NBER WORKING PAPER SERIES

LOCATION EFFECTS OR SORTING? EVIDENCE FROM FIRM RELOCATION

Pauline Carry
Benny Kleinman
Elio Nimier-David

Working Paper 33779
http://www.nber.org/papers/w33779

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2025, Revised July 2026

We are grateful to Joseph Altonji, Hadar Avivi, Michèle Belot, Marianne Bertrand, Zach Bleemer, Stephane Bonhomme, Janet Currie, Ellora Derenoncourt, Jonathan Dingel, Gilles Duranton, Cecile Gaubert, Matthew Gentzkow, Joshua Gottlieb, John Grigsby, Rick Hornbeck, Anders Humlum, Simon Jäger, Philipp Kircher, Patrick Kline, Michal Kolesar, Francis Kramarz, Ilyana Kuziemko, Thibault Lamadon, Jeremy Lise, Jack Mountjoy, Matthew Notowidigdo, Ezra Oberfield, Eleonora Patacchini, Evan Riehl, Roman Rivera, Richard Rogerson, Evan Rose, Esteban Rossi-Hansberg, Jesse Rothstein, Karthik Sastry, Benjamin Schoefer, Isaac Sorkin, Gabriel Swagel, Chad Syverson, Owen Zidar and Eric Zwick for useful feedback. We would like to thank conference and seminar participants at NBER SI Labor Studies, Cowles, SOLE, Chicago-Princeton Conference on Spatial Economics, the University of Chicago, Cornell University, Princeton University, UC Berkeley, HEC Paris, University of Houston, Boston University, Columbia University, the University of Toronto, and AREUEA 2026 for their comments. We thank Michael Amior and Ferdinando Monte for their thoughtful discussions. We thank Cedric Christensen, Adrien Mangold, 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). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Location Effects or Sorting? Evidence from Firm Relocation
Pauline Carry, Benny Kleinman, and Elio Nimier-David
NBER Working Paper No. 33779
May 2025, Revised July 2026
JEL No. F0, J0, R0

ABSTRACT

What role do firms play in geographic wage disparities? This paper exploits establishment mobility as a novel source of identification to separate firm sorting from location treatment effects, i.e., the local wage premium for a given worker at a given firm (due to, e.g., geography, infrastructure, amenities). Using data from France and the U.S., we first document key facts about establishment relocation: 4%of establishments move each year, retaining their activity and structure while adjusting their workforce and wages. Combining establishment and worker mobility in France, we estimate a model that decomposes wage variation due to workers, firms, and locations. Spatial wage differences are primarily driven by sorting: worker composition accounts for 30%, firm composition for 17%, and their co-location for 34%. Residual location effects explain only 2–4%. Establishment sorting also accounts for most of the urban wage premium and the higher wage growth in larger cities.

Pauline Carry
Princeton University
pauline.carry@princeton.edu

Elio Nimier-David
Cornell University
elio.nimier.david@cornell.edu

Benny Kleinman
Stanford University
Department of Economics
and NBER
bennyk@stanford.edu

1 Introduction

Why are wages in cities like New York or Paris higher than elsewhere? Explanations fall into two broad categories. One view emphasizes the treatment effect of a location (henceforth, a "location effect"): the local wage premium for a given worker at a given firm, reflecting a bundle of regional characteristics such as amenities, productivity, and cost of living. From this perspective, the hypothetical relocation of a firm and its workers from a low-wage city to Paris would lead to higher wages. The alternative attributes the wage premium to spatial sorting, that is, differences in the composition of firms and workers across locations. Under this view, the same relocation would leave wages unchanged. Separating these two forces remains empirically challenging, yet this distinction is central for policy. Governments allocate billions to attract firms through local business incentives and tax breaks (Bartik, 2020; Slattery and Zidar, 2020), often under the assumption that high-paying firms can raise local wages. It also matters for how economists model cities, distinguishing models built on spatial sorting from those built on differences in local fundamentals and spillovers.

This paper develops an approach to separately estimate the contributions of firms, workers, and locations to spatial wage disparities. We exploit *establishment mobility* as a source of identification to disentangle firm sorting from residual location effects. Combined with worker mobility across firms and space, this *double-mover design* quantifies two-sided spatial sorting. We proceed in two steps. First, we document stylized facts about the understudied phenomenon of establishment relocation in France and the U.S., which we rationalize through the lens of a simple spatial equilibrium model. Second, we use our double-mover design to decompose spatial wage disparities in France. We highlight three main findings. First, spatial sorting accounts for most wage disparities, while location effects net of composition account for only 2–4%. Second, firm premia are a central driver of regional wages. Third, the urban wage premium – both its static and dynamic components – largely reflects the concentration of high-paying establishments in large cities.

Studying establishment 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 requires establishments to report any change of address.[1] Uniquely, these data record the establishment's identity, its old location, and its new location on the same form, guaranteeing that we track the same establishment across space.[2] Combining

[1]To maintain 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.

[2]Throughout the paper, we measure the relocation of individual establishments and use the terms

this relocation registry with linked employer-employee data, firm balance sheet, and a survey of entrepreneurs, we conduct a series of validation tests to confirm that relocations align with the intuitive notion of a change in the establishment's location of activity. To ensure our characterization of firm relocation is not specific to France or a particular definition of relocation, we reproduce our main facts using U.S. commercial data from Dun & Bradstreet and data on the headquarters locations of U.S. publicly listed firms.

We begin by presenting key facts about establishment relocation. First, we show that relocation is common: every year, approximately 4% of all establishments change their address and 1% change their commuting zone, spanning all establishment sizes and sectors. Second, we explore the geography of relocation. Establishment mobility is characterized by a gravity structure, with substantial bidirectional flows between location pairs, suggesting a moderate role for shared location factors in driving destination choices. Third, we find evidence of spatial sorting: larger and higher-wage establishments are more likely to relocate to larger and higher-wage destinations.

We next examine what happens to establishments following a relocation to a different local labor market (commuting zone). We find no evidence of changes in industry, legal status, or input shares, and relocating entrepreneurs report no more disruptions or problems than non-movers, confirming we observe the same establishment after the move. At the same time, relocating establishments replace much of their workforce with local employees at the destination, and their average wage adjusts: moving to a market where wages are higher by 1% increases the average wage of the establishment by 0.15%. This small elasticity points to a large establishment component of wages. Our decomposition further shows that this modest adjustment mostly reflects changes in worker composition.

We develop a simple spatial equilibrium framework to rationalize this descriptive evidence and to motivate our wage decomposition. Entrepreneurs with heterogeneous productivity select locations to maximize profits, subject to idiosyncratic location preferences. Workers with heterogeneous abilities choose locations and match with firms in frictional local labor markets. The model predicts that higher-productivity firms tend to relocate to areas with higher regional productivity or more productive workers, due to complementarities in the production technology. The model also delivers a simple log-linear wage equation featuring distinct location, firm, and worker components, as one potential interpretation for our wage decomposition exercise.

We then leverage the widespread mobility of establishments to examine the role of firms in shaping spatial wage disparities. We conduct a wage decomposition analysis in the spirit of Abowd et al. (1999) (hereafter AKM), disentangling the contributions of

establishment relocation and *firm relocation* interchangeably. For multi-establishment firms, we track the relocation of all establishment types in the French data.

worker and establishment composition from those of location effects. In our baseline specification, we estimate a three-way model with worker, establishment, and location fixed effects, identified using both establishment mobility across locations and worker mobility across establishments and locations. We then consider extensions to study the role of non-linearities, as well as heterogeneous and dynamic location effects.

Our main identification assumption is that, conditional on the three fixed effects, the residual wage component is orthogonal to mobility decisions. We formalize this assumption by extending the potential outcomes framework to our double-mover design (Kline, 2024), which clarifies the conditions under which fixed effects recover causal treatment effects. We present four tests supporting this assumption. First, we consider event studies of both establishment and worker wages before and after relocations, following Card et al. (2013) and Badinski et al. (2023). We find no evidence of significant wage (nor productivity) changes before the move, and observe nearly symmetric wage adjustments when moving to higher- versus lower-wage locations, consistent with mobility being orthogonal to wage shocks or match effects. Second, our design allows a new placebo test: we analyze wages of establishments that relocate within the same commuting zone. For those, wages remain flat following the move, consistent with the local labor market being unchanged. Third, mean estimated residuals of our model are systematically small across all cells of worker, establishment, and location effects. Fourth, the survey of entrepreneurs confirms that relocating entrepreneurs did not report experiencing shocks, such as increased competition in their initial location or financial difficulties, that could affect wages.

We find that the vast majority of spatial wage inequality arises from the sorting of firms and workers across space. Differences in worker composition across cities account for approximately 30% of the total variance, and differences in establishment composition for an additional 17%. In the language of Song et al. (2019), these two components of spatial sorting reflect *spatial segregation*. A further third of the variance stems from the *co-location* of high-pay establishments and high-ability workers. In contrast, residual location effects (net of worker and firm composition) account for only 2%–4% of the variance. In absolute terms, the standard deviation of location effects is only 2.3 log points, with a 90-10 percentile gap of 5 log points. Finally, 16% of the variance is accounted for by the sorting of high-type workers and establishments to higher effects locations.

We examine heterogeneity and non-linearities in location effects. Across industries, occupations, and firm size groups, we find a stable ranking of location effects and a similar contribution to wage variance. Adapting Bonhomme et al. (2019) to allow for non-linearities between location and establishment types, we find a limited role for complementarities, consistent with their findings on worker-establishment pairs.

We show that our results are not driven by artifacts of our design. First, we replicate our analysis exclusively using relocations to the entrepreneur's hometown, events that are likely driven by personal preferences and therefore independent of wage-relevant shocks. We find results very similar to our baseline, providing support for our strategy. Second, restricting identification to new hires yields similar results, ruling out wage stickiness among workers moving with their establishment as a source of bias. Third, average establishment wages adjust substantially after relocation – even within occupations – alleviating concerns about national wage-setting policies. This adjustment operates mainly through changes in worker composition. Moreover, excluding multi-establishment firms yields the same results, ruling out uniform pay grids across locations. Fourth, restricting to long-distance moves, large enough to represent a genuine change in the local labor market; leaves the results unchanged. Fifth, we also find similar results when excluding young establishments, whose relocations may be confounded with life-cycle dynamics or driven by weak local attachment.

Spatial sorting of workers and firms explains not only the variance of wages across space, but also most of why wages rise with local density. To show this, we revisit the urban wage premium – a key piece of evidence for the productivity advantages of denser markets in the agglomeration literature. We find a wage-density elasticity of 0.007, well below the 0.02–0.04 range obtained once worker composition alone is controlled for (Glaeser and Maré, 2001; Combes et al., 2008; De la Roca and Puga, 2016), a benchmark we replicate before accounting for establishment sorting. While prior estimates inform about the causal wage gain for a worker moving to a larger city, our results additionally separate the effect of density itself from the tendency of high-paying firms to concentrate in dense locations. We then decompose the skill-premium gradient across city sizes. We find that it is driven by worker composition within skill groups and by the sorting of workers across establishments, rather than heterogeneous effects of locations or establishments by skill.

Beyond the static location size premium, we also find that most of the dynamic wage gains from working in larger cities reflect firm composition. Prior research indicates that a substantial share of the urban wage premium is dynamic, attributed to the value of experience accumulated in denser cities (Baum-Snow and Pavan, 2012; D'Costa and Overman, 2014; De la Roca and Puga, 2016). We find that 65% of the dynamic gains are due to the sorting of establishments offering higher premia and steeper wage trajectories.

Beyond density, location effects may capture other local forces that affect wages. A Shapley-value-based variance decomposition indicates a strong role for geography – market access and distance to Paris – and housing prices, while compensating differentials for regional amenities account for a more modest share of wage variation.

Our findings have three main implications. First, firm composition is a central determinant of regional wages, consistent with policymakers' view that attracting high-paying firms can shape local labor markets. Our results also suggest that local human capital – captured by regional worker effects – is consequential for firms' location choices. Second, two-sided sorting is a key driver of spatial wage inequality, reflected in the co-location of high-paying establishments and high-type workers. Third, our results help distinguish between different mechanisms for agglomeration effects: while consistent with mechanisms that materialize through firm and worker composition (e.g., labor market pooling), they point to a more limited role for mechanisms that imply a location treatment effect (e.g., shared infrastructure or suppliers and knowledge spillovers from surrounding firms).[3]

This paper contributes to several strands of the literature. First, we relate to work on the drivers of spatial wage disparities. A large literature has leveraged worker mobility to estimate wage dispersion across locations while controlling for worker heterogeneity (Glaeser and Maré, 2001; Combes et al., 2008; Moretti, 2011; De la Roca and Puga, 2016; Dauth et al., 2022; Porcher et al., 2023; Moretti and Yi, 2024; Card et al., 2025).[4] We also relate to the literature documenting higher wages in larger and denser cities (Di Addario and Patacchini, 2008; Glaeser and Gottlieb, 2009; Melo et al., 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 contribute by quantifying the spatial sorting of establishments and separating it from residual location effects. Our results also address a distinct set of questions: the contribution of firms to regional wages beyond worker composition, the effect of changing location for a firm, the magnitude of two-sided sorting, and the relative importance of different agglomeration channels.

Second, we relate to the theoretical literature in spatial economics that studies firm sorting as a driver of 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), Chikis et al. (2025), and Lhuillier (2025).[5] Our results inform the construction of spatial equilibrium models and offer a set of empirical moments to discipline the

[3]Notably, treatment effects that arise due to exposure to surrounding workers and firms – as in the agglomeration spillover literature – are reflected in our location effects.

[4]As a benchmark, we replicate the methods of Glaeser and Maré (2001); Combes et al. (2008); Card et al. (2025) in our sample. Notably, we find estimates very close to what Card et al. (2025) obtain for the U.S. 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).

[5]See Diamond and Gaubert (2022) and Diamond and Suárez Serrato (2025) for a survey.

importance of sorting relative to location fundamentals and agglomeration forces.[6]

Methodologically, our double-mover design relates to Badinski et al. (2023) in the context of variation in healthcare utilization. Our context is regional wage determination, and we use firm and worker mobility. We also relate to papers that use the structure of multinationals or multi-region firms to analyze dispersion across countries or regions in other settings, assuming common firm effects across all establishments – e.g. Alviarez et al. (2023) who study variation in firms' output across countries, and Bilal (2023) who studies variation in unemployment across local labor markets.[7] We also connect to a related 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 role of firm-level wage policies. By focusing on establishments, we can account for both the prevalence of single-establishment firms and pay variation across establishments within a firm. We find that the limited role of location effects is not driven by firm-wide pay policies: establishments adjust wages after relocation, but primarily through changes in worker composition.

Finally, we contribute to the small literature on firm mobility. Existing work studies firm location decisions in specific contexts: Duranton and Puga (2001) documents the relocation of business services from diversified to specialized cities; Voget (2011) studies multinational headquarters in response to taxation; Bergeaud and Ray (2020) and Hu et al. (2024) examine the role of real estate and housing costs, respectively; and Bryan and Guzman (2023) tracks the migration of high-potential startups. Building on these studies, we provide a comprehensive view of establishment relocation across the economy, covering its prevalence (in France and the U.S.), its geography, and how establishments evolve following a move. Crucially, we use matched employer-employee data to study how firms' workforce and wages change as they relocate, and exploit this to quantify the importance of firms in shaping regional wages.

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 econometric framework and the empirical model used to decompose spatial wage disparities in Section 5. We discuss the results in Sections 6 and 7. Section 8 concludes.

[6]Our estimates for location effects, average local worker and firm effects will be made available as a supplement to this paper. Readers interested in using those estimates may contact the authors.

[7]Another related line of research studies the importance of firm granularity for spatial differences in productivity, as in Schoefer and Ziv (2022); and spillovers between proximate firms as a determinant of firm productivity, as in Baum-Snow et al. (2024)

2 Context and Data

This section presents the data, with an emphasis on measuring establishment relocation in France and the U.S. Throughout the paper, we treat location as an endogenous, time-varying attribute of an establishment rather than as part of its fixed identity. An establishment is defined by other features such as activity, administrative structure, technology, and its entrepreneur. In Section 3 we find that proxies for those features remain stable across a change in location.

2.1 Process to Relocate an Establishment in France

An entrepreneur relocating their establishment must report a change of address to the local commercial court within three months of the move. The court then updates their business license ("Kbis") with the new address. This license is the sole document proving the existence of an operating establishment and is required to conduct business (e.g., sign contracts or open bank accounts). The address change requires three documents: (i) proof of closure of the establishment at the former location, (ii) evidence of the establishment's reopening at the new location (e.g., the new lease), and (iii) the Form M2 stating both the former and new addresses (see Figure A.1 for this form). The French statistical institute, INSEE, uses those forms to construct the registry of establishment relocations described in Section 2.2. It defines a 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*."

This formal reporting process leaves little room for misreporting. Entrepreneurs have no incentive to declare a relocated establishment as a new one, since this would involve more bureaucratic procedures. Nor are new establishments likely to be mistaken for relocations, as entrepreneurs are required to prove they closed the establishment in the last three months. In Section 3.3, we show that establishments retain the same activity and organizational structure after relocating, confirming that the same establishment is observed at a new location.

When a move exceeds 10 kilometers ($\sim$ 6.2 miles), the law mandates a renegotiation of employment contracts, creating scope for wage adjustments reflecting new local labor market conditions. Workers may decline to move with their establishment, but doing so constitutes a voluntary quit. Firm-initiated separations are strictly regulated: employers must provide just cause for dismissal, and establishment relocation does not qualify. As a result, firms have limited discretion over which employees relocate. Finally, employment-related institutions such as collective bargaining agreements are set nationally, so they do not vary across locations.

2.2 French Administrative and Survey Data

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

Registry of Establishment Relocation. In many countries, establishment identifiers are location-specific, preventing researchers from tracking their mobility across space. We overcome this limitation by leveraging the SIRENE register that provides a link between the establishment identifier at the destination and the initial location, building on Form M2 (see Section 2.1). This registry starts in 1993. It covers all establishments (e.g., headquarters, production plants) in the private sector, excluding those in agriculture and finance, and all types of moves (including within cities). The registry is specific to relocations; other changes (e.g., a merger) would require different forms.

Linked Employer-Employee Data. We use comprehensive employment records (DADS Postes) to study workers' compensation and mobility. These data, based on mandatory employer declarations, include gross wages, occupation codes, and hours worked during the year for all employees subject to French payroll taxes. 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 commuting zones (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 individuals employed 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 covering the universe of workers over the period 2002–2016.[8] For comparison, prior research has typically relied on a restricted panel version covering approximately $1/25^{th}$ of workers (see, for example, Abowd et al. (1999)). We restrict the sample to employees aged 23-60 working in metropolitan France and retain only their highest-paying job during the year. We focus on workers' gross hourly wages, converted to 2018 euros.[9]

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

[8]We end our period of consideration in 2016 because the data collection changed in 2017.

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

utilizing them.

Survey of Entrepreneurs. We leverage the SINE survey to study the characteristics of relocating companies and potential drivers of the relocation. This survey provides detailed information on the characteristics of entrepreneurs, firms' activities, and their 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 half of the survey year (about 24,000 in 2014). We focus on companies that remained active at the time of the final interview and link this survey to our relocation register using firm identifiers. In Section 3, we compare entrepreneurs that relocated at least one establishment within their first five years of activity to those that did not.

2.3 Measures of Establishment Relocation in the United States

Additionally, we provide evidence on firm relocation in the United States. We use establishment-level commercial data from Dun & Bradstreet covering the entire economy, and firm-level data on headquarters of publicly listed firms from firm filings with the U.S. 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. These linkages include headquarters identifiers for establishments within multi-unit firms and parent identifiers for subsidiaries. Unlike the French data, relocation events are not explicitly recorded, so we must construct them from observed changes in location. We focus exclusively on single-establishment firms or headquarters of multi-establishment firms, as these can be uniquely identified over time. We define relocation events at the firm level, matching firms over time based on their name, creation year, industry, and ownership structure. We restrict our analysis to firms observed for at least three consecutive years with these characteristics held constant. We then define a relocation as a single change in the firm's location. Since street addresses are reported with considerable noise in these data, we focus on moves across five-digit zip codes.

