Firm Localness and Labour Misallocation *

Click here for the latest version.

Gabriele Guaitoli†

University of Warwick, Department of Economics

January 14, 2024

Abstract

Limitations to workers' spatial job mobility reduce access to productive jobs, misallocating labour and lowering output and welfare. Several policies aim to mitigate this misallocation by bringing workers closer to productive firms. Nevertheless, they substantially differ in how they affect the local costs firms face. For example, reducing planning regulations in productive locations lowers local rental costs. In contrast, making productive locations more attractive increases them by fostering congestion. I show that the effectiveness of these policies crucially depends on which firms are more sensitive to changes in local costs. I call this sensitivity *localness*. Using UK microdata, I find that productivity and localness are negatively correlated. I evaluate the effectiveness of different types of policies using a spatial general equilibrium model that, as a novelty, accounts for the observed joint distribution of productivity and localness. I find that accounting for localness heterogeneity dampens, by up to 35%, the aggregate welfare gains from policies that decrease costs in productive locations. Intuitively, lower rental costs lead to the creation of low-productivity jobs rather than productive ones, as productive firms are less local. Conversely, policies that indirectly increase local costs in productive locations are more effective, since not many productive jobs are destroyed. Finally, I show that localness heterogeneity has broader implications for how these policies shape the distribution of wages.

Keywords: labour misallocation, firm heterogeneity, spatial mobility

JEL Classification: E23, J6, R1, R23, R31

*I am extremely grateful to my supervisors Roberto Pancrazi, Thijs van Rens, and Federico Rossi. I thank Christine Braun, James Fenske, Elisa Giannone, Amrita Kulka, Iacopo Morchio, Dennis Novy, Tommaso Porzio, Marta Santamaria, Marija Vukotic, Conor Walsh, and all other professors and colleagues that provided feedback on this project. All errors remain my own.

†Contact: g.guaitoli@warwick.ac.uk.

1 Introduction

Limitations to workers' spatial mobility, such as commuting costs, have detrimental effects on the economy: they reduce access to productive jobs (Hsieh and Moretti, 2019), create congestion (Schmutz and Sidibé, 2021) and increase local monopsony power (Heise and Porzio, 2023). Hence, they misallocate labour, reducing output and welfare.

Three policies are commonly proposed to mitigate these issues: attracting workers to productive regions by reducing land use regulations (Hsieh and Moretti, 2019), attracting productive firms towards deprived regions (Moretti, 2022), and directly reducing mobility costs, such as commuting costs (Monte et al., 2018). These policies affect workers' labour market outcomes both directly and by changing the local demand and costs firms face. Intuitively, attracting more consumers to a location increases local congestion, affecting firms' revenues and costs (e.g. rents) and, thus, employment decisions. However, changes in congestion may not affect all firms within a location equally.

In this paper, I argue that evaluating the effectiveness of these policies requires understanding *which* firms are the most sensitive to local congestion. I call this sensitivity *"localness"*. Firms differ considerably in how much they benefit from *local* demand (e.g. restaurants versus global tech firms) or suffer from higher *local* costs, such as rents (e.g. manufacturing plants versus offices). Therefore, firms differ in how they respond to changes in local congestion. Hence, understanding the consequences of the distribution of localness across firms is paramount when discussing policies that tackle spatial labour misallocation.

I fill this gap in two steps. First, I introduce new empirical measures of localness at the firm level. Using UK microdata, I find that low-productivity firms are approximately three times more sensitive (*"more local"*) to local demand and rental costs than high-productivity ones. Second, I quantify the effectiveness of the aforementioned policies given the observed distribution of localness across firms and locations. For this purpose, I build a spatial general equilibrium model that accounts for the negative correlation between productivity and localness. Accounting for the joint distribution of productivity and localness considerably dampens (by up to 35%) the welfare gains from policies that reduce congestion costs in productive locations, such as reducing land use regulations (Hsieh and Moretti, 2019) or moving productive firms to deprived areas. Intuitively, since productive firms have low localness, these policies have small positive externalities, limiting the creation of new productive jobs. Conversely, directly reducing spatial frictions (such as making commuting cheaper and faster) has larger welfare gains: although high-amenity regions (which, in the UK, are also high-productivity ones) would attract more residents, not many productive jobs would be lost to congestion thanks to the low localness of productive firms.

My first contribution is empirical. I use UK firm microdata to unpack localness. I focus on two components of localness that are particularly salient for spatial frictions: the exposure of a plant to local demand and the importance of local rental costs relative to value added. These dimensions are highly relevant to the study of spatial frictions, as workers' location choices contribute to local demand and rental prices through local congestion externalities.[1] I show that these components of localness are highly heterogeneous across firms and negatively correlated to productivity. Sorting firms by mean wages per worker, I find that firms representing the top 25% of employment are 70% less likely than those in the bottom 25% to mainly sell at the regional level or below. Moreover, their floorspace costs are 60% smaller, relative to their total wage bill (10% vs 28%).[2] This suggests that productive firms may be less affected by changes in the local environment.

My second contribution is theoretical. To study the labour misallocation effects of spatial frictions in the presence of congestion externalities on firms, I augment the spatial General Equilibrium search framework (Schmutz and Sidibé, 2018; Heise and Porzio, 2023),[3] with independent residence-work location choices to separate the direct labour market and indirect congestion effects of workers' location choices. Moreover, I allow firms to be heterogeneous in both

[1] Moreover, accounting for these dimensions provides a good benchmark, as they are often jointly used in spatial general equilibrium models to account for local congestion effects (on either residents, or both residents and firms).

[2] Similar results are derived using firms' size, which in most search models maps one-to-one with productivity and gross value added per worker.

[3] Lindenlaub et al. (2023) follow a similar approach, but with a continuum of location and free mobility.

productivity and localness. Hence, firms differ in how spatial frictions affect both their employment (from differences in monopsony power) and, due to congestion effects, their value added per worker (from differences in agglomeration effects, non-tradeable revenue, and floorspace costs).

I prove that the novel unobservable parameters of interest, i.e. those that characterise firms' productivity and localness, are identified given estimates of spatial frictions, search friction and amenities. As standard, I obtain those estimates from the spatial distribution of workers, job moves, wage gains, and value added per worker. Importantly, this step does not require any assumptions about productivity and localness. Then, I prove that the joint distribution of localness and value added observed in the data and the model market clearing conditions can identify the productivity and localness parameters.

My third contribution is quantitative. I estimate the model using UK data. I find that when localness heterogeneity is taken into account, spatial frictions reduce output by 26% more (1.1 percentage points of GDP) and welfare gains by 1.3 percentage points more, relative to assuming all firms have the same localness. Highly productive businesses are less exposed to congestion effects. Accounting for this fact yields smaller negative externalities from congestion in the housing market of productive locations and, thus, larger aggregate gains.[4]

Finally, the most important contribution goes towards understanding the importance of accounting for localness heterogeneity in the evaluation of public policies. I show that some policies are less effective at increasing welfare, despite the larger total misallocation of labour. Reducing housebuilding regulations (as in Hsieh and Moretti (2019)), a policy discussed in both California and the UK, leads to 35% (0.5 pp.) *smaller* gains in welfare when accounting for localness heterogeneity. Similarly, reducing the productivity gap between London and the poorest areas of the UK by one-third, a policy experiment reflected in "levelling-up" in the UK and other

[4] Since London and the South-East have very high amenities (according to both the structural estimation and data on the concentration of cultural and commercial activities), relaxing mobility constraints leads to more people choosing to live in high-amenity locations while holding jobs somewhere else, increasing local house and business floorspace prices. However, this congestion is less costly when accounting for the fact that productive businesses have low localness.

place-based policies around the World, improves welfare by 25% (0.4 pp.) less. The intuition is that policies that generate positive externalities by "decongesting" productive locations (lowering local rental costs) deliver smaller gains. This is because productive businesses are not very local and thus do not increase their labour demand even as local costs fall. The newly created jobs are, instead, low-productivity ones. On the other hand, policies that create congestion in cities deliver larger gains, as the associated negative externalities are small. In fact, policies that directly reduce the spatial frictions faced by workers, such as reducing commuting costs (e.g. through commuting subsidies or infrastructure building), are 10% more effective at improving welfare when accounting for localness heterogeneity, as they cause smaller negative externalities.

The main takeaway is that measuring and accounting for localness heterogeneity is important for correctly gauging the effects of congestion externalities. Firstly, the *aggregate* effects on labour misallocation are not obvious, both in size and sign, as they depend on the joint distribution of productivity and different dimensions of localness. Importantly, accounting for *within*-location localness heterogeneity is crucial for my results. Accounting only for between-location differences in localness misses half of the dampening of congestion externalities. Secondly, the effects of congestion on other relevant dimensions, such as the wage distribution, can be highly heterogeneous *within* and *across* locations when firms differ in localness. For example, when accounting for localness heterogeneity, high land use regulations considerably depress wages at the bottom and middle of the wage distribution.

Relation to the Literature The main contribution of this paper is to the growing literature on the misallocation of labour due to frictions that limit the spatial mobility of workers (Frey, 2009; Marinescu and Rathelot, 2018; Hsieh and Moretti, 2019; Herz and van Rens, 2019; Elsby and Gottfries, 2021; Bilal et al., 2022; Martellini, 2022; Bilal, 2023; Auerbach et al., 2023).[5] I show - through an agnostic decomposition - how assessing the costs of spatial frictions requires

[5]The literature finds that spatial mismatch unemployment is small (Şahin et al., 2014; Manning and Petrongolo, 2017; Marinescu and Rathelot, 2018; Herz and van Rens, 2019). However, large effects are found in the literature accounting for the mismatch into *less productive* jobs due to spatial differences in the distribution of productivity Hsieh and Moretti (2019); Elsby and Gottfries (2021); Auerbach et al. (2023); Heise and Porzio (2023).

accounting for how spatial frictions change the value added per worker and size (employment) of firms. This is reminiscent of the concepts of spillovers and congestion in Fajgelbaum and Gaubert (2020), but at the firm level. That is, I allow for heterogeneity in congestion effects *within* locations, a potentially relevant margin of misallocation when firms differ in productivity.[6]

Heise and Porzio (2023) show how labour is misallocated within locations due to heterogeneity in local monopsony power across firms ("size effect", in terms of my decomposition). My and their results are closely related to contributions that study the effects of local monopsony power (Azar et al., 2019; Bamford, 2021; Datta, 2023; Caldwell and Danieli, 2023). I build on their spatial general equilibrium framework with job-to-job transitions (based on Burdett and Mortensen (1998)), showing how (heterogeneous) congestion spillovers can further misallocate labour between and within locations by affecting firms' value added per worker, and thus profitability and labour demand along the productivity distribution.[7]

My work is also related to the literature studying how spatial frictions affect inequalities by shaping the spatial allocation of skills and firms (Moretti, 2013; Diamond, 2016; Schmutz and Sidibé, 2021; Diamond and Gaubert, 2022; Card et al., 2023).[8] I show how firms mainly employing low-skilled workers have different localness than those mainly employing high-skilled workers. This fosters the spatial sorting of jobs, as the negative congestion effects of cities affect productive low-skilled jobs more than high-skilled ones. These results also relate to the urban-biased growth literature, showing how productive manufacturing employment has shrunk

[6]However, in this paper, I focus on the role of spatial frictions with respect to workers' mobility rather than frictions with respect to shipping of goods and barriers (or congestion effects) to regional trade.

[7]Mechanically, other contributions fail to account for this channel for two reasons. Either due to assuming zero effects of congestion on firms' value added, as in Heise and Porzio (2023), or due to assuming identical effects along the productivity distribution, as in Bilal (2023) (who uses a Cobb-Douglas for land). van der List (2023) uses, in an extension, a location-specific, but identical across firms, land cost per employee. Similarly, Fajgelbaum and Gaubert (2020) use a model of between-location congestion spillovers, but no localness heterogeneity. Exceptions are contributions that discuss "dilation" effects of agglomeration economies, such as De La Roca and Puga (2016) and Combes et al. (2012). Lindenlaub et al. (2023) assume firms use one unit of land each, thus creating a negative correlation. I complement these contributions by discussing heterogeneous value-added effects arising from susceptibility to local revenues and cost and calibrating the distribution of localness to the data.

[8]Even when not explicitly stated, assuming that workers need to live where they work, hence having to face higher local prices to work in big cities, implicitly assumes infinitely high commuting costs, hence frictional worker mobility.

in cities but not in less congested locations (Chen et al., 2023).

Finally, I contribute to the literature studying policies aimed at improving the allocation of labour across and within locations (Saks, 2008; Ahlfeldt and Feddersen, 2017; Herkenhoff et al., 2018; Monte et al., 2018; Hsieh and Moretti, 2019; Fajgelbaum and Gaubert, 2020; Gaubert et al., 2021; Moretti, 2022). I find that the relative effectiveness of different policies is profoundly affected by localness heterogeneity. Relative to models with congestion spillovers on firms, I find that some policies become considerably less effective at reducing misallocation, as very productive firms have low localness. However, these results should be interpreted in the opposite way when compared to a model where congestion does not affect value added, as congestion has *some* effects on productive businesses too. In this sense, this paper provides not only quantitative estimates of relative policy effectiveness, but also a firm-level mental framework to think about the misallocation effects of spatial frictions.

Outline of the paper. The rest of the paper is organised as follows. In Section 2, I provide empirical evidence of the relevance of spatial frictions. I then define localness, and its relevance for understanding spatial labour misallocation, and provide empirical evidence of its heterogeneity. In Section 3, I rationalise this evidence using a structural model. In Section 4, I discuss the identification and estimation of the model and present the results in Section 5. Section 6 discusses the policy implications of this framework. Section 7 takes stock of the findings.

2 Data and Motivating Facts

2.1 Data and Definitions

I employ three main data sources from the UK: the Annual Survey of Hours and Earning (ASHE); the Annual Business Survey (ABS), and the National Employer Skill Survey. All of these datasets are linkable through firm identifiers. I now describe in detail each of these datasets and their use in this work.

ASHE. ASHE[9] is a matched employer-employee panel dataset, with sample size equivalent to 1% of the population.[10] It provides detailed information on pay and hours, occupation, contract type (part-time, apprenticeship, youth pay rates, covered by collective agreements) and demographic information about employed individuals. Importantly, it contains information on both the location of residence and employment. I use ASHE to track workers' earnings across space and their spatial mobility patterns.

ABS. ABS[11] is composed of two parts. The first is a yearly dataset of the universe of plants (local units), with information on employment, revenues, and activity status. The other is a large yearly survey (60,000 responding units) with detailed information on firms' activities, such as employment costs, value added, and taxation. I use this dataset to obtain additional information about the employers in ASHE, as well as to estimate the value added and floorspace costs of each firm.[12]

NESS. NESS[13] is a survey of individual plants, with around 90,000 plants being interviewed every odd year. The survey collects detailed information on employment and vacancies by occupation. Half of these plants are asked questions about the nature of their activity, including whether they obtain their revenue from local sales.

2.2 Firms and wages across space

Among high-income countries, the UK has one of the most unequal spatial distributions of income and production (Arnold et al., 2019). As illustrated in Figure 1, the GDP per capita of the North of England is 50% smaller than London's and one-third smaller than that of the South-

[9] Office for National Statistics, released 11 May 2023, ONS SRS Metadata Catalogue, dataset, Annual Survey of Hours and Earnings Longitudinal - UK, https://doi.org/10.57906/nz2h-kc10

[10] The panel is selected according to the last two digits of the national insurance number (NIN). All individuals born in England receive a NIN on their 16th birthday. Immigrants are included if they have applied for a NIN.

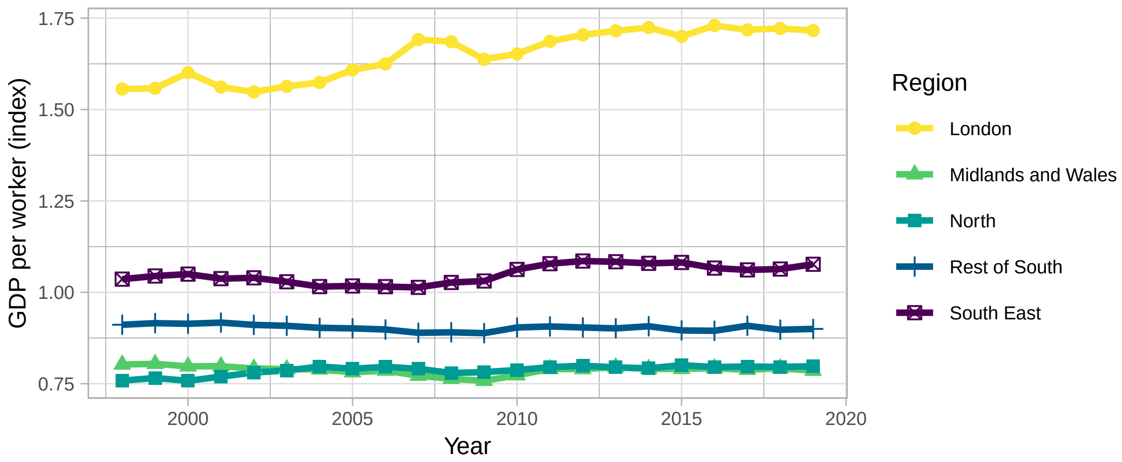
[11] Office for National Statistics, released 27 July 2023, ONS SRS Metadata Catalogue, dataset, Annual Business Survey - UK, https://doi.org/10.57906/ks2s-qx24

[12] The floorspace is obtained from the "non-domestic (business) rates" tax, a tax paid in proportion to the current market value of rental prices for that type of firm in a narrow geographical area.

[13] Department for Education, released 06 March 2023, ONS SRS Metadata Catalogue, dataset, Employer Skills Survey and Investment in Training - UK, https://doi.org/10.57906/zs9p-gw83

East region. Furthermore, these regional gaps have not narrowed for the last 25 years.

Figure 1: GDP per capita (England and Wales = 1)



Note: regional GDP per capita (in volume at chained 2019 prices), relative to the England and Wales mean. The graph covers 1998 to 2019. "North" includes "North East", "North West", and "Yorkshire and The Humber" regions. "Midlands and Wales" includes "East Midlands", "West Midlands" regions and "Wales". "Rest of South" includes "East" and "South-West" regions.

These phenomena have been linked to the presence of limitations to workers' spatial mobility ("spatial frictions"), such as commuting or relocation costs. Even though part of the observed inequalities reflects sorting,[14], spatial frictions prevent workers from moving to better but distant jobs (Hsieh and Moretti, 2019) and further reduce average productivity in less productive regions (Heise and Porzio, 2023).

I visualise one of the indirect consequences of spatial frictions by estimating the compensating differential of longer commutes. I regress the wage gains of job movers i, moving at time t, against the change in their commuting distance:

$$\Delta \ln(w_{it}) = \sum_{s \in \mathbf{S}} \beta^s d^s_{it} + BX_{it} + \varepsilon_{it}$$

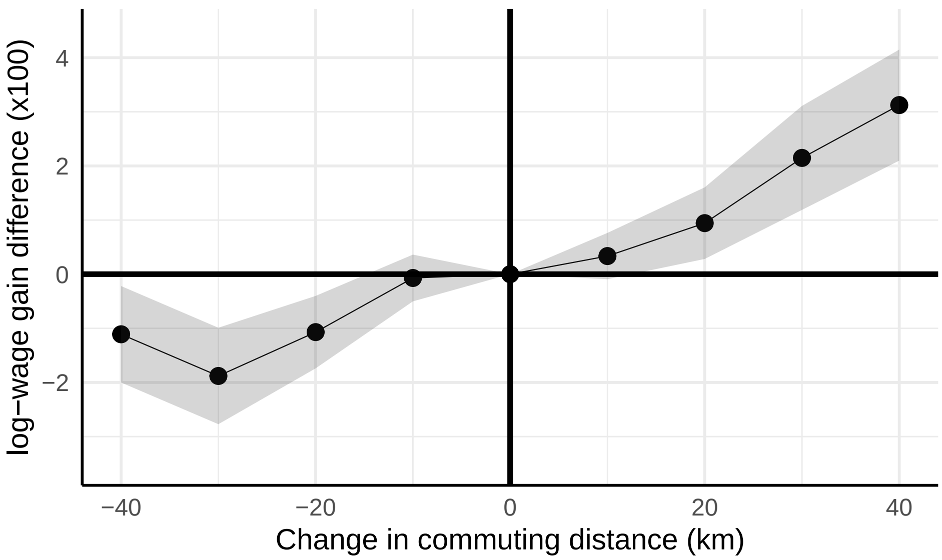
where d^s_{it} is a set of dummies representing nine groups of commuting distance changes (no change, 0-10km, 10-20km, 20-30km, 30-40km, and the equivalent negative intervals), and X_{it} are controls.[15] The coefficients of interest, $\{\beta^s\}_{s \in S}$, represent the mean wage premium of

[14]See for example models of natural advantage such as Ciccone and Hall (1996) and model of endogenous spatial sorting (Diamond, 2016; Bilal, 2023; Kuhn et al., 2022; Lindenlaub et al., 2023; Schmutz and Sidibé, 2018)

[15]Commuting distance is measured as the great circle distance between work and residence MSOAs. The controls include a sex dummy, a set of age dummies in 5-years intervals, an interaction between pre- and post-move 2-digits occupation, a dummy indicating whether the worker also moved residence during the job move, and

changing commuting distance relative to the omitted category ($d^s_{it} = 0$). I plot the results in Figure 2. On average, increasing commuting distance by 40km is associated with a 0.031 log-points higher wage gain, relative to workers who do not change commuting distance.

