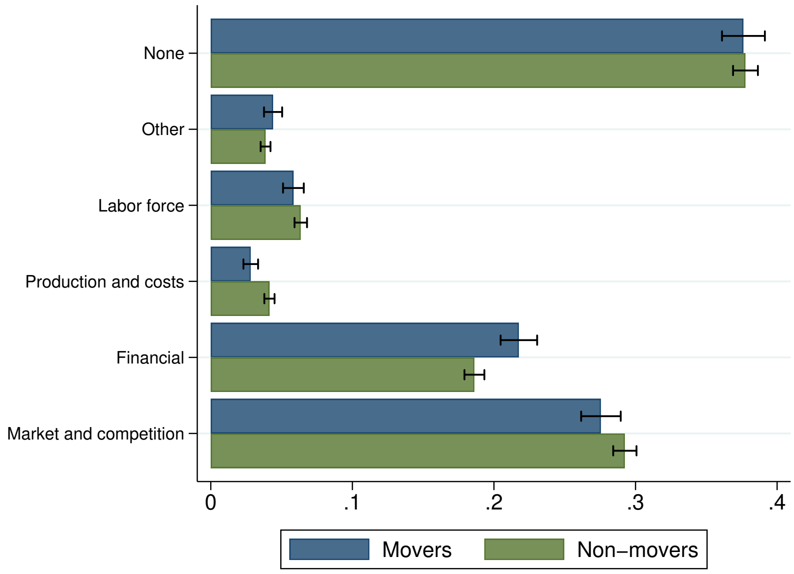
Notes: Figure 4 shows the correlation between an establishment characteristic in a given year (x-axis) and the same characteristic the subsequent year (y-axis), for two continuous variables. In panel (a), we consider the share of low-skill workers employed in the establishment, defined as blue-collars and low-skill white collars. In panel (b), we focus on 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: linked employer-employee data and FICUS-FARE. Go back to main text

Figure 5: Changes Faced According to Entrepreneurs

(a) Major Changes Experienced Since Creation



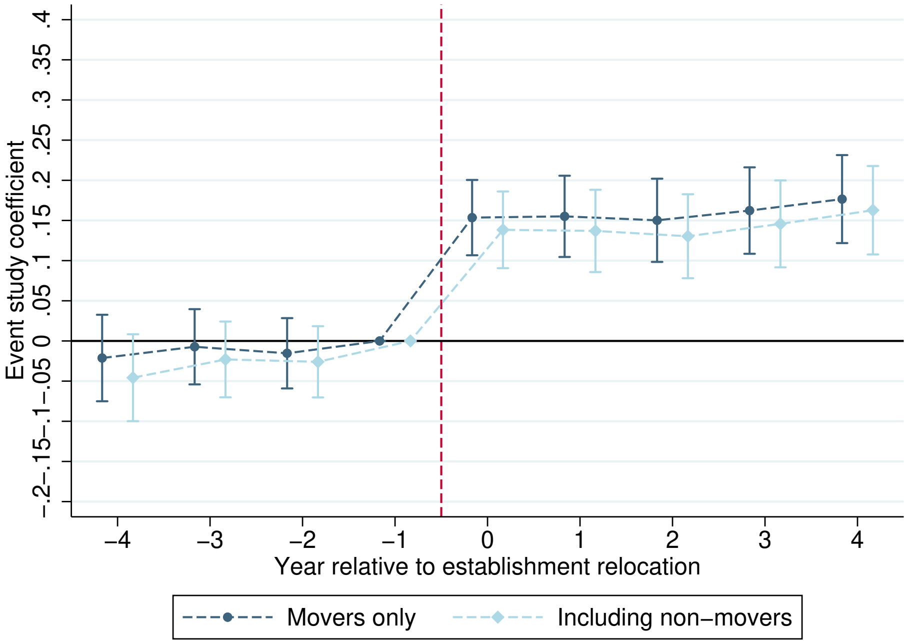
(b) Problems in the First Two Years



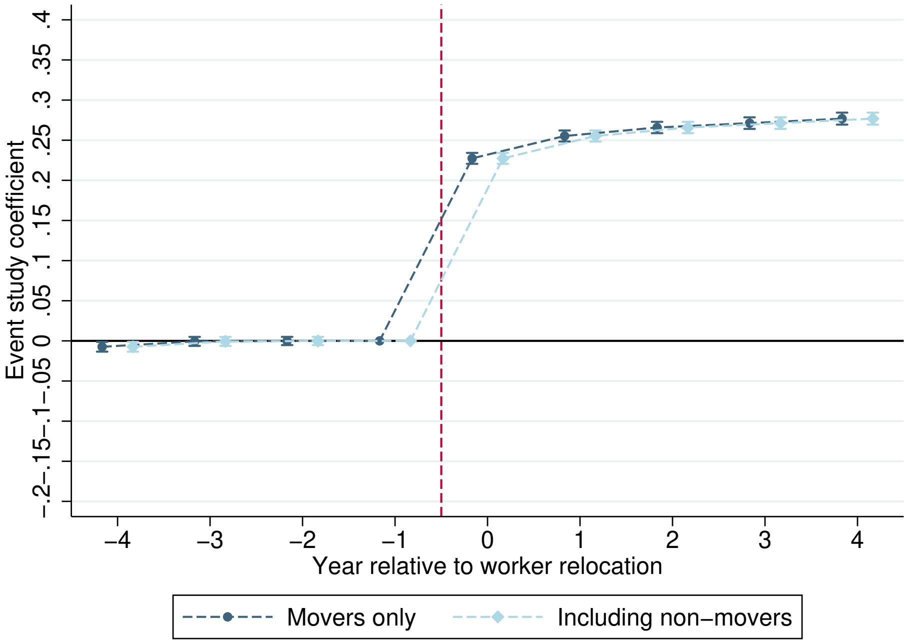
Notes: Figure 5 compares answers from entrepreneurs who relocated at least one establishment during their first five years of economic activity, with answers of entrepreneurs who did not. Panel (a) shows answers to the question "**Did you experience any major change since the creation of the firm?**". Entrepreneurs can report experiencing multiple changes. Panel (b) reports answers to the question **"Since the creation of your company, what has been THE MAIN OBSTACLE to its development?"**. "Labor force" includes recruitment, training, and conflicts. "Other" obstacles include legal challenges and natural disasters. 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

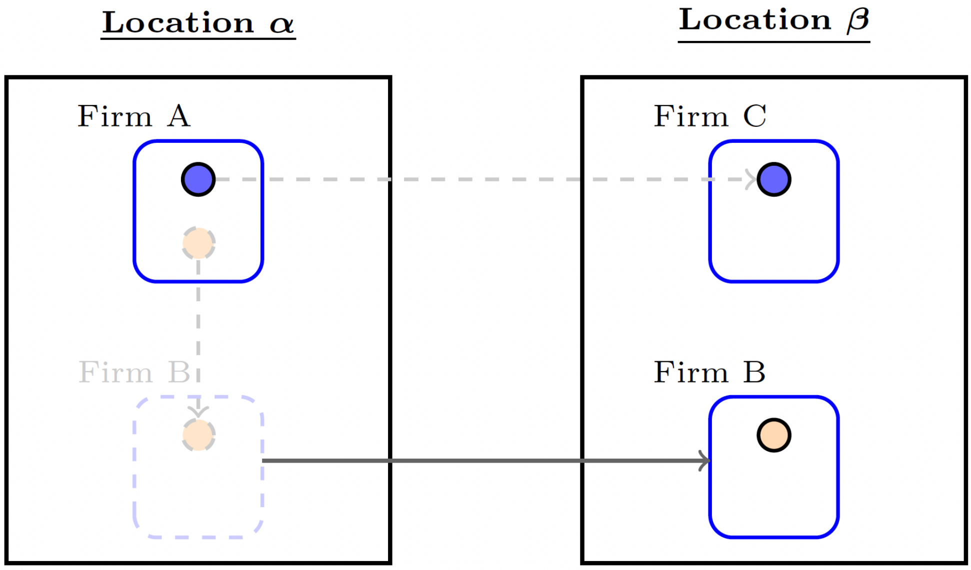


(b) Worker Relocating (and Changing Establishment)



Notes: Figure 6 plots estimates of θ in Equation (4) for two types of changes. Panel (a) considers establishments relocating to a different commuting zone. Panels (b) focuses on workers changing establishment between commuting zones. Year 0 is the first year of the change (relocation or change in employer). Each panel reports estimates on two different samples: the one of establishments or workers ever making a change, and for the whole sample (including non movers). Standard errors are clustered at the establishment (respectively worker) level. Data: linked employer-employee data and relocation register. Go back to main text

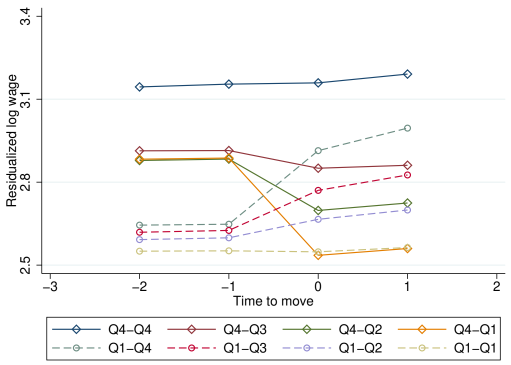
Figure 7: Identification with the Double-Mover Design