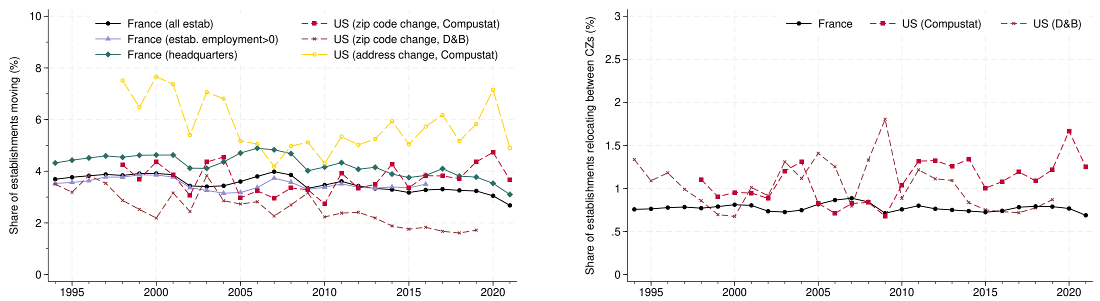
Data on U.S. public corporations. We further complement our analysis using relocation events for publicly listed firms. We obtain firms' annual 10-K reports from the SEC EDGAR (Electronic Data Gathering, Analysis, and Retrieval) system for the years 1993–2021 and extract each firm's headquarters location from its mailing address. This mailing address reflects the administrative headquarters, distinct from the incorporation address, which typically does not correspond to a physical operating site. We then merge this information with firm-level balance sheet data from Compustat, restricting to firms reporting positive employment to ensure the sample reflects firms with active operations. 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.[10]

3 Characterization of Firm Relocation

We now turn to an empirical characterization of establishment relocations, a phenomenon much less explored than worker mobility. We examine its prevalence, geographic patterns, and what happens to establishments after relocating. We document these facts for France and, where data permit, for the United States.

Figure 1: Share of Establishments Relocating in France and in the U.S.

(a) All Moves (b) Moves Between Commuting Zones



Notes: Figure 1 plots, for each year between 1994 and 2021, the number 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 U.S., the dark red dashed line with crosses includes all headquarters of firms observed in the Dun & 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 & Bradstreet data, Compustat.

3.1 How Prevalent is Firm Relocation?

Figure 1 presents the rate of establishment relocations in France and the U.S. between 1994 and 2021. 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%, followed by a modest decline since the Great Recession. Headquarters exhibit a slightly higher

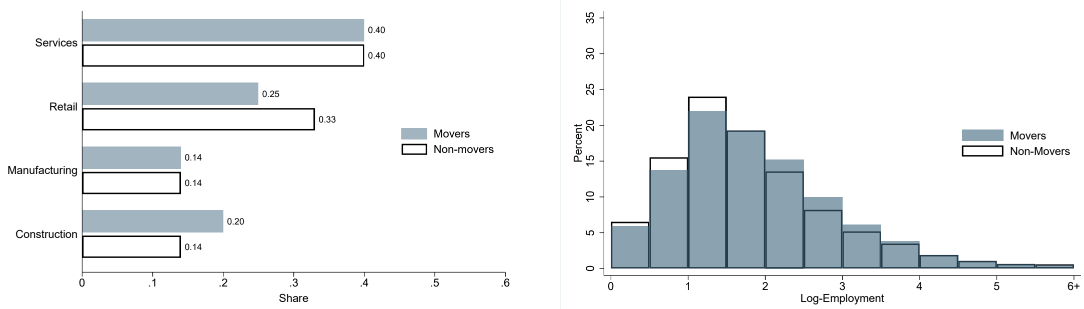
[10] While Compustat provides current addresses for firms, it does not retain historical addresses, making it insufficient for directly identifying relocations over time. Thus, address information from the annual 10-K filings is essential to identify firm relocations.

propensity to relocate, with rates between 4% and 5%. In the U.S., relocation rates are of similar magnitudes, with 2–5% of single-unit firms and of multi-unit firms' headquarters changing zip codes in both the Compustat sample and the D&B data. As expected, changes of address, observed for the Compustat sample, occur slightly more frequently, at around 5–7%. Establishments that relocate account for approximately 2.5% of the French workforce every year (Figure B.1).

Figure 2: Characteristics of Relocating Establishments in France

(a) Main Sectors

(b) Size



Notes: Figure 2 compares the main characteristics of establishments that relocate between CZs and those that do not relocate, in France. Panel (a) plots the distribution of establishments across major industries separately for movers and non-movers. Panel (b) depicts the establishment (log) employment distribution. Data: linked employer-employee data and relocation register.

Figure 1(b) shows that in both France and the U.S., 20-25% of establishment relocations involve a change of commuting zone.[11] We focus on those as they involve a change of labor market. Establishment mobility across CZs occurs across all major industries and establishment size categories, as shown by Figure 2. At a more granular level, relocation rates tend to be higher in business services and lower in hospitality and agrifood industries (Figure B.1).

3.2 Geography of Firm Relocation

We now address two key questions about the geography of firm mobility: (1) Where do establishments relocate? (2) Which establishments move where?

Relocation Distances. In line with patterns observed in worker migration, we document a strong gravity structure for establishment relocation rates. To this end, we estimate the following model:

$$log(\#\text{ moves}_{ijt}) = \eta_{it} + \mu_{jt} + \delta\, log(\text{Distance}_{ij}) + \epsilon_{ijt}, \tag{1}$$

where the dependent variable is the log of the number of establishments moving from location i to location j (where locations are Provinces in France and CZs in the U.S.) at

[11]Figure B.2 presents similar time series using different definitions of geographic boundaries. 65% of relocations involve changing municipality.

time t.[12] η_{it} and μ_{jt} are time-varying origin and destination fixed effects, respectively. $Distance_{ij}$ is the number of kilometers between locations i and j. Figure B.3 depicts the results for France and the U.S. We estimate negative elasticities of -0.96 for France and -0.42 for the U.S., suggesting a strong role for spatial frictions in shaping relocation decisions.[13]

Gross and net flows. A key characteristic of worker mobility across space is the prevalence of bidirectional flows with gross migration flows far outweighing net flows, even between very different local markets. This pattern often motivates models of location choice to allow for substantial dispersion in idiosyncratic preferences. We find a similar pattern for firm relocations. For each pair of locations, we compute the following index akin to the Grubel-Lloyd index in international trade:

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

where # Outflows represents the number of establishments moving 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, e.g., as many relocations from Paris to Marseille as there are from Marseille to Paris. Figure 3 shows the distribution of $\text{Index}^{\text{Directionality}}$ for France and the U.S.[14] The large mass of observations close to zero suggests that gross flows are much larger than net flows. Although flows are not perfectly symmetric, with some observations far from zero, the overall distribution suggests that relocation decisions are not primarily driven by shared location-specific factors, such as market size or taxes, that would make certain destinations systematically more attractive.

Sorting associated with relocation. We now examine which establishments move to which destinations, by comparing the characteristics of establishments relocating from the same origin at the same age, but to different destinations. To do so, we use the following specification:

$$x_{C(J,t+1),t} = \eta_{C(J,t)} + \gamma_1 x_{J,t} + \gamma_2 Age_{J,t} + u_{J,t}, \tag{3}$$

where $x_{J,t}$ represents the characteristics of the establishment before the relocation (e.g., log size) and $x_{C(J,t+1),t}$ denotes the characteristics of the destination commuting zone, measured before the move. We control for establishment age to highlight sorting patterns

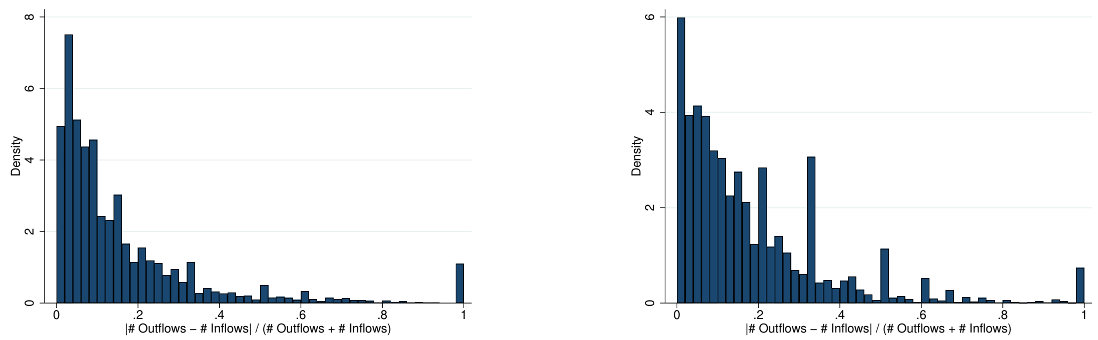
[12] A Province in France corresponds to a *département*. There are 96 départements in metropolitan France, with an average population of about 650,000 inhabitants in 2020.

[13] In France, the mean and median relocation distance for an establishment are 151 and 42 km, respectively. It is 178 and 69 km for workers (see Table B.1). Figure B.4 plots the distribution of relocation distances for establishments, highlighting the existence of relocations spanning several hundred kilometers.

[14] We compute the index over four sub-periods of approximately seven years each, so that offsetting shifts in mobility patterns do not cancel each other out over the full sample.

Figure 3: Ratio of Net and Gross Relocation Flows for Every Pair of Locations

(a) France (b) U.S.



Notes: Figure 3 plots the distribution of the ratio between net and gross flows, for every pair of provinces (France) and commuting zones (U.S.). 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 U.S. over 1994-2019. Data: relocation register and Dun & Bradstreet.

that are not driven by life-cycle dynamics. We also control for the commuting zone of origin, $\eta_{C(J,t)}$, to avoid capturing systematic differences in firms' original location. Figure 4 presents the results for establishment size in France (Panel (a)) and the U.S. (Panel (b)), and for hourly wages in France (Panel (c)). First, we find that larger establishments relocate to larger CZs, with elasticities of 0.086 (standard error: 0.008) for France and 0.057 (standard error: 0.01) for the U.S. 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 high-paying establishments into high-paying locations. Section 6.1 replicates this analysis with estimated fixed effects: sorting reflects high-fixed-effect establishments relocating to locations with higher average worker fixed effects, suggesting local human capital as a driver of relocation choices.

3.3 What Happens to an Establishment that Relocates?

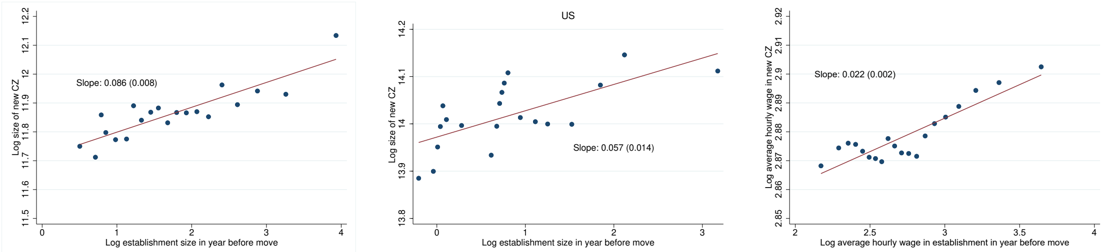
We consider two main margins of adjustment following a relocation: structural changes in activity and organization, and labor adjustments to the new local labor market. This analysis allows us to evaluate the persistence of establishment characteristics following relocation in France.

3.3.1 Activity and Organization

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

Figure 4: Correlation Between Establishment and Destination Size and Pay

(a) Size (France) (b) Size (USA) (c) Wage (France)



Notes: Figure 4 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 $x = \{\text{log employment, log average hourly wage}\}$. 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. Panels (a) and (b) focus on log size for France and the U.S. respectively, while Panel (c) depicts estimates for the log hourly wage. Data: linked employer-employee data, relocation register, and Dun & Bradstreet.

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 unchanged the characteristic indicated in the first row, between year t and year $t + 1$. These shares are presented separately for establishments experiencing a relocation between two consecutive years, and for establishments 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.

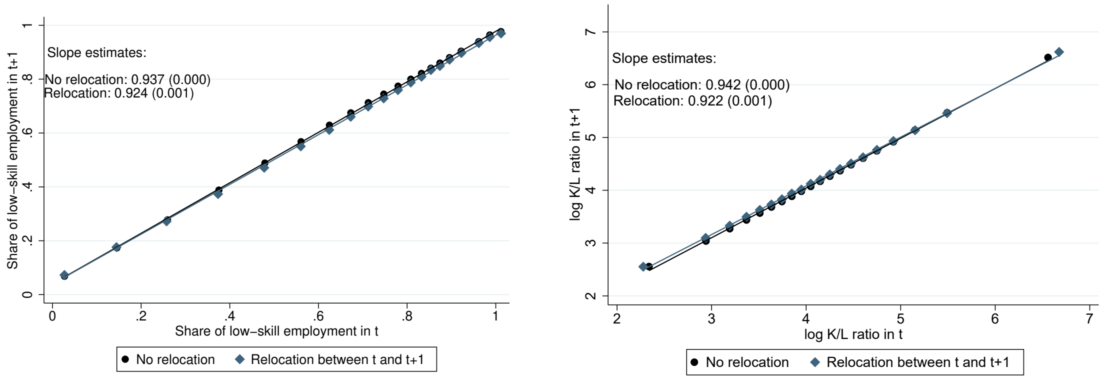
headquarters status, and 96% continue under the same legal category (e.g., simplified corporation or limited liability company). The last column indicates that the top manager remains the same in 67% of relocating establishments, compared to 76% of non-relocating establishments.[15] These figures suggest that relocation does not involve significant changes in business activity or organizational structure. In addition, entrepreneurs must report updated information when completing the change-of-address form, making it more likely that changes are recorded for 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 across all the variables, confirming that only a small fraction of establishments undergo significant changes after relocating.

Second, we use linked employer-employee and balance-sheet data to track the evolution of key inputs after relocation. Figure 5 plots the correlation between input ratios in year t and year $t + 1$, separately for establishments that relocated and those that did not. We examine the share of low-skill workers in the establishment (Panel (a)) and the capital-to-labor ratio (Panel (b)). For movers, we estimate slopes of 0.924 and 0.922,

[15]To proxy for the identity of the top manager, we use the highest-paid employee.

Figure 5: Evolution of Input Composition for Movers and Non-Movers

(a) Skill Composition (b) Capital-to-Labor Ratio

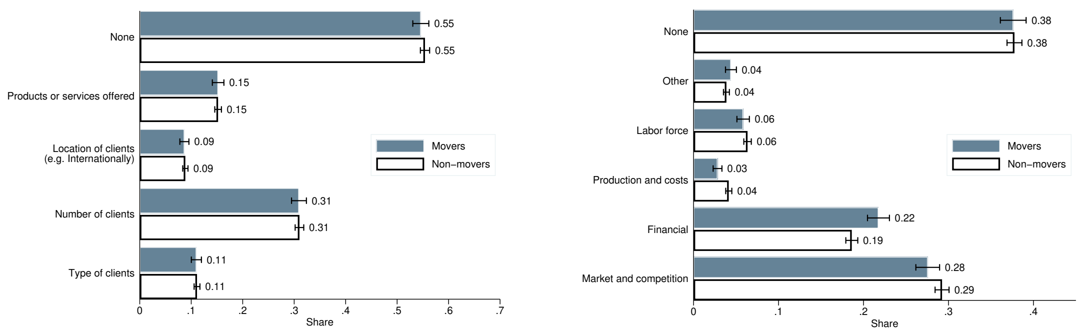


Notes: Figure 5 shows the correlation between an establishment characteristic in a given year (x-axis) and the same characteristic the subsequent year (y-axis). In Panel (a), we consider the share of low-skill workers employed in the establishment, defined as blue-collar and low-skill white-collar workers (clerks). 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 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.

respectively, indicating that input composition remains highly persistent after relocation.

Figure 6: Changes Reported by Entrepreneurs

(a) Major Changes Since Creation (b) Problems in the First Two Years



Notes: Figure 6 compares answers from entrepreneurs who relocated at least one establishment during their first five years of economic activity, with answers of entrepreneurs who did not. Panel (a) shows answers to the question "**Did your firm experience any major change since its creation?**" Entrepreneurs can report that their firm experienced 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. Figure B.5 shows additional questions on changes encountered. Data: SINE survey of entrepreneurs (2014) and relocation register.

We complement these findings using the SINE survey, which asks entrepreneurs about changes since firm creation and challenges encountered. We match the survey to administrative relocation data to compare movers and non-movers. Figure 6 reports responses to: "Did your firm experience any major change since its creation?" and "Since the creation of your company, what has been the main obstacle to its development?" Responses are strikingly similar across movers and non-movers. In both groups, 55% report no major changes and 38% no difficulties in the first two years. Movers and non-movers also show similar rates of changes in services offered (15%), client geography (domestic vs. international, 9%), number of clients (31%), and client type (11%). Beyond confirming that establishments do not adjust more upon relocating, these results suggest relocation is not driven by external shocks. For example, both groups report similar difficulties related to labor constraints (6%) and competition (28% vs. 29%). Figure B.5 further supports the absence of specific shocks as primary drivers of relocation using additional survey questions. To conclude, our administrative and survey data indicate that relocating establishments do not significantly alter their operations, the goods and services they produce, nor the type of clientele they serve. This indicates the identity of the establishment is preserved throughout relocation.

3.3.2 Labor Adjustments

Although establishments do not appear to undergo structural changes after relocation, we next examine how they adjust their labor inputs: whether they hire workers from the new location and whether compensation aligns with local wages.

First, we find that about half of an establishment's workforce is replaced by workers from the new location around the time of the move. Figure B.6 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)). After relocation, employment shifts sharply toward workers living in the destination and away from those in the origin.

Having shown that both location and workforce composition change with relocation, we now examine how establishment wages evolve. Specifically, we test whether establishments relocating to higher-wage areas raise their average wage. We compare wage trajectories across establishments, based on the gap in average wage in 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, excluding the relocating establishment. 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, $\theta_{r(J,t)}$, capture interactions between relative year indicators and the difference in average wage between the destination and origin locations. They reflect how an establishment's average wage evolves in the years before and after relocation, scaled by the destination-origin wage gap.

We estimate two versions of Equation (4). The first includes both movers and non-movers, with the relative year and Δ_J set to zero for non-movers. The second restricts to relocating establishments and omits the relative year fixed effects (x_{Jt}). In both cases, we focus on relocating establishments observed for four years before and after the move, 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 still represents the wage gap between destination and origin CZs (Figure B.7).[16]

Figure 7 presents the results for establishment relocations (Panel (a) for the main coefficients, and Panel (b) splitting between relocations to a higher- or lower-paying CZ). 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 conditions (e.g., productivity effects) 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.

Figure B.8 reports estimates of $\theta_{r(J,t)}$ for log establishment size (Panel (a)) and log value added per hour (Panel (b)). Relocating to a CZ with 1% higher wages is associated with an increase in size by 0.09–0.21% and productivity by 0.09–0.14%.

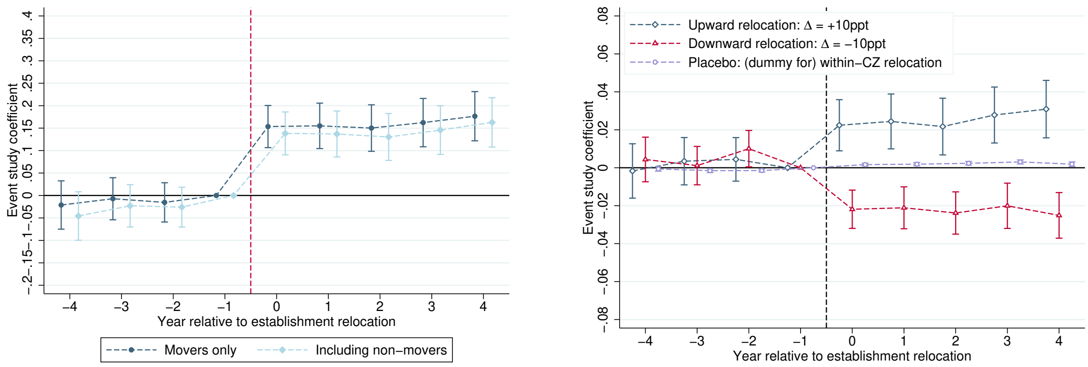
Taking stock. Establishment mobility is prevalent and can be systematically tracked in our data. Consistent with treating location as an attribute of an establishment, we find that its other key characteristics – including activity, organization, and input composition – remain highly persistent after a move. At the same time, relocating establishments substantially adjust their workforce and wages, suggesting that relocation is best understood as a change in the local labor market in which an ongoing establishment operates. These findings support using establishment mobility to study spatial wage disparities, as it lets us observe the same production unit in two locations.

[16] Panel (a) of Figure B.7 reports the results for workers changing location. Panel (b) reports the event studies for workers changing establishment within CZ, where Δ_J is then the wage gap between the new and former establishments.