Figure 2: Wage gains of job movers, by change in commuting distance



Note: The figure plots the (residualised) mean log-wage gains of job movers by their change in commuting distance after the move. The mean estimate for the omitted category ($x = 0$) is 0.018 log-points. Grey bands represent 95% confidence intervals. Each point represents one β^s. The horizontal positions of the points represent the lowest value in the commuting distance change bin.

This evidence is consistent with large frictions to workers' spatial mobility. However, to understand how spatial frictions impact the *aggregate* allocation of labour, it is necessary to understand how such choices affect firms. For example, through their residence location choices, individuals can increase local residential and business rents, demand for local services and local labour supply. Hence, by affecting local demand, prices and costs, spatial frictions can affect firms' profitability and employment decisions. In the next section, I present a general approach to assess the misallocation effects of spatial frictions.

2.3 Decomposing labour misallocation

I propose a decomposition of the misallocation of labour induced by workers' frictional spatial mobility. Dividing the economy into $j \in \mathscr{L}$ locations, aggregate value added Y can be expressed

year fixed effects.

as the sum of local value added Y_j:

$$Y \equiv \sum_{j \in \mathscr{L}} Y_j.$$

Assuming there are N_j firms in location j, indexed by $n \in \{1, \dots, N_j\}$, each with value added per worker $\tilde{p}_{j,n}$ and employment $e_{j,n}$ (so that $Y_{j,n} = \tilde{p}_{j,n} e_{j,n}$),[16] Y can be expressed as:

$$Y = \sum_{j \in \mathscr{L}} \sum_{n=1}^{N_j} \tilde{p}_{j,n} e_{j,n}.$$

Hence, the question of whether labour is misallocated can be answered by asking: (how much) does value added change when making infinitesimal changes in spatial friction x (for example, commuting costs)?

Assume $\tilde{p}_{j,n}(x)$ and $e_{j,n}(x)$, the functions mapping spatial frictions x into the value added per worker and size of firm (j,n), are differentiable in $x \in X \subseteq R$. Then, the change in Y following an infinitesimal change in spatial friction x is[17]

$$\frac{\partial \ln(Y)}{\partial x} = \sum_{j \in \mathscr{L}} \sum_{n=1}^{N_j} \underbrace{\frac{Y_{j,n}}{Y}}_{\substack{\text{Firm} \\ \text{weight}}} \left(\underbrace{\frac{\partial \ln(\tilde{p}_{j,n}(x))}{\partial x}}_{\substack{\text{Value added} \\ \text{effect}}} + \underbrace{\frac{\partial \ln(e_{j,n}(x))}{\partial x}}_{\substack{\text{Employment} \\ \text{effect}}} \right). \tag{1}$$

This decomposition implies that understanding $\frac{\partial \ln(Y)}{\partial x}$ requires knowing, for *each* location, the *whole* joint distribution of firm value added $Y_{j,n}$ and the semielasticities $\frac{\partial \ln(\tilde{p}_{j,n}(x))}{\partial x}$ (per-worker value added) and $\frac{\partial \ln(e_{j,n}(x))}{\partial x}$ (employment). I now unpack these black-box elasticities into objects connected to how firms respond to changes in their economic environment due to changes in spatial frictions.

Assumption 1. *Throughout the following definitions, I assume all functions are differentiable in their arguments.*

Definition 1. Global Factors. Define $\mathbf{g}$ as the vector of global factors $g_k, k \in G = \{1, 2, \dots\}$ that: i) directly affect firms' value added per worker and/or size ($\frac{\partial \tilde{p}}{\partial g_k} \neq 0$ or $\frac{\partial e}{\partial g_k} \neq 0$), ii) are affected by changes in spatial friction x ($\frac{\partial l_k}{\partial x} \neq 0$), and iii) do not vary across locations for any value of spatial frictions x.

[16] For clearness of exposition, assume N_j includes all the firms that *could* be active in location j, but currently have size ≈ 0.

[17] For $N_j \to \infty$, see Appendix A.1.

Definition 2. **Local Factors.** Define $\mathbf{l_j}$ as the vector of local factors $l_{jk}, k \in L = \{1,2,\ldots\}$ that: i) directly affect firms' value added per worker and/or size ($\frac{\partial \tilde{p}}{\partial l_{jk}} \neq 0$ or $\frac{\partial e}{\partial l_{jk}} \neq 0$), ii) are affected by changes in spatial friction x ($\frac{\partial l_k}{\partial x} \neq 0$), and iii) are not global.[18]

Assumption 2. *Assume spatial frictions x only affect $\tilde{p}$ and e through* $\mathbf{g}$ *and* $\mathbf{l_j}$.

Using definitions 1 and 2, assumption 2, and applying the chain rule yields:

$$\begin{aligned} \frac{\partial \ln(\tilde{p}_{j,n}(x))}{\partial x} &= \nabla_{\mathbf{l_j}} \ln(\tilde{p}_{j,n}) \frac{\partial \mathbf{l_j}}{\partial x} + \nabla_{\mathbf{g}} \ln(\tilde{p}_{j,n}) \frac{\partial \mathbf{g}}{\partial x}. \\ \frac{\partial \ln(e_{j,n}(x))}{\partial x} &= \nabla_{\mathbf{l_j}} \ln(e_{j,n}) \frac{\partial \mathbf{l_j}}{\partial x} + \nabla_{\mathbf{g}} \ln(e_{j,n}) \frac{\partial \mathbf{g}}{\partial x}. \end{aligned} \tag{2}$$

That is, the semielasticities of value added per worker and size to spatial frictions can be expressed as how local (and global) factors are affected by changes in spatial frictions and how, in turn, these changes affect firms' value added and employment decisions.[19]

Equation 2 shows how, in order to assess the effects of spatial frictions on misallocation (or wages), we need to understand how local (and global) factors affect each firm. This requires two steps. First, understanding what factors matter for different firms. Second, understanding how these factors are affected by spatial frictions or policies aimed at tackling labour misallocation.

In this paper, I will focus on the role of local factors.[20] For this scope, it is useful to define the concept of "localness" of a firm to factor l_j as follows:

Definition 3. Localness.

Call $\nabla_{\mathbf{l_j}} \ln(\tilde{p}_{j,n})$ "localness of value added per worker". A firm n in j is defined as "*more local* for local factor k" in $\tilde{p}$ than another firm n' in j' when $|\frac{\partial \tilde{p}_{jn}}{\partial l_{jk}}| > |\frac{\partial \tilde{p}_{j'n'}}{\partial l_{j'k}}|$.

Call $\nabla_{\mathbf{l_j}} \ln(e_{j,n})$ "localness of firm employment". A firm n in j is defined as "*more local* for local factor k" in e than another firm n' in j' when $|\frac{\partial e_{jn}}{\partial l_{jk}}| > |\frac{\partial e_{j'n'}}{\partial l_{j'k}}|$.

In principle, localness and its heterogeneity across firms can be associated with several existing

[18] This condition can be formalised as $\exists j \neq j' \in \mathscr{L}$, $\exists k \in L$, $\exists x \in X : l_{jk}(x) \neq l_{j'k}(x)$.

[19] A way to reformulate Equation 2 would be $\frac{\partial \ln(\tilde{p}_{j,n}(x))}{\partial x} = \sum_{k \in L} \frac{\partial \tilde{p}_{j,n}}{\partial l_{jk}} \frac{\partial l_{jk}}{\partial x} + \sum_{k \in G} \frac{\partial \tilde{p}_{j,n}}{\partial g_{jk}} \frac{\partial g_{jk}}{\partial x}$.

[20] I will account for global factors only insofar as they are affected by general equilibrium effects of changes in local factors.

concepts (factors), such as: openness and tradeability of goods (Woodland, 2017; Monte et al., 2018); local monopsony power (Caldwell and Danieli, 2023; Heise and Porzio, 2023; Datta, 2023); local vacancy filling rates (Bilal, 2023; Kuhn et al., 2022); congestion and firm selection (Fajgelbaum and Gaubert, 2020; Bilal, 2023; Lindenlaub et al., 2023); local taxation (van der List, 2023), or agglomeration (Combes et al., 2012; Martellini, 2022). While my quantitative estimation will include several of the aforementioned factors, I will focus on two important components of localness in value added per worker that are directly affected by *workers'* location choices.[21]

2.4 Firm Localness Heterogeneity

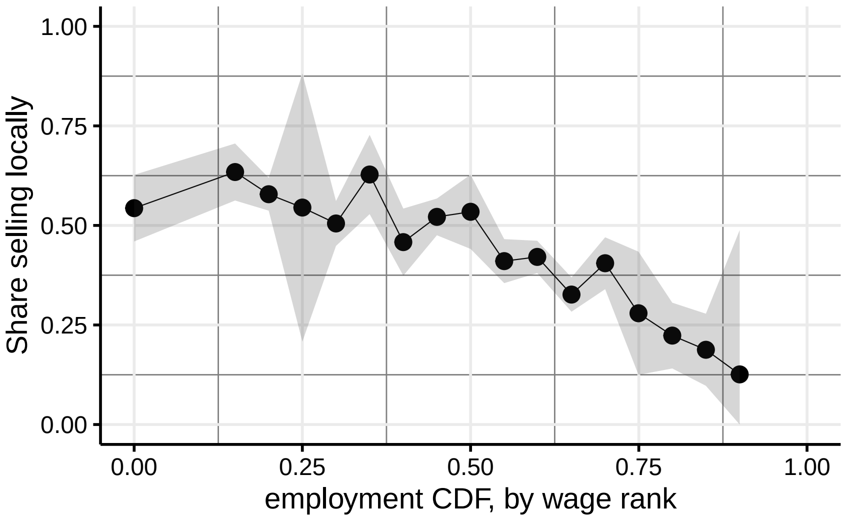
I provide direct evidence that firms differ in their value-added per worker exposure to two dimensions directly connected to the location of consumers and, hence, to the effects of spatial frictions. The first dimension relates to the geographical origin of the demand for a plant's products and services (local or global). This dimension is connected to workers' location choices through how they shape the spatial distribution of "local" demand for goods and services. The second dimension regards the costs of floorspace (rents plus taxes). These are closely connected to local land and housing prices. Since the local prices of non-tradeables and housing directly depend on workers' location choices as consumers, these dimensions are crucial for understanding the labour misallocation arising from spatial frictions.

To understand what local factors matter for what firms, I use firm microdata to measure the firm-level relevance of these two dimensions and their heterogeneity along the productivity distribution. This will address the first step: understanding what local factors matter for what firms. Then, using a spatial general equilibrium model, I put structure on both $\nabla_{\mathbf{l_j}} \ln(\tilde{p}_{j,n})$ and $\frac{\partial \mathbf{l_j}}{\partial x}$ by modelling how local demand, prices, and labour supply respond in equilibrium to workers' location choices. This is to assess the misallocation costs of spatial frictions and the effectiveness of different policies.

[21] An important literature considers, instead, the role of local and regional transportation costs for *goods*. In this paper, I focus on the *worker* margin, as a commuter and job mover. These two approaches complement each other.

Sales location. First, I document how firms that pay lower average wages are more likely to sell their products locally. Using plant-level NESS data, I compute the share of workers of a firm employed in plants that mainly sell at the local level.[22] In Figure 3, I plot the average of this figure across quantiles of employment, ranked by the mean wage at the firm, obtained by merging NESS and ABS. Over 50% of workers at below-median firms are employed at plants that declare that their revenue comes from selling their goods and services mainly at the local or regional level. However, this share declines to as little as 12.6% for firms in the 90-95th percentile of employment by wage rank. I interpret this evidence as low-wage firms being more exposed to local, rather than national or global, demand for goods and services.

Figure 3: Share of employees in plants mainly selling locally



Note: the figure shows the share of firms reporting that they "mainly sell" their products at the local or regional level, across quantiles of the per-worker employment cost distribution. Employment costs include wages and social contributions. For multi-plant firms, the figure is weighted according to the share of workers in each plant that report selling locally, rather than at the national or global level. Bars indicate 95% confidence intervals.

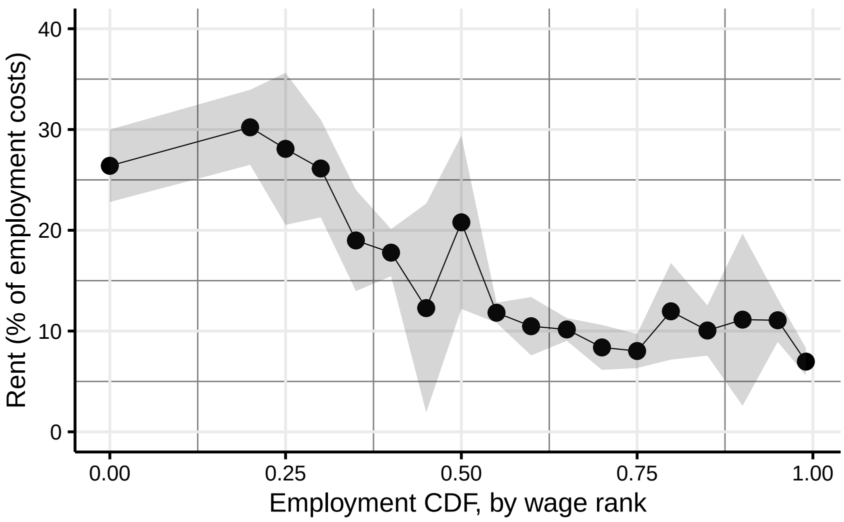
Floorspace costs. Second, firms lower in the wage rank also spend more on floorspace, relative to employment costs. Using data on firms' business rate bills (a tax paid by businesses as a proportion of the mean market rent of the property they occupy, even if they do not own it), I estimate firms' yearly floorspace costs.[23] I plot the distribution of per-worker floorspace

[22] I derived this from one of the survey's questions, asking the geographical level at which a plant mainly sells its products at.

[23] In the UK, properties are taxed at the *occupier* level. For individuals, this is the Council Tax, which funds local services. Businesses pay a similar tax, called "non-domestic rates", based on the mean open-market rental

costs against the CDF of employment, ordered by wage rank of firms, in Figure 4. On average, spending on floorspace is about 16% of spending on employment, although the figure varies considerably across firms, ranging from averages of 30.2% to 7.0% across different quintiles of the wage distribution. Firms paying higher wages per worker tend to pay lower rents relative to employment costs (wage plus social security bills), although the relationship flatlines at the top 40% of employment. This relationship is robust also within sectors, with the ratio of rents to wages falling by 0.78 log-points (95% CI. 0.71 - 0.85) for each one-point increase in log-wages per worker even after controlling for 3-digit industry fixed effects.[24]

Figure 4: Ratio of floorspace expenditure to employment cost



Note: the figure shows rents and business rates per worker as a share of employment costs per worker, across quantiles of the employment distribution, ordered by wage rank. Grey bands are 90% confidence intervals.

Taking stock

Firms are heterogeneous in terms of whether they mainly sell locally and in their floorspace costs. This suggests that productive firms are less "local" than low-productivity ones in both dimensions.

However, these stylised facts do not clarify two important elements necessary to understand

value of premises in the area that they occupy. In Appendix B, I discuss the advantages of this approach, the required imputations, and robustness checks.

[24]This means that larger and more productive businesses use more floorspace or are located in more expensive neighbourhoods. I discuss how this is indeed the case in Appendix B.2.

how spatial frictions generate misallocation through localness. First, this analysis cannot assess whether high-productivity firms are less affected by higher congestion. In principle, the higher revenues and costs from local congestion effects of workers' location choices could offset each other. Hence, it is unclear whether firms that are highly local in both dimensions benefit or lose from congestion. Second, these facts do not tell us *where* and *how much* changes in spatial frictions affect local factors. Hence, we cannot assess yet their quantitative effects.Ultimately, answering these questions requires a structural general equilibrium model.

3 A Model of Spatial Misallocation of Labour

In this section, I build a structural model that can account for how, by changing workers' mobility and location choice incentives, spatial frictions shape local congestion and how this affects - through firm localness - the allocation of labour. First, I present the problem of each agent. Then, I characterise the equilibrium wage schedule. Finally, I show how the semielasticity of value added per worker - a crucial element for the estimation of labour misallocation, as shown in Section 2.3 - is characterised in this model.

3.1 Model summary

The model is a generalisation of the Burdett-Mortensen (Burdett and Mortensen, 1998) framework to multiple locations (as in Heise and Porzio (2023)) and skills. In particular, I build on this framework by explicitly accounting for the heterogeneity in firm localness, and by decoupling the location choice of workers into a residence and a workplace choice.

The first key ingredient of the model is the interaction between search and spatial frictions. The presence of search frictions creates a job ladder and monopsony power for more productive firms. Spatial frictions make the job ladder more "local", creating a wedge between acceptable offers across space. Workers' locations and mobility change the distribution of local monopsony power (by making low-productivity businesses less exposed to competition when far away from productive jobs), and hence the local distribution of wages and employment across firms with heterogeneous productivity.

The second, and novel, feature of the model is that firms' value added per worker depends on their localness. Changes in local non-tradeables' and housing prices affect the firms' revenue and costs. In turn, this affects profitability, and thus wage and vacancy posting decisions of the firms. Most importantly, I allow localness to be heterogeneous across firms. Firms differ in localness due to different shares of labour allocated to the production of non-tradeables, different use of floorspace per worker in the production process, and different sensitivity of productivity to agglomeration economies.

The key mechanism through which spatial frictions and localness interact is congestion. Congestion due to higher demand from residents (and firms) leads to higher local prices. This affects firms' value added per worker according to their localness in non-tradeables and floorspace. In turn, this shapes their wage and vacancy posting decisions. Importantly, a firm's wage posting decision (and thus employment) depends on the *whole* distribution of firms' decisions, as in Burdett and Mortensen (1998), as - due to search frictions - firms compete with each other to poach and recruit workers. Hence, understanding how localness acts across the productivity distribution can be crucial to assess aggregate output losses.

3.2 Fundamentals

Locations. Time is continuous. There is a countable set $\mathscr{L}$ of locations. Distances between locations are given by the matrix $D_{\mathscr{L}}$. Each element d_{hj} represents the distance between locations h and j, with the largest ones normalised to 1. "h" will indicate residence locations, while "j" will indicate workplace locations. h' and j' will indicate the mobility choice of workers for their residence and workplace, respectively.

Workers. There is a mass 1 of workers. Workers can be of two types, indexed by $i \in I = \{L,H\}$ (for Low- and High-skilled). Each type has mass $M_i^w : \sum_{i \in I} M_i^w = 1$ and ability α^i, with $\alpha^H > \alpha^L$.

Firms. There is a mass 1 of firms. Firms are indexed by their type $i \in I = \{L, H\}$, and i-type firms only hire i-type workers. All firm heterogeneity can be summed up through their productivity p, which is distributed with CDF $F^i(p)$. The CDF of productivity of i-type firms within a location is $F^i_j(p)$ with pdf $f^i_j(p)$. The mass of firms of i-type in location j is M^f_{ij} such that $\sum_{j\in\mathscr{L}}\sum_{i\in I} M^f_{ij} = 1$. Firms cannot move.

Spatial Frictions. There are three sources of spatial frictions. First, commuting costs (both monetary d_0 and non-monetary d_1). Second, the cost of relocating one's residence, κ. Third, additional job search frictions when applying to distant locations, c_a.

I now present the problem of each agent. Then, I will characterise the steady state equilibrium of this model.