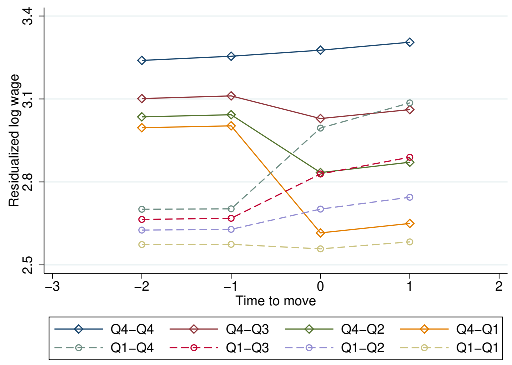
Notes: Figure 7 illustrates the double-mover design used to estimate Equation (5). Rectangles with straight angles represent locations α and β. Rounded rectangles correspond to establishments A, B and C. Finally, the arrows represent the mobility of agents (workers and establishments) between establishments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

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

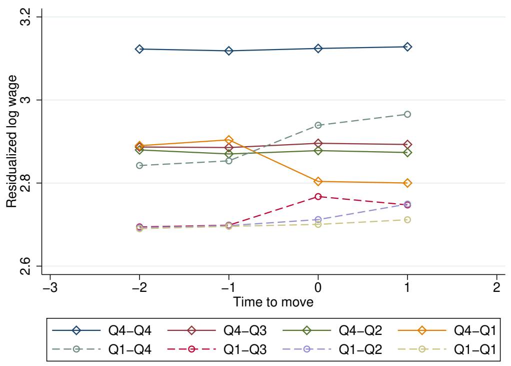
(a) Workers Changing Establishment Within Location



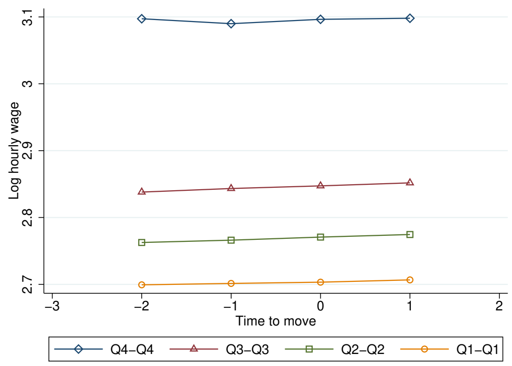
(b) Workers Changing Establishment and Location



(c) Establishments Changing Location



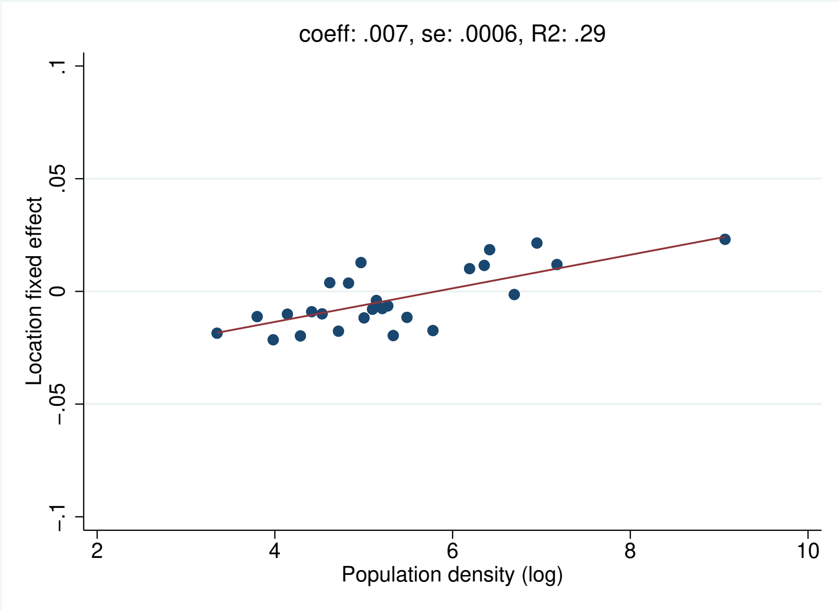
(d) Placebo: Establishments Moving Within Location



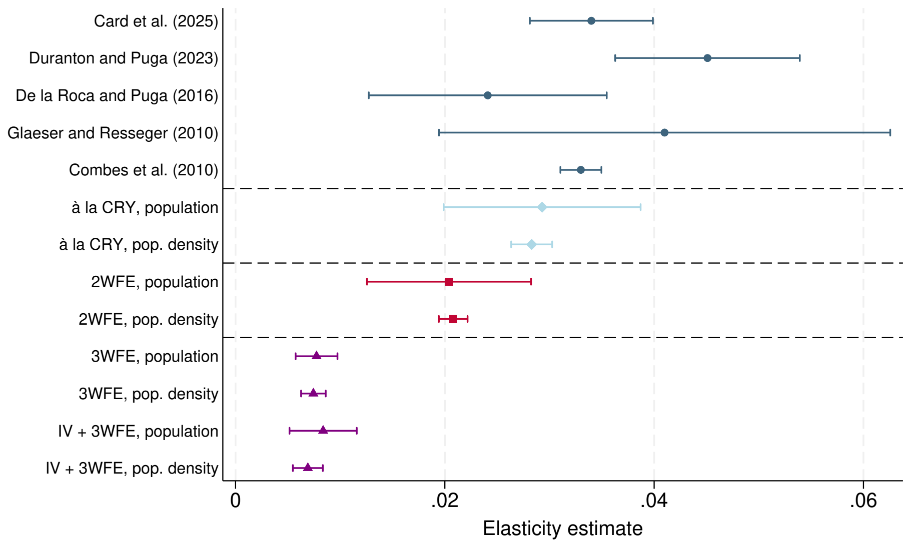
Notes: Figure 8 plots the evolution of the average (log) hourly wages earned by workers and paid by establishments around a relocation, separately for different quartiles of origin and destination fixed effects. Panel (a) focuses on workers moving between high and low paying establishments, within the 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: linked employer-employee data and relocation register. 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 our results to alternative approaches, for several measures of 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 with worker and location fixed effects. The bottom set of estimates are obtained from our main specification in Equation (5), for different measures of size. The last two estimates report a 2SLS version where we instrument city size by historical population from the 1789 census (see also Table D.5). Data: linked employer-employee data, relocation register, census data. Go back to main text

Tables

Table 1: Establishments Maintain 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	0.764

Notes: Table 1 shows the share of establishments 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 establishments relocating keep the same two-digit industry code the year after their relocation, and 99.0% of establishments that do not relocate keep the same code between two consecutive years. Data: linked employer-employee data, relocation and firm registers. 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	0.36	0.32	0.41	0.31
Worker Age	40.89	38.48	37.31	39.09	39.32
Gross hourly Wage					
Mean	18.24	18.84	19.20	16.32	19.13
S.D.	10.10	10.38	10.70	6.17	8.08
Establishment Size					
Mean	351.67	288.47	248.41	14.54	14.32
S.D.	1273.99	1101.49	913.46	66.79	44.28
Occupations					
Blue Collar	0.40	0.34	0.32	0.32	0.35
Clerks	0.27	0.27	0.25	0.37	0.23
Supervisors	0.18	0.20	0.22	0.18	0.21
Managers	0.13	0.18	0.20	0.11	0.18
Executives	0.01	0.01	0.01	0.02	0.03
Industries					
Manufacturing	0.32	0.23	0.21	0.14	0.14
Retail	0.22	0.24	0.24	0.33	0.25
Services	0.35	0.43	0.45	0.40	0.40
Construction	0.11	0.09	0.10	0.14	0.20
Number of Units					
Total Units	22,992,426	11,707,015	5,607,244	1,905,832	38,967
Unit-Year Obs.	77,172,570	96,248,977	45,844,597	11,596,827	333,383
% Changing Firms					
1 Time	.	0.46	0.30	.	.
2 Times	.	0.26	0.28	.	.
3+ Times	.	0.27	0.40	.	.
% Changing Location					
1 Time	.	0.27	0.57	.	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 over the period 2002–2016. Column (1) includes workers who never change location nor employer. Column (2) corresponds to workers who change employers and column (3) shows only shows those switching commuting zones. Columns (4) displays establishments that remain in the same location while column (5) focuses on establishments that relocate. 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). Data: linked employer-employee data and relocation register. 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	0.113
Number of worker-year observations	161,287,397	161,408,784	161,069,306
Number of workers	22,666,846	22,686,678	22,634,794
Number of plants	1,765,604	1,867,873	1,730,061
Number of locations	304	304	304
Std. dev. of worker effect	0.062	0.063	0.063
Std. dev. of plant effects	0.047	0.041	0.047
Std. dev. of location effects	0.023	0.018	0.022
Std. dev. of Xb	0.003	0.002	0.003
Correlation of worker/plant effects	0.733	0.972	0.723
Correlation of plant/location effects	0.286	0.638	0.297
Correlation of worker/location effects	0.507	0.454	0.493
Share of variance of log wages due to:			
Person effects	0.302	0.313	0.306
Plant effects	0.174	0.133	0.169
Location effects	0.042	0.024	0.039
Covariance of worker and plant effects	0.336	0.396	0.338
Covariance of plant and location effects	0.049	0.072	0.055
Covariance of worker and location effects	0.114	0.079	0.118
Xb and other covariances	-0.017	-0.017	-0.017
Residual	0.000	0.000	0.000
RMSE of model	0.16	0.16	.
Adjusted R2	0.86	0.86	.