Figure 7: Event-Study Estimates of Establishment Wage Responses to Relocation

(a) Establishment Wage Response to the Local Wage Gap

(b) Positive versus Negative Local Wage Gaps, with Placebo



Notes: Figure 7 reports estimates of θ in Equation (4) for relocating establishments. Panel (a) uses all CZ changes. The estimates are displayed for two samples: establishments ever making a change, and the whole sample (including non-movers). Δ_J is in raw units, so a value of one corresponds to a 100 log point gap between the origin and destination locations. Panel (b) splits estimates between relocations with a positive Δ_J and those with a negative one. In the latter case, the opposite of the coefficient is displayed. Δ_J is rescaled so that a unit change equals 10 log points, the order of magnitude of the average relocation. Finally, the lavender line with circles corresponds to a placebo event study that considers a dummy for establishments relocating within the same CZ. Standard errors are clustered at the establishment level. Data: linked employer-employee data and relocation register.

4 A Model of Firm Relocation

In this section, we develop a simple model of firm location decisions with two aims. First, it provides a stylized framework that rationalizes the key patterns of firm relocation documented in Section 3. Second, it offers a theoretical setting consistent with the assumptions underlying our wage decomposition in Section 5, and provides a potential interpretation of the estimated worker, firm, and location effects as reflecting productivity differences. We focus on a simple setting to fix ideas, while emphasizing that our reduced-form location effects may, in principle, capture forces beyond local productivity, such as amenities and local competition. We present a decomposition of location effects into such different forces in Section 7.3.

4.1 Model Setup

We consider a static model with N locations, a unit mass of firm owners, and L workers, each initially located at $o \in \{1, \dots, N\}$. Workers and firm owners can relocate to any location $n \in \{1, \dots, N\}$. All agents consume freely traded goods.

Workers. Each worker is endowed with an innate ability α, representing the efficiency units of labor they supply to firms. This ability is drawn from a cumulative distribution function (CDF) $F_\alpha(\cdot)$. Let $w_n(\alpha)$ denote the expected wage for a worker of type α if they choose location n. Each worker selects the location n that maximizes:

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

where κ_{on}^{w} is a bilateral migration friction between o and n (equal to zero if $n = o$), and ϵ_n^{w} is an idiosyncratic preference shock for location n, drawn from a Type-I extreme value distribution with scale parameter ξ^{w}.

Firm Owners. Each firm owner is endowed with productivity level ϕ, drawn upon entry from a CDF $F_\phi(\cdot)$. Let $\pi_n(\phi)$ denote the expected profits of a type-ϕ firm when operating from location n, which is endogenous to the local composition of workers (see Appendix C). A firm starting in location o chooses the location n that maximizes:

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

where κ_{on}^{f} is a bilateral relocation friction between o and n, and ϵ_n^{f} is an idiosyncratic preference shock for location n, drawn from a Type-I extreme value distribution with scale parameter ξ^{f}. Relocations reflect a potential trade-off between profits and idiosyncratic preferences.

Production. The output of a firm is linear in the output of its workers. Each worker's output is supermodular in their individual ability and the firm's productivity:

$$y_n(\phi, \alpha, x) = \psi_n \, \phi \, \alpha \, x,$$

where ψ_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 $F_x(\cdot)$ with mean 1. We allow ψ_n to potentially depend on the mass and composition of local workers and firms, capturing agglomeration forces in addition to exogenous regional fundamentals. For instance, regional productivity may follow the standard functional form from the agglomeration literature, $\psi_n = \bar{\psi}_n L_n^{\gamma}$.

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 vacancy is matched with a random worker at a 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 in n to the mass of workers choosing n – and η represents the elasticity of the matching function with respect to the mass of workers choosing n. Accordingly, the share of job-seekers matched with vacancies is $p_n = \vartheta_n^{1-\eta}$.

When a firm meets a job seeker, it decides whether to hire and train the worker. The hiring cost c is drawn from CDF $F_c(\cdot)$. The firm observes the worker's type α, but not the match-specific productivity shock x, which is realized at production. If the firm hires, output is produced and a constant share $1 - \beta$ is paid as wages. This can be interpreted as Nash bargaining, given zero outside options after c is sunk. Under these assumptions, the probability that a firm of type ϕ hires a worker of type α in location n, conditional on a match, is $\mathcal{P}_n(\phi,\alpha) = F_c\left(\beta\, y_n(\phi,\alpha)\right)$, where $y_n(\phi,\alpha)$ is expected output, integrating over x.[17]

4.2 Model Implications

We summarize key implications from the model, leaving derivations for Appendix C.

Proposition 1. ***Characteristics of Firm Relocation and Wages in the Model.***
Let $s^f_{on}(\phi)$ denote the share of type-ϕ firms moving from o to n, and $M_{on}(\phi)$ their mass.

(a) ***Gravity structure for firm relocation rates:*** *the share of type-ϕ 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:*

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

(b) ***The model features gross bilateral relocation flows:*** *if κ^f_{on} and κ^f_{no} are finite, then $s^f_{on}(\phi) > 0$ and $s^f_{no}(\phi) > 0$ for all o, n, ϕ. 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, $\frac{|M_{on}(\phi)-M_{no}(\phi)|}{M_{on}(\phi)+M_{no}(\phi)}$.*

(c) ***Firm spatial sorting:*** *suppose that the inverse of the hiring and training cost, c, is distributed Pareto, and that the cost of posting vacancies is a power function: $H(v) = \frac{v^{1+\delta}}{1+\delta}$. Then, $s^f_{on}(\phi)$ is log-supermodular in the firm's productivity ϕ and the location productivity ψ_n. Conditional on the origin location o, higher productivity firms select locations with higher productivity (ψ_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 α from location o chooses location n, $s^w_{on}(\alpha)$, is log-supermodular in the worker's ability α and the location productivity ψ_n. Conditional on the origin location o, higher ability workers select locations with higher productivity (ψ_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 $\zeta\mu > \sigma$, the probability that a*

[17]Note that we abstract from modeling workers' outside options. The firm pays c anticipating a share β of the surplus, as c is sunk at bargaining.

random match is turned into an employment relationship is log-supermodular in the firm productivity and the worker's ability:

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

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 α who is employed by a firm of type ϕ in location n. The wage paid to this worker is given by* $(1 - \beta)\, \psi_n \, \phi \, \alpha \, x$.

The model's predictions align closely with observed patterns of worker and plant mobility in France and the U.S. (Section 3). In particular, Propositions 1(a)-(c) are consistent with the gravity structure in Figure B.3, the large bilateral gross flows in Figure 3, and the sorting of higher-paying firms into higher-paying locations in Figure 4. Propositions 1(d)-(e) align with worker sorting and assortative matching patterns documented in the AKM literature (e.g., Card et al. (2013); Song et al. (2019); Card et al. (2025)). Although our empirical analysis does not rely on the parametric assumptions used to derive the wage equation, the model demonstrates that the log-linear wage equation is consistent with positive sorting (within and across locations) and complementarities between worker, firm, and location productivity.

5 Empirical Model and Identification

This section introduces our empirical model and its main identification assumptions. We then test these assumptions, describe the sample, and lay out the variance decomposition framework.

5.1 Statistical Model

Our main analysis is based on the following model:

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

where Y_{it} is the log gross hourly wage of employee i, working in establishment $J(i,t)$, located in commuting zone $C(i,t)$, observed at time t. The worker fixed effect, α_i, captures 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. The location fixed effect, $\psi_{C(i,t)}$, captures compensation differences between locations while controlling for worker and establishment heterogeneity. A positive $\psi_{C(i,t)}$ implies that similar workers and establishments experience higher wages in this location compared to the average location. This parameter bundles differences in infrastructure, amenities, agglomeration forces, and local competition. Notably, it also captures treatment effects that arise due to exposure to surrounding workers and firms – as in the agglomeration spillover literature. In Section 7.3 we decompose its drivers. We define locations at the level of the commuting zone in our main analysis and present a robustness check at the province level.[18]

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. We later also consider complementarities (Section 6.1) and dynamic location effects (Section 7.2).

5.2 Assumptions and Identification

We use the potential outcome framework (Rubin, 1974) to state under which identification assumptions our estimates carry a causal interpretation, extending Kline (2024) to our double-mover context. Appendix D.1 provides additional details and proofs.

Setup and notations. Let $J(i,t)$ denote the establishment employing worker $i \in \{1,...,N\} \equiv [N]$ in period $t \in \{1,...,T\} \equiv [T]$, and $C(i,t)$ their location, with establishments indexed by $\{1,...,J\} \equiv [J]$ and locations by $\{1,...,C\} \equiv [C]$. Let $L(j,t)$ denote the location of establishment j at time t, since the same establishment can be observed in different CZs. By construction, $C(i,t) = L(J(i,t),t)$. For exposition, we consider two periods and treat wages as residualized from time-varying covariates (e.g., worker and establishment age) in a previous step.

A worker's potential wage could in principle depend on their entire assignment path, and the location paths of all their employers. Following Kline (2024), we let $Y_{it}\big(\{J(i,k),\, C(i,k),\, L(J(i,1),k),\, L(J(i,2),k)\}_{k=1}^{2}\big)$ denote the potential log wage of worker i in period t as a function of the full profile of assignments in both periods $k \in \{1,2\}$: the worker's own establishment $J(i,k)$ and location $C(i,k)$, together with the locations $L(J(i,1),k)$, $L(J(i,2),k)$ of their two establishments.

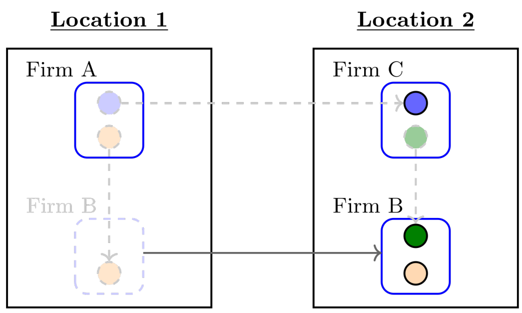
Sources of variation. Identification relies on a double-mover design, illustrated in Figure 8. The first source of variation is *worker mobility*: workers who change establishment between periods, either within a commuting zone ($J(i,1) \neq J(i,2)$, $C(i,1) = C(i,2)$, *within-location movers*) or across commuting zones ($J(i,1) \neq J(i,2)$, $C(i,1) \neq C(i,2)$, *cross-location movers*). As explained below, within-location movers identify differences in establishment pay policies, while cross-location movers provide

[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.

additional variation spanning both establishment and location differences. Worker mobility alone, however, cannot separate establishment from location effects, since a worker changing location faces a joint change in employer and location.

We leverage *establishment relocation* – establishments operating in different locations across periods ($L(j,1) \neq L(j,2)$) – to separately identify location premia from establishment pay policies. Relocation generates two distinct sources of identification. First, some workers remain at the same establishment as it relocates ($J(i,1) = J(i,2)$, $C(i,1) \neq C(i,2)$, $L(J(i,1),1) \neq L(J(i,1),2)$), referred to as *co-movers*. In Figure 8, Firm B relocates from Location 1 to Location 2 while retaining some employees, for whom the only change is the location effect. This source may, however, be affected by wage rigidity since workers may not experience a wage decrease even when their establishment relocates to a lower-paying location.[19] To address this concern, we leverage a second source of variation that does not rely on incumbents moving with their establishment: worker mobility connecting relocating establishments to the other firms in both the origin and destination locations. In Figure 8, some workers transition to or from Firm B while it is in Location 1, and others connect Firm B to the rest of the employers after it relocates to Location 2. Comparing the wage gains of workers hired at the relocating establishment in the origin versus the destination location identifies the location effect without relying on incumbents. We first estimate the model leveraging all sources of variation, then re-estimate excluding co-movers to rule out biases from wage rigidities, and find similar results.

Figure 8: Identification in a Double-Mover Design



Notes: Figure 8 illustrates the double-mover design used to identify parameters of Equation (5). Rectangles with square corners represent locations 1 and 2. Rounded rectangles correspond to establishments A, B, and C. Finally, the arrows represent the mobility of agents (workers and establishments) between establishments and locations.

Assumptions. We now discuss our five identification assumptions. Empirical tests are provided in Section 5.3.

[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.

Assumption 1 (Exclusion):

$Y_{it}(\{J(i,k), C(i,k), L(J(i,1),k), L(J(i,2),k)\}_{k=1}^{2}) = Y_{it}(J(i,t), C(i,t)), \quad \forall t \in \{1,2\}.$

Assumption 1 implies that wages are determined solely by current assignments, i.e., the establishment and location where individuals work.[20] In Section 7, we relax this assumption and allow past locations and employers to affect wages through accumulated experience.

We collect the realized assignments of a worker i into the conditioning event $S_i \equiv \{J(i,1),\ J(i,2),\ C(i,1),\ C(i,2)\}$, and we denote $s = (j, j', c, c')$ a realization.

Assumption 2 (Parallel Trends):

$$E[Y_{i2}(j,c) - Y_{i1}(j,c) \mid S_i = s] = 0, \quad \forall\, s \ \text{ s.t. } \ j \neq j' \vee c \neq c'.$$

This assumption implies that, among workers switching establishments ($j \neq j'$), average wages at their origin establishment would have been constant absent the move, irrespective of the establishment's current, past or future location. Since we include year fixed effects and an age polynomial, this amounts to requiring that potential wages at origin and destination establishments follow a common time trend. It also implies that the wage paid by a relocating establishment ($j = j'$, $c \neq c'$) to a given worker would have remained constant absent the relocation, irrespective of the worker's past and future assignments.

Assumption 3 (Stationarity):

$$E[Y_{i1}(j',c') - Y_{i1}(j,c) \mid S_i = s] = E[Y_{i2}(j',c') - Y_{i2}(j,c) \mid S_i = s],$$
$$\forall\, s \ \text{ s.t. } \ j \neq j' \vee c \neq c'.$$

The average treatment effect of moving is time-invariant, whether the move is (i) across establishments within a location ($j \neq j'$, $c = c'$), (ii) across establishments and locations ($j \neq j'$, $c \neq c'$), or (iii) within establishment across locations ($j = j'$, $c \neq c'$). This holds irrespective of the past or future location of establishments. In Section 3, we show that there is a stable establishment component upon relocation, as activity and organization remain the same.[21]

Assumption 4 (No Selection on Treatment Effects):

$E[Y_{it}(j',c') - Y_{it}(j,c) | S_i = s] = E[Y_{it}(j',c') - Y_{it}(j,c)], \quad \forall\, s \ \text{ s.t. } \ j \neq j' \vee c \neq c'.$

The treatment effect of moving between any pair of establishments, or within an establishment across locations, is independent of who selects into that move, ruling out selection

[20] Di Addario et al. (2023) show that previous employers have little influence on wages for Italian workers. Although they do not consider explicitly geography, their sample also includes workers changing locations.

[21] Lachowska et al. (2023) provide empirical evidence that establishment effects are stable over medium-run horizons.

on match-specific wage components and on establishments' location trajectories. It also implies that establishments do not select into relocating based on a specific wage effect of the location change: the treatment effect of operating in c' rather than c is independent of the relocation decision. The assumption still permits different types of establishments to move to different types of locations, and allows different types of workers to sort across establishments and locations (e.g., assortative matching).

Assumption 5 (No interactive treatment effects):
(i) $E[Y_{it}(j',c)-Y_{it}(j,c)] = E[Y_{it}(j',c')-Y_{it}(j,c')] \quad \forall(j,j') \in [J]^2, \forall(c,c') \in [C]^2,$
(ii) $E[Y_{it}(j,c')-Y_{it}(j,c)] = E[Y_{it}(j',c')-Y_{it}(j',c)] \quad \forall(j,j') \in [J]^2, \forall(c,c') \in [C]^2.$

These conditions state that the pay premium associated with each component (establishment and location) is portable across the other dimension.[22] Together, they imply that differences in expected potential wages are additively separable, as shown in Appendix D.1.[23] In Section 6.1, we allow for departures from this assumption and find a very limited role for complementarities between establishment types and locations in logs.

Appendix D.1 discusses the network conditions required for identification and introduces the normalization. First, identification requires the standard AKM connectivity condition, applied to establishment × location pairs: worker mobility must connect these pairs. Second, identifying location effects separately from establishment effects requires that relocating establishments connect locations: for each location, at least one relocating establishment must have both its establishment × location pairs in the largest connected set. In practice, we observe many such relocation within the largest connected set, ensuring that all 304 local labor markets are connected.

Proposition 2. ***Identification of causal estimands.***
Under Assumptions 1-5 and connectivity of the mobility network, establishment switches and relocations identify average treatment effects of establishments and locations separately. For any two establishments $(j,j') \in [J]^2$ and locations $(c,c') \in [C]^2$:

$$E[Y_{i2} - Y_{i1} \mid S_i = s] = E[Y_i(j',c') - Y_i(j,c)] = \phi_{j'} - \phi_j + \psi_{c'} - \psi_c,$$

where the treatment effect equals $\phi_{j'} - \phi_j + \psi_{c'} - \psi_c$ for cross-location movers; $\phi_{j'} - \phi_j$ for within-location movers ($j \neq j'$, $c = c'$); and $\psi_{c'} - \psi_c$ for co-movers ($j = j'$, $c \neq c'$).

Proof. See Appendix D.1. □

[22] In particular, relocating establishments are atomistic and do not affect location pay premia. This is consistent with their small average size (14.3 employees) relative to commuting zones.

[23] Section 4 shows that this assumption is supported by a simple search and matching model with worker and firm mobility, in which the production function features complementarities in levels for Y_{it}.

Proposition 2 also permits identification excluding co-movers. Appendix D.1 shows we can exploit establishments that relocate and hire workers in both locations. To illustrate, consider establishment j' relocating from c to c', and let worker i transition from j to j' before j' relocates, while worker i' transitions from j to j' after j' relocates. Contrasting wage changes for workers i and i' identifies the location effect: $E[Y_{i'}(j', c') - Y_{i'}(j, c)] - E[Y_i(j', c) - Y_i(j, c)] = \psi_{c'} - \psi_c$. In practice, workers i and i' need not come from the same origin establishment j, as long as locations are connected through establishment relocations and these relocating establishments are connected to the rest of the network through worker mobility in both origin and destination.

5.3 Empirical tests

We provide five sets of empirical tests in support of Assumptions 1–5.

Test 1: Event study evidence. We analyze wage dynamics before, during, and after worker and firm moves through two exercises. First, we follow Badinski et al. (2023) and study establishment (worker) wage change associated with moving to a higher-paying location, using Equation (4). Figure 7 reports the results for establishments, and Figure B.7 shows estimates for workers. We confirm the results using firm value added per hour in Figure B.8(b). Second, we stratify moves by quartiles (and by vigintiles) of mean co-worker wages, following Card et al. (2013, 2016). Figure D.1 plots residualized log hourly wages by origin and destination quartiles for workers changing establishments within the same location (Panels (a) and (b)) and across locations (Panels (c) and (d)). We revisit this quartile-by-quartile design at the establishment level around relocations (Figure D.2), where quartiles are based on the leave-one-out local average wage.

Three patterns emerge. First, across all worker and establishment moves, wages are flat prior to the move, supporting the parallel trend assumption (A2). Labor productivity is also flat prior to relocation, indicating that wages adjust in response to the location change rather than to productivity shocks or trends. Second, the adjustment is immediate at the time of the move, with minimal post-move dynamics, consistent with the exclusion and stationarity assumptions (A1, A3). Third, wage changes are approximately symmetric between moves to higher- and lower-paying locations (or establishments), supporting the no selection on treatment effects assumption (A4). Panels (b) and (d) of Figure D.1 report wage changes across the coworker wage distribution for upward and downward moves; the estimated slope is indistinguishable from one in absolute value, consistent with our log-additive model and a limited role for match effects. Although establishment-level symmetry is not required by our worker-level model, which permits heterogeneous workforce changes, we still find approximately symmetric changes: Figure 7(b) estimates Equation (4) separately for moves to higher versus lower-paying locations and yields similar coefficients in absolute terms (confirmed across quartiles in Figure D.2(b)). Establishments moving in either direction also retain a similar share of their workforce (Appendix Figure B.6).

Finally, while worker-level changes around job switches can be large – up to thirty log points – establishment-level changes are smaller, consistent with greater heterogeneity across establishments than across locations.

Test 2: Within-CZ placebo. Our framework permits a novel placebo test: examining wage dynamics for establishments that relocate *within* the same commuting zone. Panel (b) of Figure 7 reports the event-study results (see also Panel (c) of Figure D.2 for results disaggregated by quartiles of leave-one-out local average wage). Average establishment wages remain remarkably stable before, during, and after the move, consistent with establishments facing the same location effect and local workforce. This supports the parallel trend assumption (A2) and indicates that relocations are neither driven by unobserved wage shocks nor affect wages themselves.

Test 3: Model fit and residuals. Estimating Equation (5) yields an adjusted R^2 of 0.86, indicating that the model provides a good approximation of the data. We also examine the distribution of residuals across cells of worker, establishment, and location effects. Consistent with the no interactive treatment effects assumption (A5), mean residuals are systematically small, uniformly below 2% in magnitude (Figure D.3). Section 6 provides additional evidence of a limited role for complementarities.