3.3 Workers

Individuals reside in location $h \in \mathscr{L}$ (*h*ome), potentially different from their workplace $j \in \mathscr{L}$. They consume three types of goods: i) tradables (c), ii) non-tradeables (l), and iii) housing H. Non-tradeable goods l and housing H cannot be traded across locations, and have local prices $P_{l,h}$ and $P_{H,h}$ respectively. Tradeable goods are the numeraire, with their unique global price normalised to 1. Workers are born unemployed and live forever,[25] with lifetime utility

$$\mathscr{U}^i_t = \int_t^{\infty} e^{-\rho(x-t)} E_t\left[U^i_{h(x)j(x)}(c(x), l(x), H(x))\right] dx.$$

At each instant, workers derive utility from living in h, working in j and consuming a bundle (c, l, H), chosen to maximise instantaneous utility:

$$\max_{c,l,H} U^i_{hj}(c, l, H) = \frac{\tau_h}{d^i_{1,hj}} c^{\eta_1} l^{\eta_2} H^{1-\eta_1-\eta_2} \quad s.t. \quad \begin{cases} c + P_{l,h} l + P_{H,h} H \leq y - d_{0,hj} \\ c \geq 0, \;\; l \geq 0, \;\; H \geq 0 \end{cases}$$

where τ_h represents amenities derived from living in h. y is income, equal to $\alpha^i w$ when employed (for w being the per-ability unit rate of pay) and b^i_h when unemployed. $d_{0,hj}$ is the monetary cost of commuting (working in j while living in h), and $d^i_{1,hj}$ the non-monetary cost.

[25] An equivalent interpretation would be that workers die at rate $\rho - r$, with $r \in [0, \rho]$, and new unemployed are born in the same location where workers die.

A worker of type i living in h and working in j has indirect utility

$$V^i_{hj}(y) = \frac{\tau_h\left(y - d_{0,hj}\right)}{P_h d^i_{1,hj}}, \tag{3}$$

where P_h is the local optimal price index, equal to $P_h = p_c^{\eta_1} P_{l,h}^{\eta_2} P_{H,h}^{1-\eta_1-\eta_2} = P_{l,h}^{\eta_2} P_{H,h}^{1-\eta_1-\eta_2}$.

Search. Individuals can be employed or unemployed. When employed, workers are randomly unmatched according to an idiosyncratic shock with Poisson arrival rate δ^i. Individuals must spend search effort s_x to send $a^i_{hx}\frac{s_x^{1+\varepsilon_0}}{1+\varepsilon_0}$ jobs applications to location x. The rate at which effort is translated into applications, a^i_{hx}, is allowed to depend on both h and x due to the presence of spatial frictions, which reduces the applications filed per unit of effort when $x \neq h$. In particular, it depends on a "fixed" component a^i, and a distance discount $c^i_{a,hx}$, so that $a^i_{hx} = a^i c^i_{a,hx}$. This captures how applying and finding a job can be harder far away from home, for example due to a lack of social connections or delays in conducting interviews. Total effort $S = \sum_{x\in\mathscr{L}} s_x$ has cost $\psi(S)$, parameterised as:

$$\psi(S) = \begin{cases} \dfrac{S^{1+\varepsilon_1}}{1+\varepsilon_1} & \text{if employed;} \\ \nu^{-\frac{\varepsilon_1-\varepsilon_0}{1+\varepsilon_0}} \dfrac{S^{1+\varepsilon_1}}{1+\varepsilon_1} & \text{if unemployed.} \end{cases} \tag{4}$$

I assume $\varepsilon_1 > 0$, and $\varepsilon_0 \in (-1,0)$. The former captures the increasing disutility from spending more effort applying for jobs, while the latter captures the decreasing return from spending effort in a location. The parameter $\nu > 0$ captures differences in the employed and unemployed search efficiency, as in Moscarini and Postel-Vinay (2016).

Matching. Vacancies of firms in location j with productivity p of type i are denoted as $v^i_j(p)$. Firms post wages, and there is no bargaining. Thus, each (i,j,p) vacancy is associated with the same wage offer $w^i_j(p)$. Within each (i,j) market, total vacancies $\bar{v}^i_j$ and total applications $\bar{a}^i_j$ are matched according to the matching function $M(\bar{a}^i_j,\bar{v}^i_j) = (\bar{a}^i_j)^{\chi}(\bar{v}^i_j)^{-1-\chi}$. Denoting market tightness as $\theta_{i,j} = \bar{v}^i_j/\bar{a}^i_j$, vacancies are matched at rate $\theta_{i,j}^{-\chi}$, and applications at rate $\theta_{i,j}^{1-\chi}$.

Worker choice. Consider a worker residing in h and working in j for wage w. When they receive an offer from a firm in location j' at wage rate w', they choose whether to reject or accept the offer (and, in the latter case, whether to move residence) according to the rule:

$$(h^*, j^*, w^*) = \arg\max \left\{ W^i_{hj}(w) + \varepsilon,\ \max_{h' \in \mathscr{L}} W^i_{h'j'}(w') + \varepsilon' + \varepsilon_{h'} - \kappa^i_{hh'} \right\}, \tag{5}$$

where $\kappa^i_{hh'}$ represents relocation costs, equal to zero whenever $h = h'$ and ≥ 0 otherwise, capturing not only the pecuniary costs of moving, but also the disutility from looking for a house, or moving away from friends.

I assume $\varepsilon, \varepsilon', \{\varepsilon_{h'}\}_{h' \in \mathscr{L}}$ are distributed according to a Generalised Extreme Value distribution. Thus, the solution of workers' offer acceptance problem (and residence location choice) follows a nested logit.[26]

3.4 Firm Problem

Production. All firms of type i with productivity p in location j produce both (or either of) the tradeable and the non-tradeable good according to the production functions:

$$\text{Tradeables: } T^i_j(p,n) = A^T_j \beta^i(p) pn,$$

$$\text{Non-Tradeables: } N^i_j(p,n) = A^N_j (1 - \beta^i(p)) pn,$$

where $\beta^i(p) \in [0,1]$ represents the amount of labour a firm allocates to produce tradeables, and n is the amount of labour hired.[27] Floorspace does not directly enter the production function but represents a cost $P_{H,j}\phi^i(p)$, where $\phi^i(p)$ is the floorspace employed for each pn units of effective labour hired.[28] For the sake of exposure, will I assume throughout that $\beta^i(p)$ is an exogenous characteristic, but provide a microfoundation and additional results in Appendix K.1.

[26] The upper-level nest has choices {reject offer; accept offer}. The lower level of the "accept" nest involves choosing the worker's residence over all $h' \in \mathscr{L}$. The "reject" nest is degenerate. In Appendix C.2, I derive the offer acceptance rates and residence move rates, and the workers' value function.

[27] This setup is justified by the fact that the workers' job ladder is more contained within occupations rather than sectors. For example, a solicitor could work in-house at a retail firm (SIC code 45), or a law firm firm (SIC code 81).

[28] This is identical to assuming a Leontieff production function for both tradeables and non-tradeables, with floorspace representing a minimum requirement for a given production process, but unproductive if bought in excess. van der List (2023) adopts a similar approach, although without firm heterogeneity in $\phi(\cdot)$.

Profit maximisation. Firms maximise total profits $\pi_j^i(p,w,v)$. Expressing the number of hires per vacancy as $n(w)$, profits per vacancy $\tilde{\pi}_j^i(p,w)$ are:

$$\begin{aligned}\tilde{\pi}_j^i(p,w) &= A_j^T \beta^i(p) pn(w) + P_{l,j} A_j^N (1-\beta^i(p)) pn(w) - wn(w) - P_{H,j}\phi(p) pn(w) \\ &= (\tilde{p}-w)n(w).\end{aligned}$$

$$\text{With } \tilde{p} = p \underbrace{\left[A_j^T(p)\beta^i(p) + P_{l,j}A_j^N(1-\beta^i(p)) - P_{H,j}\phi^i(p)\right]}_{\text{Value added multiplier} = \varphi_j^i(p)}.$$

That is: profits per vacancy are log-linear in the amount of labour hired per vacancy $n(w)$, and in the profits per labour unit $(\tilde{p}-w)$. The value added per worker $\tilde{p}$ depends on productivity p, and a *value added multiplier* capturing the relative importance of revenue from tradeables $A_j^T(p)\beta^i(p)$ and non-tradeables $P_{l,j}A_j^N(1-\beta^i(p))$, and floorspace costs $P_{H,j}\phi^i(p)$.[29]

Wages are chosen to maximise per-vacancy profits:

$$w^*(\tilde{p}) = \arg\max_w (\tilde{p}-w)n(w).$$

Finally, given the optimal wage $w^*(p)$, firms choose vacancies to maximise total profits given a convex vacancy cost function $\zeta_{ij}(v)$:

$$v_{ij}^*(\tilde{p}) = \arg\max_v \pi_{ij}(\tilde{p}, w^*, v) = (\tilde{p}-w^*)n(w^*)v - \zeta_{ij}(v).$$

3.5 Housing

Local housing services H_j, used by both workers (as H_h) and firms (as floorspace $\phi^i(p)pn$), are produced with a Cobb-Douglas production function employing land (in fixed supply L_j in each location j) and capital owned by non-residents, supplied elastically at price r:

$$H_j = A_j {L_j}^{b_j} K^{1-b_j} = \tilde{A}_j K^{1-b_j}.$$

Where the second equality, with $\tilde{A}_j = A_j {L_j}^{b_j}$, follows from land being in fixed supply. b_j is allowed to differ across locations to capture differences in local housing regulations and geography that affect housing price elasticity of demand (Diamond, 2016).

[29] p can be interpreted as the "exogenous" productivity/quality level of a firm, which differs from value added due to: i) local conditions increasing overall productivity (e.g. $A_j^T > 1$), ii) nominal revenue effects from $P_{l,j}$, and iii) "exogenous" floorspace requirements to set up production with productivity p, ϕ^i, shifted by a iv) nominal cost effect $P_{H,j}$.

3.6 Stationary Equilibrium

I study the model in its stationary equilibrium.

Definition 4. Stationary Equilibrium in the Labour Market. A stationary equilibrium in the labour market consists of a set of wages and vacancy policies $\{w^i_j(p), v^i_j(p)\}_{j\in\mathscr{L},i\in I}$, search efforts $\{s^{E,i}_{jhx}, s^{U,i}_{hx}\}_{j,h,x\in\mathscr{L},i\in I}$, wage offer distributions $\{F^i_j(p)\}_{j\in\mathscr{L},i\in I}$, acceptance probabilities $\{\mu^{E,i}_{jhx}(w,w'), \mu^{U,i}_{hx}(w')\}_{j,h,x\in\mathscr{L},i\in I}$, move probabilities $\{m^{E,i}_{jhxh'}(w,w'), m^{U,i}_{hxh'}(w')\}_{j,h,x,h'\in\mathscr{L},i\in I}$, labour per vacancy $\{n^i_j(w)\}_{j\in\mathscr{L},i\in I}$, unemployment $\{u^i_h\}_{h\in\mathscr{L},i\in I}$, and market tightness $\{\theta_{ij}\}_{j\in\mathscr{L},i\in I}$ such that

1. workers file applications, accept offers, and move their residence location to maximise their expected present discounted value of work and unemployment, taking as given market tightness $\{\theta_{ij}\}_{j\in\mathscr{L},i\in I}$ and the wage offer distribution $\{F^i_j(w)\}_{j\in\mathscr{L},i\in I}$
2. firms set wages to maximise vacancy profits, and choose vacancies to maximise firm profits, taking as given the function mapping wages to per-vacancy size $\{n^i_j(w)\}_{j\in\mathscr{L},i\in I}$
3. the arrival rates of offers and the wage distributions are consistent with aggregate applications, wage policies, and vacancy posting behaviour
4. firm sizes and worker distributions across workplace and residence locations satisfy the stationary equations
5. each local non-tradeable good market and housing market clears.

Equilibrium Characterisation

I now characterise the stationary equilibrium. Define $\overline{a}^i_j$ the mass of applications of skill i towards j. Then, the mass of labour units per vacancy hired by a firm of type i in j offering wage w is

$$\dot{n}^i_{jh}(w) = \alpha^i \theta_{ij}^{-\chi} P^i_{jh}(w) - Q^i_{hj}(w(p))n^i_{jh}(w),$$

where $n^i_j(w) = \sum_{x\in\mathscr{L}} n^i_{jx}(w)$. $P^i_{jh}(w)$ is the probability that a worker matched with a vacancy in j (at rate $\theta_{ij}^{-\chi}$) accepts the offer and moves to h, while $Q^i_{hj}(w)$ is the flow rate at which workers who live in H and work in j leave their current job. In Appendix C.3 I characterise the functional forms for $P(\cdot)$ and $Q(\cdot)$. In steady state, $\dot{n}^i_{jh}(w) = 0$, so that:

$$n^i_{jh}(w) = \frac{\alpha^i \theta_{ij}^{-\chi} P^i_{jh}(w)}{Q^i_{hj}(w)}. \tag{6}$$

Proposition 1. *If the CDF of $\tilde{p}$ is differentiable, the wage policy of the Stationary Equilibrium is characterised as the following set of differential equations and initial conditions:*

$$\begin{aligned}
w^i_j(\tilde{p}) &= w^i_j(\underline{\tilde{p}}) + \int_{\underline{\tilde{p}}_j}^{\tilde{p}} \frac{\partial w^i_j(\tilde{p})}{\partial \tilde{p}} dF^i_j(p), \\
w^i_j(\underline{\tilde{p}}_j) &= \max\left\{R_j, \arg\max_{\hat{w}} (\underline{\tilde{p}}_j n^i_j - \hat{w} n^i_j(\hat{w}))\right\}, \\
\frac{\partial w^i_j(\tilde{p})}{\partial \tilde{p}} &= \frac{(\tilde{p} - w)\left[\alpha^i \sum_{h\in\mathscr{L}} \left(\frac{\frac{\partial \tilde{P}^i_{hj}(\tilde{p})}{\partial \tilde{p}} \tilde{Q}^i_{hj}(\tilde{p}) + \frac{\partial \tilde{Q}^i_{hj}(\tilde{p})}{\partial \tilde{p}} \tilde{P}^i_{hj}(\tilde{p})}{\left(\tilde{Q}^i_{hj}(\tilde{p})\right)^2}\right) \theta_{ij}^{-\chi}\right]}{\alpha^i \sum_{h\in\mathscr{L}} \frac{\tilde{P}^i_{hj}(\tilde{p})}{\tilde{Q}^i_{hj}(\tilde{p})} \theta_{ij}^{-\chi}}.
\end{aligned} \tag{7}$$

With $\tilde{Q}^i_{hj}(\tilde{p}) = Q^i_{hj}(w(\tilde{p}))$, and $\tilde{P}^i_{hj}(\tilde{p}) = P^i_{hj}(w(\tilde{p}))$.

Proof. See Appendix C.4. □

3.7 Spatial frictions and misallocation

To show the intuition for how the model can pin down the effects of spatial frictions on the allocation of labour through localness, I rewrite Equation 1 in terms of the model's parameters. The value added per worker at each point of the FTP distribution depends on three components:

$$\frac{\partial \ln(\tilde{p}^i_j(p))}{\partial x} = \underbrace{\Phi^i_j(p)\frac{\partial \ln(A^T_j)}{\partial x}}_{\text{Agglomeration}} + \underbrace{\Phi^i_{l,j}(p)\frac{\partial \ln(P_{l,j})}{\partial x}}_{\text{Non-tradeable congestion}} - \underbrace{\Phi^i_{H,j}(p)\frac{\partial \ln(P_{H,j})}{\partial x}}_{\text{Floorspace congestion}}.$$

$$\text{Where } \Phi^i_j(p) = pA^T\beta(p)\left[\tilde{p}^i_j(p)\right]^{-1}, \tag{8}$$

$$\Phi^i_{l,j}(p) = pP_{l,j}A^N(1-\beta^i(p))\left[\tilde{p}^i_j(p)\right]^{-1},$$

$$\Phi^i_{H,j}(p) = pP_{H,j}\phi^i(p)\left[\tilde{p}^i_j(p)\right]^{-1}.$$

The first term represents the effects of spatial frictions on local agglomeration effects in the production of tradeable goods. The second term represents the change in value added arising from changes in the price of local non-tradeable goods. Finally, the third term captures the effect of changes in floorspace prices. Each term is simply the derivative of the corresponding local element (agglomeration, non-tradeable prices, and floorspace prices), weighted by their importance for total value added per (exogenous) productivity per worker p. In the baseline results, I will exclude agglomeration effects (which I do not observe directly).

In the following sections, I calibrate the model to UK data and use this general equilibrium framework to derive the effects of changes in spatial friction on the allocation of labour across firms and across space. I later discuss the policy implications.

4 Bringing the Model to the Data

I estimate the model's parameters using data for England and Wales for the period 2011-2018. In the rest of this section, I describe the estimation procedure and results.

Functional forms I assume the following functional forms for commuting costs, spatial search decay, relocation costs, and vacancies:

Table 1: Functional forms for quantitative calibration

Description	Functional form
Non-pecuniary commuting cost	$d^i_{1,hj} = e^{d^i_1 d_{hj}}$
Pecuniary commuting cost	$d^i_{0,hj} = d^i_0 d_{hj}$
Search efficiency decay	$c^i_{d,hj'} = e^{c^i_d d_{hj'}}$
Relocation costs	$\kappa^i_{hh'} = \kappa^i + e^{\kappa^i_d d_{hh'}}$
Vacancy posting costs	$\zeta^i_j(v,p) = (\zeta^i_{0,j})^{-\zeta^i_1} \overline{\pi}^i_j \frac{v^{1+\zeta^i_1}}{1+\zeta^i_1}$

Notes: The table shows how the parameters of the model related to spatial frictions and vacancies are specified for the quantitative estimation.

For the non-monetary costs of commuting $d^i_{1,hj}$ and spatial search decay $c^i_{d,hj'}$, I assume exponential functional forms. These are commonly used in the economic geography and labour literature. For the monetary cost of commuting $d^i_{0,hj}$, I assume it scales linearly with distance (reflecting the fact that the marginal cost of one extra mile of travel is mostly linear). Relocation costs depend on a fixed cost κ^i, plus a term that depends on the distance of the move. For the cost of vacancies $\zeta^i_j(v,p)$, I follow Heise and Porzio (2023) in scaling the cost by average profits per vacancy across firms in a location ($\overline{\pi}^i_j$) to reduce the effects of increases in profitability on vacancy posting.[30]

Locations. I estimate the model using five locations. The first, "London", includes Greater London and a few Local Authority Units within its commuting zone. The second, "Bristol", includes all LAUs above the 25th percentile of average gross value added per worker outside of the South-East and London. These represent highly productive locations, such as several urban agglomerates in the Midlands (Milton Keynes, Warwickshire), South-West (Bath, Bristol), and some areas of the North-West (Manchester). The third, "Liverpool", includes all the locations within 50km of a "Bristol" location that are not located in London or the South-East. These represent locations that are not highly productive themselves but are within commutable distance of a productive location. The fourth, "Hull", includes all the remaining locations outside of the South-East. The fifth and final location includes all LAUs in the South-East statistical

[30] This can be rationalised as vacancies being generated by "Human Resources" workers, whose pay scales with the average profits of the firms in a location.

region.

Externally Calibrated Parameters. Table 2 summarises the parameters calibrated directly to the data. Three parameters represent normalisations.[31] I assume five parameters. I set the discount rate to 10% (similarly to Kuhn et al. (2022)); the share of tradeable goods in consumption to 0.38 (Diamond, 2016), and the minimum wage to 98.2% of the value added per worker of the least productive and active firms (Heise and Porzio, 2023). Since no data on workers' applications across space are available, I assume $\varepsilon_0 = -0.5$. Finally, I assume value added per worker is distributed across skill and location according to a Beta distribution, with first parameter $g_{1,ij}$. The remaining parameters listed in Table 2 are calibrated to various data sources, described in Appendix D.

[31] χ represents how many matches happen, given applications and vacancies. Since both of are endogenous, χ can be normalised to 0.5. $\alpha^L = 1$ normalises income relative to wage rates. $\tau_1 = 1$ normalises total utility, given wages.

Table 2: Externally Calibrated Parameters

Parameter	Description	Source	Value
χ	Matching efficiency	normalisation (Şahin et al., 2014)	0.5
α^L	Low skill value	normalisation	1
τ_1	Amenities in London	normalisation	1
ρ	Yearly discount rate	Kuhn et al. (2022)	0.1
η_1	Tradeables share	Diamond (2016)	0.38
ι	Minimum wage to $\underline{p}$	Heise and Porzio (2023)	0.982
ε_0	Effort return curvature	assumption	-0.5
$g_{1,ij}$	Distribution of $\tilde{p}$, first parameter	Bilal (2023)	1.3
δ^i	Exogenous unmatching, by i	Labour Force Survey	Appendix D.1
η_2	Non-Tradeables share	CPIH index weights *(ONS)*	0.27
α^H	High skill value	AKM on ASHE (Appendix D.2)	1.29
I^H	Mass of high-skilled	ASHE	0.62
M^i_j	Mass of employers, by i,j	ABS	Appendix D.4
$\underline{p}$	Minimum productivity	ABS	Appendix D.5
$\mathbf{D}$	Distance Matrix, by $\mathscr{L} \times \mathscr{L}$	Google Maps travel times	Appendix D.6
b_j	Elasticity of housing price, by j	House transactions data	Appendix D.7
c^i_d	Search efficiency decay, by i	ASHE	Appendix D.8
$g_{2,ij}$	Distribution of $\tilde{p}$, second parameter, by i,j	ABS	Appendix D.9
ν	Inverse of unemployed search cost	LFS (Appendix D.10)	4.13

Notes: The table lists all the parameters that are directly calibrated to empirical moments or assumed. The first three represent normalising assumptions. The second group represents assumptions from the literature. The third group regards calibrated parameters. For each of them, I list a description of their nature, the source dataset (or paper), and their value (or, due to space limitations, a reference to the section of the Appendix where they are discussed).