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. Data: linked employer-employee data and relocation register. Go back to main text

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

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Baseline	Controlling for Plant Age	Controlling for Plant Age (Clustered)	Without Stayers (Clustered)	Alternative Clustering	Clustered (100)	Two Steps
Standard deviation of log wages	0.113	0.116	0.116	0.105	0.113	0.113	0.113
Number of worker-year observations	161,287,397	152,483,623	152,604,961	146,175,449	161,367,746	161,408,784	161,264,138
Number of workers	22,666,846	21,838,330	21,858,355	21,564,702	22,681,298	22,686,678	22,663,762
Number of plants	1,765,604	1,746,225	1,848,261	1,851,499	1,831,864	1,867,873	1,763,568
Number of locations	304	304	304	304	304	304	304
Std. dev. of worker effect	0.062	0.063	0.065	0.058	0.063	0.062	0.062
Std. dev. of plant effects	0.047	0.047	0.043	0.039	0.042	0.045	0.046
Std. dev. of location effects	0.023	0.024	0.018	0.017	0.017	0.016	0.025
Std. dev. of Xb	0.003	0.004	0.003	0.003	0.002	0.002	0.003
Correlation of worker/plant effects	0.733	0.736	0.975	0.967	0.971	0.969	0.712
Correlation of plant/location effects	0.286	0.277	0.658	0.632	0.618	0.566	0.307
Correlation of worker/location effects	0.507	0.511	0.487	0.431	0.429	0.371	0.561
Share of variance of log wages due to:							
Person effects	0.302	0.297	0.313	0.308	0.313	0.302	0.302
Plant effects	0.174	0.166	0.134	0.136	0.136	0.154	0.162
Location effects	0.042	0.042	0.023	0.027	0.024	0.020	0.048
Covariance of worker and plant effects	0.336	0.327	0.399	0.396	0.401	0.418	0.315
Covariance of plant and location effects	0.049	0.046	0.073	0.077	0.070	0.063	0.054
Covariance of worker and location effects	0.114	0.114	0.083	0.079	0.074	0.058	0.135
Xb and other covariances	–0.017	0.007	–0.025	–0.024	-0.017	–0.017	–0.017
Residual	0.000	0.000	0.000	0.000	0.000	0.000	0.000
RMSE of model	0.16	0.16	0.16	0.16	0.16	0.16	.
Adjusted R2	0.88	0.86	0.85	0.86	0.85	0.85	.

Notes: Table 4 compares the results of our main specification, 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). Column (4) re-estimates the main model, with 10 clusters, 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 (5) and (6) investigate alternative clustering approaches. First, column (5) uses mean and standard deviation of wages as clustering variables, instead of percentiles of the wage distribution. Second, column (6), replicates the standard clustering approach with 100 clusters instead of 10. Finally, column (7) provides a decomposition of between CZ variance of wages using a two-step estimation approach (see footnote 22). Data: linked employer-employee data and relocation register. Go back to main text

Table 5: Comparison with Alternative Approaches

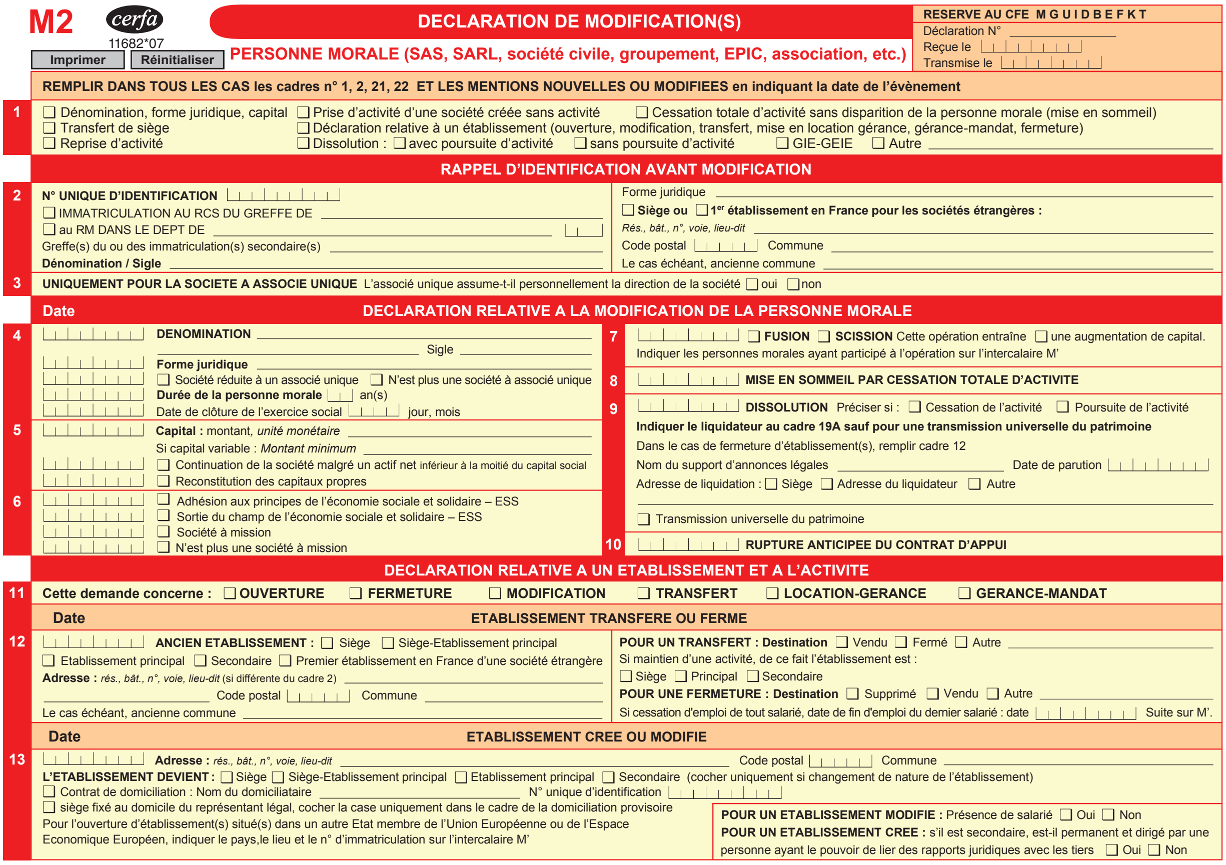
	(1)	(2)	(3)	(4)	(5)
	Baseline	TWFE Worker mobility	CRY (2025)	à la CRY (2025)	TWFE Plant mobility
	FRA	FRA	USA	FRA	FRA
Standard deviation of log wages	0.113	0.113	0.147	0.113	0.095
Number of observations	161,287,397	161,408,784	2,523M Q	161,264,166	16,583,493
Number of workers	22,666,846	22,686,678	112M	22,663,762	.
Number of plants	1,765,604	.	.	2,023,329	2,053,805
Number of locations	304	304	741	304	304
Std. dev. of worker effect	0.062	0.078	0.081	0.062	.
Std. dev. of plant effects	0.047	.	.	.	0.062
Std. dev. of location effects	0.023	0.043	0.079	0.058	0.046
Std. dev. of Xb	0.003	0.002	0.006	0.003	0.008
Correlation of worker/plant effects	0.733	.	.	.	.
Correlation of plant/location effects	0.286	.	.	.	0.604
Correlation of worker/location effects	0.507	0.753	0.642	0.796	.
Share of variance of log wages due to:					
Person effects	0.302	0.479	0.303	0.302	.
Plant effects	0.174	.	.	.	0.429
Location effects	0.042	0.144	0.293	0.265	0.237
Covariance of worker and plant effects	0.336	.	.	.	.
Covariance of plant and location effects	0.049	.	.	.	0.385
Covariance of worker and location effects	0.114	0.395	0.382	0.450	.
Xb and other covariances	-0.017	-0.018	0.022	-0.017	-0.050
Residual	0.000	0.000	0.000	0.000	0.000
Adjusted R2	0.86	0.84	.	0.86	0.75
RMSE of model	0.16	0.17	.	0.16	0.18

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 denote as the 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 relocation register. 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