Test 4: Survey evidence. We leverage the SINE survey of entrepreneurs to assess whether relocating establishments face specific challenges or shocks. Section 3 already shows that movers do not undergo more changes than non-movers. Panel (b) of Figure 6 shows that relocating establishments were no more likely than non-relocating ones to experience operational difficulties in their first two years, with reported challenges–workforce issues, production costs, and financial constraints–strikingly similar across movers and non-movers. Panel (a) of Figure B.5 further shows that both groups report comparable changes in local competition, with 58% of movers and 55% of non-movers indicating no change. Panel (b) confirms that movers and non-movers faced similar obstacles when opening their establishments. These results suggest that relocations are not primarily driven by shocks affecting wages or by binding constraints, consistent with the parallel trend assumption (A2).

Test 5: Subset of exogenous relocations. We show that our main results are robust to re-estimating model (5) using only establishments that relocate to their top manager's hometown province. We interpret these moves as primarily driven by managers' personal preferences rather than economic motives affecting wages, consistent with the no selection on treatment effects assumption (A4). Section 6.2 details our approach. We obtain similar results to those from the main sample.

5.4 Sample

We estimate Equation (5) on the universe of private-sector companies with at least one full-time equivalent employee over 2002–2016. Table D.1 reports summary statistics by mover status. Although movers and stayers are not identical, they overlap substantially in observable characteristics: relocating establishments average 14.3 employees (versus 14.5 for non-movers), and span all major industries. They employ a slightly higher share of skilled workers and men, with similar average worker age. About 90% of establishments relocate only once over the period.

For identification, we restrict the analysis to the largest connected set of workers, establishments, and locations (see Appendix D.1.2 for a discussion). This set comprises more than 161 million worker-year observations, covering approximately 1.8 million establishments, 22.7 million workers, and 304 commuting zones.

5.5 Variance Decomposition Framework

To decompose disparities in wages across commuting zones, we first compute averages at the commuting-zone level for 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. This approach allows us to assess the contribution of these components to wage dispersion *across* commuting zones. We then implement the following variance decomposition:

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

The variance of average wages across locations can be decomposed into the variances of the average worker, firm, and location effects, plus their covariances. We interpret the variance components as capturing (a) differences in local characteristics and intrinsic productivity (ψ_c), and (b) differences in the local composition of workers and establishments ($\overline{\alpha_c}$, $\overline{\phi_{J(c)}}$). The covariance terms capture worker and establishment co-location and sorting between locations. A positive $\operatorname{cov}(\psi_c, \overline{\phi_{J(c)}})$ (respectively, $\operatorname{cov}(\psi_c, \overline{\alpha_c})$) suggests that high-paying establishments tend to locate in high-effect areas (respectively, high-ability workers sort into high-effect locations). Finally, $\text{cov}(\overline{\alpha_c}, \overline{\phi_{J(c)}})$ reflects the degree of co-location of high-paying establishments and high-ability workers, irrespective of the location type. In what follows, we present the contribution of each component as a fraction of the total variance in average wages across commuting zones, $Var(\overline{\log(Y_c)})$.

Existing studies using fixed-effects models emphasize that limited mobility can bias estimates of variance and covariance terms (Abowd et al., 2004; Andrews et al., 2008; Kline et al., 2020). However, this concern is mitigated in our case for two reasons. First, a substantial volume of worker and establishment moves. Second, we decompose differences in wages *across* a small number of commuting zones. As discussed, we aggregate estimated fixed effects at the CZ level over a large number of workers and plants before computing the variance. With enough observations per area, the noise around CZ-level averages should be close to zero. Table D.2 reports the average number of unique workers and establishments per CZ, along with the number of movers.

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 groups using a k-means algorithm based on percentiles of the within-establishment wage distribution. Second, we apply a variance correction method using a repeated split-sample approach, described in Appendix D.4.

6 Results

6.1 Decomposing Spatial Wage Disparities

Table 2 reports the results of the variance decomposition. Most spatial variation in wages is driven by differences in establishment and worker composition. Average worker effects ($\overline{\alpha_c}$) account for 30% of the variance between CZs, and average establishment effects ($\overline{\phi_{J(c)}}$) for a further 17%. Both exhibit substantial dispersion: the 90–10 percentile gap is 18 log points for workforce composition (standard deviation 0.062) and 13 log points for establishment composition (standard deviation 0.047). By contrast, pure location effects (net of worker and firm composition) account for only 4.2% of the cross-CZ variance in log wages (2.4% in the clustered version), with a 90-10 gap of 5 log points and a standard deviation of 0.023 (see Table E.1). Figure E.1 maps average wages and the estimated fixed effects across France: both worker and establishment components closely mirror the spatial distribution of gross hourly wages, concentrating in Paris and other large cities such as Lyon, Grenoble, and Toulouse. Location-specific characteristics thus contribute little to spatial wage disparities beyond worker and establishment composition.

Covariances account for the remaining half of the variance. The co-location of high-wage workers with high-paying establishments — irrespective of the location type

Table 2: Decomposition of Spatial Wage Disparities

	(1)	(2)	(3)
	Baseline	Clustered	Variance corrected
Standard deviation of log wages	0.113	0.113	0.113
Number of observations	161,287,397	161,408,784	161,280,009
Number of workers	22,666,846	22,686,678	22,665,560
Number of establishments	1,765,604	1,867,873	1,763,172
Number of locations	304	304	304
Std. dev. of average worker effect	0.062	0.063	0.063
Std. dev. of average establishment effect	0.047	0.041	0.047
Std. dev. of location effects	0.023	0.018	0.022
Std. dev. of $X'\beta$	0.003	0.002	0.003
Correlation of average worker/establishment effects	0.733	0.972	0.725
Correlation of average establishment/location effects	0.286	0.638	0.299
Correlation of average worker/location effects	0.507	0.454	0.483
Share of variance of log CZ-wages explained by:			
Average worker effect	0.302	0.313	0.306
Average establishment effect	0.174	0.133	0.170
Location effects	0.042	0.024	0.039
Covariance of average worker and establishment effects	0.336	0.396	0.337
Covariance of average establishment and location effects	0.049	0.072	0.056
Covariance of average worker and location effects	0.114	0.079	0.116
$X'\beta$ and other covariances	-0.017	-0.017	-0.017
Residual	0.000	0.000	0.000
Adjusted R^2	0.86	0.86	.
RMSE of model	0.16	0.16	.

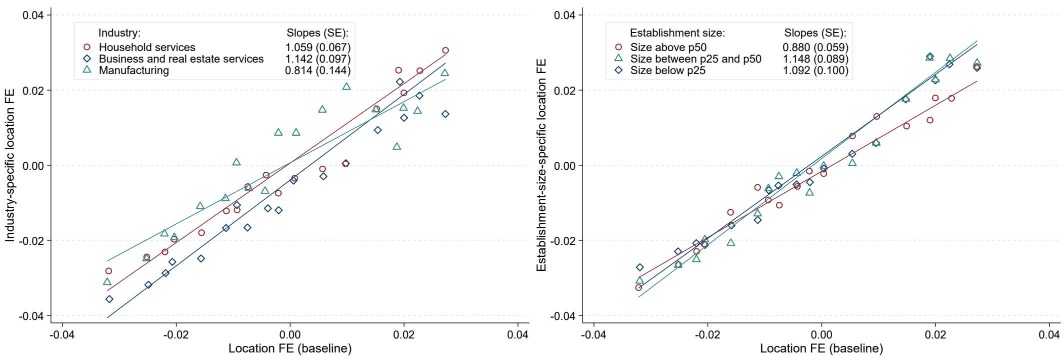
Notes: Table 2 provides the results of the variance decomposition of the average log gross hourly wage between commuting zones based on Equation (6). Column (1) reports the results of the plug-in estimator (see Equation (5)), while column (2) provides results of the model with clustered establishment 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, implemented across five independent random splits of the worker sample. X denotes time-varying observable characteristics and β the associated coefficients. Covariance terms are multiplied by two. Data: linked employer-employee data and relocation register.

— explains roughly one-third of the cross-CZ variance in log wages, making it the largest component of the decomposition. These strong co-location patterns underscore the importance of two-sided spatial sorting as a driver of wage disparities. In addition, both high-wage workers and high-paying establishments sort disproportionately into high-premium locations: the covariance between average worker and location effects explains 11% of the variance, and that between average establishment and location effects a further 4.9%. Figure E.2 illustrates the correlation between worker, plant, and location effects. The co-location of worker and firms, and their sorting into high-premium locations, may arise through three channels: (i) good locations may themselves produce high-ability workers and entrepreneurs, for example through local education or access to financing; (ii) high-ability individuals may sort into better locations when they first enter the labor market or create their firm; and (iii) workers and establishments may relocate toward better places after entry, as described in Section 3. Although pure location effects alone explain only a small share of spatial wage disparities, the sorting of workers and firms substantially amplifies these differences. These results are stable across subperiods (Figure E.3).

Section 3 showed that larger, higher-paying establishments tend to relocate to larger, higher-paying CZs. Using our estimated fixed effects, Figure E.4 plots the correlation between relocating establishments' fixed effects and the origin–destination gap in (i) location effects, (ii) average worker effects, and (iii) average establishment effects. High-fixed-effect establishments relocate primarily to areas with more skilled workers and higher-fixed-effect plants, suggesting that local human capital acquisition is an important driver of relocation decisions.[24]

Figure 9: Location Effects Allowing Heterogeneity by Establishment Characteristics

(a) By Industry (b) By Establishment Size



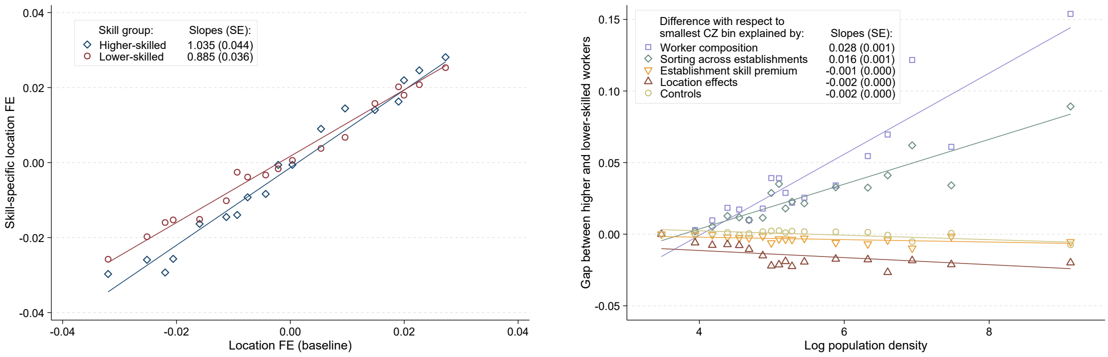
Notes: Figure 9 contrasts our baseline location estimates to industry-specific (Panel (a)) and establishment-size-specific (Panel (b)) location effects. The 25$^{\text{th}}$ percentile of establishment size is 10, and the median is 39. We report the estimated slope from regressing the characteristic-specific location effects on the baseline estimates for each establishment type. Business services are consulting, marketing, accounting and real estate. Household services include retail and accommodation services. Data: linked employer-employee data and relocation register.

Complementarities between establishment types and locations. We extend the analysis to allow location effects to vary with the size and industry of the establishment. Figure 9 shows that complementarities play a limited role along these dimensions. First, we allow for location effects specific to manufacturing, business services, and household services (Panel (a)). In all three industries, we cannot reject a slope of one with our baseline location effects: high-paying cities pay more across the board. Table E.2 reports the results of the decomposition by industry. The contribution of location effects remains consistently below 5%, although sorting patterns differ: manufacturing exhibits strong co-location of high-ability workers and high-premium establishments, but little sorting into high-premium locations, whereas in services both workers and establishments sort into high-fixed-effect areas. Second, we estimate establishment-size-specific location effects

[24]This is consistent with Nimier-David (2023), which finds that the "massification" of higher education in France in the 1990s led to the entry of many productive firms.

Figure 10: Wage Gap between High- and Lower-Skilled Workers Across Space

(a) Skill-Specific Location Effects (b) Skill Premium Gradient by Location Size



Notes: Figure 10 reports estimates of Equation (5) run separately for high-skilled workers (professionals, managers, engineers, technicians) and lower-skilled workers (clerks, blue-collar occupations). Panel (a) compares skill-specific location fixed effects with the pooled baseline. Panel (b) decomposes the gradient of the skill premium with respect to log population density, i.e., the "skill-premium density gradient." Quantities are relative to the smallest CZ, netting out sample-specific normalizations. *Worker composition* is the gap in average worker effects between skill groups. *Establishment skill premium* is the gap in average establishment effects, at lower-skill sorting. *Sorting across establishments* reflects the differential sorting, at high-skill establishment effects. *Location effects* is the gap in location effects. Components sum to the total skill premium in each location (relative to the smallest CZ). Slopes are from regressions of each component on log density; standard errors in parentheses. Data: linked employer-employee data and relocation register.

(Panel (b)). Again, we obtain a similar ranking of location effects across establishment sizes. The standard deviation of the location effect is consistently below 0.023. Figure 10 documents similar results for occupation-specific location effects.

To test for complementarities between establishments and locations driven by unobserved characteristics, we adapt Bonhomme et al. (2019) to our setting. We first estimate a two-way linear model to recover establishment-by-location effects net of worker fixed effects. We then group establishments and locations into clusters. Finally, we estimate a model with complementarities across clusters. Figure E.5 reports the results. First, wages are relatively flat across location types for all establishment types, confirming the small role of locations beyond worker and establishment composition. Second, the location gradient is remarkably parallel across establishment types, indicating limited scope for complementarities.

6.2 Robustness

We assess the robustness of our variance decomposition. First, we re-estimate the model using a set of plausibly exogenous establishment relocations. Second, we rule out potential factors that could attenuate the dispersion in location effects. Third, we explore alternative models and samples. Tables 3 and E.3 present the results.

Subset of exogenous relocations. First, we re-estimate Equation (5) using relocations plausibly driven by non-economic motives: moves to the entrepreneur's province of birth. These relocations are likely motivated by personal ties to their hometown rather than economic considerations. Because birthplace is observed only at the province level, we run the analysis across 96 provinces and group establishments into 10 clusters. For comparison, we also re-estimate the baseline model using all province-level moves. Panels (1) and (2) of Table 3, show that the province-level decomposition closely tracks its CZ counterpart, and that restricting to hometown moves leaves it essentially unchanged. Location effects account for 2.3% of spatial wage disparities using hometown relocations, 2.2% using all province-level moves, compared to 2.4% in the baseline CZ decomposition with clusters.

Factors that could mute location effects. We then implement several checks to ensure that the small variance in location effects we estimate is not an artifact of our design. First, we rule out uniform pay-setting across locations (e.g., firms set wages according to a standardized grid). Figure 7 shows that a given establishment's mean wage significantly changes after a relocation, in line with local wage levels. Figure B.8(c) confirms this finding separately for higher- and lower-skilled occupations, rejecting the idea of a fixed establishment-level pay grid across areas. Wages do adjust, but most of the response reflects changes in workers' composition (besides occupations). To address the concern that multi-establishment firms are more likely to set wages nationally (e.g., through firm-level bargaining), we also re-estimate the model using only single-establishment firms. Column (3) of Table 3 reports similar results, with location effects explaining 4.1% of wage disparities. Second, we verify that our results are not driven by wage rigidities for workers who remain in relocating establishments. We re-estimate our model on new hires, thus excluding co-movers (Column (4)). As discussed in Section 5.1, our parameters remain identified through establishment mobility between locations and worker mobility between establishments in origin and destination areas. This also addresses potential bias from compensation for longer commutes (Verdugo and Kandoussi (2024)). The clustered specification yields 2.7%, again close to the baseline. Third, we re-estimate our model using exclusively longer-distance moves (at least 50km) to ensure there is a meaningful change in labor market and find similar results (Column (7) of Table E.3). By construction, these relocations involve fewer co-movers, so this exercise also confirms that our estimates are not driven by the cost of replacing a larger share of the workforce. Finally, our approach would overstate the contribution of location effects if establishments raised their pay premium when moving to higher-paying areas and lowered it when moving to lower-paying ones. The small location effects and large establishment contribution we estimate suggest this is not the case: pay policies are preserved across relocations.

Alternative controls and clustering. Table E.3 presents variance decompositions for alternative specifications. First, controlling for a polynomial in establishment age,

Table 3: Robustness: Alternative Samples and Specifications for the Decomposition

	(1)	(2)	(3)	(4)
	Baseline Provinces	Relocation to Hometown Province	Single Establishment Firms	Without co-Movers
Standard deviation of log wages	0.112	0.112	0.105	0.105
Number of observations	156,600,282	156,444,436	82,409,166	146,175,449
Number of workers	22,351,675	22,195,829	14,125,902	21,564,702
Number of establishments	1,833,469	1,831,480	1,383,457	1,851,499
Number of locations	96	96	304	304
Std. dev. of average worker effect	0.063	0.063	0.048	0.058
Std. dev. of average establishment effect	0.041	0.041	0.042	0.039
Std. dev. of location effects	0.017	0.017	0.021	0.017
Std. dev. of $X'\beta$	0.002	0.002	0.002	0.003
Correlation of average worker/establishment effects	0.976	0.975	0.958	0.967
Correlation of average establishment/location effects	0.655	0.654	0.778	0.632
Correlation of average worker/location effects	0.486	0.481	0.576	0.431
Share of variance of log location-wages explained by:				
Average worker effect	0.315	0.313	0.211	0.308
Average establishment effect	0.132	0.132	0.163	0.136
Location effects	0.022	0.023	0.041	0.027
Covariance of average worker and establishment effects	0.398	0.396	0.356	0.396
Covariance of average establishment and location effects	0.071	0.072	0.128	0.077
Covariance of average worker and location effects	0.081	0.082	0.107	0.079
$X'\beta$ and other covariances	–0.018	–0.018	–0.006	–0.024
Residual	0.000	0.000	0.000	0.000
Adjusted R^2	0.85	0.85	0.84	0.86
RMSE of model	0.16	0.16	0.16	0.16

Notes: Table 3 reports robustness checks on our main variance decomposition. Column (1) reproduces the baseline specification with provinces, rather than commuting zones, as the location unit. Column (2) repeats this exercise using only establishments relocating to their top manager's province of birth for identification. We define the top manager as the highest-paid employee within the highest-ranked occupation group present in the establishment, where one-digit occupation groups are ranked as follows: (i) executives, (ii) managers and engineers, and (iii) supervisors and skilled technicians, excluding non-managerial occupations. We then exploit the province of birth provided in the linked employer-employee data. We exclude relocations when (a) the manager's province of birth is missing, (b) the move originates from the province of birth, and (c) the destination is not the province of birth. Column (3) excludes firms with multiple establishments. Column (4) excludes workers who remain in the same establishment after its relocation, such that fixed effects are identified only through establishment relocation and worker mobility across establishments. X denotes time-varying observable characteristics and β the associated coefficients. In all specifications, we cluster establishments using a K-means algorithm. Table E.3 presents results for additional specifications and samples. Covariance terms are multiplied by two. Data: linked employer-employee data and relocation register.

motivated by life-cycle shifts in pay-setting (Brown and Medoff, 2003; Babina et al., 2019), has negligible impact (columns (1) and (2)).[25] We also re-estimate our model excluding young establishments (below age 5 at the move) to make sure our results are not driven by establishments with low local attachment or systematic life-cycle responses and find similar results (Column (6)).[26] Second, the clustering procedure does little work. Clustering on means and standard deviations rather than percentiles (Column (3)) leaves the decomposition essentially unchanged. Varying the number of clusters from 10 to 100 (Column (4)) slightly increases the contribution of establishments and reduces that of locations, but the overall pattern is stable. Third, instead of jointly estimating all parameters, we use a two-step procedure. We first estimate worker and establishment-by-location effects via a standard AKM, then decompose the establishment-by-location

[25]E.g., young firms may back-load wages due to financial constraints (Michelacci and Quadrini, 2005).

[26]Duranton and Puga (2001) build a model in which relocations are associated with life-cycle transitions in establishments' activity, which also motivates this exercise.

effects using establishment mobility across locations.[27] Column (5) yields results nearly identical to the baseline.

Finally, we consider two alternative benchmarks. We start by extending the standard worker and location fixed-effects model to control for observable establishment characteristics (i.e., industry, age, size). Table E.4 shows that establishment characteristics explain only 0.5% of spatial disparities, indicating that firm sorting is based on unobservables and confirming the need for our design. Second, we consider a model with establishment and location effects that omits worker heterogeneity. The results are reported in column (6) of Table 4. The contribution of establishments to the total variance more than doubles, and that of location is multiplied by about 6, reinforcing the importance of accounting for both establishment and worker composition.