Estimated Parameters. This leaves 39 parameters and two distributions to be estimated internally, which I report in Table 3. I prove that the parameters can be identified in two steps.

1. I estimate the set of 19 parameters connected to amenities, search frictions, and spatial frictions by minimising a loss function calculated from over 300 moments, given the distribution of value added per worker $\tilde{p}$ (between and within locations). The intuition of the estimation procedure is that within-location moments of job flows and wage gains map heavily onto search frictions, while between-location moments map onto spatial frictions (Heise and Porzio, 2023). I discuss the details of the choice of targeted moments in Appendix E.1, and the algorithm in Appendix G.

2. The remaining 20 parameters (representing vacancy costs and productivity) and two dis-

tributions (representing firm localness in non-tradeables and floorspace costs) are exactly identified from i) matching the distribution of localness, and ii) 20 additional moments, together with applying market clearing conditions. I provide details of the set of simultaneous equations that identify the second block's parameters in Appendix E.2.

The key to separating the estimation into two blocks is knowing the distribution of value added per worker $\tilde{p}$ across locations. Given this distribution, the estimation of the first block does not depend on knowing p, β, ϕ, A^T or A^N, as these are components of value added $\tilde{p}$ itself. The parameters are estimated by Simulated Method of Moments. Then, I identify p, β, ϕ and A^N by perfectly matching another set of moments (prices, value added, and localness), given the market clearing conditions and the parameters estimated in the first block.

Finally, I exploit the knowledge of the empirical spatial distributions of vacancy mass by skill, $\bar{v}^i_j$, to identify the 10 parameters representing vacancy costs $\zeta^i_{0,j}$. I take as given the distribution of vacancies in the first block and then identify $\zeta^i_{0,j}$ by perfectly matching the vacancy mass given the distribution of equilibrium firm profits and the estimated ζ^i_1.

Table 3: Estimated Parameters

I. First Block

Parameter	Description	Number	Parameter	Description	Number
- Utility			*- Search Frictions*		
τ_h	Amenities relative to London	4	ε_1	Effort cost curvature	1
σ_i	Offer taste shock, by i	2	a^i	Efficiency of search, by i	2
σ_i^m	Residence taste shock, by i	2			
- Spatial frictions			*- Firms*		
κ^i	Moving cost (fixed), by i	2	ζ_1^i	Vacancy cost curvature, by i	2
κ_d	Moving cost (distance)	2			
d_1^i	Commuting cost, by i	2			

II. Second Block

Parameter	Description	Number	Parameter	Description	Number
$\zeta^i_{0,j}$	Vacancy cost, by i,j	10	$\beta^i(p)$	Tradeable production distribution	.
A^H_j	Productivity of housing, by j	5	$\phi^i(p)$	Floorspace use distribution	.
A^T_j	Productivity of non-tradeables, by j	5			

Notes: The table lists all the parameters to be endogenously estimated, as they depend on equilibrium objects. The first block refers to the moments being targeted in the MCMC algorithm to estimate the first 21 parameters. The second block refers to the moments being targeted, and perfectly fitted, in the second block to identify localness and local productivity.

Overall, I target 504 parameters, listed in Table 4. 324 are used in the first block and 180 in the second block. The first block is overidentified (324 moments for 20 parameters), while the second block is exactly identified (20 moments for 20 parameters and 160 moments to fit 160 points of two distributions). For each moment, I list the dimensions for which it is computed (h for residence, j for workplace, i for skill/type, and h', j' for post-move locations), the key parameters they help identify.[32]

Table 4: Targeted Moments

Moment	Dimension	Key Parameters	Number
Residents	h,i	τ	10
Jobs	h,i	τ	10
Commuters	h,j,i	d^i	50
Job Movers, within location	j,i	a	10
Job Movers, between location	j,j',i	a^i, c_d^i	40
Wage gains, within location	j,i	σ^i, a^i	10
Wage gains, between location	j,j',i	α_f^i	40
Wage gains, new commuters	j,i	d^i	10
Prob(negative wage gain)	j,i	σ^i	10
Residence movers	h,h',i	$\sigma_m^i, \tilde{\kappa}^i, \kappa_d$	50
Firm size distribution	j,i, deciles	ζ_1^i	60
Relative wages, by skill	i	$\tilde{\mu}$	2
Relative gva of firms	j	$\tilde{\mu}, \zeta_1^i$	5
Relative wages per worker	i,j	p_μ	10
Unemployment rate	h	v	5
dln(search effort)/dln(w)	i	ε_1	2
Vacancy mass	i,j	$\zeta_{0,j}^i$	10
House prices	h	A_H	5
Local good prices	h	A^N	5
Share of local sales to wages	i,j, deciles	β^i	80
Floorspace costs to wages	i,j, deciles	ϕ^i	80
Total			504

Notes: The table shows the moments being targeted in the estimation. The first block refers to the moments being targeted in the MCMC algorithm to estimate the first 20 parameters. The second block refers to the moments being targeted in the second block to identify localness and local productivity.

[32]See Appendix E.1 for a theoretical discussion of what parameters map more heavily in what moments, and Appendix G.1.2 for the empirical Jacobian of the model at the estimated parameters, showing how the theoretical intuition is reflected in the data.

4.1 Parameter Estimates

I report details on the estimated spatial frictions' parameters. Relocation costs are estimated to be, when moving to the locations, 62% of the average value of unemployment. This corresponds, in consumption equivalent terms, to £150,000 (in 2015 British pounds). The minimum moving cost between regions is £96,000. Giannone et al. (2023) finds for Canada an estimate of approximately £109,000 for the average mover ($217,000 in 2016 CAD). However, the comparison should be taken with a grain of salt, as i) the two models do not nest each other, and ii) they compare different countries. The spatial decay in search efficiency for distance $d_{hj'}$, the inverse of $\exp(c_d^i d_{hj'})$, is estimated to be approximately 83% (on average across the two skills). This number is slightly smaller than the 95% of Heise and Porzio (2023), although comparable.

Table 5: Estimated parameters: spatial frictions

Parameter	Description	Estimate	S.E.	Monetary equivalent
Relocation cost *(share of value of unemployment, poorest location; b/w furthest locations)*				
κ^H	Relocation cost (high-skilled)	0.51	(0.05)	£133,730
κ^L	Relocation cost (low-skilled)	0.83	(0.06)	£171,756
Commuting costs *(share of instantaneous utility; b/w furthest locations)*				
d_1^H	Utility cost of commute (high-skilled)	0.52	(0.03)	£13,623-33,423†
d_1^L	Utility cost of commute (low-skilled)	0.52	(0.04)	£10,843-22,341†
Spatial search frictions *(share of search efficiency; b/w furthest locations)*				
c_d^H	Spatial search efficiency decay (high-skilled)	0.80	(0.04)	
c_d^L	Spatial search efficiency decay (low-skilled)	0.87	(0.03)	

†: first figure calculated at minimum wage in "Hull" location, second figure calculated at the best job available in the economy.

Notes: The table shows the estimates and standard errors of selected parameters from the model. The parameters are rescaled to represent the cost paid to travel from the two furthest away locations (London to "Hull"). Relocation costs are expressed as a share of the value of unemployment in the poorest location. Commuting costs are a share of the instantaneous utility they are applied to. Search efficiency decay is a share of the efficiency of applying to the location where one lives.

5 The cost of spatial frictions

To understand the role of firms' heterogeneous localness in determining the output, wage and welfare losses from spatial friction to workers' mobility, I study two counterfactuals where I shut down commuting and search costs.

Fixed Value Added Effect Counterfactual. In the first counterfactual, I shut down the heterogeneous value added effect of spatial frictions by fixing $\beta_j^i(p) = \overline{\beta}$ and $\phi_j^i(p) = \overline{\phi}$ for all firms.[33] In this case, the value added per worker reacts in the same proportion to a change in spatial frictions, and thus local prices, along the whole productivity distribution.[34]

Heterogeneous Localness Counterfactual. In the second counterfactual, I account for heterogeneity in the localness of the value added per worker of firms. I fit $\beta(p)$ and $\phi(p)$ to their empirical distribution across locations and employment of high and low-skilled workers. This makes it possible to account for how more productive firms are less local than less productive ones.

Results. I plot the key result from the two counterfactuals in Table 6. By removing commuting costs and spatial search frictions, real GDP increases by 4.3% in the Fixed VA effect counterfactual and 5.4% in the Heterogeneous one (26% more). Since the profits from the housing sector accrue to non-resident agents, it is useful to look also at firms' value added. Accounting for localness heterogeneity delivers a 1.4 pp. (£41bn, or \$50bn) larger increase in value added. Similar figures can be found for real wages (+1.7 pp.) and welfare (+1.3 pp.).

[33] The constants are obtained so that: i) $\overline{\phi} = \frac{\sum_i \alpha^i \sum_j M_j^i P_{H,j} \int s\phi_j^i(s) dE_j^i(s)}{\sum_i \alpha^i \sum_j M_j^i P_{H,j} \int s dE_j^i(s)}$, and ii) $\overline{\beta} = \frac{\sum_i \alpha^i \sum_j M_j^i \int s\beta_j^i(s) dE_j^i(s)}{\sum_i \alpha^i \sum_j M_j^i \int s dE_j^i(s)}$. That is, total expenditure on floorspace and the share of effective units of labour employed in the production of the tradeable good are identical across models. In practical terms, as these terms depend on knowing p, I obtain $\overline{\phi}$ by matching the share of expenditure to wages across all firms. $\overline{\beta}$ is recovered by using $\tilde{p} + \phi(p)P_H$ as a variable of integration, which is the best an econometrician without access to firm microdata, but only mean summary statistics would be able to estimate. Using p would not be feasible to an econometrician who does not know either $\overline{\beta}$, or at least one $A_{l,j}$.

[34] Notice that, due to changes in monopsony power, localness in size is allowed to vary across firms as in Heise and Porzio (2023).

Table 6: Key variables, real change relatively to baseline

	Het. VA (%)	Fixed VA (%)	Difference (p.p.)
GDP	5.37	4.26	1.11
VA of firms	1.24	-0.11	1.35
VA of firms: per worker	-1.44	-2.78	1.34
Wages	1.66	-0.08	1.74
Wages: per worker	-1.03	-2.75	1.72
Welfare	30.88	29.6	1.28

Notes: Standard errors in parentheses. All variables are expressed in real terms. Nominal GDP is adjusted to real terms using a price deflator at the product level. All other variables are adjusted using the average price index of consumption of all agents. "GDP" is the total value of production, including housing. "VA of firms" is the total value added of firms. "Wages" are the total wage bill, and "Wages: per workers" are wages by the mass of employed individuals. "Welfare" is the mean value of the value function of all individuals. The "Het. VA" and "Fixed VA" column shows the differences between the "Heterogeneous VA effect" and the "Fixed VA effect" counterfactuals and the baseline calibration, in percentage points. The last column shows the difference, in percentage points, between the two variations.

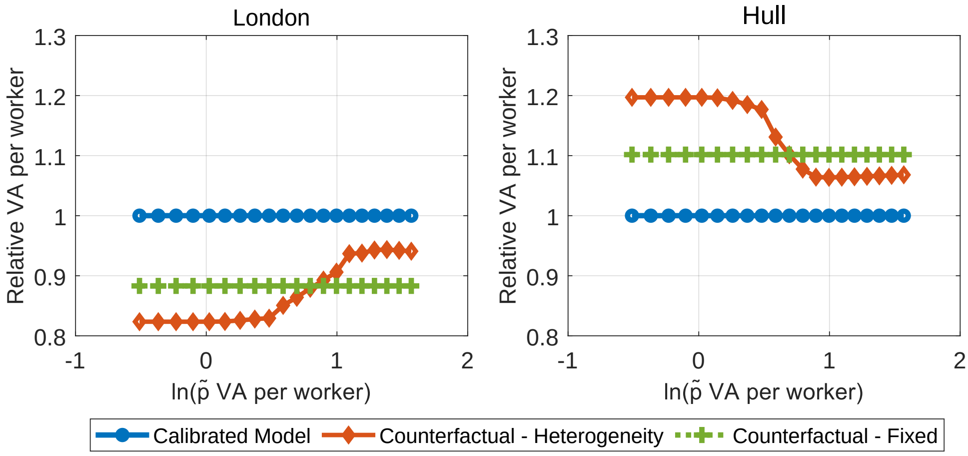
Interpretation. To understand the source of the larger misallocation implied by accounting for firm localness heterogeneity, I propose two exercises. First, I show how the value added of firms changes, in relative terms, across the two counterfactuals. This makes it possible to visualise who are the winners and losers in each counterfactual. Then, I perform other counterfactuals where I suppress part of the heterogeneity, in order to show whether this arises from between-location, skill, or within-location differences in localness.

First, in Figure 5, I plot the grid of value added per worker $\tilde{p}$ of the baseline calibration on the horizontal axis against its relative value in: i) the baseline estimate (blue line), ii) the fixed VA effect counterfactual (green line), and iii) the heterogeneous VA counterfactual (red line). To put the graph in the context of the earlier discussion on the decomposition of the effects of spatial frictions, this is equivalent to plotting $\frac{\partial \ln(\tilde{p})}{\partial x}\Delta x$, the semi-elasticity of value added per worker to spatial frictions from Equation 8, scaled by the change in frictions. In the fixed VA counterfactual (green line with $+$ marker), all firms are - within each location - equally impacted in their VA per worker $\tilde{p}$, as they have the same localness. This scenario would be consistent with models such as Heise and Porzio (2023) and Bilal (2023) or Monte et al. (2018).[35] However,

[35] In Bilal (2023) and Heise and Porzio (2023), firms' value added coincides with productivity and is exogenous. In their quantitative calibration, Bilal (2023) extends their model to a $\phi(p)$ floorspace cost that is constant, in relative terms, across the value added per worker distribution. Monte et al. (2018) assume identical firms, hence no heterogeneity in localness of value added per worker.

when I introduce heterogeneity in $\tilde{p}$ localness, the results change substantially. High-$\tilde{p}$ firms experience smaller losses (or gains) than low-$\tilde{p}$ firms since localness is negatively correlated to value added per worker. Differences in the average localness across locations also shift the average gain/losses within each location.

Figure 5: Relative VA, baseline (=1) vs counterfactuals



Second, I calibrate an identical heterogeneous localness counterfactual, but modify the fixed VA counterfactual in two ways. In the first, I calibrate $\phi(p)$ and $\beta(p)$ to their location averages instead of the economy-wide averages. In the second, I allow for heterogeneity in localness across locations and skill groups but not within each of these cells. In the third, I allow for heterogeneity also within a location but impose that the negative correlation between value added per worker and localness must be the same within each location. I plot the results, relative to the fixed VA counterfactual, in Figure 6. Approximately half of the gains in GDP and wages, and two-thirds of those of other economic aggregates, arise from between-location misallocation. Adding between-skill heterogeneity in localness of $\tilde{p}$ is essentially inconsequential for GDP but has minor effects on the total wages and value added (although in the opposite direction). The remaining contribution arises from the heterogeneity in localness *within* location and skill, although failing to account for differences in the negative correlation between localness and value added per worker within each location would overshoot the total impact.

Figure 6: Effects of heterogeneity in localness relative to Fixed VA counterfactual

Note: the figure shows the change, in percentage points, in key aggregates of the economy under different counterfactuals when removing spatial frictions, as a difference from the change under the "Fixed VA" counterfactual. "Between-Location Only" refers to a counterfactual calibrated with only between-location heterogeneity in localness in $\tilde{p}$. "Skill and Between-Location" refers to a counterfactual calibrated with heterogeneity in localness in $\tilde{p}$ across locations and skills, but not within each of these cells. The "Full Heterogeneity" counterfactual refers to the full counterfactual with heterogeneity within and between locations and skills. The "Same Slope Within" counterfactual accounts for both within and between-locations differences in localness, but assumes all firms with the same value added have the same localness across locations (equal to the mean of the localness levels in the "Full Heterogeneity" counterfactual), instead of calibrating to the actual data.

6 Policy

Although we can quantify the cost of misallocation by modelling the removal of spatial frictions, it does not represent a feasible policy. I analyse three policies that have been proposed or implemented to mitigate the effects of spatial frictions: i) a reduction in housing planning restrictions (Hsieh and Moretti, 2019), ii) a place-based policy Moretti (2022) aimed at increasing the average firm productivity of the poorest location by 5 percentage points ("levelling-up"), and iii) a reduction in commuting costs (Monte et al., 2018).

Reducing planning restrictions (such as high-rise bans, or minimum lots' size) aims to bring more workers to productive locations by reducing equilibrium housing prices, making cities cheaper to live in. Moving firms to deprived areas is often proposed as a way to reduce inequalities, but can also have efficiency implications by reducing local monopsony power and mismatch unemployment. Finally, reducing commuting costs directly lowers the costs of spatial frictions for workers, allowing them to climb the job ladder faster or to live in high-amenity but further-away locations.

In the rest of the paper, I focus on highlighting the role of localness heterogeneity for the predicted output and welfare gains of these policies. I show that localness heterogeneity can have profound effects on their relative effectiveness. This is due to different policies acting through different channels. Each of these channels, by changing workers' location choice incentives, changes the spatial distribution (and amount) of congestion in markets for local non-tradeable goods and housing/floorspace. Thus, accounting for the distribution of localness (firms' sensitivities to these changes in local congestion) can be crucial for assessing these policies.

For an example of why localness heterogeneity can be important, consider how reducing floorspace costs in London would make the city cheaper, but would be unlikely to change the size of a large tech firm's workforce due to paying somewhat cheaper rents. Accounting for localness heterogeneity incorporates this intuitive dynamic in the model.

Table 7 summarises the previous paragraphs, together with a preview of whether accounting for localness heterogeneity increases (+) or decreases (−) their effectiveness at increasing total value added and welfare. Reducing planning regulations and levelling-up policies deliver lower gains. The reason is that - through their general equilibrium effects[36] - they reduce congestion in productive locations (floorspace costs) more than they reduce revenues from non-tradeable production. Accounting for localness heterogeneity makes it possible to recognise that these positive externalities do not lead to the creation of productive jobs, as productive firms are less local. Hence, total gains are lower. Conversely, reducing commuting costs brings more residents to high-amenity locations. This has negative externalities on local firms, as congestion makes costs increase faster than local revenues. Accounting for localness heterogeneity makes it possible to recognise how this additional congestion *does not* destroy many productive jobs, as productive firms are less local. Hence, total gains are larger.

[36]Reducing planning restrictions brings more consumers in cities by lowering local rental prices. Levelling-up directly moves jobs away from productive locations.

Intuitively, the aggregate effects of the externalities running from workers' location choices to firms' productivity and employment decisions (due to changes in local costs and demand) are dampened due to two facts. First, the negative correlation between localness and productivity. This is true both for positive ones (such as lower rents due to lower planning regulations) and negative ones (such as those arising from lower commuting costs, as more workers move to London and the South-East for their better amenities).

Table 7: Policy exercise summary

Policy	Reference	Aim	Counterfactuals' difference Value added	Welfare
Lower planning restrictions	Hsieh and Moretti (2019)	workers to cities	−	−
Levelling-up	Moretti (2022)	firms to workers	−	−
Reduce commuting costs	Monte et al. (2018)	increase mobility	+	+

Notes: The table shows the three policies taken into consideration. "Lower planning regulations" refers to a policy that relaxes regulations that cause lower elasticity of housing supply to demand (equivalently, higher elasticity of prices), as in Hsieh and Moretti (2019). "Levelling up" refers to a place-based policy aimed at bringing more productive firms in deprived areas. "Reduce commuting costs" refers to a policy aimed at reducing the cost of commutes. In the third column, I briefly list the aim of the policy. The last two columns represent the difference between the heterogeneous localness counterfactual, against the Fixed VA counterfactual. A downward-point arrow signifies that the policy improves the aggregate of interest *less* when considering localness heterogeneity.