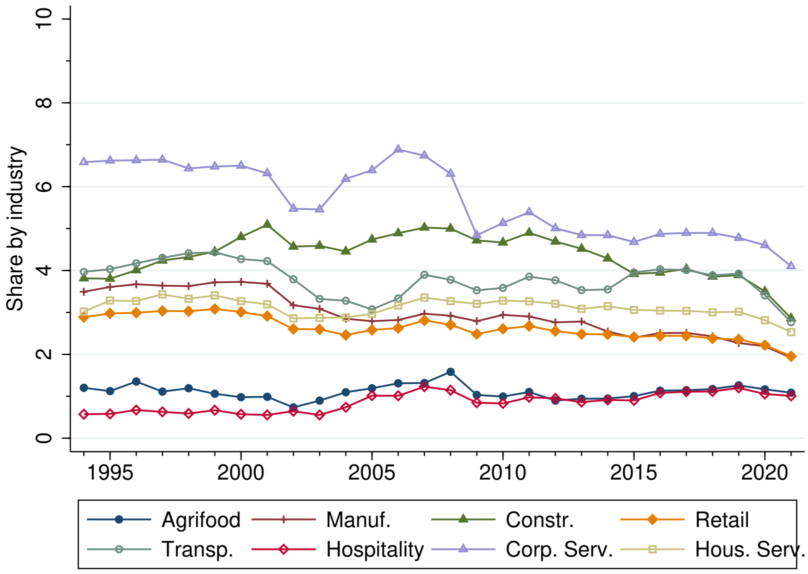
Notes: Figure A.1 presents the one-page form "M2", which entrepreneurs are required to fill when relocating their establishment. The form mandates the completion of three key sections: boxes (1), (2), and (13). Box (1) specifies the type of change being reported, where entrepreneurs must check the option "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 to 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 (e.g., the new lease). Go back to main text

B Additional Results on Establishment Relocation

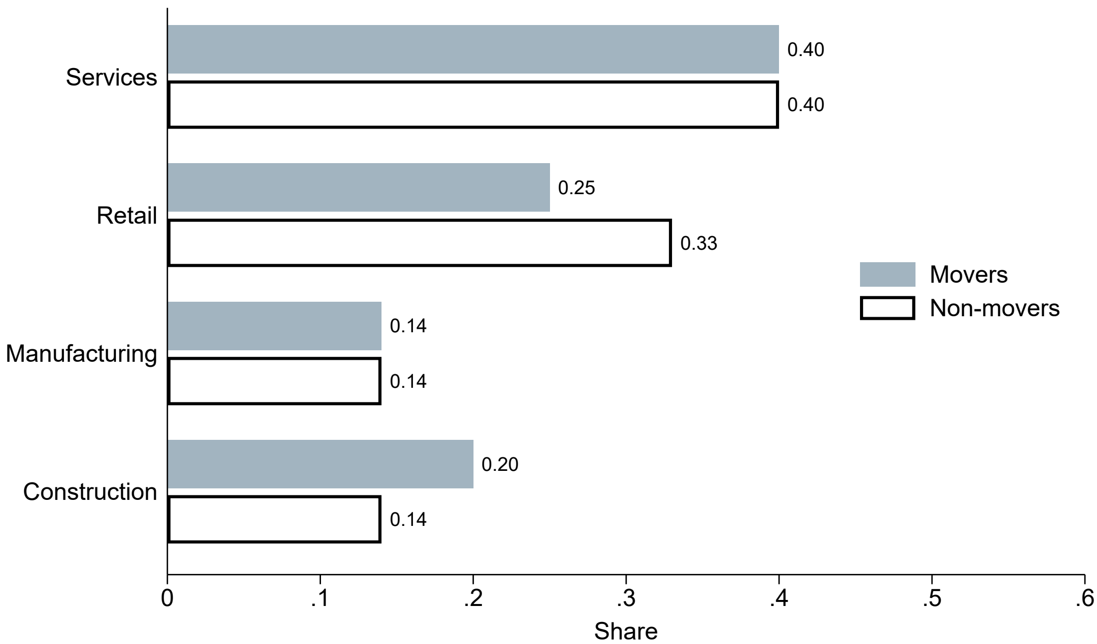
B.1 Prevalence of Establishment Relocation

Figure B.1: Establishment Relocation and Industry Composition

(a) All Moves, by Industry



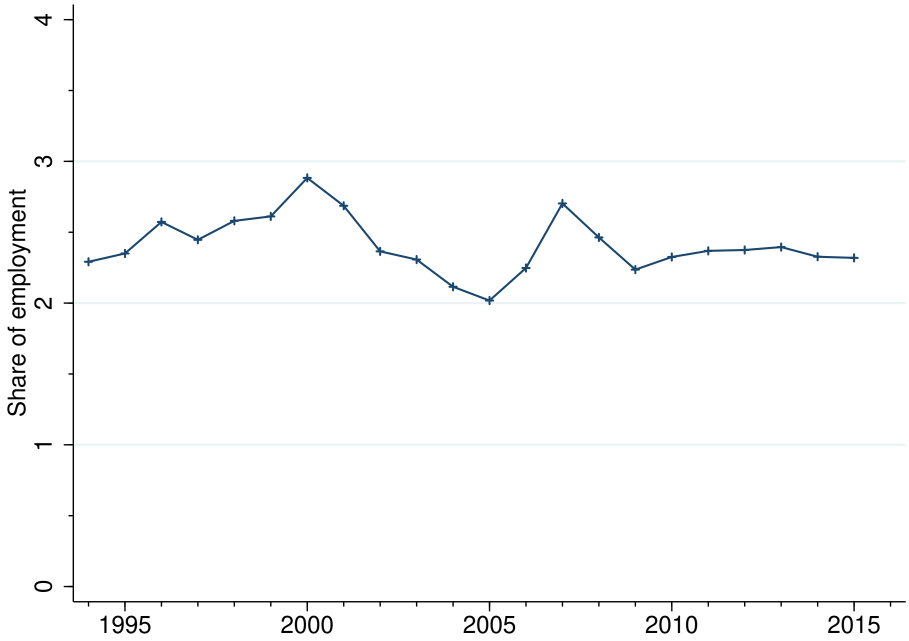
(b) Sectoral Distributions for Movers Between CZs and Non-movers



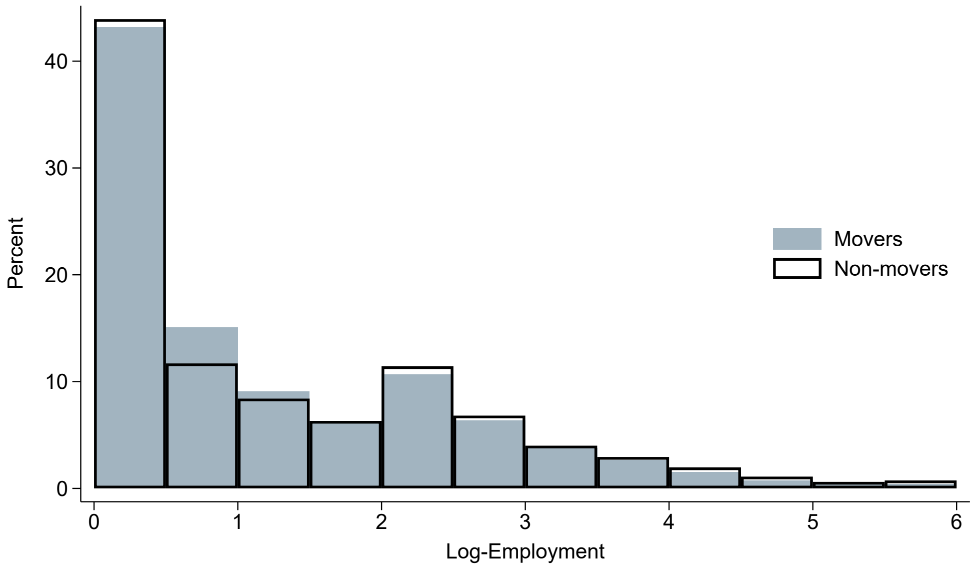
Notes: Figure B.1 plots the share of establishments moving every year by major industry. Panel (a) considers all moves, while Panel (b) focuses on relocations between commuting zones. Data: linked employer-employee data and relocation register. Go back to main text