6.3 Connection to Prior Work

Our paper relates to two strands of research. While the existing literature asks whether a given worker would earn different wages across cities, our estimates capture a different object and allow us to tackle a different set of questions. Below, we discuss how previous approaches differ from ours, both qualitatively and quantitatively, and clarify the respective thought experiments. Finally, Table 4 shows that replicating these approaches in our sample provides comparable results despite the differing context, suggesting that our results are not specific to France.

First, the seminal approach of Glaeser and Maré (2001) and Combes et al. (2008) quantifies the average change in wages when workers move across cities, irrespective of the establishment in which they work. They use a two-way fixed effects model with worker and location effects, relying on worker mobility for identification. Column (2) reports these estimates from our sample: location-specific factors (which include establishment composition) explain 14% of the variance across CZs, while worker composition accounts for 48%, and sorting for the remaining 40%.[28]

Second, recent work by Dauth et al. (2022) for Germany and Card et al. (2025) for the U.S. extends this approach to account for changes in establishment hierarchy upon relocation. As Card et al. (2025) shows, upward mobility across locations typically coincides with downward mobility across establishments (i.e., their rank on the local job ladder). Their approach builds on the standard AKM model with worker and establishment fixed effects, the latter being aggregated at the CZ level. Their estimates hence capture the wage change for a worker moving from an average establishment in one location to an average establishment in another. Columns (3) and (4) reproduce their

[27]Step 1: $Y_{it} = \alpha_i + \phi_{J(i,t),C(i,t)} + X'_{it}\beta + \epsilon_{it}$. Step 2: $\phi_{J,C_{(J,t)}} = \phi_J + \psi_{C_{(J,t)}} + \eta_{Jt}$.

[28]We estimate the following model: $Y_{it} = \alpha_i + \psi_{C_{(i,t)}} + X'_{it}\beta + \epsilon_{it}$.

original estimates, while column (5) replicates their methodology in our sample. Despite strong differences in labor markets and geography, the results are remarkably similar across the three countries: average location-specific factors explain about 25–30% of wage disparities.

The location effect identified by our double-mover design is conceptually distinct. In both approaches above, the estimates combine the pure location effect, establishment composition, and the sorting of establishments across locations. In our model, the location effect, which we estimate to contribute only 2-4% to spatial disparities, is net of worker and firm composition. It reflects the causal effect on wages of changing location for a given worker *in the same establishment*.

Table 4: Comparison with Alternative Approaches

	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline	TWFE Worker mobility	Dauth et al. (2022)	Card et al. (2025)	à la Card et al. (2025)	TWFE Plant mobility
	FRA	FRA	DEU	USA	FRA	FRA
Standard deviation of log wages	0.113	0.113	0.098	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	12,105,712	112M	22,663,762	.
Number of establishments	1,765,604	.	880,656	.	2,023,329	2,053,805
Number of locations	304	304	204	741	304	304
Std. dev. of average worker effect	0.062	0.078	0.061	0.081	0.062	.
Std. dev. of average establishment effect	0.047	.	.	.	.	0.062
Std. dev. of location effects	0.023	0.043	0.047	0.079	0.058	0.046
Std. dev. of $X'\beta$	0.003	0.002	.	0.006	0.003	0.008
Correlation of average worker/establishment effects	0.733	.	.	.	.	.
Correlation of average establishment/location effects	0.286	.	.	.	.	0.604
Correlation of average worker/location effects	0.507	0.753	.	0.642	0.796	.
Share of variance of log CZ-wages explained by:						
Average worker effect	0.302	0.479	0.398	0.303	0.302	.
Average establishment effect	0.174	.	.	.	.	0.429
Location effects	0.042	0.144	0.236	0.293	0.265	0.237
Covariance of average worker and establishment effects	0.336	.	.	.	.	.
Covariance of average establishment and location effects	0.049	.	.	.	.	0.385
Covariance of average worker and location effects	0.114	0.395	0.417	0.382	0.450	.
$X'\beta$ and other covariances	-0.017	-0.018	-0.051	0.022	-0.017	-0.050
Residual	0.000	0.000	0.000	0.000	0.000	0.000
Adjusted R^2	0.86	0.84	.	.	0.86	0.75
RMSE of model	0.16	0.17	.	.	0.16	0.18

Notes: Table 4 compares our main estimates based on Equation (5) (column (1)) with alternative approaches. Column (2) displays the results of a two-way fixed effects model with workers and locations. The model is estimated using worker mobility between commuting zones. Column (3) reproduces the results from the Dauth et al. (2022) paper using IAB data from Germany for the period 2008-2014. Column (4) reproduces the results from the Card et al. (2025) paper based on LEHD data in the U.S. Column (5) 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 establishment fixed effects at the commuting zone level, which we refer to as the location fixed effect. Column (6) displays the results of a two-way fixed effects model with establishment and location effects. X denotes time-varying observable characteristics and β the associated coefficients. Covariance terms are multiplied by two. Data: linked employer-employee data and relocation register.

In addition to unpacking the location effect faced by workers, our approach speaks to a distinct set of questions: How important are firms for local wages, conditional on the local workforce? How large is the role of two-sided sorting in explaining spatial disparities? It also sheds light on the channels underlying agglomeration effects. Our findings suggest a limited role for mechanisms that imply a location treatment effect conditional on worker and firm composition – such as shared inputs and infrastructure, or knowledge spillovers from nearby firms – while leaving room for mechanisms that operate through changes in firms' worker composition.

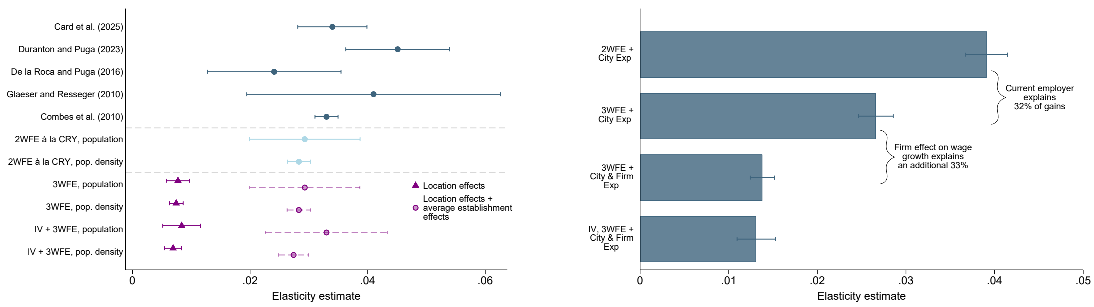
7 Size and Other Predictors of Location Effects

Our estimated location effects reflect the treatment of a location on wages, conditional on the worker and establishment identity. They can capture various theoretical forces: exogenous differences in productivity and amenities, infrastructure, compensating differentials, and agglomeration spillovers that do not operate through worker or firm composition (such as knowledge spillovers). In this Section, we first revisit the role of city size, a key suggestive piece of evidence for the productivity effects of density in the agglomeration literature. We then examine the role of other local characteristics.

Figure 11: Location-Size Premium

(a) Estimates of the City Size Premium

(b) Decomposition of Dynamic Effects



Notes: Figure 11(a) plots the slopes from regressions of estimated fixed effects on the log of city size for several measures of location size, and compares our results with alternative approaches. The first four estimates are taken from the literature. The "CRY" estimates result from our own computations, using the method in Card et al. (2025). The bottom set of estimates is obtained from our main specification in Equation (5). The triangles correspond to the pure location effects, while diamonds sum the location and average establishment effects. The last two estimates report a 2SLS version where we instrument city size with historical population from the 1789 census (see also Table F.2). Panel (b) decomposes the dynamic gains from density into the contributions of location effects and establishment sorting. These gains capture the sum of the static premium from working in a larger city and the steeper wage growth of workers in larger cities, evaluated at 7.72 years of experience following De la Roca and Puga (2016). Data: linked employer-employee data, relocation register, and census data.

7.1 Revisiting the Location Size Premium

A well-established finding in urban economics is that larger locations tend to offer higher wages, even after controlling for worker composition (Glaeser and Maré, 2001; Combes et al., 2008; De la Roca and Puga, 2016; Duranton and Puga, 2023; Card et al., 2025).[29] A key question is the extent to which this size effect arises from the local composition of establishments – e.g., workers in larger cities may work in intrinsically more productive firms, 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.

[29] See also Glaeser and Gottlieb (2009); Glaeser and Resseger (2010); Baum-Snow and Pavan (2012).

We start by estimating the raw correlation (Table F.1): regressing average local log wages on log population density, we estimate an unconditional elasticity of 0.058 (s.e. 0.0024). We then decompose this unconditional elasticity by regressing each component from Equation (5) – averaged at the CZ level – on density. The elasticity of the pure location effect is much smaller: 0.007 (s.e. 0.0006). This implies that city size positively affects wages even after controlling for the type of establishments where workers are employed, but to a much lesser extent than raw wage differences suggest. Most of the wage-density relationship reflects differences in worker and establishment composition across locations, with respective elasticities of 0.031 and 0.021. These results are robust to alternative definitions of CZ size and to instrumenting location size using historical population from the 1789 census, following Ciccone and Hall (1996) and Combes et al. (2010) (Table F.2).

Figure 11(a) compares our estimated location elasticity to results from the literature. Existing estimates of the urban wage premium range from 3–5%. Applying the Card et al. (2025) 2WFE methodology to our data yields an elasticity of 0.028, in line with the literature and three to four times larger than our estimate that controls for firm composition. Consistently, regressing the sum of our location effect and average local plant effect from our 3WFE yields an elasticity of 0.028, which confirms the major role played by establishment composition for spatial disparities.

A related key finding in the literature is a positive relationship between the college premium and city size (Davis and Dingel, 2019). Our design allows us to decompose the gradient of this premium across location density. Since education is not observed in our data, we classify workers by occupation. Panel (b) of Figure 10 shows that the skill-premium density gradient is driven by two channels: (i) the within-skill composition differs across cities (worker composition), and (ii) high-skill workers are more likely to work in more productive establishments in larger cities (assortative matching). Differences in establishment or location effects across skills contribute very little (slightly negatively) to this gradient, consistent with our main approach in which establishment and location effects are portable across workers.[30]

7.2 Dynamic Effects

Recent advances in the literature on the urban wage premium highlight the importance of dynamic effects (Baum-Snow and Pavan, 2012; De la Roca and Puga, 2016) – cities are characterized by higher wage growth rates, not just higher levels. We extend our approach to investigate the drivers of this higher growth rate and, in particular, the role played by local employers. We ask whether dynamic location premia reflect access

[30]Conclusions are similar when allowing for dynamic effects using the model described in Section 7.2.

to establishments offering higher initial wages and steeper wage profiles, rather than experience in larger locations being inherently more valuable conditional on individual and establishment type.[31] We then extend our three-way fixed effects model to capture experience accumulated in different types of locations and firms:

$$Y_{it} = \alpha_i + \phi_{J(i,t)} + \psi_{C(i,t)} + \left(\sum_{l=\{S,M,L\}} \sum_{f=1}^{K} \delta_{lfC_{(i,t)}} e_{ilft} \right) + X_{it}'\beta + \epsilon_{it} \tag{7}$$

where e_{ilft} is experience accumulated in a type-f establishment in a type-l location (small, medium or large) up to time t. We allow the return to experience, $\delta_{lfC_{(i,t)}}$, to depend on the type of establishment (f), the size of the city where it was accumulated (l), and where it is used.[32] We also control for occupation and establishment tenure.

We proceed in three steps and report estimates of dynamic elasticities—i.e., the sum of the static premium of working in a larger city and steeper wage growth—in Figure 11(b).[33] First, we estimate a two-way fixed effects model (workers and locations) with experience by city size, that abstracts from the role of employers, as in De la Roca and Puga (2016). We estimate a dynamic elasticity of 0.039 (s.e. 0.001). Second, we extend the model to account for the local composition of employers by including an establishment fixed effect. The elasticity drops by 32% to 0.027 (s.e. 0.001). This likely represents a lower bound on the contribution of employers to the city-size premium, since the inclusion of establishment effects simultaneously corrects for hierarchy bias (Card et al., 2025), which works in the opposite direction. Finally, our full model (Equation (7)) allows establishments to differ not only in pay levels but also in wage profiles, by clustering them into four groups to reduce dimensionality and allowing the return to experience to vary with establishment type, location size, and where the experience is used. Once we account for the fact that larger locations host establishments offering both higher wages and steeper wage profiles, the elasticity falls to 0.014 (s.e. 0.001) – only 35% of the initial premium. We conclude that the majority of the dynamic gains from working in larger cities come from the sorting of employers that offer higher wages and steeper wage profiles, irrespective of their location. These results explain why workers face a steeper wage ladder in larger locations (Lindenlaub et al., 2022; Lhuillier, 2025).

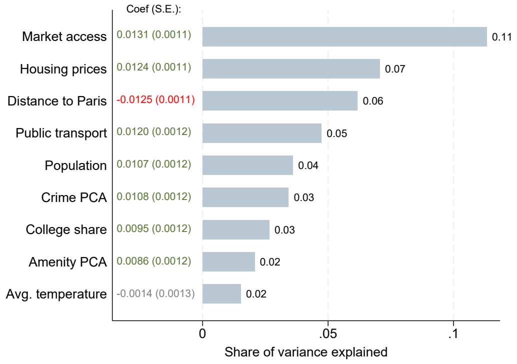
[31]To measure the number of years spent in different locations and firms, we focus on individuals younger than 25 when they enter our panel. As a first step, we re-estimate our baseline (static) model on this sample and find a comparable elasticity of 0.008 (s.e. 0.0007).

[32]Large locations are the 20 CZs of the Parisian region, and medium locations are those of Marseille, Lyon, Lille, Roubaix-Tourcoing, Toulouse and Bordeaux. About 27% of worker-years are observed in large locations and 14% in medium-sized locations, comparable to De la Roca and Puga (2016).

[33]We evaluate the gains at eight years, following De la Roca and Puga (2016).

7.3 Other Predictors of Location Effects

Figure 12: Shapley Decomposition of Location Effects



Notes: Figure 12 reports the Shapley decomposition of the R^2 from a regression of estimated location fixed effects on nine standardized regional characteristics, weighted by the number of observations per commuting zone. Continuous variables are log-transformed before standardization; the crime and amenity indices are PCA composites standardized directly. Each bar represents the fair contribution of a predictor to the overall R^2, averaging over all possible orderings of predictors. The annotation at the left of each bar reports the coefficient and standard error from a univariate weighted regression of the location effect on each predictor. Green indicates a significant positive coefficient, red a significant negative coefficient, and gray indicates insignificance (at the 10% level). Data: linked employer-employee data, relocation register, and census data.

Beyond city size, we examine which regional characteristics predict location effects using a Shapley decomposition of the R^2 from a weighted regression of estimated location effects on nine standardized (log) predictors: market access, housing prices, distance to Paris, public transport accessibility, population, college share, average temperature, and composite indices for crime and amenities.[34] While most of these characteristics are jointly determined in equilibrium, this exercise emphasizes alternative forces and provides suggestive evidence for which components are likely to drive location treatment effects.

Figure 12 reports the Shapley values, which allocate the overall R^2 across predictors accounting for collinearity. Geography is the dominant force: market access and distance to Paris are the first and third strongest predictors, accounting for 11% and 6% of the variance in location effects, respectively. This is consistent with a central prediction of the economic geography literature, where access to markets is a key determinant of local wages. Housing prices are the second strongest predictor (7%), correlating positively with location effects, which suggests that part of the location premium compensates for higher local costs (Moretti, 2013). Public transport accessibility and population also enter positively, consistent with the density premium results above and with the role of local infrastructure. Finally, location effects partly reflect compensating differentials, but to a smaller extent: crime enters with a positive coefficient and average temperature with a negative (insignificant) one. Notably, when higher wages compensate for higher

[34] The crime and amenity indices are constructed using principal component analysis. The amenity index includes per-capita levels of restaurants, schools, daycares, and classroom sizes.

housing prices and other compensating differentials have limited explanatory power, then the contribution of productivity forces to wages is bounded above by our estimated location effects.

8 Conclusion

Firm relocation is widespread and a key mechanism allocating economic activity across space. Combining linked employer-employee records with relocation data, we examine a central question in labor and spatial economics: to what extent do spatial wage differences reflect "location effects" – such as geography, infrastructure, compensating differentials, and various forms of agglomeration spillovers – rather than the sorting of workers and firms? We first characterize relocation decisions using data from France and the U.S. We then develop an approach that separately identifies the contribution of workers, firms, and locations, by leveraging worker and firm mobility in France. Beyond worker sorting, our design allows us to quantify the contribution of establishment sorting and the co-location patterns of workers and firms.

Our findings carry several implications. First, firm composition is central to local wages, consistent with policymakers' efforts to attract high-paying firms. Second, observed relocation patterns underscore the importance of local human capital in firms' location choices. Third, our results highlight the value of modeling two-sided sorting when studying spatial disparities. Fourth, they speak to the relative magnitude of agglomeration channels, suggesting small role for mechanisms that imply location treatment effects conditional on worker and firm composition, such as shared inputs and infrastructure, and knowledge spillovers from surrounding firms and workers.

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 Sébastien Pérez-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.

Alviarez, Vanessa, Javier Cravino, and Natalia Ramondo, "Firm-Embedded Productivity and Cross-Country Income Differences," *Journal of Political Economy*, 2023, *131* (9), 2289–2327.

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.

__ , Nicolas Gendron-Carrier, and Ronni Pavan, "Local productivity spillovers," *American Economic Review*, 2024, *114* (4), 1030–1069.

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.

Chikis, Craig A, Benny Kleinman, and Marta Prato, "The Geography of Innovative Firms," Technical Report, National Bureau of Economic Research 2025.

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.

Dauth, Wolfgang, Sebastian Findeisen, Enrico Moretti, and Jens Suedekum, "Matching in Cities," *Journal of the European Economic Association*, August 2022, *20* (4), 1478–1521.

Davis, Donald R. and Jonathan I. Dingel, "A Spatial Knowledge Economy," *American Economic Review*, January 2019, *109* (1), 153–70.

D'Costa, Sabine and Henry G. Overman, "The urban wage growth premium: Sorting or learning?," *Regional Science and Urban Economics*, 2014, *48*, 168–179.

Di Addario, Sabrina and Eleonora Patacchini, "Wages and the City. Evidence from Italy," *Labour Economics*, 2008, *15* (5), 1040–1061.

__ , Patrick Kline, Raffaele Saggio, and Mikkel Sølvsten, "It ain't where you're from, it's where you're at: Hiring origins, firm heterogeneity, and wages," *Journal of Econometrics*, 2023, *233* (2), 340–374.

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.

Hu, Maggie Rong, Desmond Tsang, and Wayne Xinwei Wan, "Corporate Relocation and Housing Market Spillovers," *Management Science*, 2024, *71* (5), 4344–4376.

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

Kline, Patrick M., "Firm Wage Effects," in "Handbook of Labor Economics," Vol. 5, Elsevier, 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.

Lachowska, Marta, Alexandre Mas, Raffaele Saggio, and Stephen A. Woodbury, "Do firm effects drift? Evidence from Washington administrative data," *Journal of Econometrics*, 2023, *233* (2), 375–395.

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.

Melo, Patricia C., Daniel J. Graham, and Robert B. Noland, "A meta-analysis of estimates of urban agglomeration economies," *Regional Science and Urban Economics*, 2009, *39* (3), 332–342.

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.

__ , "Real Wage Inequality," *American Economic Journal: Applied Economics*, January 2013, *5* (1), 65–103.

__ **and Moises Yi**, "Size Matters: Matching Externalities and the Advantages of Large Labor Markets," NBER Working Papers 32250 March 2024.

Nimier-David, Elio, "Local Human Capital and Firm Creation: Evidence from the Massification of Higher Education in France," 2023.

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.

Rubin, Donald B., "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies," *Journal of Educational Psychology*, 1974, *66* (5), 688–701.

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.

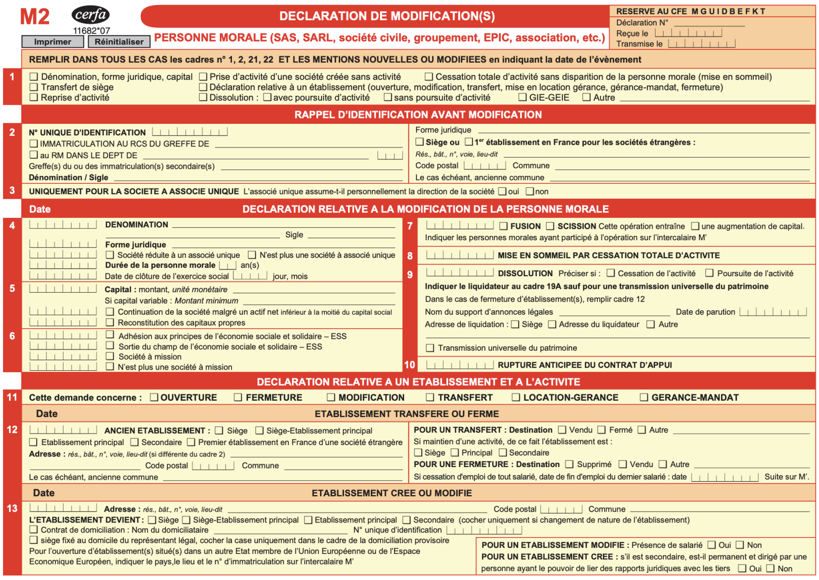
Voget, Johannes, "Relocation of Headquarters and International Taxation," *Journal of Public Economics*, 2011, *95* (9-10), 1067–1081.