6.1 Lower planning restrictions

Planning restrictions increase equilibrium house prices at each demand level by reducing land supply or increasing the costs of, or banning, high-rise buildings. When productive locations are also very regulated (such as the Bay Area in California or London), reducing regulations can improve the allocation of labour. First, it can reduce between-location misallocation, making more people move to markets with higher average productivity. Second, it can reduce business floorspace costs, making businesses more profitable and thus increasing employment.

Studying this type of policy is particularly relevant for two reasons. Not only it is at the centre of policymaking discussions both in the UK and in several areas of the US (in the form of housebuilding targets, removing bans on high-rise buildings, and reducing local councils' discretionary powers in planning), but also represents a good setting to study the effects of localness heterogeneity in the context of the existing literature (Hsieh and Moretti, 2019; Diamond, 2016).

I simulate a policy bringing planning regulations to the level of the median Local Authority. To identify the effects of regulations on house price elasticities, I follow the approaches of Hilber and Vermeulen (2016) and Diamond (2016). If the housing sector follows a Cobb-Douglas production function in land and capital,[37] the equilibrium house prices $P_{H,jt}$ can be written as a function of interest rates r_t, local construction costs C_{jt}, and housing expenditure H_{jt}^D:

$$\begin{aligned} \ln(P_{H,jt}) &= \ln(r_t) + b_j \ln(H_{jt}^D) + \ln(C_{jt}) + \varepsilon_{jt}, \\ b_j &= b + b^{\text{geo}} x_j^{\text{geo}} + b^{\text{reg}} x_j^{\text{reg}}. \end{aligned} \tag{9}$$

Here, b_j represents the elasticity of local price to local demand. I assume it can be decomposed into a constant factor b, a geography factor $b^{\text{geo}} x_j^{\text{geo}}$, and a regulation factor $b^{\text{reg}} x_j^{\text{reg}}$. The goal is to estimate $b^{\text{reg}} x_j^{\text{reg}}$.

Estimation. I estimate $b^{\text{reg}} x_j^{\text{reg}}$ by following Hilber and Vermeulen (2016). They use the share of rejected planning applications on large projects (10 houses or more) as a proxy for x_j^{reg}, instrumented using a 2002 policy that briefly changed the incentives of restrictive planning authorities to reject applications.[38] I estimate the regression using data for 1998 to 2008 at the Local Authority Unit level. I report the results in Table 8. In my preferred IV specification (Column 2), a one standard deviation increase in regulations (captured by "ln(income) × Refusal share") increases the elasticity of house prices to demand by 0.173 log-points. The estimates are stable across other IV specifications.[39]

[37] This equation originates from the Cobb-Douglas production function of housing, standard in the literature and shared by my and their model. However, it can considered as a first-order approximation to several different formulations of the housing sector and its equilibrium.

[38] The source of endogeneity Hilber and Vermeulen (2016) worry about is the fact that developers may apply less in authorities known for being strict. Hence, this would reduce rejection rates, suggesting that OLS estimates are downward biased. Moreover, strict authorities may further "dilate" response times to postpone rejections, and hence appeals. The instrument takes care of these dimensions. OLS are lower than IV estimates in 8, as the direction of endogeneity would suggest.

[39] In Appendix I, I provide additional robustness checks, using real estate expenditure (which includes business floorspace rents), and - following Diamond (2016) - shift-share instruments.

Table 8: Estimation of regulation impact on housing price elasticity

	ln(house price)			
	OLS	IV		
	(1)	(2)	(3)	(4)
ln(income)	0.049*	0.038	0.036	0.039
	(0.026)	(0.061)	(0.069)	(0.060)
ln(income) × Refusal share	0.112***	0.173***	0.188***	0.171***
	(0.013)	(0.008)	(0.017)	(0.009)
Observations	3,444	3,444	3,444	3,444
R^2	0.788	0.732	0.702	0.735
Within Adjusted R^2	0.565	0.451	0.389	0.456
Kleibergen-Paap		95.3	145.9	58.0
Year fixed effects	✓	✓	✓	✓
Instruments		All	Delay	Labour votes
Endogenous		Refusal	Refusal	Refusal

Notes: p-values: *** ≤ 0.01, ** ≤ 0.05, * ≤ 0.10. Standard errors are Conley spatial standard errors, with a cut-off of 100km. Observations are Local Authority Unit × year, for the period 1998-2019. The dependent variable is ln(house price) in each specification. Controls include: Share of land built in 1990 within the LAU, delta elevation of terrain within the LAU, ln(income) × Share of land built in 1990, and ln(income) × delta elevation.

Model calibration. I average the LAU-level estimates across the model's five locations, weighting by population. The most regulated location is the South-East, with a regulation-induced additional elasticity $b^{\text{reg}}x_j^{\text{reg}}$ of 0.179 (one-third of its total elasticity), followed by London with 0.158 (one-fourth of its total elasticity). My policy counterfactual is to bring the regulation of all above-median Local Authority Units to the median level (0.136). In Table 9, I compare the baseline elasticities against the counterfactual values.

Table 9: Estimated house price elasticity from regulations ($b^{\text{reg}}x_j^{\text{reg}}$)

	Model Location					Median LAU
	London	Bristol	Liverpool	Hull	South-East	
Baseline	0.158	0.106	0.123	0.127	0.179	0.136
	(0.023)	(0.025)	(0.024)	(0.023)	(0.023)	(0.025)
Counterfactual	0.126	0.098	0.113	0.114	0.132	

Notes: Standard errors estimated from the IV specification reported Column 2 in Table 8. The first row shows the baseline estimation. The second row shows the counterfactual where the $b^{\text{reg}}x_j^{\text{reg}}$ parameter, capturing the elasticity of house price to demand, of each Local Authority Unit part has its regulatory elasticity set to the median value (0.136), if higher than their baseline value.

Results. I show the results of the two counterfactual exercises in Table 10. Accounting for heterogeneity in localness (first column) leads to lower gains in all dimensions. Relaxing housing regulations increases GDP by 0.6% (0.5 pp. less than without localness heterogeneity), value added of firms by 0.9% (0.4 pp. less), and real wages per worker by 1.1% (0.5 pp. less). Not accounting for heterogeneity in localness leads to overestimating welfare gains by over 40% (0.46 percentage points out of 0.97).

Table 10: Housing regulation policy effects

	Het. VA (%)	Fixed VA (%)	Difference (p.p.)
GDP	0.58	1.09	-0.52
VA of firms	0.88	1.29	-0.41
VA of firms: per worker	0.84	1.25	-0.40
Wages	1.13	1.63	-0.50
Wages: per worker	1.09	1.58	-0.50
Welfare	0.97	1.43	-0.46

Notes: All variables are in real terms, adjusted by the average price index of consumption of all agents. "GDP" is the total value of production, including housing. "VA of firms" is the total value added of firms. "Wages" are the total wage bill, and "Wages: per workers" are wages by the mass of employed individuals. "Welfare" is the mean value of the value function of all individuals. The "Het. VA" and "Fixed VA" column shows the differences between the "Heterogeneous VA effect" and the "Fixed VA effect" counterfactuals and the baseline calibration in percentage points. The last column shows the difference, in percentage points, between the two counterfactuals.

Why is the sign of the relative gain between the two counterfactuals flipped, relative to simply removing spatial frictions (Section 5)? Figure 7 provides a first intuitive explanation. Ignoring the negative correlation between productivity and localness leads to overestimating the effects of lowering local rental prices on the value added of productive firms, in particular in London (where the drop in prices is largest). Hence, employment shifts less towards higher-productivity firms (Figure 8). This implies smaller aggregate output, welfare, and wage gains in the economy.

Conceptually, the average productivity of the jobs created for location movers is lower when I account for localness heterogeneity, as low-productivity businesses expand their employment and wages (Figure 9) more than high-productivity ones in response to the lower local floorspace costs.

Figure 7: Relative VA, baseline (=1) vs counterfactuals

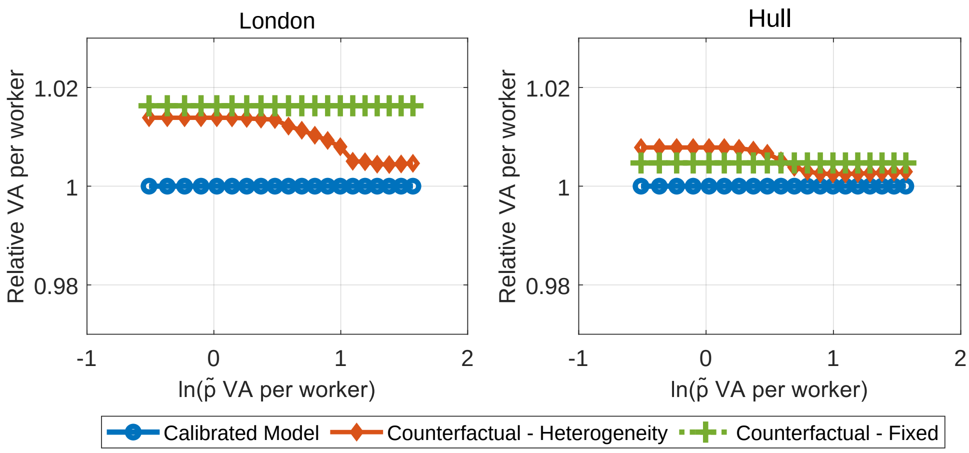


Figure 8: Relative employment, baseline (=1) vs counterfactuals

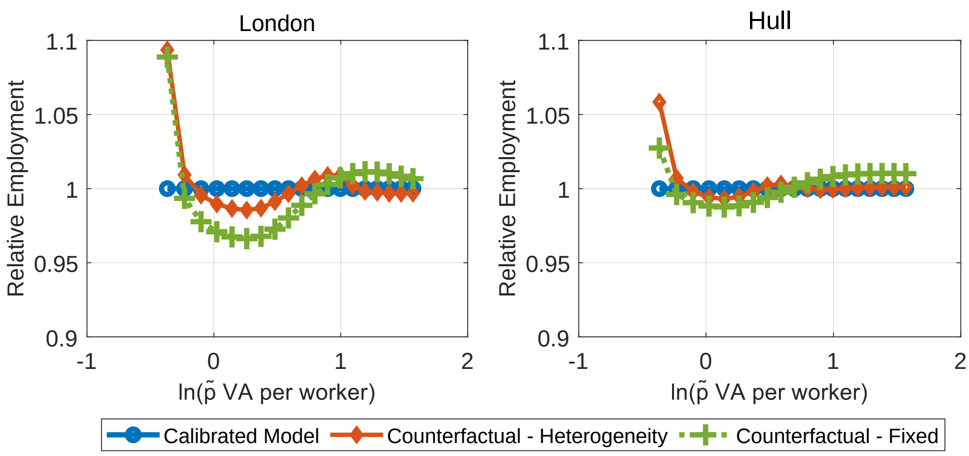
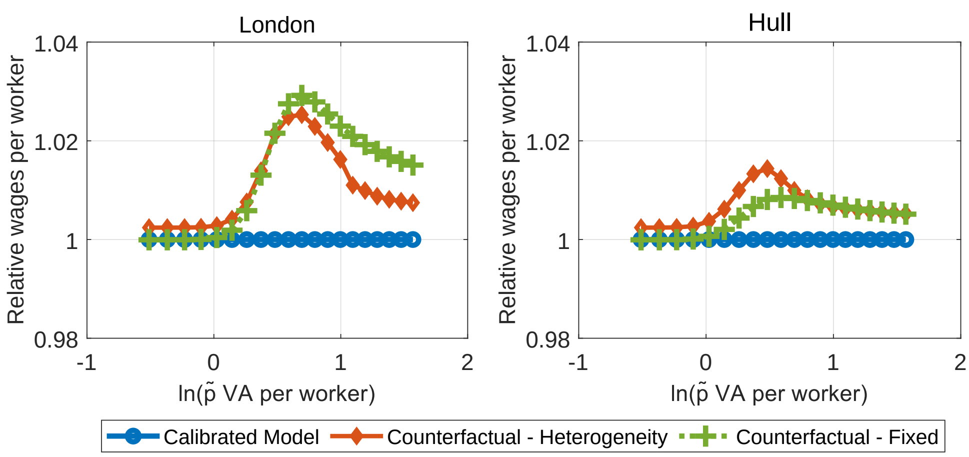


Figure 9: Relative wages, baseline (=1) vs counterfactuals



6.2 Levelling-up

I consider a policy where I increase the average productivity of high-skilled firms in "Hull" by 5 percentage points by moving firms from London. This is equivalent to reducing the 28 percentage point gap of average value added per worker of high-skilled firms by one-third.[40] This is reminiscent of the UK Government efforts to "spread out" skilled civil service jobs, but also policies that tend to favour deprived areas in the allocation of economic incentives and tax breaks.

The rationale for this policy is twofold. For policymakers, this type of policy is often associated with reducing inequalities or providing employment in areas with a high number of job seekers. From an economic point of view, moving jobs closer to people can deliver efficiency gains by i) reducing commuting costs and ii) reducing monopsony power in areas where there are few (Caldwell and Danieli, 2023; Datta, 2023) or bad (Heise and Porzio, 2023) outside options, due to being too far away from productive locations, and iii) reduce inefficient (rather than "high") unemployment (Kline and Moretti, 2013).

For this counterfactual, I recalibrate the model as follows. I fit the distribution of firms' productivity in "Hull" to a Beta distribution with a second parameter chosen to that the average productivity of firms is 5 percentage points higher than in the baseline. Then, I perform the procedure for London, reducing the average productivity of local firms by an amount such that the average productivity across all firms (across all locations) is unchanged relatively to the baseline.[41]

Results. I show the results in Table 11. Levelling up increases GDP, value added, real wages, and welfare. However, these gains are between 15% (value added and wages) and 42% smaller (welfare) when accounting for localness heterogeneity.

The reason for these results is similar to the insight derived for the housing regulation policy.

[40] Notice that this figure refers to the gap in average value added per worker of *firms*. Since more productive firms tend to be larger, the gap in the average productivity of *jobs* is considerably larger.

[41] This corresponds to reducing London's average productivity by $x^i = \frac{5}{100} \frac{M^i_{\text{London}}}{M^i_{\text{Hull}}}$.

By moving some productive firms from London to Hull, local demand and rental costs fall as - due to commuting costs - some residents move towards Hull, where now there are more jobs. However, congestion increases in Hull. Accounting for localness heterogeneity implies that, in the quantitative exercise, the lower congestion in London favours the creation of low-productivity jobs. On average, productivity (and thus wages and welfare) increase less.

Table 11: Levelling-up policy effects

	Het. VA (%)	Fixed VA (%)	Difference (p.p.)
GDP	0.26	0.30	-0.04
VA of firms	0.86	1.02	-0.17
VA of firms: per worker	0.88	1.03	-0.16
Wages	0.87	1.14	-0.27
Wages: per worker	0.89	1.15	-0.26
Welfare	0.43	0.76	-0.32

Notes: All variables are in real terms, adjusted by the average price index of consumption of all agents. “GDP” is the total value of production, including housing. “VA of firms” is the total value added of firms. “Wages” are the total wage bill, and “Wages: per workers” are wages by the mass of employed individuals. “Welfare” is the mean value of the value function of all individuals. The “Het. VA” and “Fixed VA” column shows the differences between the “Heterogeneous VA effect” and the “Fixed VA effect” counterfactuals and the baseline calibration, in percentage points. The last column shows the difference, in percentage points, between the two counterfactuals.

6.3 Commuting cost reduction

I study a policy reducing commuting costs by 21%, similar to the 75th percentile of bilateral county-level commuting cost reductions the US experienced between 1990 and 2010 (Monte et al., 2018). I interpret this reduction in commuting costs in the sense of lower pecuniary commuting costs (d_0), such as making public transport cheaper, utility commuting costs (d_1), such as making commuting faster, and distance-related relocation costs (κ_d), as making commuting cheaper/faster can also reduce the utility costs of moving residence (for example, due to being able to see distant family and friends, or search for new accommodation, at a lower pecuniary and time cost).

Results. Reducing commuting costs increases welfare by 0.75 percentage points, 9% more (0.06 pp.) than when not accounting for localness heterogeneity.[42] GDP falls slightly as the outside option of workers increases, but by half as much with localness heterogeneity. The value added of firms grows by just 0.12 percentage points, but this is 33% larger than the gains obtained without localness heterogeneity. The discrepancy between wages and welfare arises from the fact that less costly commutes allow firms to pay lower nominal wages (as making the job attractive to commuters becomes less expensive), while still delivering welfare gains for the average worker due to a fall in commuting costs and an increase in employed persons.

Table 12: Commuting cost reduction counterfactual

	Het. VA (%)	Fixed VA (%)	Difference (p.p.)
GDP	-0.11	-0.27	0.15
VA of firms	0.12	0.09	0.03
VA of firms: per worker	0.02	-0.01	0.02
Wages	-0.29	-0.38	0.09
Wages: per worker	-0.39	-0.48	0.09
Welfare	0.75	0.69	0.06

Notes: All variables are in real terms, adjusted by the average price index of consumption of all agents. "GDP" is the total value of production, including housing. "VA of firms" is the total value added of firms. "Wages" are the total wage bill, and "Wages: per workers" are wages by the mass of employed individuals. "Welfare" is the mean value of the value function of all individuals. The "Het. VA" and "Fixed VA" column shows the differences between the "Heterogeneous VA effect" and the "Fixed VA effect" counterfactuals and the baseline calibration, in percentage points. The last column shows the difference, in percentage points, between the two counterfactuals.

7 Conclusions

Frictions limiting workers' mobility have been shown to prevent workers from obtaining better jobs far away from home, reducing both output and welfare. Several policies have been proposed to tackle the costs of spatial frictions.

In this paper, I have shown that the relative effectiveness of these policies depends on which firms react to changes in local congestion. Using UK microdata, I have contributed to unpack

[42] This differs from Monte et al. (2018), who finds a 6 pp. increase in the US from a similar policy, because - having calibrated the model to five locations - I am considering interregional commuting, rather than county-level. That is, I am missing how improvements in commuting costs would actually benefit a much larger share of workers who commute within locations.

the black box of congestion externalities and found that high-productivity firms are less likely to be exposed to changes in local demand and rental costs. That is, these firms have low "localness".

Accounting for localness heterogeneity in a spatial general equilibrium model yields 26% larger estimates of output losses (1.1 percentage points). Welfare losses are 1.3 percentage points larger. Moreover, accounting for localness heterogeneity considerably affects the relative effectiveness of different policies. Reducing land use regulations (as in Hsieh and Moretti (2019)) or "levelling-up" deprived areas yields 35% and 25% lower welfare gains, respectively. Reducing commuting costs (as in Monte et al. (2018)) yields 10% larger gains. The reason is that the former two policies reduce congestion in productive locations, while the latter increases it (since productive locations are also high-amenity ones). Since the low localness of productive businesses dampens congestion effects towards zero, lowering (increasing) congestion does not have large positive (negative) aggregate effects.

This paper has three main takeaways. First, the effects of congestion are much more unequal *within* locations. For example, when accounting for localness heterogeneity, high land use regulations considerably depress wages in the bottom half of the wage distribution. This is due to low-productivity firms having low monopsony power: higher costs lead to lower employment and wages rather than lower profits.

Second, the losses arising from the congestion externalities on firms are small in *aggregate*. Hence, failing to account for the heterogeneous effects of congestion may lead to severely overestimate, or underestimate, the effects of policies with considerable general equilibrium effects on workers' location choices. This holds true also for realistic values of "dilation" effects of agglomeration (Combes et al., 2012).

Third, accounting for within-location firm heterogeneity is crucial for these results. Calibrating firm localness to location averages misses around half of the dampening of congestion externalities relative to a counterfactual where all firms have the same localness. Hence, the availability of detailed microdata is crucial when deviating from assumptions of zero congestion spillovers to firms' activity or when giving structural interpretations to reduced-form estimates of congestion effects.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999): "High Wage Workers and High Wage Firms," *Econometrica*, 67, 251–333, doi:https://doi.org/10.1111/1468-0262.00020.

Ahlfeldt, G. M. and A. Feddersen (2017): "From periphery to core: measuring agglomeration effects using high-speed rail," *Journal of Economic Geography*, 18, 355–390, doi:10.1093/jeg/lbx005.

Arnold, T., C. Wong, M. Baker, C. Koksal, A. Schulze Baing, and W. Zheng (2019): "Measuring Spatial Inequality in the UK: What We Know and What We Should Know?" Working paper.

Auerbach, J. U., E. Keller, J. Neira, and R. Singhania (2023): "Spatial Misallocation of Native Labor and Immigration," .

Azar, J., S. Berry, and I. E. Marinescu (2019): "Estimating Labor Market Power," SSRN Scholarly Paper ID 3456277, Social Science Research Network, Rochester, NY.