Figure B.2: Size of Establishments Relocating in France

(a) Establishment Relocation Weighted by Employment Share



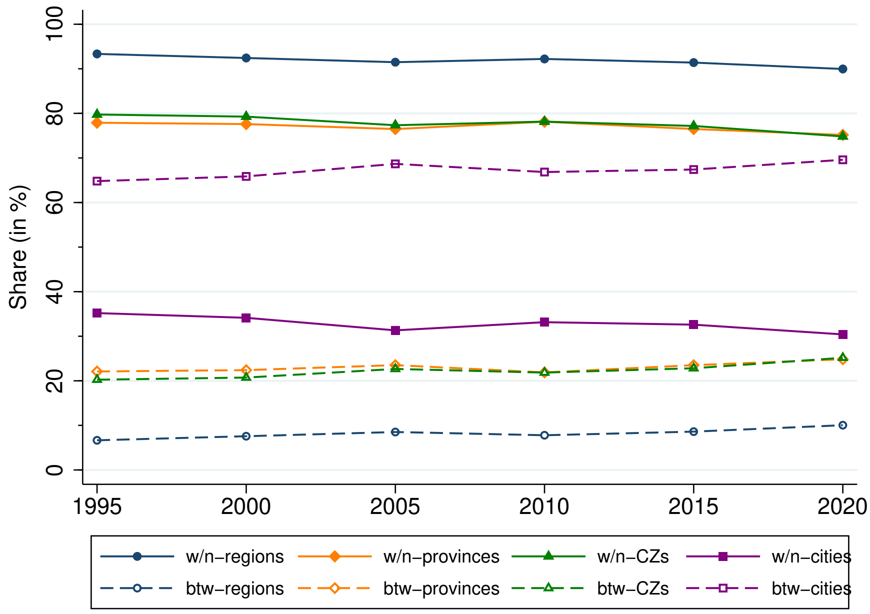
(b) Size Distribution



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

B.2 Distances of Relocations

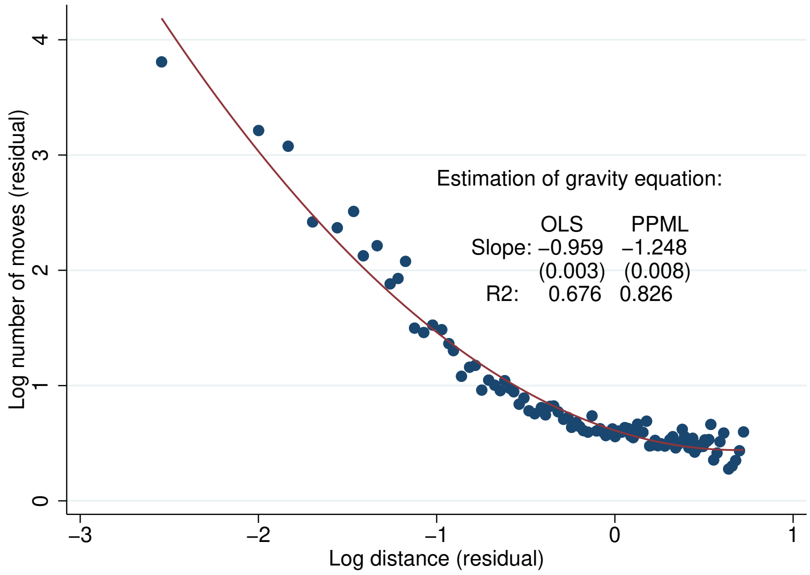
Figure B.3: Establishment Relocation by Geographical Unit



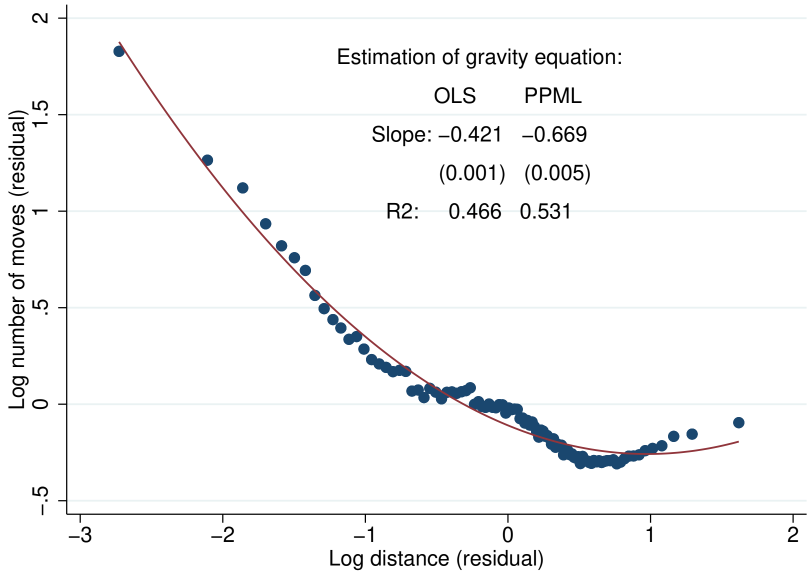
Notes: Figure B.3 reports the share of moves that took place *within* vs. *between* locations, for four types of administrative boundaries. Solid lines indicate relocations within a location, while dashed lines represent mobility across locations. The administrative units considered include 22 regions, 96 provinces (départements), 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



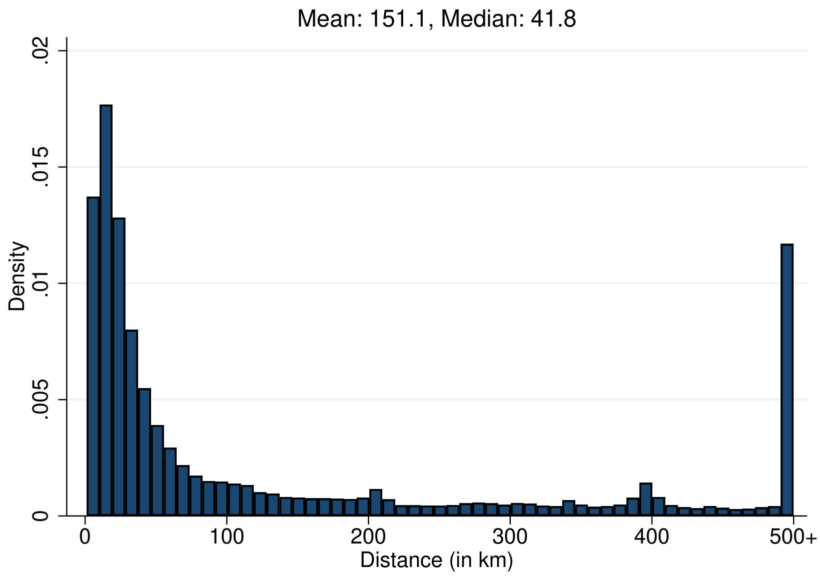
Notes: Figure B.4 is a binscatter that plots the relationship between the (log) number of establishments relocating between every pairs of location and the (log) distance between these two locations. Both quantities are first residualized for origin by year and destination by year fixed effects to account for systematic differences in market access. The table reports the coefficients of a log-log regression model (column 1) and a pseudo-poison model (column 2). The corresponding specification is: $log(\#moves_{ijt}) = log(distance_{ij}) + \alpha_{it} + \gamma_{jt} + \epsilon_{ijt}$. Panel (a) is for France and focuses on moves between every two pairs of provinces (there are 96 of them). Panel (b) is for the US and considers pairs of commuting zones. Data: relocation register and Dun and Bradstreet. 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. Data: linked employer-employee data and relocation register. 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 establishments moving between different commuting zones. Half of relocations involve a move further than 41.8 kilometers. Relocations exceeding 500 kilometers are grouped into a single category. Data: relocation register. Go back to main text

B.2.1 Post-relocation Evolution of Activity

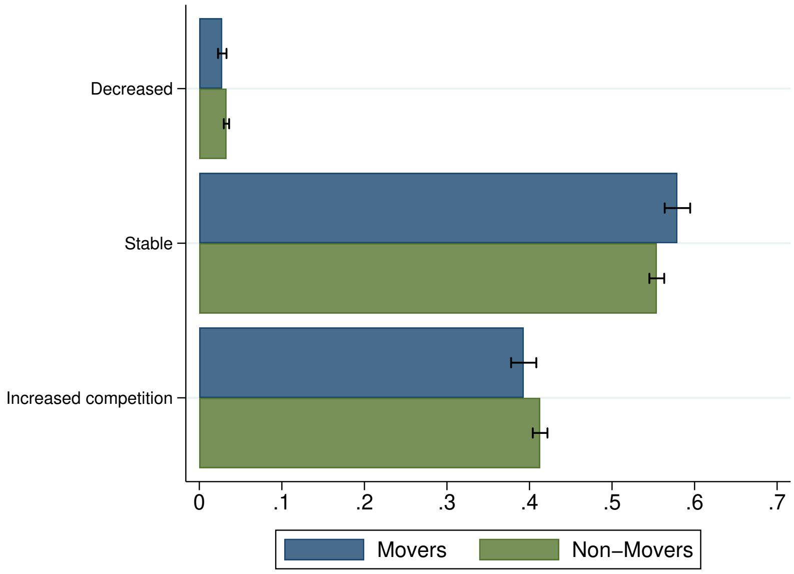
Table B.2: Difference-in-Differences: Establishments Preserving their Activity and Organization

	Same Industry	Preserve HQ Status	Same Legal Category	Same Top Manager
Mover x year post move	-0.044	-0.007	0.020	-0.084
	(0.000)	(0.000)	(0.000)	(0.001)
Mover	-0.002	-0.003	0.025	-0.005
	(0.000)	(0.000)	(0.000)	(0.001)
Year post move	-0.003	0.000	-0.004	-0.014
	(0.000)	(0.000)	(0.000)	(0.001)
Year F.E.	Yes	Yes	Yes	Yes