Online Appendix

Contents

A Context

Figure A.1: "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

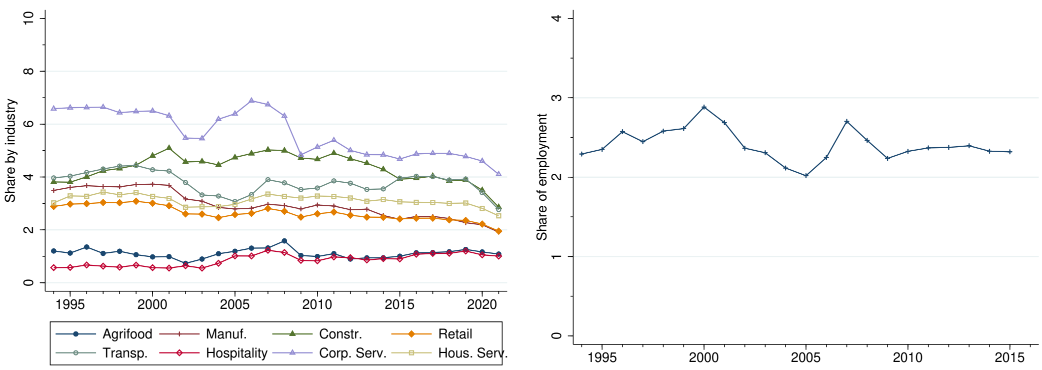
Notes: Figure A.1 presents the one-page form "M2", which entrepreneurs are required to fill out when relocating their establishment. The form mandates the completion of three key sections: boxes (1), (2), and (12). Box (1) specifies the type of change being reported, where entrepreneurs must check the 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 (12) provides the new address. Entrepreneurs must submit this form to the local commercial court, as part of the relocation process, within three months of the relocation. Additionally, they are required to include 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 by Detailed Industry and Employment

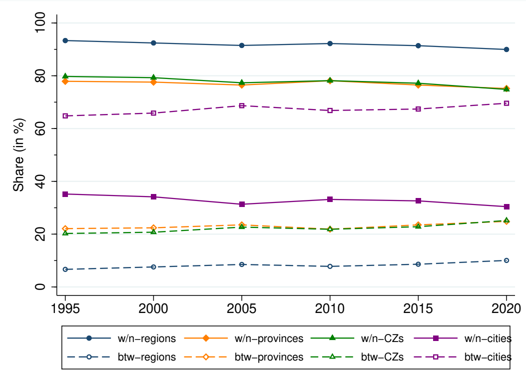
(a) All Moves, by Industry (b) Relocation Weighted by Employment



Notes: Figure B.1 plots the share of establishments moving every year by major industry (Panel (a)) and the average share of establishments moving weighted by their size (Panel (b)). Establishment size is measured the year before the move Data: linked employer-employee data and relocation register. Go back to main text

B.2 Distances of Relocations

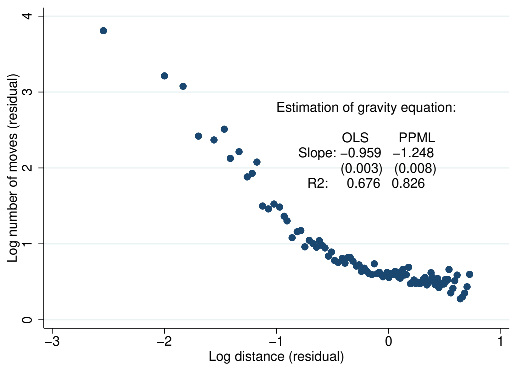
Figure B.2: Establishment Relocation by Geographical Unit



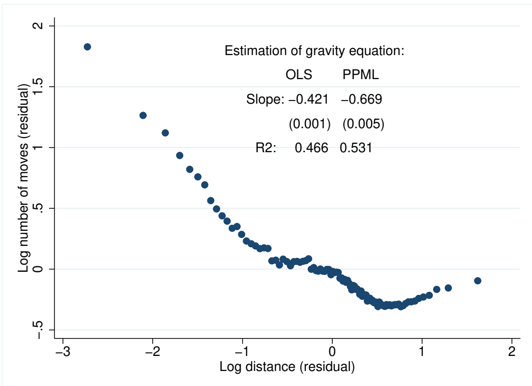
Notes: Figure B.2 reports the share of moves that took place *within* vs. *between* locations, for four types of administrative boundaries. Solid lines indicate relocations within a location, while dashed lines represent mobility across locations. The administrative units considered include 22 regions, 96 provinces (départements), 304 commuting zones, and about 35,000 municipalities. Data: relocation register. Go back to main text

Figure B.3: Gravity Patterns

(a) France



(b) U.S.



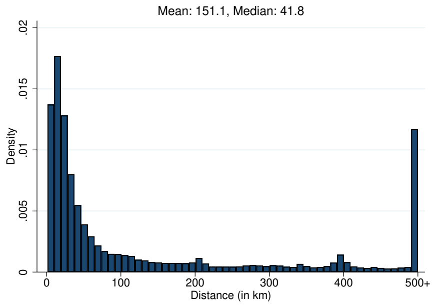
Notes: Figure B.3 is a binscatter that plots the relationship between the (log) number of establishments relocating between every pair of location and the (log) distance between these two locations. Both quantities are first residualized for origin by year and destination by year fixed effects to account for systematic differences in market access. The table reports the coefficients of a log-log regression model (column 1) and a pseudo-Poisson model (column 2). The OLS specification is: $log(\#moves_{ijt}) = \eta_{it} + \mu_{jt} + \delta\, log(distance_{ij}) + \epsilon_{ijt}$. 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 & 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
Distance to the closest CZ (in km)	18	21	26	31	32	39	45	49

Notes: Table B.1 reports the distribution of distances traveled by establishments and workers moving between commuting zones. The third row reports, for each commuting zone in France, the distribution of distances to the nearest commuting zone. This 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.4: Distribution of Distances: Between CZ Moves in France



Notes: Figure B.4 depicts the distribution of relocation distances for establishments moving across 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.3 Post-relocation Evolution of Activity

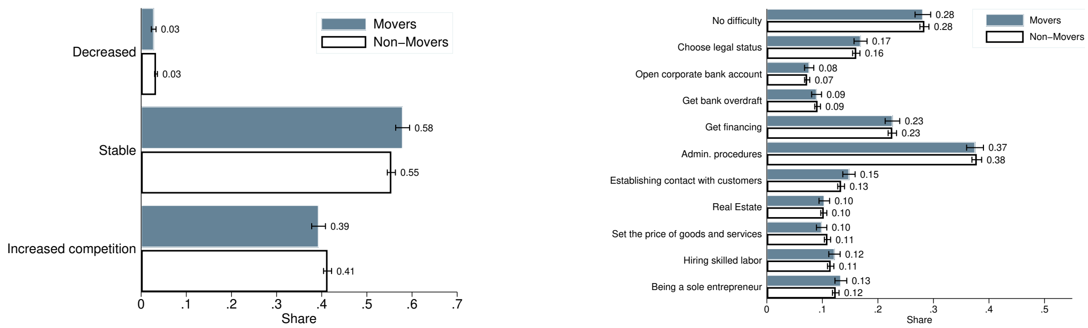
Table B.2: Establishments Preserve their Activity and Organization

	Same Industry	Preserve HQ Status	Same Legal Category	Same Top Manager
Mover x year post move	-0.047	-0.001	-0.005	-0.090
	(0.000)	(0.000)	(0.000)	(0.001)
Mover	-0.000	-0.005	0.061	-0.007
	(0.000)	(0.000)	(0.000)	(0.001)
After	-0.002	0.000	-0.057	-0.027
	(0.000)	(0.000)	(0.000)	(0.001)
Year F.E.	Yes	Yes	Yes	Yes

Notes: Table B.2 reports the estimates of a difference-in-differences specification. Each outcome, indicated in the first row, is a dummy equal to one if the establishment keeps the given characteristic from one year to the next. The specification is: $y_{it} = \delta_t + \beta_1 \text{Mover}_i \text{ x Year Post Move}_t + \beta_2 \text{Mover}_i + \beta_3 \text{After}_t + u_{it}$, where δ_t are calendar year fixed effects, Year Post Move$_t$ equals one in the year of relocation, and After$_t$ equals one in every subsequent year. For example, an establishment relocating is 4.7 percentage points less likely to keep an identical industry code the year after relocating. We focus on establishments relocating once. The standard errors are clustered at the establishment level and are reported in parentheses. Data: linked employer-employee data and relocation register. Go back to main text

Figure B.5: Additional Changes Declared by Entrepreneurs

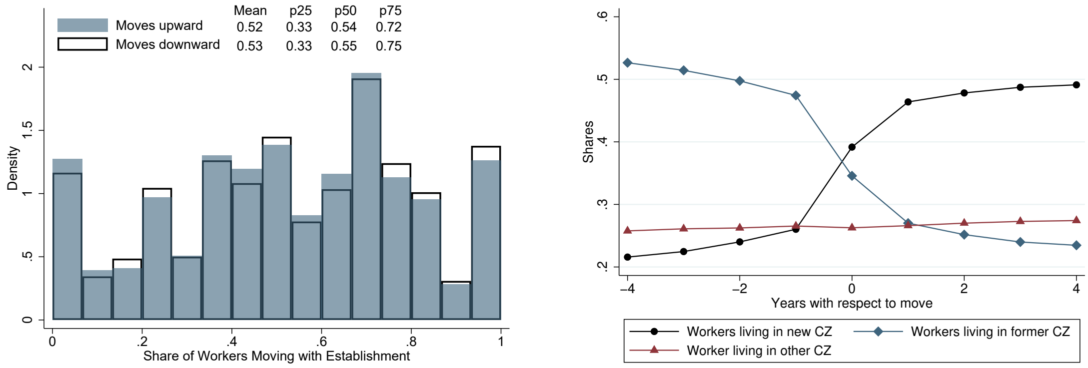
(a) Changes in Local Competition (b) Problems During Creation



Notes: Figure B.5 compares answers from entrepreneurs who relocated at least one establishment during their firm's first five years of activity, with answers of entrepreneurs who did not. We report two additional questions beyond those in Figure 6. 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.6: Composition of Workforce in Establishment Around Relocation

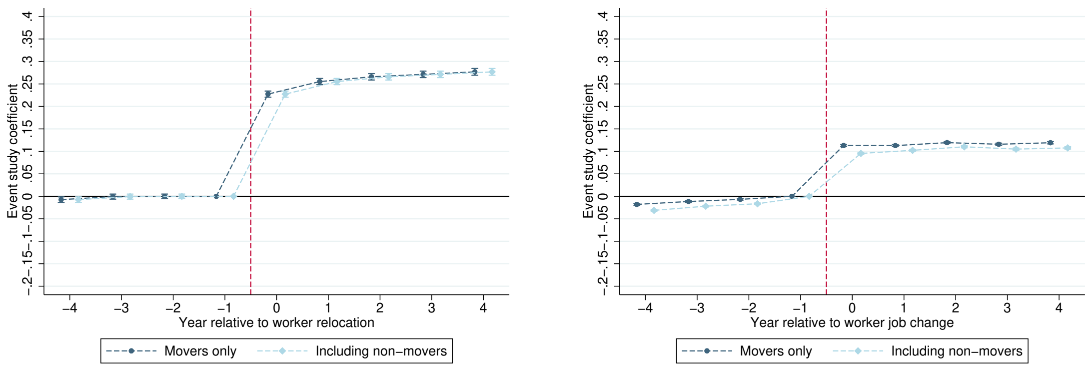
(a) Share of Workers Staying in Establishment (b) Workforce Place of Residence



Notes: Panel (a) of Figure B.6 displays the distribution of the share of workers who stay with their establishment following a move, separately for relocations to higher and lower-paying locations (relative to the CZ of origin). The share of workers who stay is measured as 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 origin, destination, 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.7: Wage Changes for Workers Changing Establishment

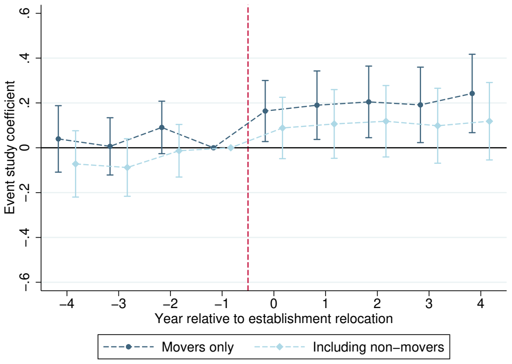
(a) Worker Relocating (and Changing Establishment) (b) Worker Changing Establishment, Within Location



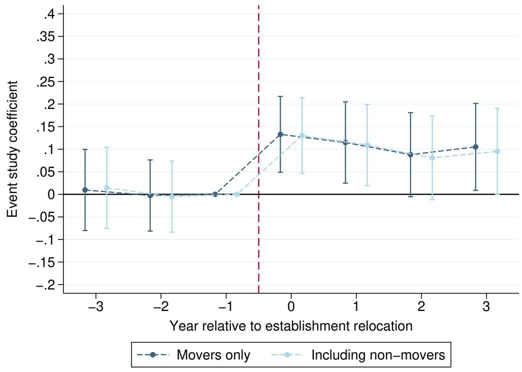
Notes: Figure B.7 plots estimates of θ in Equation (4) at the worker level, for two types of changes. The outcome is the log hourly wage of the worker. Panel (a) considers workers changing establishment to another location, and Δ is the wage gap between destination and origin CZs. Panel (b) reports workers changing establishment within a location, and Δ is the wage gap between the new establishment and the former one. Year 0 is the first year of the employer change. Each panel reports estimates on two 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

Figure B.8: Event Studies for Establishment Relocating: Size, Labor Productivity and Wages by Occupation

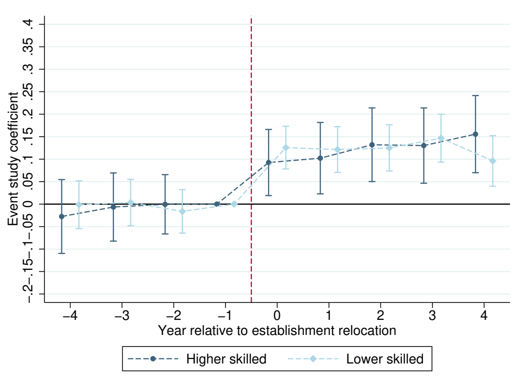
(a) Employment



(b) Labor Productivity



(c) Wages by Occupation



Notes: Figure B.8 reports estimates of θ in Equation (4), for additional outcomes: log employment (Panel (a)), log value added per hour (Panel (b)), and log occupation-specific average wage (Panel (c)). Year 0 denotes the year of the relocation. Standard errors are clustered at the establishment level. In Panels (a) and (b), estimates are reported both for the sample of relocating establishments and for a sample that also includes non-movers. Estimates in Panel (b) are obtained on the sample of single-establishment firms: because value added is reported at the firm level in our data, restricting to single-establishment firms ensures that firm-level productivity reflects the relocating establishment alone, rather than a mix of relocating and non-relocating establishments within the same firm. Finally, Panel (c) is obtained on the sample of relocating establishments separately by skill, while Figure 7 shows the corresponding results for the average wage pooling all occupations. Data: linked employer-employee data, Ficus-Fare, and relocation register. Go back to main text

C Model appendix

Firm Location Decisions and Firm Flows. Recall the firm's location decision problem:

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

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

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

Taking logs, we obtain a gravity structure:

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

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

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

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

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

Worker Location Decisions and Worker Flows. Recall the worker's location decision problem

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

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

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

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

$$L_n = \sum_{o=1}^{N}\bar{L}_o\int_0^\infty s^w_{on}(\alpha')\,\mathrm{d}F_\alpha(\alpha').$$

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

$$F_{\alpha,n}(\alpha) = \sum_{o=1}^{N}\frac{\bar{L}_o}{L_n}\int_0^\alpha s^w_{on}(\alpha')\,\mathrm{d}F_\alpha(\alpha').$$

Search and Matching. The mass of job-seekers in market n is L_n. The mass of vacancies posted in market n is $V_n = M_n\int_\phi v_n(\phi)\,\mathrm{d}F_{\phi,n}(\phi)$, where $v_n(\phi)$ is the chosen number of vacancies by firms of type ϕ 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 $q_n = \vartheta_n^{-\eta}$.

Firm Expected Profits. The expected profits to the firm from operating in market n are given by

$$\pi_n(\phi) = \max_v\left\{vq_n\int_\alpha \mathcal{P}_n(\phi,\alpha)\left(\beta y_n(\phi,\alpha)-\mathbb{E}\left[c|c<\beta y_n(\phi,\alpha)\right]\right)\,\mathrm{d}F_{\alpha,n}(\alpha)-H(v)\right\},$$

where $F_{\alpha,n}(\alpha)$ 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 $q_n \int_\alpha \mathcal{P}_n(\phi,\alpha)\, \mathrm{d}F_{\alpha,n}(\alpha)$. For each worker of type α that the firm hires, it gets a constant share of output, $\beta y_n(\phi,\alpha)$, minus the expected hiring and training cost, $\mathbb{E}\left[c | c < \beta y_n(\phi,\alpha)\right]$.

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, $\pi_n(\phi)$ equals to

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

where π_0 is a constant subsuming various parameters, and $\mathcal{A}_n \equiv \left(\int_\alpha \alpha^{1+\theta}\, \mathrm{d}F_{\alpha,n}(\alpha)\right)^{\frac{1}{1+\theta}}$ is a measure of location-n human capital, captured by a power-mean of worker 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 ϕ, the location productivity ψ_n, the endogenous local stock of human capital $\mathcal{A}_n$, and the matching rate q_n.

Worker Expected Wages. The expected wage of a worker of type α from choosing location n is $w_n(\alpha) = p_n(1-\beta)\int_\phi \mathcal{P}_n(\phi,\alpha)\, y_n(\phi,\alpha)\, \mathrm{d}F^v_{\phi,n}(\phi)$, where $F^v_{\phi,n}$ is the CDF of productivity ϕ across vacancies in location n, given by $F^v_{\phi,n}(\phi) = \frac{\int_0^\phi v_n(x)\, dF_{\phi,n}(x)}{\int_0^\infty v_n(x)\, dF_{\phi,n}(x)}$.

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,

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

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

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

where π_0 is a constant subsuming various parameters, and $\mathcal{A}_n \equiv \left(\int_\alpha \alpha^{1+\theta}\, \mathrm{d}F_{\alpha,n}(\alpha)\right)^{\frac{1}{1+\theta}}$ is a measure of location-n human capital. In this case, firm expected profits are a power function of the firm's productivity ϕ, the location productivity ψ_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

$$(1-\beta)\frac{\int_\alpha \mathcal{P}_n(\phi,\alpha)\, y_n(\phi,\alpha)\, \mathrm{d}F_{\alpha,n}(\alpha)}{\int_\alpha \mathcal{P}_n(\phi,\alpha)\, \mathrm{d}F_{\alpha,n}(\alpha)} = (1-\beta)\,\psi_n \phi \frac{\int_\alpha \alpha^{\theta+1}\, \mathrm{d}F_{\alpha,n}(\alpha)}{\int_\alpha \alpha^\theta\, \mathrm{d}F_{\alpha,n}(\alpha)}.$$

Also, $w_n(\alpha)$ equals to

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

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

$$\mathcal{Z}_n \equiv \left(\int_\phi \phi^{1+\theta} \, \mathrm{d}F_{\phi,n}(\phi) \right)^{\frac{1}{1+\theta}}.$$

In this case, we have

$$s^f_{on}(\phi) = \frac{\exp\left(\frac{1}{\xi^f} \left(\pi_0 q_n^{\left(\frac{1}{\delta}+1\right)} (\psi_n \mathcal{A}_n \phi)^{(\theta+1)\left(\frac{1}{\delta}+1\right)} - \kappa^f_{on} \right) \right)}{\sum_{n'=1}^N \exp\left(\frac{1}{\xi^f} \left(\pi_0 q_{n'}^{\left(\frac{1}{\delta}+1\right)} (\psi_{n'} \mathcal{A}_{n'} \phi)^{(\theta+1)\left(\frac{1}{\delta}+1\right)} - \kappa^f_{on'} \right) \right)},$$

and

$$s^w_{on}(\alpha) = \frac{\exp\left(\frac{1}{\xi^w} \left((1-\beta)\,\beta^\theta p_n (\psi_n \mathcal{Z}_n \alpha)^{(\theta+1)} - \kappa^w_{on} \right) \right)}{\sum_{n'=1}^N \exp\left(\frac{1}{\xi^w} \left((1-\beta)\,\beta^\theta p_{n'} (\psi_{n'} \mathcal{Z}_{n'} \alpha)^{(\theta+1)} - \kappa^w_{on'} \right) \right)}.$$

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,

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

Then, if $\zeta\mu > \sigma$,

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

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

D Empirical model and sample

D.1 Potential outcome framework

D.1.1 Additive model

Under Assumption 5(i), the wage premium of establishment j over j' is the same regardless of which location they operate in: establishment pay policies are independent of geography. Assumption 5(ii) states that the wage difference across two locations is the same regardless of the establishment at which it is evaluated.