Bamford, I. (2021): "Monopsony Power, Spatial Equilibrium, and Minimum Wages," Mimeo.

Barrero, J. M., N. Bloom, and S. J. Davis (2023): "The Evolution of Working from Home," .

Bick, A., A. Blandin, and K. Mertens (2023): "Work from Home before and after the COVID-19 Outbreak," *American Economic Journal: Macroeconomics*, 15, 1–39, doi:10.1257/mac.20210061.

Bilal, A. (2023): "The Geography of Unemployment*," *The Quarterly Journal of Economics*, 138, 1507–1576, doi:10.1093/qje/qjad010.

Bilal, A., N. Engbom, S. Mongey, and G. L. Violante (2022): "Firm and Worker Dynamics in a Frictional Labor Market," *Econometrica*, 90, 1425–1462, doi:https://doi.org/10.3982/ECTA17955.

Burdett, K. and D. T. Mortensen (1998): "Wage Differentials, Employer Size, and Unemployment," *International Economic Review*, 39, 257–273.

Caldwell, S. and O. Danieli (2023): "Outside Options in the Labor Market," *Review of Economics Studies*.

Card, D., J. Rothstein, and M. Yi (2023): "Location, location, location," Nber working paper.

Chen, N., D. Novy, C. Perroni, and H. C. Wong (2023): "Urban-Biased Structural Change," *CEPR Discussion Paper No. 18522*.

Chi, B., A. Dennett, T. Oléron-Evans, and R. Morphet (2021): "Shedding new light on residential property price variation in England: A multi-scale exploration," *Environment and Planning B: Urban Analytics and City Science*, 48, 1895–1911, doi:10.1177/2399808320951212.

CICCONE, A. AND R. E. HALL (1996): "Productivity and the Density of Economic Activity," *The American Economic Review*, 86, 54–70.

COMBES, P.-P., G. DURANTON, L. GOBILLON, D. PUGA, AND S. ROUX (2012): "The Productivity Advantages of Large Cities: Distinguishing Agglomeration From Firm Selection," *Econometrica*, 80, 2543–2594, doi:https://doi.org/10.3982/ECTA8442.

DATTA, N. (2023): "Local Monopsony Power," Working paper.

DE LA ROCA, J. AND D. PUGA (2016): "Learning by Working in Big Cities," *The Review of Economic Studies*, 84, 106–142, doi:10.1093/restud/rdw031.

DIAMOND, R. (2016): "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000," *American Economic Review*, 106, 479–524, doi:10.1257/aer.20131706.

DIAMOND, R. AND C. GAUBERT (2022): "Spatial Sorting and Inequality," *Annual Review of Economics*, 14, 795–819, doi:10.1146/annurev-economics-051420-110839.

ELSBY, M. W. L. AND A. GOTTFRIES (2021): "Firm Dynamics, On-the-Job Search, and Labor Market Fluctuations," *The Review of Economic Studies*, 89, 1370–1419, doi:10.1093/restud/rdab054.

ENGBOM, N. AND C. MOSER (2022): "Earnings Inequality and the Minimum Wage: Evidence from Brazil," *American Economic Review*, 112, 3803–47, doi:10.1257/aer.20181506.

FAJGELBAUM, P. D. AND C. GAUBERT (2020): "Optimal Spatial Policies, Geography, and Sorting." *Quarterly Journal of Economics*, 135.

FREY, W. (2009): "The great American migration slowdown," *Brookings Institution, Washington, DC*.

GAUBERT, C., P. M. KLINE, AND D. YAGAN (2021): "Place-Based Redistribution," NBER Working Papers 28337, National Bureau of Economic Research, Inc.

GIANNONE, E., Q. LI, N. PAIXAO, AND X. PANG (2023): "Unpacking moving," .

HANSEN, S., P. J. LAMBERT, N. BLOOM, S. J. DAVIS, R. SADUN, AND B. TASKA (2023): "Remote Work across Jobs, Companies, and Space," Working Paper 31007, National Bureau of Economic Research, doi:10.3386/w31007.

HEISE, S. AND T. PORZIO (2023): "Labor Misallocation Across Firms and Regions," Working Paper 28792, National Bureau of Economic Research, doi:10.3386/w28792.

HERKENHOFF, K., L. OHANIAN, AND E. PRESCOTT (2018): "Tarnishing the golden and empire states: Land-use restrictions and the U.S. economic slowdown," *Journal of Monetary Economics*, 93, 89–109.

HERZ, B. AND T. VAN RENS (2019): "Accounting for Mismatch Unemployment," *Journal of the European Economic Association*, 18, 1619–1654, doi:10.1093/jeea/jvz018.

Hilber, C. A. L. and W. Vermeulen (2016): "The Impact of Supply Constraints on House Prices in England," *The Economic Journal*, 126, 358–405, doi:https://doi.org/10.1111/ecoj.12213.

Hsieh, C.-T. and E. Moretti (2019): "Housing Constraints and Spatial Misallocation," *American Economic Journal: Macroeconomics*, 11, 1–39, doi:10.1257/mac.20170388.

Jarosch, G. (2023): "Searching for Job Security and the Consequences of Job Loss," *Econometrica*, 91, 903–942, doi:https://doi.org/10.3982/ECTA14008.

Kline, P. and E. Moretti (2013): "Place Based Policies with Unemployment," *American Economic Review*, 103, 238–43, doi:10.1257/aer.103.3.238.

Kuhn, M., I. Manovskii, and X. Qiu (2022): "The Geography of Job Creation and Job Destruction," ECONtribute Discussion Papers Series 122, University of Bonn and University of Cologne, Germany.

Lindenlaub, I., R. Oh, and M. Peters (2023): "Firm sorting and spatial inequality," Tech. rep., National Bureau of Economic Research.

Manning, A. and B. Petrongolo (2017): "How Local Are Labor Markets? Evidence from a Spatial Job Search Model," *American Economic Review*, 107, 2877–2907, doi:10.1257/aer.20131026.

Marinescu, I. and R. Rathelot (2018): "Mismatch unemployment and the geography of job search," *American Economic Journal: Macroeconomics*, 10, 42–70.

Martellini, P. (2022): "Local labor markets and aggregate productivity," Working paper.

Monte, F., S. J. Redding, and E. Rossi-Hansberg (2018): "Commuting, Migration, and Local Employment Elasticities," *American Economic Review*, 108, 3855–90, doi:10.1257/aer.20151507.

Moretti, E. (2013): "Real Wage Inequality," *American Economic Journal: Applied Economics*, 5, 65–103, doi:10.1257/app.5.1.65.

——— (2022): "Place-based policies and geographical inequalities," *The IFS Deaton Review*.

Moscarini, G. and F. Postel-Vinay (2016): "Did the job ladder fail after the Great Recession?" *Journal of Labor Economics*, 34, S55–S93.

Saks, R. E. (2008): "Job creation and housing construction: Constraints on metropolitan area employment growth," *Journal of Urban Economics*, 64, 178–195, doi:https://doi.org/10.1016/j.jue.2007.12.003.

Schmutz, B. and M. Sidibé (2018): "Frictional Labour Mobility," *The Review of Economic Studies*, 86, 1779–1826, doi:10.1093/restud/rdy056.

——— (2021): "Search and Zipf: A model of Frictional Spatial Equilibrium," Working Papers 2021-01, Center for Research in Economics and Statistics.

van der List, C. (2023): "How do Establishments Choose their Location? Taxes, Monopsony, and Productivity," Tech. rep.

Woodland, A. D. (2017): *Tradable and Non-tradable Commodities*, London: Palgrave Macmillan UK, 1–7.

Şahin, A., J. Song, G. Topa, and G. L. Violante (2014): "Mismatch Unemployment," *American Economic Review*, 104, 3529–64, doi:10.1257/aer.104.11.3529.

A Additional details on decomposition

A.1 Continuum of firms

For $N_j \to \infty$ the problem can be easily recasted as an integration problem. Suppose all firm heterogeneity can be summarised by A. For example, if firms are identical given their productivity, A can be though as productivity. I will proceed with this example for readability, but the problem can be similarly specified by integrating across any number of types of firm heterogeneity that determine value added and size regardless of spatial frictions, and a dimension that summarises their susceptibility along this joint distribution.

Suppose there is a continuum of firms, heterogeneous in productivity A, and with value added per worker $\tilde{p}_j(A)$. Suppose $E_j(A)$ is the cumulated mass of employment in j up to productivity A, and that it can be written as the product of a function $e_j(A)$ and the CDF of local productivity $F_j(A)$: $E_j(A) = e_j(A)F_j(A)$. Call $VA_j(A) \equiv \tilde{p}_j(A)e_j(A)$. Then,

$$VA = \sum_{j\in\mathscr{L}} \int_A \tilde{p}(x)e_j(x)dF_j(x). \tag{10}$$

The derivative of VA to s is then

$$\frac{\partial VA}{\partial s} = \sum_{j\in\mathscr{L}} \left[\underbrace{\int_A \frac{\partial \ln(e(x))}{\partial s} VA_j(A)dF_j(x)}_{\substack{\text{Size}\\ \text{effect}}} + \underbrace{\int_A \frac{\partial \ln(\tilde{p}(x))}{\partial s} VA_j(A)dF_j(x)}_{\substack{\text{Value added}\\ \text{effect}}} \right]. \tag{11}$$

A.2 Wage decomposition

A similar decomposition can be carried out for wages. Call $\omega_{j,n}$ the wages paid by firm n in location j as a proportion of value added per worker.[43] The average wage per employee is

$$w \equiv \frac{1}{\sum\limits_{j\in\mathscr{L}} E_j} \sum_{j\in\mathscr{L}} \sum_{n=1}^{N_j} \omega_{j,n}\tilde{p}_{j,n}e_{j,n} = \frac{1}{\sum\limits_{j\in\mathscr{L}} E_j} \sum_{j\in\mathscr{L}} \sum_{n=1}^{N_j} \underbrace{w_{j,n}e_{j,n}}_{=W_{j,n}}.$$

Then,

[43] This can also be interpreted as the reciprocal of wage markdown: $\omega_{j,n} = 1 - \text{MD}_{j,n}$.

$$\frac{\partial w}{\partial x} = \sum_{j \in \mathscr{L}} \sum_{n=1}^{N_j} \left(\underbrace{\frac{\partial \ln(\omega_{j,n}(x))}{\partial x} W_{j,n}}_{\substack{\text{Wage markdown} \\ \text{effect}}} + \underbrace{\frac{\partial \ln(\tilde{p}_{j,n}(x))}{\partial x} W_{j,n}}_{\substack{\text{Value added} \\ \text{effect}}} + \underbrace{\frac{\partial \ln(e_{j,n}(x))}{\partial x} W_{j,n}}_{\substack{\text{Size} \\ \text{effect}}} \right).$$

B Additional facts about firms

B.1 Business Rates data

Nature of the tax. "Business rates" are a tax charged on non-domestic properties. The tax burden falls on the occupier, rather than the owner, and is calculated using open market rental values by the Valuation Office Agency (VOA). Then, these open market values are multiplied by a coefficient (on average 0.48 for the 2011-2018 period). A small discount ($\approx$ 0.02 coefficient points) is given to small businesses (defined as businesses with a rateable rate below £18,000 in 2021). Very small businesses (e.g. with a rateable value below £12,000 in 2023) receive a total exemption, which is gradually phased out until £15,000 (over which the full tax is paid). Several other discounts exists, for example for small pubs.

From tax to floorspace costs. Define floorspace cost as the rental price of floorspace, plus the business rate tax paid on the premises. Since the tax is (almost) a direct proportion of the average rents per square meter of similar properties in the area, it is possible to back out the approximate total floorspace costs of a firms by: i) dividing the reported business rate by the correct multiplier (to obtain the approximate rents), and ii) adding business rates themselves. However, since the tax has different bands and a no-tax area, I need to perform a few imputations. I provide details in the next section.

Imputation. For each business, I estimate the rateable value by dividing the declared tax by the yearly multiplier. I select the correct multiplier by checking whether, when dividing by the standard multiplier, the rateable value is above, or below, the small business threshold for the year. If below, I select the rateable value given by dividing the rateable value by the small business rate.

For businesses for which I obtain a rateable value below the yearly full-tax threshold (e.g. £15,000), I proceed as follows. For those reporting *some* tax that would give a rateable value below £15,000 (using the standard procedure), I calculate the implied rateable value by solving the equation (Observed Tax) $=$ (rate $\times$ rateable value) $\times \frac{\text{rateable value} - \text{NoTaxThreshold}}{\text{DiscountThreshold} - \text{NoTaxThreshold}}$. For zero values, which imply a rateable value below the yearly no-tax threshold, I impute an average rateable value of £7,000 in 2015 British Pounds. This is equivalent to a £538/month rental cost. For comparison, small retail units near the high street of Blackpool, one of the most deprived cities in England, were up for rent at no less than £650 as of September 2023.

B.2 Robustness checks for floorspace costs

I validate my preferred measure of floorspace costs against different data sources.

Using LSOA data on business rates per square meter, floorspace, additional rates, employment, and gross value added, I replicate results similar to Figure 4 with three caveats: i) I use as denominator GVA per worker, ii) I use as unit of observation the LSOA, rather than the business, and iii) I do not plot the data against the CDF of employment, but the CDF of LSOA-level value added per worker. I plot the results in Figure B.1. The results replicate the downward slope, although the scale is somewhat different: productive locations see businesses spending around 4% of their GVA on rents. Businesses in the least productive locations spend around 9.5% of the average GVA rents. Conversely, in the microdata firms representing the top 20% of employment spend around 5.2% of their GVA on rents, and firms representing the bottom 20% spend around 17%. These differences can be rationalised by three facts:

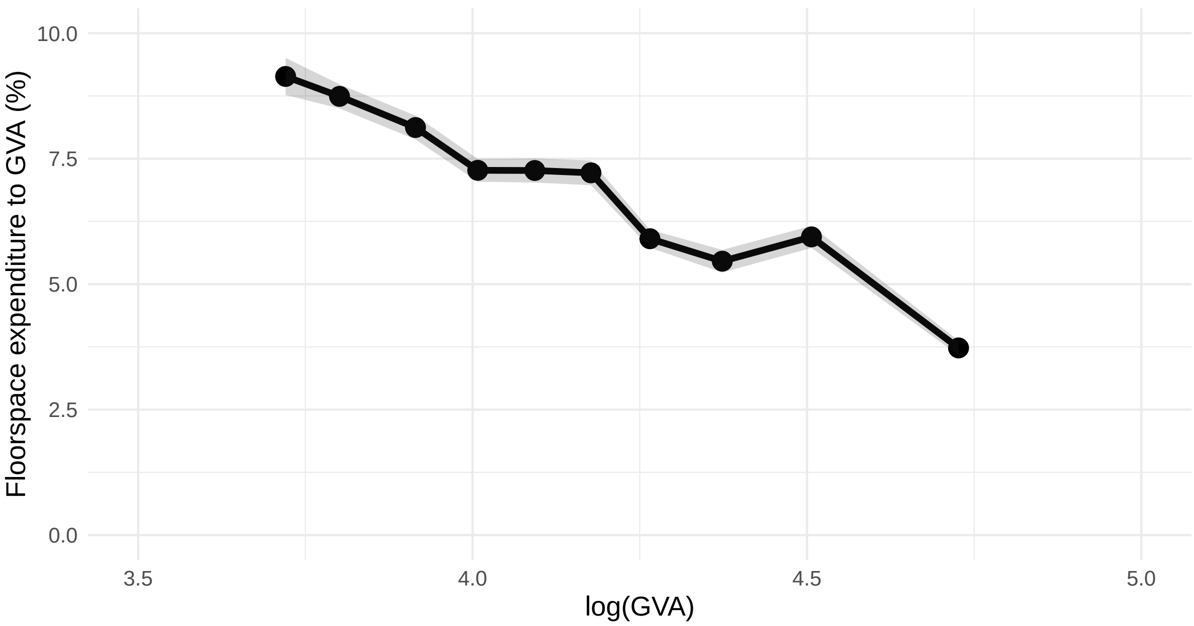
1. ABS data do not include financial firms. However, I am unable to exclude them in the LSOA-level analysis. Since these are high-value-added firms, they are likely to dampen the expenditure estimates across all locations. Notice how accounting for financial firms in the distribution of floorspace expenditure to GVA would *strengthen* the results presented in the main text.

2. LSOA floorspace data are severely censored, with approxmately half of the observations reporting missing data. This happens because areas with only a few units are excluded from the counts for privacy reasons. However, areas with very few units may be residential areas with either a few commercial activities (which pay high rents relative to their

GVA), or areas hosting large plants (which usually rent at a premium, relatively to small light manufacturing units). Hence, this would bias my estimates down across the whole distribution, and in particular in the left tail (mom-and-pops businesses). This bias would be consistent with the discrepancies observed between ABS microdata and LSOA-level estimates.

3. Locations with low GVA per worker may also host *some* productive businesses, lowering the average rent-to-GVA relative to the microdata estimates for low-productivity businesses. In fact, among the non-missing LSOAs, the 10% least productive locations has a mean GVA per worker of £33,000. For comparison, the bottom 10% of businesses has a GVA per worker of just £15,000 or less. Again, this factor would be consistent with the discrepancies between microdata and aggregate estimates, showing the robustness of the former.

Figure B.1: Floorspace costs as share of GVA, by LSOA deciles



Note: The figure plots the ratio between local business rates and local Gross Value Added (GVA) at ten deciles of the LSOA-level distribution of GVA per worker. LSOAs are statistical units representing areas of 1000-3000 residents.

Finally, I provide third-party evidence for the fact that low-productivity businesses do indeed spend large shares of their value added on rents by looking at evidence from sectoral studies. The 2017 report by Property Industry Alliance (PIA),[44] a consortium of real estate funds associations and the Urban Land Institute research centre, illustrates how retail businesses' floorspace costs (rents plus business rates) amount to approximately 50% of retail values of the retail sector, but only 12.5% of the "office sector". These data are compatible, assuming retail businesses are low-wage and the "office sector" medium to high wage, with the ones described in Figure 4, despite the rental costs in the PIA report are obtained from a different source.

B.3 Floorspace prices and house prices

In the model, I assume housing and floorspace costs are derived from the same good ("housing services"). However, in practice, houses and offices (or manufacturing plants) are likely to be different goods. Hence, their price does not need to covary substantially. For example, planning regulations involving residential and non-residential buildings may follow different processes. I deal with with potential concern in two ways:

1. I calibrate the model's key parameter for floorspace, $\phi()$, to the average expenditure in the economy. In this way, both housing and floorspace are calibrated to match total expenditure, and not physical consumption of floorspace. This captures how the *physical* space fitted for offices, manufacturing, or residential purposes can have very different costs. This choice is important also for the interpretation of what "consuming more housing services" means. Consuming more housing services does not necessarily translate into moving to larger houses, as consumers can also choose between better neighbourhoods, better built houses, and so on and so forth.

2. I show that - given the difference in *average* price (discussed above) - house prices and business rents are highly correlated at the local level, meaning that expensive locations for workers are also expensive for firms. In Figure B.2 I plot the average residential rent of

[44] Available at https://bpf.org.uk/media/3278/bpf-pia-property-report-2017-final.pdf, page 13.

each Local Authority Unit in 2019 against the average industrial business rates per square meter for the same year. The correlation is very high, both on values (Pearson correlation 0.91) and ranks (Spearman correlation 0.87).

Separately accounting for different types of industrial properties (industrial, retail, offices), as in Figure B.3, provides similar results. Hence, the overall correlation does not depend on property-type composition effects. These facts justify treating floorspace as having the same nature as housing. This is not fully surprising, as land itself represents a major component of property prices.