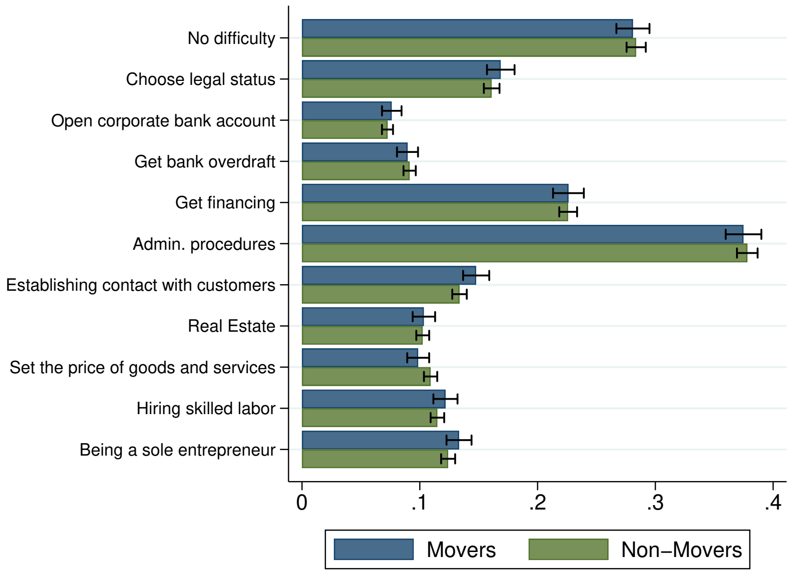
Notes: Table B.2 reports the estimates of a difference-in-difference specification in which the first row indicates the outcome variable. Each outcome is a dummy variable equal to one if the establishment keeps a given characteristic identical from one year to the other. The specification is the following: $y_{it} = \delta_t + \beta_1 \text{Mover}_i \text{ x Year Post Move}_t + \beta_2 \text{Mover}_i + \beta_3 \text{Year Post Move}_t + u_{it}$. For an establishment relocating, the variable Year Post Move_t equals one only the first year post-relocation. δ_t accounts for calendar year fixed effects. 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. Data: linked employer-employee data and relocation register. Go back to main text

Figure B.6: More Questions: Changes Faced by Entrepreneurs Relocating their Establishments

(a) Changes in Local Competition



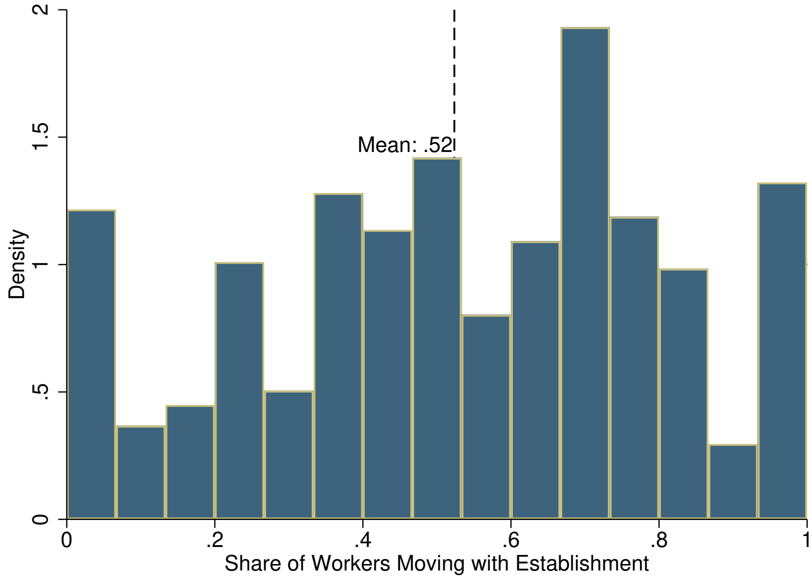
(b) Problems During Creation



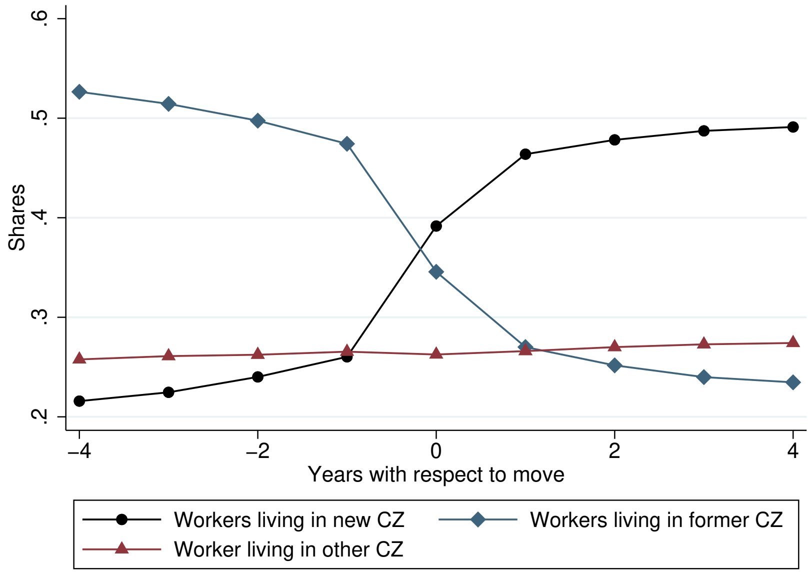
Notes: Figure B.6 compares answers from entrepreneurs who relocated at least one establishment during their firm first five years of economic activity, with answers of entrepreneurs who did not. There are two questions of interest (on top of the two questions already depicted in Figure 5). Panel (a): **"Since the creation of your firm, would you say that the level of direct competition: (i) increased, (ii) is stable, (iii) decreased?"**. Panel (b): "**When you set up your company, what were the MAIN CHALLENGES you encountered?**". 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 Place of Residence



Notes: Panel (a) of Figure B.7 displays the distribution of the share of workers staying with their establishment following a move. It is the ratio between the number of workers observed 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 other commuting zones. Averages are computed for each year before and after relocation. Year 0 corresponds to the year of the move. Data: linked employer-employee data and relocation register. Go back to main text

Figure B.8: Wage Changes for Workers Changing Establishment

(a) Worker Changing Establishment, Any Location

(b) Worker Changing Establishment, Within Location

Notes: Figure B.8 plots estimates of θ in Equation (4) for two types of changes. Panel (a) considers workers changing establishment, either within or between locations. Panel (b) reports workers changing establishment within a commuting zone. Year 0 is the first year of the employer change. Each panel reports estimates on two different samples: the one of workers ever making a job change, and the whole sample. Standard errors are clustered at the worker level. Data: linked employer-employee data 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:

$$\max_{n\in\{1\ldots N\}} \pi_n(a) - \kappa_{on}^f + \epsilon_n^f.$$

Since ϵ_n^f 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

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

Taking logs, we obtain a gravity structure:

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

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

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

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

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

C.2 Worker Location Decisions and Worker Flows

Recall the worker's location decision problem

$$\max_{n\in\{1\ldots N\}} w_n(z) - \kappa_{on}^w + \epsilon_n^w.$$

Since ϵ_n^w 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

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

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

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

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

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

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

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

and

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

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: Average Number of Unique Worker and Establishment by Commuting Zone

	Workers	Establishments
All	114,143	7,328
Movers	28,751	137

Notes: Table D.1 reports the average of the number of unique observations per CZ (Panel A) and number of movers by CZ (Panel B) in our estimation sample. Data: linked employer-employee data and relocation register. 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:

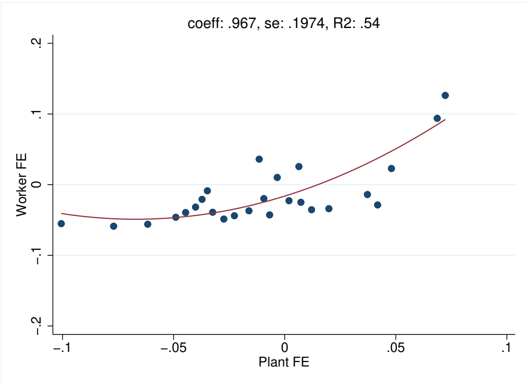
$$\begin{aligned} Var(\overline{\alpha_c}) &= Cov(\bar{\hat{\alpha}}^{A(B)}, \bar{\hat{\alpha}}^A) \\ Var(\overline{\phi_{J(c)}}) &= Cov(\bar{\hat{\phi}}^A, \bar{\hat{\phi}}^B) \\ Var(\psi_c) &= Cov(\hat{\psi}^A, \hat{\psi}^B). \end{aligned}$$

   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: $Cov(\psi_c, \overline{\phi_{J(c)}}) = Cov(\hat{\psi}^A, \bar{\hat{\phi}}^B)$.