Together, these conditions imply that differences in expected potential wages are additively separable. Fixing reference establishment $j = 1$, reference location $c = 1$, and reference worker $i = 1$, we can define $\phi_j = E[Y_{it}(j,c) - Y_{it}(1,c)]$, $\psi_c = E[Y_{it}(j,c) - Y_{it}(j,1)]$, and $\alpha_i = E[Y_{it}(j,c) - Y_{1t}(j,c)]$, each of which is well-defined as a scalar under the respective no interactive effects condition. It then follows that:

$$E[Y_{it}(j,c)] = \mu + \alpha_i + \phi_j + \psi_c \tag{A.1}$$

where $\mu = E[Y_{1t}(1,1)]$. The portability of establishment effects is supported by

Section 3.3, where we show that establishments' defining characteristics – activity, organization, and inputs – remain stable before and after relocation. Portability of location effects is tested through the homogeneity exercises in Section 6, where we find little evidence that location premia vary systematically across worker or establishment characteristics.

D.1.2 Mobility network

To characterize the mobility network, it is useful to first decompose the estimation into two steps and then contrast this two-step procedure with our baseline one-step estimation. In the first step, we estimate a standard AKM model with worker and establishment × location effects. For relocating establishments, this yields one fixed effect per location. Identification follows from worker mobility across establishment × location pairs, as in the standard AKM model. In the second step, we use establishment relocation across commuting zones to decompose each estimated establishment × location effect into an establishment and a location effect. Section 6.2 implements this two-step approach and shows that it delivers results similar to those of our baseline joint three-way fixed-effects regression.

In this context, the first step is akin to a standard AKM estimation, and the largest connected set of establishment × location pairs is defined accordingly (Kline, 2024). The only difference is that an establishment observed in two locations, because of a relocation, is treated as two distinct units. As a result, the same establishment may belong to the largest connected set in one of its locations but not in the other. Two cases determine whether both pairs of a relocating establishment are in the largest connected set. If at least one worker relocates with the establishment, the two pairs are directly linked in the worker-mobility graph; both enter the largest connected set as soon as either one connects through worker flows. If no worker relocates with the establishment, both pairs belong to the largest connected set only if each is independently connected to it through worker flows.

The second step decomposes the estimated establishment × location effects into establishment and location components, using the largest connected set of pairs from Step 1. Locations are connected by establishments that relocate between them. Because we observe a large number of relocations across only 304 commuting zones, all 304 CZs belong to the largest connected set of locations across every sample and specification we consider. The second step therefore imposes no additional sample restriction.

The largest connected set of the two-step procedure is contained in that of the one-step estimation. The two differ only for relocating establishments whose two establishment × location pairs are split by the first step—that is, exactly one pair lies in the largest connected set and the other does not. In the two-step procedure, the workers attached to the disconnected pair are dropped. In the one-step procedure, by contrast, the establishment effect is identified from the connected pair and then used to back out the worker effects of the employees in the disconnected pair. Empirically, the joint estimation retains 4,592 more observations than the two-step procedure, out of 161,287,397.

D.1.3 Identification proof

For workers changing establishments and (or) locations, define

$$S_i \equiv \{J(i,1),\ J(i,2),\ C(i,1),\ C(i,2)\},$$

and denote $s = (j, j', c, c')$ a realization. We then have

$$\begin{aligned}
E[Y_{i2} - Y_{i1}|S_i = s] &\overset{A_1}{=} E\left[Y_{i2}(j',c') - Y_{i1}(j,c)|S_i = s\right],\\
&= E\left[Y_{i2}(j',c') - Y_{i2}(j,c) + Y_{i2}(j,c) - Y_{i1}(j,c)|S_i = s\right],\\
&\overset{A_2}{=} E[Y_{i2}(j',c') - Y_{i2}(j,c)|S_i = s],\\
&\overset{A_3}{=} E[Y_i(j',c') - Y_i(j,c)|S_i = s],\\
&\overset{A_4}{=} E[Y_i(j',c') - Y_i(j,c)],\\
&= E[Y_i(j',c') - Y_i(1,1) + Y_i(1,1) - Y_i(j,c)],\\
&= E[Y_i(j',c') - Y_i(1,c')] + E[Y_i(1,c') - Y_i(1,1)]\\
&\quad - E[Y_i(j,c) - Y_i(1,c)] - E[Y_i(1,c) - Y_i(1,1)],\\
&\overset{A_5}{=} \phi_{j'} - \phi_j + \psi_{c'} - \psi_c.
\end{aligned}$$

For workers changing establishments within a location, this simplifies to

$$E[Y_{i2} - Y_{i1}|S_i = s] = E[Y_{i2} - Y_{i1}|J(i,1) = j,\ J(i,2) = j',\ C(i,1) = c,\ C(i,2) = c]$$

$$\overset{A_1-A_4}{=} E[Y_i(j',c) - Y_i(j,c)] = E[(Y_i(j',c) - Y_i(1,c)) - (Y_i(j,c) - Y_i(1,c))]$$

$$\overset{A_5}{=} \phi_{j'} - \phi_j.$$

Let us now consider an establishment j' relocating. First, we consider the case of an establishment with no co-movers, and identification therefore relies on new hires. We extend the setup to allow for three periods, t, t' and t''.

For new hires i in the origin location, we find

$$E[Y_{it'} - Y_{it}|S_i = s] = E[Y_{it'} - Y_{it}|J(i,t) = j,\ J(i,t') = j',\ C(i,t) = c,\ C(i,t') = c]$$

$$\overset{A_1-A_4}{=} E[Y_i(j',c) - Y_i(j,c)]$$

$$\overset{A_5}{=} \phi_{j'} - \phi_j.$$

Similarly, for new hires i' in the destination location:

$$E[Y_{i't''} - Y_{i't'}|S_{i'} = s'] = E[Y_{i't''} - Y_{i't'}|J(i',t') = j,\ J(i',t'') = j',\ C(i',t') = c,\ C(i',t'') = c']$$

$$\overset{A_1-A_4}{=} E[Y_{i'}(j',c') - Y_{i'}(j,c)]$$

$$\overset{A_5}{=} \phi_{j'} - \phi_j + \psi_{c'} - \psi_c.$$

Therefore, the difference between the new hires in the destination versus the new hires in origin isolates the location effects:

$$
\begin{aligned}
& E[Y_{i't''} - Y_{i't'} | S_{i'} = s'] - E[Y_{it'} - Y_{it} | S_i = s] \\
& \stackrel{A_1-A_4}{=} E[Y_{i'}(j', c') - Y_{i'}(j, c)] - E[Y_i(j', c) - Y_i(j, c)] \\
& \stackrel{A_5}{=} (\phi_{j'} - \phi_j + \psi_{c'} - \psi_c) - (\phi_{j'} - \phi_j) \\
& = \psi_{c'} - \psi_c .
\end{aligned}
$$

In practice, new hires do not have to come from the same establishment in the origin and destination locations, as long as all locations are connected via establishment mobility and that these establishments relocating are connected to the rest of the network via worker mobility both in the origin and destination locations.

Finally, for the specific case of co-movers ($i = i'$), the problem simplifies to:

$$
\begin{aligned}
& E[Y_{i2} - Y_{i1} | J(i,1) = j, J(i,2) = j, C(i,1) = c, C(i,2) = c'] \\
& \stackrel{A_1-A_4}{=} E[Y_i(j, c') - Y_i(j, c)] = E[(Y_i(j, c') - Y_i(j, 1)) - (Y_i(j, c) - Y_i(j, 1))] \\
& \stackrel{A_5}{=} \psi_{c'} - \psi_c .
\end{aligned}
$$

D.2 Sample

Table D.2: Average Number of Unique Workers and Establishments per Commuting Zone

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

Notes: Table D.2 reports the average number of unique workers and establishments per CZ, both overall and among movers, in our estimation sample. Data: linked employer-employee data and relocation register.

Go back to main text

Table D.1: Worker-Level and Establishment-Level Summary Statistics

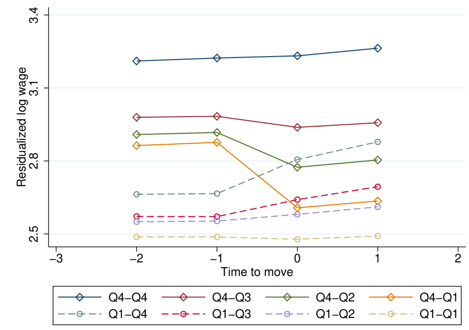
	Workers			Establishments	
	Stayers	Movers (All)	Movers (CZ)	Stayers	Movers (CZ)
Demographics					
Share Female	0.39	0.36	0.32	0.41	0.31
Worker Age	40.89	38.48	37.31	39.09	39.32
Gross hourly Wage					
Mean	18.24	18.84	19.20	16.32	19.13
S.D.	10.10	10.38	10.70	6.17	8.08
Establishment Size					
Mean	351.67	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 D.1 reports descriptive statistics for workers and establishments over the period 2002–2016. Column (1) includes workers who never change location or employer. Column (2) corresponds to workers who change employers irrespective of the location, and column (3) shows only those switching commuting zones. Column (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 establishment-averaged and then unweighted across establishments in columns (4)–(5). Data: linked employer-employee data and relocation register. Go back to main text

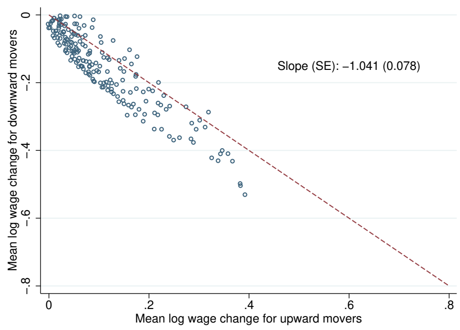
D.3 Falsification tests

Figure D.1: Change in Worker-Level Wages by Type of Move

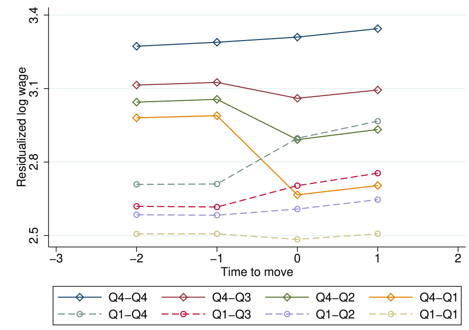
(a) Establishment Change: by Quartile of Coworker Wage



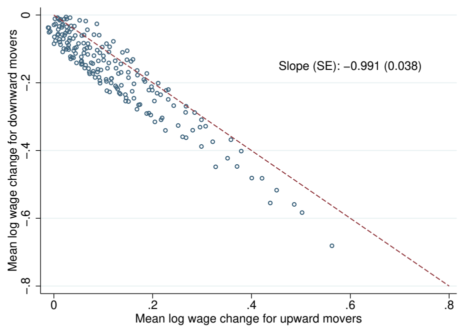
(b) Establishment Change: by Vigintile of Coworker Wage



(c) Location Change: by Quartile of Coworker Wage



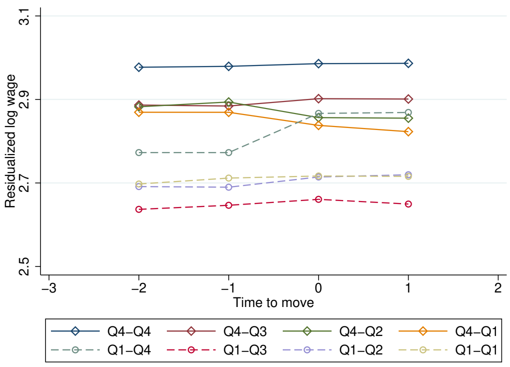
(d) Location Change: by Vigintile of Coworker Wage



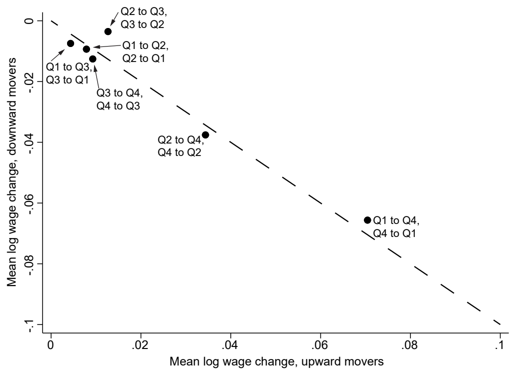
Notes: Figure D.1 plots the evolution of the average (log) hourly wage earned by workers around a job change. The wage is first residualized on year and worker age fixed effects. Panels (a) and (b) focus on workers changing establishment within the same CZ. Panels (c) and (d) focus on workers changing CZ (and establishment). We report results from two types of exercises. Panels (a) and (c) plot the average (log) wage by change in coworker-wage quartile, for a balanced sample of workers observed two years before and after the move, excluding the transition year. Time to move equal to 0 corresponds to the first full year after the move. Panels (b) and (d) plot mean (log) wage changes for job changers who move across coworker wage vigintile groups, for workers observed the year before and the year after the move. The dashed line represents symmetric changes for upward and downward movers. We report the slopes (and standard errors) weighted by the number of moves in each cell. Data: linked employer-employee data and relocation register. Go back to main text

Figure D.2: Change in Establishment-Level Mean Wages by Type of Move

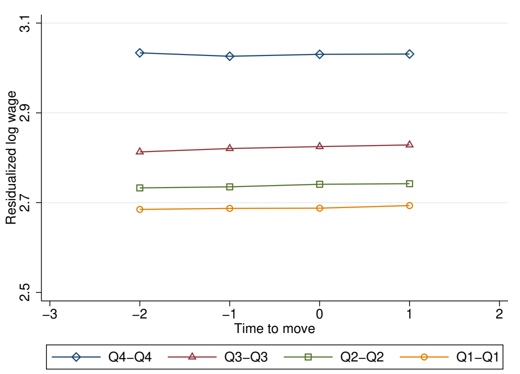
(a) Location Change: By Quartile of Leave-One-Out Local Wage



(b) Location Change: By Quartile of Leave-One-Out Local Wage

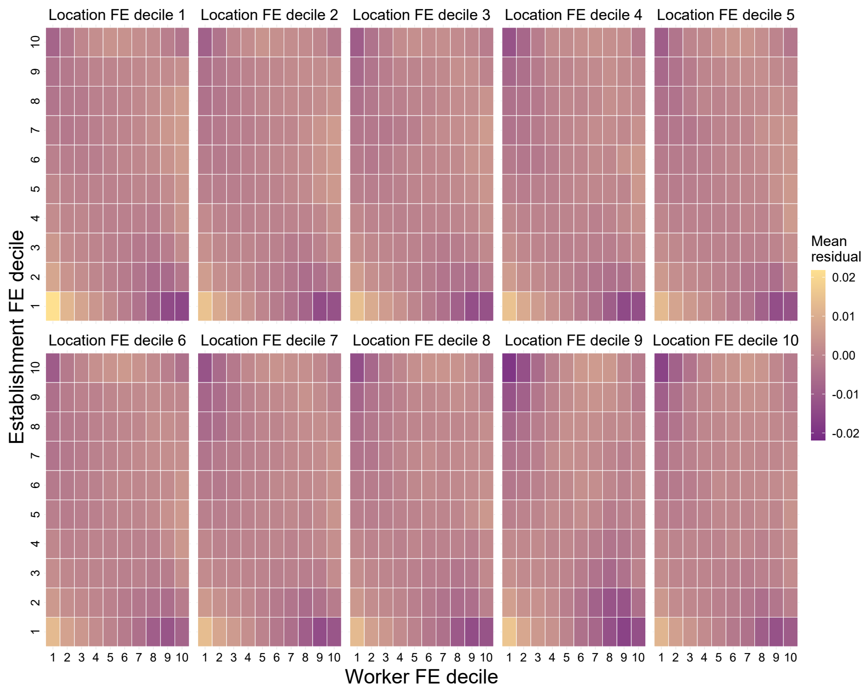


(c) Placebo: Establishments Moving Within Location



Notes: Figure D.2 plots the evolution of the average (log) hourly wage paid by establishments around a relocation. The wage is first residualized on year and worker age fixed effects before being aggregated at the establishment level. Panel (a) displays the average wage by change in quartile of leave-one-out local-wage, around a relocation. The quartiles are based on the employment-weighted leave-one-out average wage of establishments in the origin and destination CZs. Panel (b) plots the mean wage changes for establishments that relocate across the leave-one-out local wage quartile groups. The dashed line represents symmetric changes for upward and downward movers. Panel (c) focuses on establishments relocating within the same commuting zone, hence within the same quartile of leave-one-out local wage. Time to move equal to 0 corresponds to the first full year after the move. Panels (a) and (c) focus on a balanced sample of establishments observed two years before and after the move and exclude the transition year. Panel (b) includes establishments observed the year before and the year after the relocation. Data: linked employer-employee data and relocation register. Go back to main text

Figure D.3: Mean Residuals by Worker, Establishment, and Location Fixed Effect Decile



Notes: Figure D.3 shows mean residuals from our main estimates based on Equation (5) in cells defined by the intersection of estimated worker fixed effect deciles (x-axis), establishment fixed effect deciles (y-axis), and location fixed effect deciles (across panels, ordered from lowest in the top-left to highest in the bottom-right). Deciles are computed within the estimation sample, each containing an equal share of person-year observations. Data: linked employer-employee data and relocation register. Go back to main text

D.4 Variance Correction with Repeated Split Sample Approach

Variance decomposition exercises in a mover design typically overestimate the variance components and underestimate the covariances when only a moderate fraction of agents move over the period of interest. To verify that the variance and covariance components of our decomposition are not biased by limited mobility of workers and establishments, we implement a split sample approach. That said, this concern is partially mitigated in our setting, as we aggregate fixed effects to the CZ level before proceeding to the variance decomposition.

The split-sample strategy builds on Kline et al. (2020) and Babet et al. (2022). Relative to the leave-one-out estimator of Kline et al. (2020), which requires as many estimations as there are observations, the split-sample variant requires only two per split. This approach is particularly relevant in our case which features a high degree of dimensionality (three sets of fixed effects) and a large number of observations spanning several decades. Intuitively, we obtain two independent, unbiased sets of estimates for our fixed effects from two separate estimations. The covariance between these independent estimates provides information about the underlying variance or covariance. We repeat this approach five times and therefore realize five splits.

In practice, we implement this approach as follows:

1. We randomly split the 2002-2016 sample by worker identifier, so that each worker appears in only one of the two samples (A or B).

2. We estimate Equation (5) separately on each subsample, yielding estimates $(\hat{\alpha}^k, \hat{\phi}^k, \hat{\psi}^k, \hat{\beta}^k)$ for $k \in \{A, B\}$.

3. For each worker i in sample A, we construct an alternative worker fixed effect as the average residual from sample B's estimates of the remaining parameters:
$$\hat{\alpha}_i^{A(B)} = \frac{1}{n_i} \sum_t \left(Y_{it} - \hat{\phi}^B_{J(i,t)} - \hat{\psi}^B_{C(i,t)} - X'_{it}\hat{\beta}^B \right), \tag{A.2}$$
where n_i is the number of observations of worker i.

4. We aggregate the fixed effects to the CZ level by averaging them, yielding $(\bar{\hat{\alpha}}^k, \bar{\hat{\phi}}^k, \hat{\psi}^k, \bar{\hat{\beta}}^k)$ for $k \in \{A, B\}$.