Figure B.2: Residential rents vs business rates, by LAU

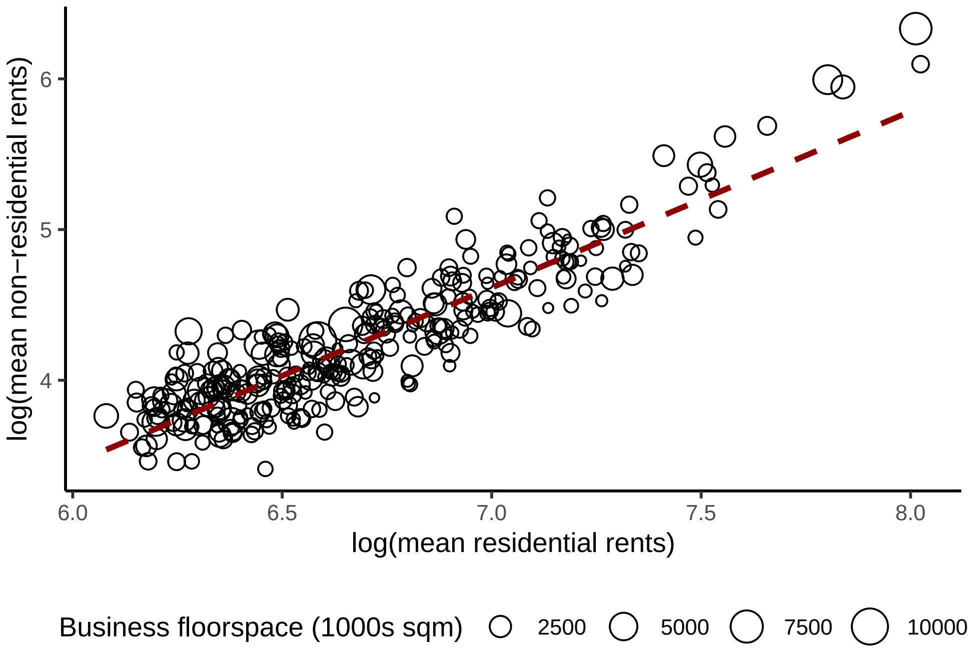
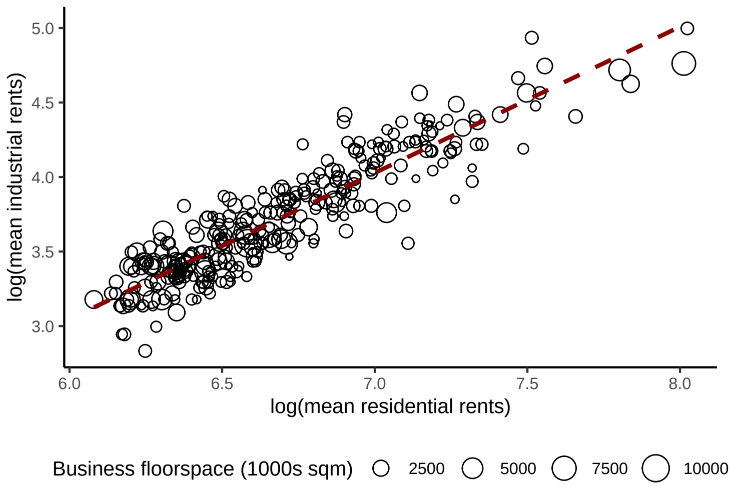
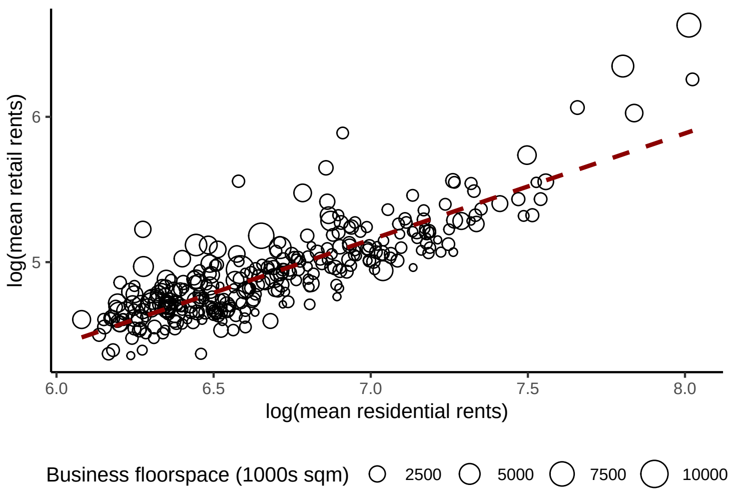


Figure B.3: Residential rents vs business rates, by LAU and space type

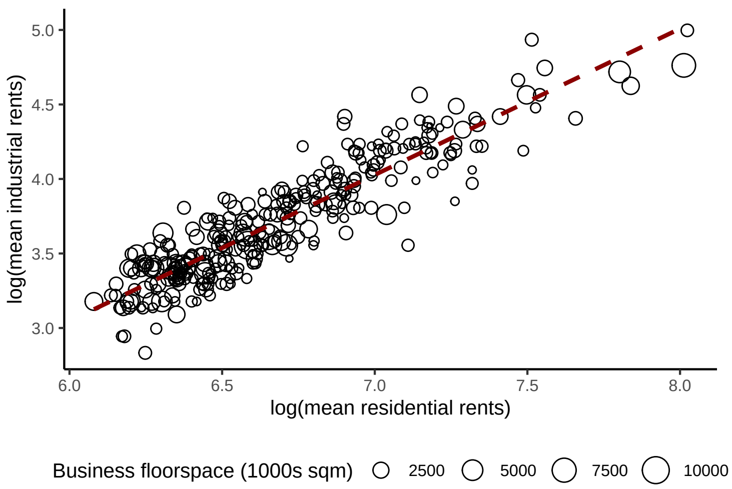
(a) Industry properties



(b) Retail properties



(c) Offices



C Additional derivations for the model

C.1 Goods market equilibrium

Define local income in j as Y_j^r, which includes: i) wages and home production of residents of j, ii) vacancy costs spent in j (recall vacancy costs can be seen as the cost of hiring local headhunters and recruitment agencies), and iii) firms' profits. The income of landowners and housing capital does not enter in this definition, as they are assumed to be foreign residents.

In the non-tradeable good markets, the total expenditure of residents on non-tradeables must be equal to the value of the production of tradeables at all firms in j:

$$(1-\eta_2)Y_j^r = P_{l,j}A_{l,j}\sum_{i\in I}M_j^i\int_p (1-\beta(s))sn^*(s)v^*(s)\gamma_j^i(s)ds.$$

This implies that the market-clearing price of non-tradeables $P_{l,j}$ is

$$P_{l,j} = \frac{\eta_1 Y_j^r}{A_{l,j}\sum_{i\in I}M_j^i\int_p (1-\beta(s))sn^*(s)v^*(s)\gamma_j^i(s)ds}. \tag{12}$$

In the housing market, demand is given by both workers and firms. Calling $H_{w,j}^D = P_{H,j}^{-1}(1-\eta_1-\eta_2)Y_j^r$ the demand for residential housing services and $H_{f,j}^D = \sum_{i\in I}M_j^i\int_p \phi(s)sn^*(s)v^*(s)\gamma_j^i(s)ds$ the demand for business floorspace, the market clearing price must satisfy

$$P_{H,j} = \frac{r}{(1-b_j)}\tilde{A}_{H,j}^{-\frac{1}{1-b_j}}\left[H_{w,j}^D + H_{f,j}^D\right]^{\frac{b_j}{1-b_j}}. \tag{13}$$

C.2 Worker choice and value function

The probability that a worker accepts an offer is

$$\mu_{hjj'}^{E,i}(w,w') = \frac{\exp\left(\sigma^{-1}IV_h^i(j',w')\right)}{\exp\left(\sigma^{-1}(W_{hj}^i(w)\right) + \exp\left(\sigma^{-1}IV_h^i(j',w')\right)}, \tag{14}$$

where $IV_h^i(j',w') = \sigma_m \ln\left(\sum_{h'\in\mathscr{L}}\exp\left(\sigma_m^{-1}(W_{h'j'}^i(w') - \kappa_{hh'})\right)\right)$.

Identical definitions can be derived for the unemployed:

$$\mu_{hjj'}^{U,i}(w,w') = \frac{\exp\left(\sigma^{-1}IV_h^i(j',w')\right)}{\exp\left(\sigma^{-1}U_h^i\right) + \exp\left(\sigma^{-1}IV_h^i(j',w')\right)}. \tag{15}$$

The probability that a worker moves from location h to location h', conditional on accepting an offer in j' is (for both employed and unemployed workers):

$$m^i_{hh'j'} = \frac{\exp\left(\sigma_m^{-1}(W^i_{h'j'}(w') - \kappa_{hh'})\right)}{\sum_{h''\in\mathscr{L}} \exp\left(\sigma_m^{-1}(W^i_{h''j'}(w') - \kappa_{hh''})\right)}. \tag{16}$$

Value Function. Call σ the inverse scale parameter of the upper ε (firm-specific idiosyncratic) shocks and σ_m the inverse scale parameter of the $\varepsilon_{h'}$ (residence-specific idiosyncratic) shocks from Equation 5. Then, the expected value of an offer is

$$V^i_{hjj'}(w,w') = \sigma \ln\left[\exp\left(\sigma^{-1}W^i_{hj}(w)\right) + \frac{\sigma_m}{\sigma}\ln\left(\sum_{h'\in\mathscr{L}} \exp\left(\sigma_m^{-1}(W^i_{h'j'}(w') - \kappa_{hh'})\right)\right)\right].$$

Thus, the value function of a worker is the sum of instantaneous utility from living in h, working in j, at wage w, plus the continuation value of search, the continuation value of unemployment, minus the cost of effort:

$$\begin{aligned}\rho W^i_{hj}(w) =& \frac{\tau_h(\alpha^i w - d^i_{0,hj})}{P_h d^i_{1,hj}} + \delta(U^i_h - W^i_{hj}(w)) \\ &+ \sum_{x\in\mathscr{L}} \max_{s_x}\left(\int_{w'} a^i_{hjx}\frac{s_x^{1+\varepsilon_0}}{1+\varepsilon_0}\theta_{ix}^{1-\chi}(V^i_{hjx}(w,s) - W^i_{hj}(w))dF_j(s)\right) - \psi(S).\end{aligned} \tag{17}$$

To obtain the value function, first define the net continuation value of an offer as $\overline{V}^{E,i}_{hjx}(w) = V^i_{hjx}(w,s) - W^i_{hj}(w)$, and then solve for the optimal s_x.

$$s_x = \frac{S^{\frac{\varepsilon_1}{\varepsilon_0}}}{\left(a^i_{hjx}\theta_{ix}^{1-\chi}\overline{V}^{E,i}_{hjx}(w)\right)^{\frac{1}{\varepsilon_0}}}. \tag{18}$$

Then, notice how the aggregate effort is

$$S^i_{hj}(w) = \left[\sum_{x\in\mathscr{L}}\left(a^i_{hjx}\theta_{ix}^{1-\chi}\overline{V}^{E,i}_{hjx}(w)\right)^{-\frac{1}{\varepsilon_0}}\right]^{\frac{\varepsilon_0}{\varepsilon_1-\varepsilon_0}}. \tag{19}$$

Plugging Equations 18 and 19 into Equation 17 and performing the appropriate algebraic simplifications leads to

$$\begin{aligned}\rho W^i_{hj}(w) =& \underbrace{\frac{\tau_h\alpha^i(w - d^1_{hj})}{P_h d^0_{hj}}}_{\text{Istantaneous utility}} + \underbrace{\delta(U^i_h - W^i_{hj}(w))}_{\text{Net value of unemployment}} \\ &+ \underbrace{\frac{\varepsilon_1-\varepsilon_0}{(1+\varepsilon_0)(1+\varepsilon_1)}\left[\sum_{x\in\mathscr{L}}\left(a^i_{hjx}\theta_{ix}^{1-\chi}\overline{V}^{E,i}_{hjx}(w)\right)^{-\frac{1}{\varepsilon_0}}\right]^{\frac{\varepsilon_0(1+\varepsilon_1)}{\varepsilon_0-\varepsilon_1}}}_{\text{Net continuation value of search}}.\end{aligned}$$

Similarly, using identical definitions for offers received while unemployed,

$$\rho U_h^i = \frac{\tau_h \alpha^i b_h}{P_h} + v \frac{\varepsilon_1 - \varepsilon_0}{(1+\varepsilon_0)(1+\varepsilon_1)} \left[\sum_{x \in \mathscr{L}} \left(a_{hx}^i \theta_{ix}^{1-\chi} \overline{V}_{hx}^{U,i} \right)^{-\frac{1}{\varepsilon_0}} \right]^{\frac{\varepsilon_0(1+\varepsilon_1)}{\varepsilon_0 - \varepsilon_1}} .$$

C.3 P and Q functional forms

Recall

$$n_{jh}^i(w) = \frac{\alpha^i P_{jh}^i(w) \theta_{ij}^{-\chi}}{Q_{hj}^i(w)}.$$

Then,

$$P_{jh}^i(w) = \frac{1}{\overline{a}_j^i} \left(\underbrace{\sum_{h' \in \mathscr{L}} \sum_{j' \in \mathscr{L}} E_{h'j'} \int_{w'} a_{h'j'j}^i(w') m_{h'j'hj}^i(w',w) \mu_{h'j'j}^{E,i}(w',w) dE_{h'j'}^i(x)}_{\text{Mass hired from job-tojob moves}} \right.$$

$$\left. + \underbrace{\sum_{h' \in \mathscr{L}} u_{h'} a_{h'j'j}^{u,i} m_{h'hj}^i(w) \mu_{h'j}^{U,i}(w)}_{\text{Mass hired from unemployment}} \right) . \tag{20}$$

$$Q_{hj}^i(w) = \delta^i + \underbrace{\sum_{j' \in \mathscr{L}} a_{hjj'}^i(w) \theta_{ij'}^{1-\chi} \underbrace{\int_{w'} \mu_{h'j'j}^{E,i}(w,x) dF_{i,j'}^w(x)}_{\text{mean acceptance rate to j'}}}_{\text{Outflow rate from job-to-job moves}} .$$

The first part of P_{jh} is the mass of already employed workers that finds a job at firms posting wage w in j, given by: i) the mass of employees in the workers' previous locations h', j', multiplying ii) the average probability with which a worker who received an offer accepts and moves to h. The second part provides a similar expression for the unemployed.

Q represents the rate at which workers quit, being the sum of: i) the exogenous unmatching rate δ^i and ii) the outflow rate from job-to-job moves, that is the rate at which workers receive offers $a_{hjj'}^i(w)\theta_{ij'}^{1-\chi}$, and the probability they accept them ($\int_{w'} \mu_{h'j'j}^{E,i}(w,x) dF_{i,j'}^w(x)$).

C.4 Equilibrium proof

Recall profits per vacancy are

$$\pi_h(\tilde{p}, w) = (\tilde{p} - w) n_j^i(w),$$

where $n_h(w) = \sum_h n^i_{jh}(w)$. Assume that $\pi(\tilde{p}, w)$ is continuous in w for a given p. Then, profit maximisation over w implies the necessary first order condition

$$(\tilde{p} - w)\frac{\partial n^i_h(w)}{\partial w} = n^i_h(w).$$

Expanding n under the derivative in the left-hand side of the equation, this can be written as

$$(\tilde{p} - w)\frac{\partial \left(\sum_h n^i_{jh}(w)\right)}{\partial w} = n^i_j(w). \tag{21}$$

Recall that in the stationary Equilibrium, $n^i_{jh}(w) = \frac{\alpha^i P^i_{jh}(w)\theta^{-\chi}_{ij}}{Q^i_{hj}(w)}$ (see discussion around Equation 6 in the main text). Hence, the derivative of n to w is

$$\frac{\partial n^i_{jh}(w)}{\partial w} = \alpha^i \left(\frac{\frac{\partial P^i_{jh}(w)}{\partial w} Q^i_{hj}(w) + \frac{\partial Q^i_{hj}(w)}{\partial w} P^i_{jh}(w)}{\left(Q^i_{hj}(w)\right)^2} \right). \tag{22}$$

Plugging Equation 22 into 21, we obtain

$$n^i_h(w) = (\tilde{p} - w)\left[\alpha^i \theta^{-\chi}_{i,j} \sum_{h \in \mathscr{L}} \frac{\overline{a}^i_{hj}}{\overline{a}^i_j} \left(\frac{\frac{\partial P^i_{jh}(w)}{\partial w} Q^i_{hj}(w) + \frac{\partial Q^i_{hj}(w)}{\partial w} P^i_{jh}(w)}{\left(Q^i_{hj}(w)\right)^2} \right) \right]. \tag{23}$$

Finally, define $\tilde{Q}^i_{hj}(\tilde{p}) = Q^i_{hj}(w(\tilde{p}))$, and $\tilde{P}^i_{hj}(\tilde{p}) = P^i_{hj}(w(\tilde{p}))$. Then, for $w^i_j(\tilde{p})$ being a continuous and differentiable function of $\tilde{p}$ mapping the effective productivity to the optimal wage in w,[45]

$$\begin{aligned} \frac{\partial \tilde{P}^i_{jh}(\tilde{p})}{\partial \tilde{p}} &= \frac{\partial P^i_{jh}(w(\tilde{p}))}{\partial w}\frac{\partial w(\tilde{p})}{\partial \tilde{p}}, \\ \frac{\partial \tilde{Q}^i_{jh}(\tilde{p})}{\partial \tilde{p}} &= \frac{\partial Q^i_{jh}(w(\tilde{p}))}{\partial w}\frac{\partial w(\tilde{p})}{\partial \tilde{p}}. \end{aligned} \tag{24}$$

Plugging 24 into 23 and bringing $\frac{\partial w(\tilde{p})}{\partial \tilde{p}}$ to the LHS,

$$\frac{\partial w^i_j(\tilde{p})}{\partial \tilde{p}} = \frac{\left(\tilde{p} - w^i_j(\tilde{p})\right)\left[\alpha^i \theta^{-\chi}_{ij} \sum_{h \in \mathscr{L}} \frac{\overline{a}^i_{hj}}{\overline{a}^i_j} \left(\frac{\frac{\partial \tilde{P}^i_{hj}(\tilde{p})}{\partial \tilde{p}} \tilde{Q}^i_{hj}(\tilde{p}) + \frac{\partial \tilde{Q}^i_{hj}(\tilde{p})}{\partial \tilde{p}} \tilde{P}^i_{hj}(\tilde{p})}{\left(\tilde{Q}^i_{hj}(\tilde{p})\right)^2} \right) \theta^{-\chi}_{ij} \right]}{\tilde{n}^i_j(\tilde{p})}.$$

Finally, this can be used to pin down $w^i_j(\tilde{p})$ as a differential equation, with initial condition given by the maximum between the local reservation wage R_j (below which $n = 0$), and the

[45] P and Q are differentiable in w, as can be seen by the definition of their components in the previous sections.

profit-maximising wage of the least profitable active firm:

$$w_j^i(\tilde{p}) = w_j^i(\underline{\tilde{p}}) + \int_{\underline{\tilde{p}}_j}^{\tilde{p}} \frac{\partial w_j^i(\tilde{p})}{\partial \tilde{p}} dF_j^i(p),$$

$$w_j^i(\underline{\tilde{p}}_j) = \max\left\{R_j, \arg\max_{\hat{w}}(\underline{\tilde{p}}_j n_j^i - \hat{w} n_j^i(\hat{w}))\right\}.$$

This pins down $w_j^i(\tilde{p})$ as a continuous function of $\tilde{p}$, consistently with the assumptions.

D Calibrated Parameters

D.1 Exit rate δ

Since I am estimating the model using full-time employees only, δ needs to be calibrated to the sum of workers: i) exiting the workforce, ii) entering unemployment, and iii) becoming involuntary part-time employees. I calculate δ^i using the Labour Force Survey (LFS). Since δ represents an arrival rate, I then transform the quarterly separation probabilities from the LFS into yearly arrival rates of separation shocks. In table D.1, I report both the estimated separation probability and the arrival rate.

Table D.1: Estimates of δ^i

Skill	δ^i	Yearly unmatch share
High	0.071	0.0704
		(0.015)
Low	0.133	0.1220
		(0.023)

D.2 Relative skill of high-skilled α^H

I estimate the relative skill of high- and low-skilled workers using an AKM estimation. While AKM is unsuitable to estimate second moments, the means of worker and firm fixed effects on wages are consistent (Abowd et al., 1999). Normalising the skill of the low-skilled to 1, I find that the worker component of the high-skilled is 29% higher than the one of the low-skilled. The results are reported in Table D.2.

Table D.2: Estimates of α^i

Skill	α^i
High	1.29
Low	1

D.3 Local Prices

Housing. I use private rental market statistics by the Office of National Statistics (ONS) to obtain an index of costs of private housing services. In order to account for differences in individual preferences and the segmentation of the housing market across locations, I exclude 4+ bedrooms, rooms, and studios from the data. Then, I compute a weighted average of all rents across location groups. Table D.3 shows the results in terms of units of mean rent/price.

I perform two robustness checks. First, I include all properties. Second, using the dataset on house price transactions by Chi et al. (2021), and motivated by their results on the importance of price per square meter in correctly assessing relative house prices, I perform an hedonic regression of cross-sectional house prices on several characteristics associated with building quality and type. I find that the difference between the simple average of rental prices and the mean residuals of the hedonic regressions are negligible, apart from the set of "Bristol" locations (where the deviation is, in any case, within 10% of the base estimate). In the estimation of my model, I will use the estimates from this last hedonic regression.