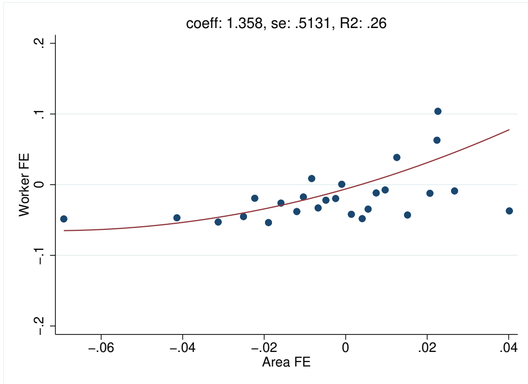
D.3 Model Estimates

Figure D.1: Correlation Between Estimated Fixed Effects

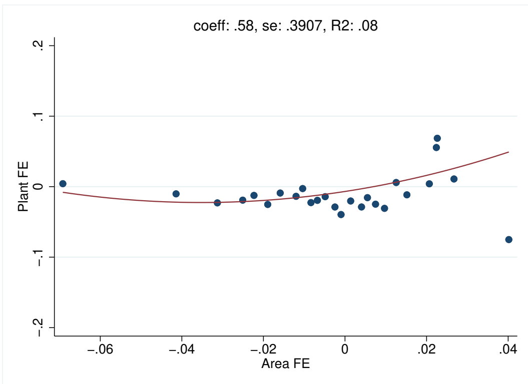
(a) Worker and Establishment



(b) Worker and Location



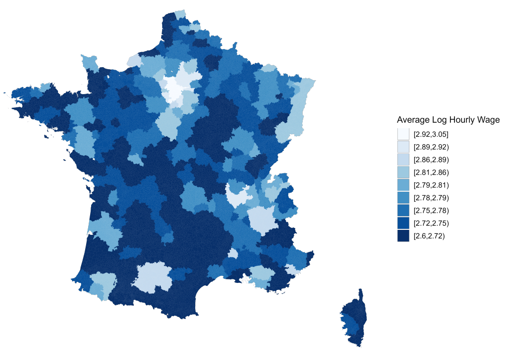
(c) Establishment and Location



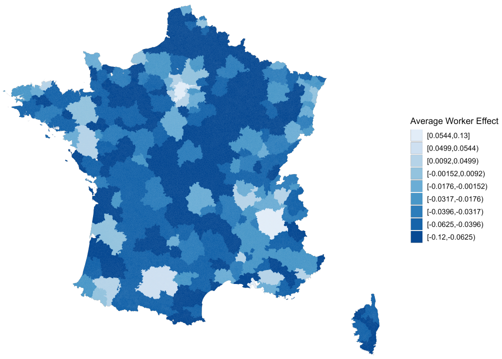
Notes: Figure D.1 shows the result of a binscatter that displays the correlation between average worker, establishment and area fixed effects, between commuting zones. The fixed effects result from the estimation of Equation (5). The coefficients, standard errors and R^2 are derived from a linear regression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

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

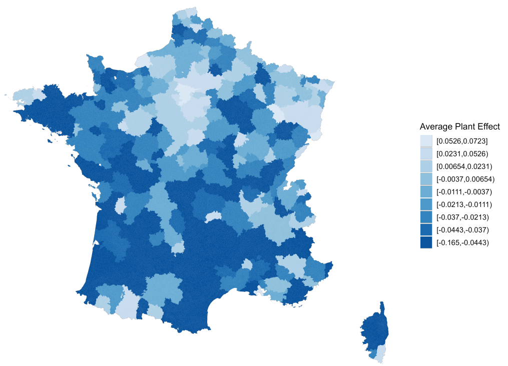
(a) Gross Hourly Wage



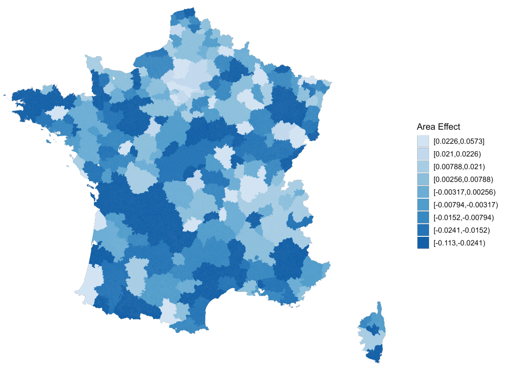
(b) Average Worker Fixed Effects



(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). Data: linked employer-employee data and relocation register. Go back to main text

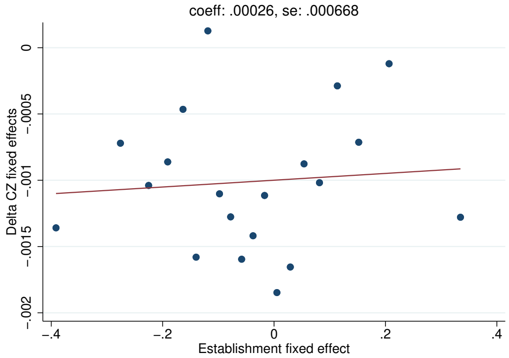
Table D.2: Three-Way Decomposition by Industry

	Baseline	Manufacturing	Higher Skill	Lower Skill
Standard deviation of log wages	0.113	0.145	0.158	0.100
Number of observations	161,287,397	44,584,953	30,622,137	67,245,702
Number of workers	22,666,846	6,141,148	5,480,448	10,807,315
Number of plants	1,765,604	226,757	364,266	1,030,670
Number of locations	304	304	304	304
Std. dev. of worker effect	0.062	0.099	0.098	0.046
Std. dev. of plant effects	0.047	0.049	0.051	0.039
Std. dev. of location effects	0.023	0.019	0.021	0.020
Std. dev. of Xb	0.003	0.003	0.006	0.002
Correlation of worker/plant effects	0.733	0.973	0.990	0.965
Correlation of plant/location effects	0.286	−0.053	0.693	0.818
Correlation of worker/location effects	0.507	−0.255	0.595	0.658
Share of variance of log wages due to:				
Person effects	0.302	0.466	0.384	0.216
Plant effects	0.174	0.115	0.103	0.149
Location effects	0.042	0.018	0.180	0.042
Covariance of worker and plant effects	0.336	0.451	0.393	0.347
Covariance of plant and location effects	0.049	-0.005	0.060	0.129
Covariance of worker and location effects	0.114	-0.046	0.099	0.125
Xb and other covariances	−0.017	0.001	-0.056	-0.008
Residual	0.000	0.000	0.000	0.000
RMSE of model	0.16	0.16	0.17	0.16
Adjusted R2	0.86	0.86	0.88	0.84

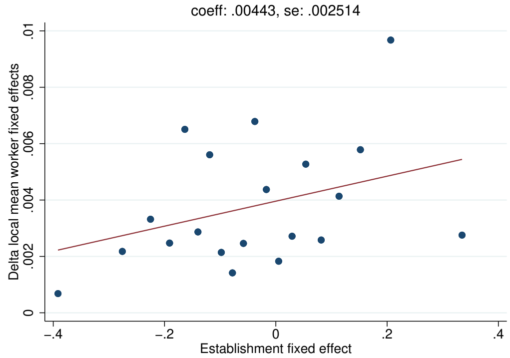
Notes: Table D.2 reproduces the variance decomposition exercise from Equation (6), separately by major industry. Lower-skill services include retail, accommodation, food services and other services to households. Higher-skill services include services to firms and real estate services. Data: linked employer-employee data and relocation register. Go back to main text

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

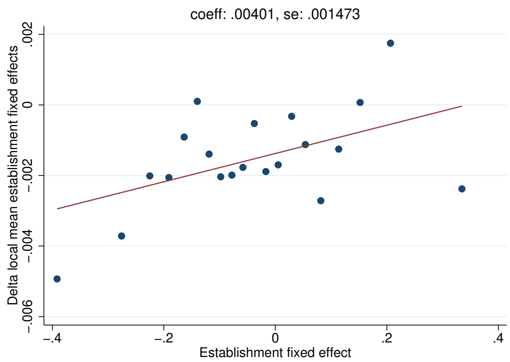
(a) Location



(b) Local Workforce

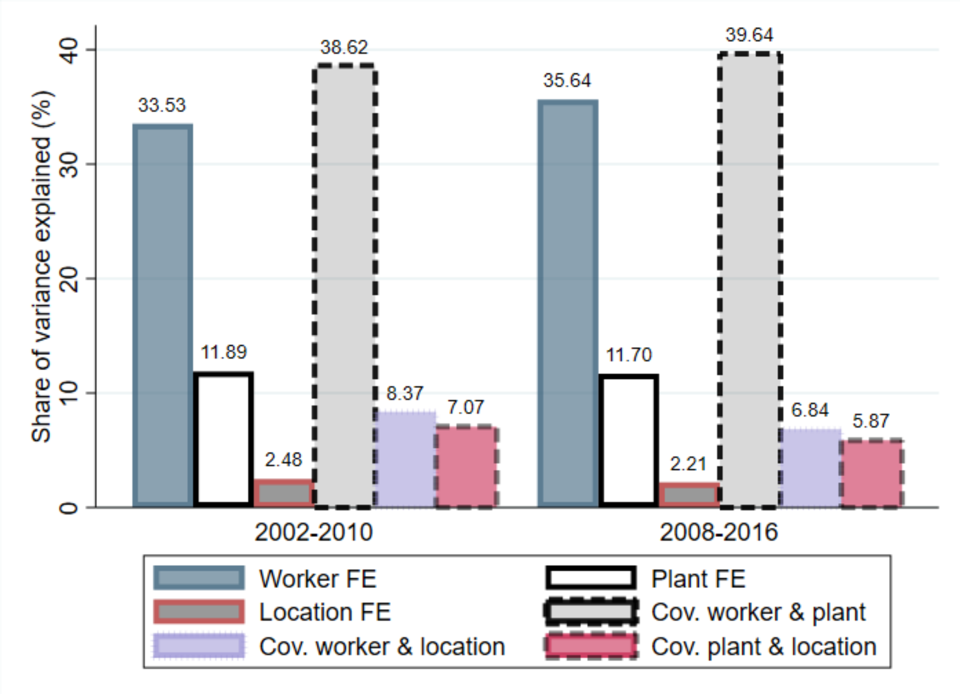


(c) Local Establishments



Notes: Figure D.3 depicts the correlation between the estimated fixed effect of the establishment relocating (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 measured 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. Data: linked employer-employee data and relocation register. 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 the periods 2001-2010 and 2008-2016. Each bar 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 33.53% of the variance, and it explains 35.64% over 2008-2016. Data: linked employer-employee data and relocation register. Go back to main text