5. We use information from both samples to estimate the variance and covariance terms as follows:
$$Var(\overline{\alpha_c}) = Cov(\bar{\hat{\alpha}}^A, \bar{\hat{\alpha}}^{A(B)})$$
$$Var(\overline{\phi_{J(c)}}) = Cov(\bar{\hat{\phi}}^A, \bar{\hat{\phi}}^B)$$
$$Var(\psi_c) = Cov(\hat{\psi}^A, \hat{\psi}^B).$$
Cross-component covariances are computed by taking each component from a different subsample. For example, the covariance between location and establishment fixed effects is computed as: $Cov(\psi_c, \overline{\phi_{J(c)}}) = Cov(\hat{\psi}^A, \bar{\hat{\phi}}^B)$.

To verify that our estimates do not depend on the particular random split of workers into samples A and B, we repeat the procedure five times with independent splits. Table D.3 reports the mean and standard deviation across iterations of standard deviations and correlations of the estimated AKM components. Table 2 column (3) reports the average across the five splits, along with the number of observations and units appearing in at least one iteration. The small standard deviations confirm that our estimates are stable across splits.

Table D.3: Stability of Estimates Across Sample Splits

	Mean	Std. dev.
Std. dev. of average worker effect	0.063	0.0000
Std. dev. of average establishment effect	0.047	0.0002
Std. dev. of location effects	0.022	0.0001
Std. dev. of $X'\beta$	0.003	0.0000
Correlation of average worker/establishment effects	0.725	0.0030
Correlation of average establishment/location effects	0.299	0.0213
Correlation of average worker/location effects	0.483	0.0099

Notes: Table D.3 reports the mean (column 1) and standard deviation (column 2) of each statistic across the five split-sample iterations. Data: linked employer-employee data and relocation register.

E Additional results

E.1 Main model

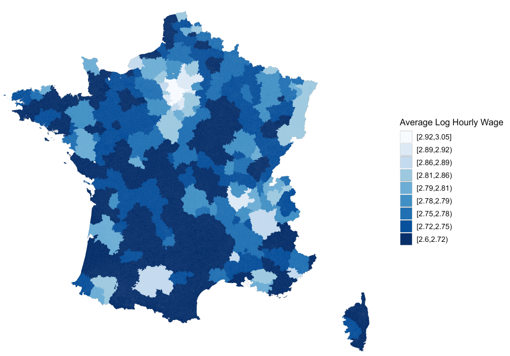
Table E.1: Distribution of Estimated Fixed Effects

	Local Worker	Local Establishment	Location
S.D.	0.062	0.047	0.023
p_5	-0.082	-0.077	-0.041
p_{10}	-0.070	-0.055	-0.027
p_{25}	-0.046	-0.035	-0.012
p_{50}	-0.016	-0.005	0.003
p_{75}	0.044	0.039	0.022
p_{90}	0.107	0.071	0.023
p_{95}	0.107	0.071	0.025

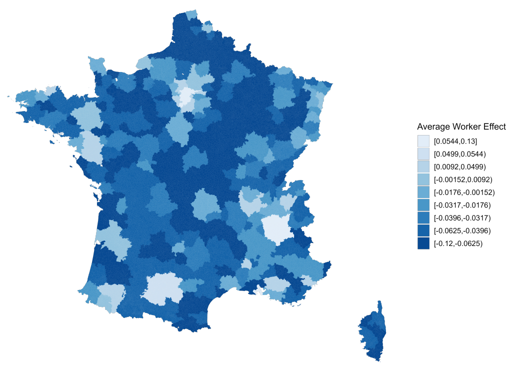
Notes: Table E.1 reports the population-weighted distribution of average estimated fixed effects for workers, establishments, and locations by commuting zone from our baseline specification. Go back to main text

Figure E.1: Spatial Disparities in Wages and Estimated Fixed Effects

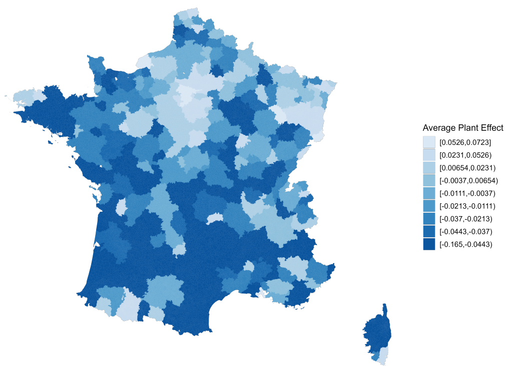
(a) Gross Hourly Wage



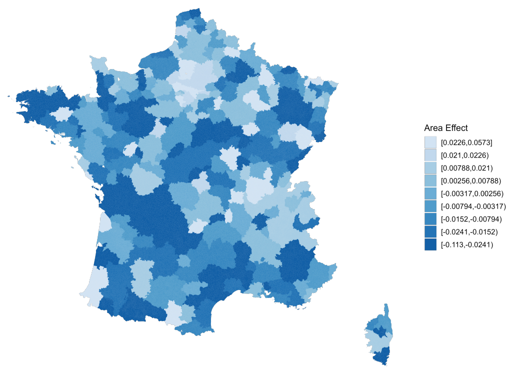
(b) Average Worker Fixed Effects



(c) Average Establishment Fixed Effects



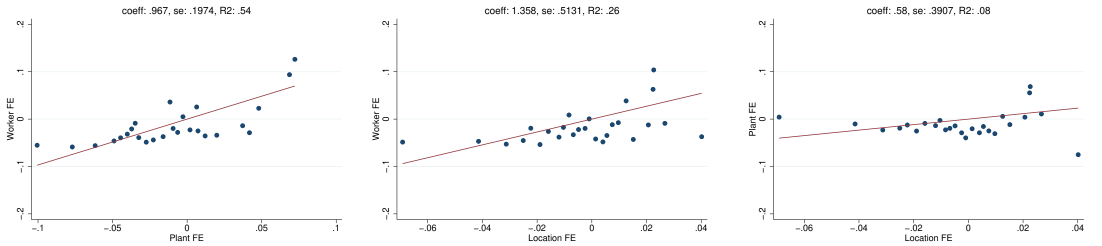
(d) Location Fixed Effects



Notes: Figure E.1 plots the distribution of wages across commuting zones. Panel (a) displays the average gross hourly wage per CZ. Panel (b), (c), and (d) plot the average worker, establishment and location fixed effects per CZ, respectively. The fixed effects are estimated based on Equation (5). Data: linked employer-employee data and relocation register. Go back to main text

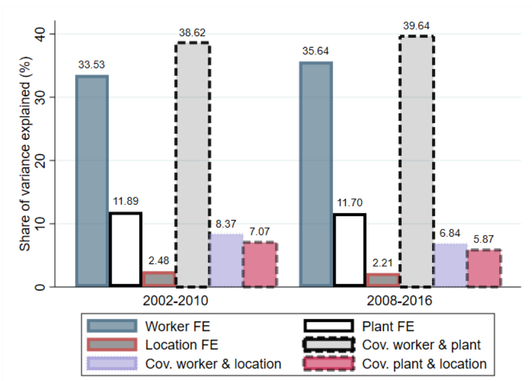
Figure E.2: Correlation Between Estimated Fixed Effects

(a) Worker and Establishment (b) Worker and Location (c) Establishment and Location



Notes: Figure E.2 shows binscatters of the correlation between average worker, establishment and location fixed effects across commuting zones. The fixed effects come from the estimation of Equation (5). The coefficients, standard errors, and R^2 are derived from a linear regression at the commuting zone level. Data: linked employer-employee data and relocation register. Go back to main text

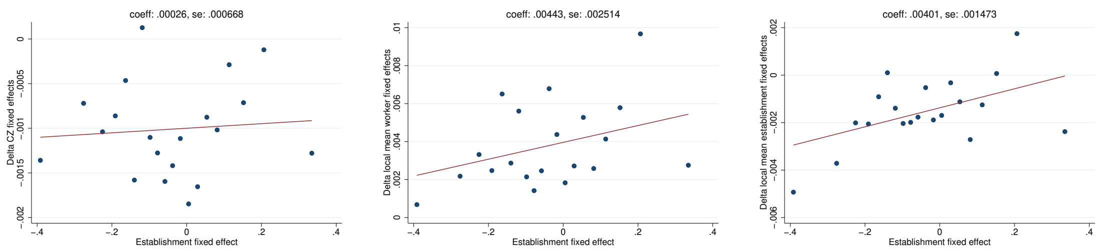
Figure E.3: Decomposition of Spatial Disparities over Two Periods



Notes: Figure E.3 presents the results of two separate decompositions, from the estimation of Equation (5) over the periods 2002-2010 and 2008-2016. Each bar depicts the share of the variance of mean wages across 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 worker composition between commuting zones explains 33.53% of the variance, and 35.64% over 2008-2016. Data: linked employer-employee data and relocation register. Go back to main text

Figure E.4: Correlation between Mover Fixed Effects and Change in Fixed Effects of:

(a) Location (b) Local Workforce (c) Local Establishments



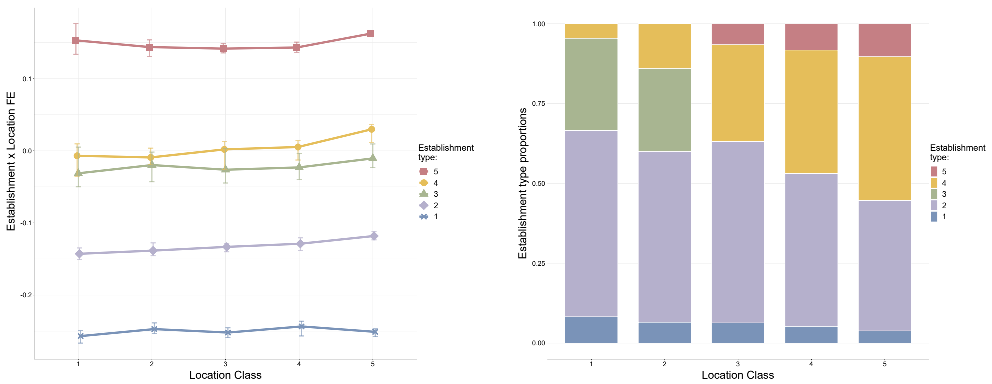
Notes: Figure E.4 depicts the correlation between the estimated fixed effect of the relocating establishment (x-axis) and the change in average fixed effects between destination and initial commuting zones (y-axis). The change considered is the difference in location fixed effects (Panel (a)), average local worker fixed effects (Panel (b)), and average local establishment fixed effects (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

E.2 Additional Robustness and Heterogeneity

Figure E.5: Assessing Establishment-Location Complementarities with BLM Estimates

(a) Average Wage Effects by Establishment Type and Location Class

(b) Proportions of Establishment Types by Location Class



Notes: Figure E.5 reports Bonhomme et al. (2019) estimation results for establishment-by-location cells. We estimate an AKM specification regressing log hourly wages on worker and establishment-by-location fixed effects along with our usual controls, before applying the static version of BLM to the estimated establishment-by-location effects, leveraging establishment mobility across locations for identification. We implement BLM's procedure with five location classes and five establishment types: locations are first grouped into classes via K-means clustering based on twenty percentiles of the within-location distribution of establishment-by-location fixed effects, and establishments are then probabilistically assigned to types using a finite-mixture model. Panel (a) reports the average log wage effect for each establishment-by-location cell, with error bars showing 95% parametric bootstrap quantile bands. Panel (b) reports the estimated proportions of establishment types within each location class. Data: linked employer-employee data and relocation register. Go back to main text

Table E.2: Three-Way Decomposition by Industry

	Baseline	Manufacturing	Business Services	Household Services
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 establishments	1,765,604	226,757	364,266	1,030,670
Number of locations	304	304	304	304
Std. dev. of average worker effect	0.062	0.099	0.098	0.046
Std. dev. of average establishment effect	0.047	0.049	0.051	0.039
Std. dev. of location effects	0.023	0.019	0.021	0.020
Std. dev. of $X'\beta$	0.003	0.003	0.006	0.002
Correlation of average worker/establishment effects	0.733	0.973	0.990	0.965
Correlation of average establishment/location effects	0.286	–0.053	0.693	0.818
Correlation of average worker/location effects	0.507	–0.255	0.595	0.658
Share of variance of log wages explained by:				
Average worker effect	0.302	0.466	0.384	0.216
Average establishment effect	0.174	0.115	0.103	0.149
Location effects	0.042	0.018	0.018	0.042
Covariance of average worker and establishment effects	0.336	0.451	0.393	0.347
Covariance of average establishment and location effects	0.049	-0.005	0.060	0.129
Covariance of average worker and location effects	0.114	-0.046	0.099	0.125
$X'\beta$ and other covariances	–0.017	0.001	-0.056	-0.008
Residual	0.000	0.000	0.000	0.000
Adjusted R^2	0.86	0.86	0.88	0.84
RMSE of model	0.16	0.16	0.17	0.16

Notes: Table E.2 reproduces the variance decomposition from Equation (6), separately by major industry. Household services include retail, accommodation, food services, and other services to households. Business services include services to firms and real estate services. X denotes time-varying observable characteristics and β the associated coefficients. Covariance terms are multiplied by two. Data: linked employer-employee data and relocation register. Go back to main text

Table E.3: Robustness Checks: Additional Specifications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Controlling for Establishment Age	Controlling for Establishment Age (Clustered)	Alternative Clustering	Clustered (100)	Two Steps	Excluding Young Establishments (Clustered)	Excluding Short Distance Moves (Clustered)
Standard deviation of log wages	0.116	0.116	0.113	0.113	0.113	0.117	0.113
Number of observations	152,483,623	152,604,961	161,367,746	161,408,784	161,282,805	132,631,776	159,878,538
Number of workers	21,838,330	21,858,355	22,681,298	22,686,678	22,666,178	20,108,520	22,565,767
Number of establishments	1,746,225	1,848,261	1,831,864	1,867,873	1,765,171	1,309,035	1,863,691
Number of locations	304	304	304	304	304	304	304
Std. dev. of average worker effect	0.063	0.065	0.063	0.062	0.062	0.066	0.063
Std. dev. of average establishment effect	0.047	0.043	0.042	0.045	0.046	0.041	0.041
Std. dev. of location effects	0.024	0.018	0.017	0.016	0.025	0.019	0.018
Std. dev. of $X'\beta$	0.004	0.003	0.002	0.002	0.002	0.003	0.002
Correlation of average worker/establishment effects	0.736	0.975	0.971	0.969	0.712	0.970	0.971
Correlation of average establishment/location effects	0.277	0.658	0.618	0.566	0.294	0.647	0.633
Correlation of average worker/location effects	0.511	0.487	0.429	0.371	0.552	0.461	0.445
Share of variance of log CZ-wages explained by:							
Average worker effect	0.297	0.313	0.313	0.302	0.302	0.321	0.312
Average establishment effect	0.166	0.134	0.136	0.154	0.164	0.124	0.133
Location effects	0.042	0.023	0.024	0.020	0.048	0.026	0.024
Covariance of average worker and establishment effects	0.327	0.399	0.401	0.418	0.317	0.387	0.395
Covariance of average establishment and location effects	0.046	0.073	0.070	0.063	0.052	0.073	0.072
Covariance of average worker and location effects	0.114	0.083	0.074	0.058	0.133	0.084	0.077
$X'\beta$ and other covariances	0.007	−0.025	−0.017	−0.017	−0.017	−0.015	−0.014
Residual	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Adjusted R^2	0.86	0.85	0.85	0.85	.	0.86	0.85
RMSE of model	0.16	0.16	0.16	0.16	.	0.16	0.16

Notes: Table E.3 reports additional robustness checks on our main variance decomposition, complementing Table 3. Column (1) reproduces the baseline specification, adding a polynomial in establishment age as a control. Column (2) repeats this exercise using clustered establishment fixed effects instead of standard establishment fixed effects. Column (3) replicates the standard clustering approach using the mean and standard deviation of wages as clustering variables, rather than percentiles of the wage distribution. Column (4) replicates the standard clustering approach with 100 clusters instead of 10. Column (5) reproduces the decomposition using a two-step estimation approach (see footnote 27). Column (6) replicates the standard clustering approach on a sample that excludes young establishments (less than 5 years old). Column (7) applies the same approach while excluding short-distance establishment relocations (moves under 50 kilometers). X denotes time-varying observable characteristics and β the associated coefficients. Covariance terms are multiplied by two. Go back to main text

Table E.4: Model with Establishment Fixed Effects vs. Establishment Characteristics

	(1)	(2)
	Baseline	TWFE with estab. controls
Standard deviation of log wages	0.113	0.116
Number of observations	161,287,397	152,604,961
Share of variance of log CZ-wages explained by:		
Average worker effect	0.302	0.481
Average establishment effect or establishment observables	0.174	0.005
Location effects	0.042	0.127
Covariance of average worker and establishment effects	0.336	0.018
Covariance of average establishment and location effects	0.049	0.009
Covariance of average worker and location effects	0.114	0.379
$X'\beta$ and other covariances	-0.017	-0.019
Residual	0.000	0.000
Adjusted R^2	0.860	0.844
RMSE of model	0.160	0.167
Implied city-size elasticity:		
Estimate	0.007	0.020
	(0.0006)	(0.0007)

Notes: Table E.4 provides a comparison of our main estimates based on Equation (5) (column (1)) with an alternative specification that replaces the establishment fixed effect with a set of controls for establishment-level observables (column (2)). The alternative model retains worker and location fixed effects, and additionally controls for two-digit industry, establishment size and firm size, and establishment age. For each specification, we report the variance decomposition of between-CZ wage disparities, along with the implied city-size premium, measured as the elasticity of the location fixed effect with respect to population density. X denotes time-varying observable characteristics and β the associated coefficients. Covariance terms are multiplied by two. Data: linked employer-employee data, relocation register, and census data. Go back to main text

F Additional Evidence on the Location Size Premium

Table F.1: 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: Local wage and size						
Log 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 effect (from 3WFE model) and size						
Average worker 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 establishment effect (from 3WFE model) and size						
Average establishment 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 F.1 reports the results of a linear regression of the wage and estimated fixed effects on different measures of location size and density: (1) log of population, (2) log of employment, (3) log of population per square kilometer, (4) log of employment per square kilometer, (5) share of college graduates, and (6) log of college graduates per square kilometer, observed in the 2015 census. The outcome is the average local wage (Panel A), the location fixed effect (Panel B), the average local worker fixed effect (Panel C) and the average local establishment fixed effect (Panel D). All fixed effects are estimated from Equation (5). The standard errors are robust to heteroskedasticity. Data: linked employer-employee data, census data, and relocation register. Go back to main text

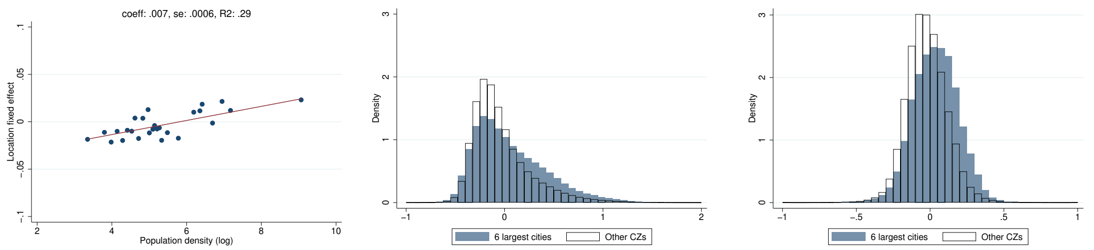
Table F.2: 2SLS Estimates of the Elasticity of the Location Effect to Density

	(1) Population (log)	(2) Employment (log)	(3) Population Density (log)	(4) 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 F.2 reports the results of a two-stage least squares regression of the estimated location effects on different measures of location size and density: (1) log of population, (2) log of employment, (3) log of population per square kilometer, and (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, relying on historical census data at the municipality level. Following the protocol outlined in Enamorado et al. (2019), we probabilistically match 1789 municipalities to today's CZs based on Jaro-Winkler distance on city names (blocking on department). All matches are manually checked. We match 198 of the 304 current CZs. Panel A reports the results of the regression of today's size on size in 1789. Panel B reports estimates of the second stage. Standard errors are robust to heteroskedasticity. Data: linked employer-employee data, census data, and relocation register. Go back to main text

Figure F.1: Location-Size Premium and Distribution of Worker and Establishment Fixed Effects by Location Size

(a) Location-Size Premium (b) Worker Fixed Effects (c) Establishment Fixed Effects



Notes: Figure F.1 relates the estimated fixed effects to commuting-zone size. Panel (a) plots the correlation between the location fixed effects estimated from Equation (5) and the log of the population per square kilometer of the commuting zone, observed in the 2015 census; the coefficient, R^2, and robust standard errors come from the corresponding linear regression. Panels (b) and (c) plot the distributions of the estimated worker and establishment fixed effects, respectively, separately for the six largest French CZs and all other CZs. Data: linked employer-employee data, relocation register, and census data. Go back to main text