Table D.3: Decomposition using Establishment Relocations to Home-Province of Top managers

	(1)	(2)
	Baseline, Provinces	Relocation to Hometown Province
Standard deviation of log wages	0.112	0.112
Number of worker-year observations	156,600,282	156,444,436
Number of workers	22,351,675	22,195,829
Number of plants	1,833,469	1,831,480
Number of locations	96	96
Std. dev. of worker effect	0.063	0.063
Std. dev. of plant effects	0.041	0.041
Std. dev. of location effects	0.017	0.017
Std. dev. of Xb	0.002	0.002
Correlation of worker/plant effects	0.976	0.975
Correlation of plant/location effects	0.655	0.654
Correlation of worker/location effects	0.486	0.481
Share of variance of log wages due to:		
Person effects	0.315	0.313
Plant effects	0.132	0.132
Location effects	0.022	0.023
Covariance of worker and plant effects	0.398	0.396
Covariance of plant and location effects	0.071	0.072
Covariance of worker and location effects	0.081	0.082
Xb and other covariances	-0.018	-0.018
Residual	0.000	0.000
Number of clusters	10	10
RMSE of model	0.16	0.16
Adjusted R2	0.85	0.85

Notes: Table D.3 reproduces the variance decomposition exercise at the province level. Column (1) includes all establishment relocations while column (2) exploits only moves back to the establishment top manager's province of birth. We define the top manager as the employee with the highest hourly wage within the highest-ranked occupation group. We consider only three one-digit occupation groups: (i) executives, (ii) managers and engineers, and (iii) supervisors and skilled technicians. We exclude non-managerial occupations. We then exploit the province of birth provided in the linked employer-employee data to identify relocations from a province different from the manager's province of birth, to their province of birth. We exclude relocations when the managers' province of birth is missing, when the move originates from the province of birth, or when the destination is not the birth province. Data: linked employer-employee data and relocation register. Go back to main text

D.3.1 Additional Evidence on the Size Premium

Table D.4: 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 hourly wage	0.06724***	0.06568***	0.05836***	0.05720***	0.00960***	0.05192***
	(0.00755)	(0.00653)	(0.00236)	(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	0.00774***	0.00763***	0.00744***	0.00731***	0.00105***	0.00644***
	(0.00102)	(0.00088)	(0.00060)	(0.00063)	(0.00014)	(0.00052)
Observations	304	304	304	304	304	304
Adj R^2	0.21	0.22	0.29	0.30	0.21	0.29
Panel C: Local worker effects (from 3WFE model) and size						
Worker fixed effect	0.03852***	0.03748***	0.03068***	0.03009***	0.00571***	0.02797***
	(0.00327)	(0.00280)	(0.00174)	(0.00149)	(0.00023)	(0.00113)
Observations	304	304	304	304	304	304
Adj R^2	0.72	0.74	0.70	0.72	0.87	0.76
Panel D: Local plant effects (from 3WFE model) and size						
Plant fixed effect	0.02159***	0.02115***	0.02087***	0.02040***	0.00292***	0.01804***
	(0.00394)	(0.00353)	(0.00095)	(0.00086)	(0.00052)	(0.00083)
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.4 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. Data: linked employer-employee data and relocation register. Go back to main text

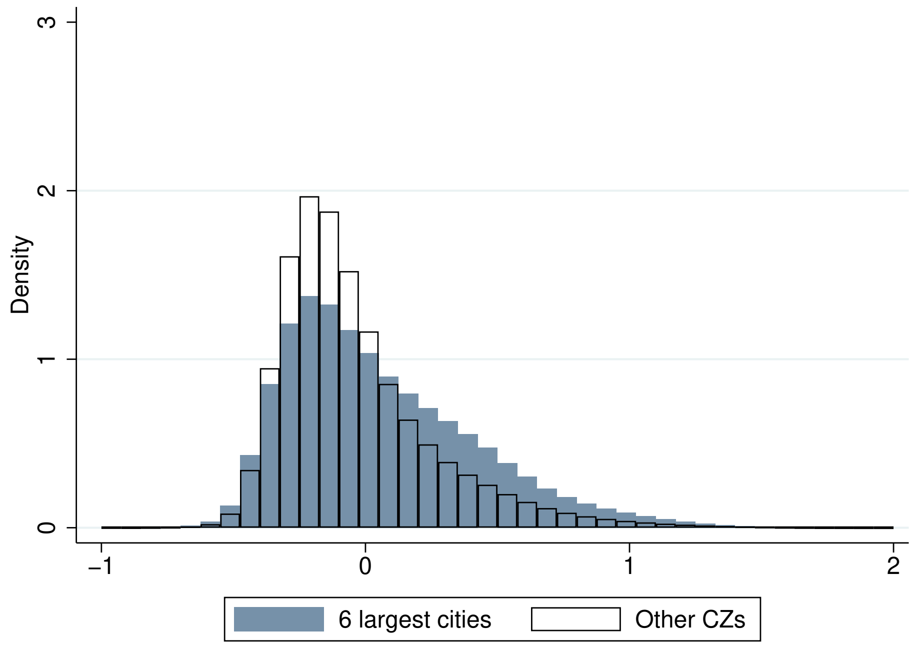
Table D.5: 2SLS Estimates of the Elasticity of the Location Effect to Density

	(1)	(2)	(3)	(4)
	Population (log)	Employment (log)	Population Density (log)	Employment Density (log)
Panel A: First Stage				
Log population in 1789	0.822***	0.861***	0.996***	1.035***
	(0.0245)	(0.0277)	(0.1415)	(0.1461)
Adj R^2	0.89	0.89	0.79	0.79
F-Stat	1124.6	967.3	49.6	50.2
Panel B: Second Stage				
Area fixed effect	0.00837***	0.00799***	0.00691***	0.00664***
	(0.00164)	(0.00153)	(0.00073)	(0.00070)
Adj R^2	0.27	0.28	0.32	0.33
Observations	198	198	198	198

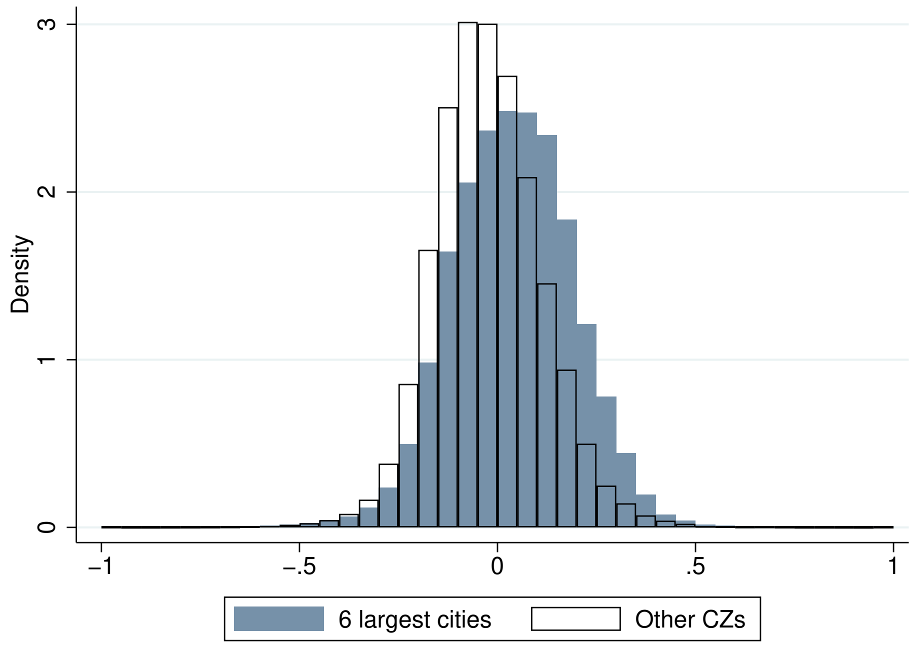
Notes: Table D.5 reports the results of a two-stage least squares regression of the estimated location effects on different measures of location density: (1) log of population, (2) log of employment, (3) log of population per meter square, (4) log of employment per square kilometer, observed in the 2015 census. The outcome of the regression, the location fixed effects, is estimated from Equation (5). We instrument the location size (as defined by the column title) by the size of the location in 1789. We rely on the 1789 census historical data, at the municipality level. Following the protocol outlined in Enamorado et al. (2019), we probabilistically matched 1789 municipalities to today's CZs based on Jaro-Winkler distance for city name (blocking on department). All matches were manually checked. We were able to match 198 of the 304 current CZs. Panel A reports the results of the regression of today's size on the size in 1789. Panel B reports estimates of the second stage. The standard errors are robust to heteroskedasticity. Data: linked employer-employee data and relocation register. 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: linked employer-employee data and relocation register. Go back to main text