Table D.3: Housing service prices, relative to mean

	Location				
	London	Bristol	Liverpool	Hull	South-East
Baseline (rents)	1.95	0.86	0.78	0.76	1.09
All properties (rents)	1.93	0.85	0.77	0.76	1.09
Hedonic regression (prices)	1.97	0.78	0.75	0.72	1.11

Non-tradeables. For local non-tradeable prices, I take estimates from regional CPI data published by ONS. These data are published at very infrequent intervals as experimental statistics. For my purpose, I use the 2010 edition of the publications.[46] Since the regional CPI does not provide a perfect mapping to my model's locations $j \in \mathscr{L}$, I calculate CPI_j as the weighted average of the regional CPIs, where the weights are given by the sum of the population weights of all Local Authority Units that are part of the (model) location j and of each administrative

[46] Available at https://www.ons.gov.uk/economy/inflationandpriceindices/articles/relativeregionalconsumerpricelevelsuk/2016, "Comparison with 2010 results" section.

region.

Since the local price index the average between the prices of tradeables and non-tradeables, but exclude all housing costs,[47] the local price of non-tradeables can be recovered as:

$$P_{l,j} = (\mathrm{CPI}_j)^{\frac{\eta_1+\eta_2}{\eta_2}}.$$

The results are presented in Table D.4.

Table D.4: Non-tradeable prices

	Location				
	London	Bristol	Liverpool	Hull	South-East
Baseline	1.19	1.02	0.99	0.96	1.05

D.4 Mass of Employers M_j^i

Using data on individual plants from the universe of firms dataset (part of ABS) and aggregating over the firm identifier, I compute the average number of firms that operate in a location during the 2011-2018 period. This gives me the total mass of firms across locations, M_j.

Using ASHE data, for each 3-digits industry code I calculate the share of high- and low-skilled workers in each location. Then, I impute M_j^i by weighting each firm according to the share of i-skill workers in location j and each firm's 3-digits industry code. That is, for N_j being the number of firms in the dataset,

$$M_j^i = \frac{1}{N_j} \sum_{n=1}^{N_j} s_{j,n}^i,$$

with $s_{j,n}^i$ being the share of workers in occupation i in the 3-digits 2007 industry code of firm n. The results are provided in Table D.5

Table D.5: Employer Mass

	London	Bristol	Liverpool	Hull	South-East
High-Skilled	0.1962	0.1102	0.1046	0.1364	0.1969
Low-Skilled	0.0507	0.0378	0.0464	0.0673	0.0534

[47] In the UK, the CPI index does not include housing costs, besides household services and basic maintenance.

D.5 Minimum productivity

I assume the minimum level of productivity in the economy is £15,000 in 2015 prices. This is compatible with the minimum wage at the time, equivalent to £13,980 per year for a full-time worker, plus social security costs of employment. In model terms, this is equivalent to 53.6% of the £28,000 average value added in the worst location for low-skilled jobs, which is normalised to 1 in the calibration. Thus, I calibrate $\underline{p}_j^L = 0.536$. Since high-skilled workers have higher productivity per employee, I calibrate $\underline{p}_j^H = \frac{\underline{p}_j^L}{\alpha^H}$ so that the two skills have the same minimum income $\alpha^i w(\underline{p}_j^i) = \iota \times 0.536 \approx 0.526$.

D.6 Distance matrix

To calculate the distance matrix between locations, I obtain data from Google Maps about car travel distance at peak hour between the population centroids of all Local Authority Units. Then, using Open Street Map, I calculate the public transport travel plus walking times using the 2018 schedule of buses and trains, obtained from public data available through the Department of Transport. I use as maximum travel time 300 minutes, and then normalise the average distance, weighted by origin population and destination jobs, between "London" locations and "Hull" locations as 1.

Table D.6: Distance Matrix

	Location				
	London	Bristol	Liverpool	South-East	Hull
London	0	0.67	0.79	0.35	1.00
Bristol	0.67	0	0.25	0.68	0.50
Liverpool	0.79	0.25	0	0.78	0.50
South-East	0.35	0.68	0.78	0	0.98
Hull	1.00	0.50	0.50	0.98	0

D.7 Elasticity of Housing Price

I estimate the elasticity of housing price to expenditure (b_j in model terms) by estimating Equation 13 in logs. I obtain house price data and local disposable income at the local authority level for the period 1998 to 2019. Assuming the share of workers' expenditure on housing follows the model, I obtain total expenditure on residential housing a $(1-\eta_1-\eta_2)$ share of disposable income. Then, I add commercial rents using business rates data to obtain expenditure E_{lt}. Then, I estimate by OLS the following expression:

$$\ln(P_{lt}) = \alpha + \sum_{j\in\mathscr{L}} b_j d_j \ln(E_{lt}) + \phi_l + \gamma_t + \varepsilon_{lt},$$

where d_j is a (model) location dummy, ϕ_l is a local authority fixed effect, and γ_t a time fixed effect. b_j represents the parameter of interest. The results are reported in Table D.7.

Table D.7: Estimated house price elasticity to demand b_j

	Location				
	London	Bristol	Liverpool	South-East	Hull
Total expenditure	0.682	0.375	0.437	0.479	0.409
	(0.066)	(0.100)	(0.133)	(0.063)	(0.101)
Housing expenditure	0.677	0.372	0.435	0.475	0.410
	(0.067)	(0.101)	(0.136)	(0.063)	(0.103)

D.8 Search decay in distance

I show how the rate of search efficiency decay with distance, c_d^i, can be estimate directly from worker's job moves microdata. The rate at which individuals living in h and working in j with wage w obtain a job in j' is

$$(\text{Move rate})^i_{hj'}(w) = \underbrace{a^i \theta_{j'}^{1-\chi} e^{c_d^i d_{hj'}}}_{\text{Match rate}} \underbrace{\sum_j \frac{\left(s^i_{hjj'}(w)\right)^{1+\varepsilon_0}}{1+\varepsilon_0}}_{\text{total effort}} \underbrace{\int \mu^{i,E}_{hjj'}(w,w')dF^f_{j'}(w')}_{\text{avg. acceptance}}.$$

Taking logs and replacing $S^i_{hjj'}(w)$ with its optimal value from the worker problem, we can express the right-hand side as the sum of a constant that is w-independent, and a function of w:

$$\ln\left((\text{Move rate})^i_{hj'}(w)\right) = \underbrace{\ln\left(\frac{\left(a^i\theta_{j'}^{1-\chi}e^{c_d^i d_{hj'}}\right)^{-\frac{1}{\varepsilon_0}}}{1+\varepsilon_0}\right)}_{\beta^i_{hj'}} + \underbrace{\ln\left[\sum_j \left(\frac{(S_{hj}(w))^{\frac{\varepsilon_1}{\varepsilon_0}}}{(\overline{V}^{i,E}_{hj}(w))^{\frac{1}{\varepsilon_0}}}\right)^{1+\varepsilon_0} \int P^{i,E}_{zjj'}(w,w')dF^f_{j'}(w')\right]}_{f^i_{hj'}(w)}. \tag{25}$$

I estimate $\beta^i_{hj'}$ by estimating Equation 25 using a Poisson regression. For each h and j', I subset my dataset to keep only h-resident workers of type i, and create a dummy variable $D_{kt}(j'(k) = j')$ if worker k moves to work in j' in period t. Then, I estimate the following equation by implementing a Poisson quasi-maximum likelihood estimation:

$$E(D_{kt}(j'(k) = j')) = \exp\left(\beta^i_{hj'} + f(w_{kt}) + \gamma_t + \beta X_k\right).$$

In principle, $f(w_{kt})$ could be any function of wage. In practice, I approximate $f(w_{kt})$ using a third degree polynomial of log-wage of individual k. γ_t is a year fixed effect and X_k is a set of individual controls.

Having obtained the empirical estimates of $\beta^i_{hj'}$, $\hat{\beta}^i_{hj'} = \beta^i_{hj'} + \varepsilon^i_{hj'}$, from each regression on (i,h,j') subsets of workers, notice that

$$\beta^i_{hj'} - \beta^i_{j'j'} = -\frac{1}{\varepsilon_0} c^i_d d_{hj'}.$$

Taking into account that we do not observe $\beta^i_{j'j'}$ and $\beta^i_{hj'}$, but their estimated counterparts $\hat{\beta}^i_{j'j'}$ and $\hat{\beta}^i_{hj'}$, we can recover c^i_d as

$$c^i_{d,hj'} = -\varepsilon_0 \frac{\beta^i_{hj'} - \beta^i_{j'j'}}{d_{hj'}} + \varepsilon^i_{c,hj'}, \tag{26}$$

where $\varepsilon^i_{c,hj'} = \varepsilon_0 \frac{\varepsilon^i_{hj'} - \varepsilon^i_{j'j'}}{d_{hj'}}$. Notice that while the model allows for only I coefficients c^i_d, Equation 26 provides $I \times Z \times Z$ estimates of $c^i_{d,hj'}$. I choose the estimator for c^i_d to be the weighted average of all $c^i_{d,hj'}$, with the weights given by the inverse of the relative precision of the estimates of $\beta^i_{j'j'} + \beta^i_{hj'}$:[48]

$$\hat{c}^i_d = \frac{1}{\sum_h \sum'_j w_{hj'}} \sum_h \sum_{j'} w_{hj'} c^i_{d,hj'}.$$

The estimates and their standard errors are provided in Table D.8.

Table D.8: Estimates of c^i_d

Skill	c^i_d
High	-1.47
Low	-2.03

[48]Consistency is easy to prove. $c^i_d = c^i_d + \frac{1}{\sum_h \sum'_j w_{hj'}} \sum_h \sum_{j'} w_{hj'} \varepsilon^i_{c,hj'}$. For $\varepsilon^i_{c,hj'}$ being i.i.d. across (h,j') pairs, for $Z \to \infty$, $\frac{1}{\sum_h \sum'_j w_{hj'}} \sum_h \sum_{j'} w_{hj'} \varepsilon^i_{c,hj'} \to E(\varepsilon^i_c) = 0$. Hence, $\hat{c}^i_d = c^i_d$.

D.9 Distribution of $\tilde{p}$

To identify the local distribution of value added, by skill, $F_{ij}(\tilde{p})$, I proceed as follows. I assume $F_{ij}(\tilde{p})$ follows a Beta distribution with parameters $g_{1,ij}$ and $g_{2,ij}$, shifted and stretched so that:

- The minimum productivity is $\underline{p}$, as calibrated in Appendix D.5.
- The maximum productivity is $\overline{p} = \underline{p}$ (equivalent to £120,000 in 2015)

This means that, in each location j and skill i, $\frac{\tilde{p}-\underline{p}}{\overline{p}-\underline{p}} \sim \text{Beta}(g_{1,ij}, g_{2,ij})$. I assume $g_{1,ij} = 1.3$ following Bilal (2023). I calibrate $g_{2,ij}$ so to match the relative mean value added per worker of firms across locations and skills, minus the worker component α_i:

$$g_{2,ij} = g_{1,ij} \frac{\left(1 - \frac{E_{ij}(\tilde{p})-\underline{p}}{\overline{p}-\underline{p}}\right)}{\frac{E_{ij}(\tilde{p})-\underline{p}}{\overline{p}-\underline{p}}} = g_{1,ij} \frac{\overline{p} - E_{ij}(\tilde{p})}{E_{ij}(\tilde{p}) - \underline{p}},$$

$$E_{ij}(\tilde{p}) = \frac{1}{\alpha^i} \sum_{n \in N_j} \omega^i_{nj} (\text{GVA per i worker})_{nj}.$$

Where N_j is the number of firms in j used in the calculation of the average value added, ω^i_{nj} the weight of firm n in the calculation of the GVA per worker in skill i, and $(\text{GVA per i worker})_{nj}$ is the GVA per worker of firm i. I estimate $E_{ij}(\tilde{p})$ using ABS data.

D.10 Unemployment search efficiency

In order to estimate the relative efficiency at searching of the unemployed v, consider how unemployed workers of type i residing in h obtain a job in j' at rate $a^i_{hj'} \frac{s_{hj'}^{1+\varepsilon_0}}{1+\varepsilon_0} \theta_{ij'}^{1-\chi} \int_{w'} \mu^{i,U}_{hj'}(x) dF^i_{j'}(x)$.

Notice that minimum-wage workers residing and working in h have, by definition, the same value of the unemployed. Hence, $\mu^{i,U}_{hj'}(x) = \mu^{i,E}_{hj'}(\underline{w}_h, x)$. Moreover, they face the same market tightness $\theta_{ij'}^{1-\chi}$ and have the same application intensity per unit of effort $a^i_{hj'}$. Hence, the ratio of the rates of UE rates of the unemployed and EE rates of minimum-wage workers is

$$\frac{UE^i_{h,j'}}{EE^i_{\underline{w},hj'}} = \frac{s^{1+\varepsilon_0}_{U,hj'}}{s^{1+\varepsilon_0}_{\underline{w},hj'}} = \frac{v\left(\frac{\left[\sum_{x\in\mathscr{L}}\left(a^i_{hjx}\theta^{1-\chi}_{ix}\overline{V}^{U,i}_{hj'}(w)\right)^{-\frac{1}{\varepsilon_0}}\right]^{\frac{\varepsilon_0}{\varepsilon_1-\varepsilon_0}}}{\left(a^i_{hj'}\theta^{1-\chi}_{ij'}\overline{V}^{U,i}_{hj'}\right)^{\frac{1}{\varepsilon_0}}}\right)^{1+\varepsilon_0}}{\left(\frac{\left[\sum_{x\in\mathscr{L}}\left(a^i_{hjx}\theta^{1-\chi}_{ix}\overline{V}^{E,i}_{hj'}(\underline{w},w)\right)^{-\frac{1}{\varepsilon_0}}\right]^{\frac{\varepsilon_0}{\varepsilon_1-\varepsilon_0}}}{\left(a^i_{hj'}\theta^{1-\chi}_{ij'}\overline{V}^{E,i}_{hj'}(\underline{w},w)\right)^{\frac{1}{\varepsilon_0}}}\right)^{1+\varepsilon_0}} = v,$$

where the last equality follows from the fact that, as discussed above, $\overline{V}^{E,i}_{hj'}(\underline{w},w) = \overline{V}^{U,i}_{hj'}(w)$.

This relationship can be directly estimated in the data, with the assumption that minimum wage workers have similar unobservable characteristics of unemployed workers, conditional on their type i. I estimate this relationship using Labour Force Survey data. I find an estimate of $v = 4.09$ for the low-skilled and $v = 4.17$ for the high-skilled. Taking an average of the two, I calibrate my model to $v = 4.13$. This is similar to both the estimate of my model when I allow v to be endogenously estimated by targeting the unemployment rate (3.97) and the estimates of Heise and Porzio (2023) for Germany (5.9).

E Parameter Estimation

In this section, I provide details on how the parameters are estimated. Section E.1 describes the mapping between the first set of parameters (search and spatial frictions, firm characteristics other than productivity) and different moments. As these parameters all depend on general equilibrium effects and no clear likelihood function is available, I employ a MCMC algorithm with a flat prior over a large domain to find the set of parameters that minimises a loss function based on a large set of empirical moments, as described in Appendix G.[49] Section E.2 described how to identify the second set of parameters related to firms' productivity and localness, and the housing sector productivity. Unlike the first set, these will be exactly identified.

E.1 First Block Parameters

I report a summary of the parameters estimated in the first block in Table E.1. For each of these, I also specify - if different from the parameter itself - whether in the actual estimation algorithm I employ the parameter itself, or a function of it and other parameters or objects.

Table E.1: First block parameters

Parameter group	Description	Number	MCMC target	Reference
- Utility				
τ_h	Amenities relative to London	4		E.1.6
σ_i	Offer taste shock, by i	2	$\sigma_i(\alpha_i \underline{w}_i)^{-1}$	E.1.4
σ_i^m	Residence taste shock, by i	2	$\sigma_i^m(W^i_{\text{Hull}}(\underline{w}))^{-1}$	E.1.5
- Search Frictions				
ε_1	Effort cost curvature	1		E.1.3
a^i	Efficiency of search, by i	2	$(a^i)^{\frac{1+\varepsilon_1}{\varepsilon_1-\varepsilon_0}}$	E.1.3
- Spatial frictions				
κ^i	Moving cost (fixed), by i	2	$\frac{\kappa}{\sigma^m}$	E.1.1
κ_d	Moving cost (distance)	2	$\frac{\kappa}{\sigma^m}$	E.1.1
d_1^i	Commuting cost, by i	2		E.1.6
- Firms				
ζ_1^i	Vacancy cost curvature, by i	2		E.1.2

[49] Notice that one of the spatial frictions, the search efficiency decay across space, can be directly estimated from the data, as described in Appendix D.8.

E.1.1 Location moving costs.

Consider two workers. Both earn wage w in j, but one resides in h and another in $h' \neq h$. Suppose you observe both of them accepting an offer at wage w' in location $j' = h$. Then, consider the probabilities of the first worker not moving his residence, $P^i_{jhhh}(w,w')$, and of the second worker moving to h, $P^i_{jh'hh}(w,w')$. Their likelihood ratio is then

$$\ln\left(\frac{P^i_{jhhh}(w,w')}{P^i_{jh'hh}(w,w')}\right) = \frac{\kappa^i_{h'h}}{\sigma^i_m}\frac{1}{\rho+\delta^i} + \ln\left(\frac{\sum_s \exp(W^i_{sh}(w') - \kappa^i_{h's})^{\frac{1}{\sigma^i_m}}}{\sum_x \exp(W^i_{xh}(w') - \kappa^i_{hx})^{\frac{1}{\sigma^i_m}}}\right). \tag{27}$$

The first term represent the present discounted value of paying the moving cost κ today to enjoy location h, while the second term represents the fact that the probability of moving to a location depends on the outside options, whose value may differ for workers residing in different locations before the move.

Intuitively, the likelihood ratio of performing the same "move" (from w to w', given a worker currently working in j) to (hh) depends on the relative moving costs and outside options. An high moving cost would make a $(jh'hh)$ move less likely than a $(jhhh)$ move by reducing the incentives of moving from h' to h (reducing $P^i_{jh'hh}(w,w')$), and from h to h' (increasing $P^i_{jhhh}(w,w')$). Hence, higher likelihood ratios across all $(j,h,h') \in \mathscr{L} \times \mathscr{L} \times (\mathscr{L} \setminus \{h\})$ would imply higher moving costs relative to the taste shock parameter σ^i_m, given the discount rate.

Since the expression is monotonically increasing in κ,[50] larger deviations from 0 of the LHS pin down a unique value for $\frac{\kappa^i_{h'h}}{\sigma^i_m}$, given the value of employment. With more than 2 locations, it is then possible to disentangle between the fixed-cost $\frac{\tilde{\kappa}^i}{\sigma^i_m}$ and the distance-cost component of moving $\frac{\tilde{\kappa}^d}{\sigma^i_m}$ by fitting all residence move rates across h,h'.

---

[50]Notice that the derivative of the second term is 0. In fact, $\frac{\partial \ln\left(\sum_s \exp(W^i_{sh}(w') - \kappa^i_{h's})^{\frac{1}{\sigma^i_m}}\right)}{\partial \kappa} = -\frac{1}{\sigma^i_m}$. The same holds for the denominator within the log. Hence, the derivative of the likelihood ratio is simply $\frac{1}{\sigma^i_m(\rho+\delta^i)}$, which is always strictly positive for $\sigma^i_m < \infty$.

E.1.2 Vacancy costs

The mass of vacancies in occupation i and location j is

$$v_j^i = \zeta_{0,j}^i M_j^i \int_{\underline{p}}^{\overline{p}} \left(\frac{\pi_j^i(p)}{\overline{\pi}_j^i} \right)^{\frac{1}{\zeta_1^i}} dF_j^i(p). \tag{28}$$

I identify ζ_1^i by targeting the distribution of employment along the productivity distribution. Intuitively, the CDF of employment depends on firm's size. Which, in turn, depends on the number of vacancies posted: call $q_{j,x}^i$ the value of productivity p at the x-th percentile of the productivity distribution. Then, the employment in all firms with up to productivity $q_{j,x}^i$ is

$$E_j^i(q_{j,x}^i) = \frac{\int_{\underline{p}}^{q_{j,x}^i} n_j^i(p) v_j^i(p) dF_j^i(p)}{\int_{\underline{p}}^{\overline{p}_j^i} n_j^i(p) v_j^i(p) dF_j^i(p)} = \frac{\int_{\underline{p}}^{q_{j,x}^i} n_j^i(p) (\pi_j^i(p))^{\frac{1}{\zeta_1^i}} dF_j^i(p)}{\int_{\underline{p}}^{\overline{p}_j^i} n_j^i(p) (\pi_j^i(p))^{\frac{1}{\zeta_1^i}} dF_j^i(p)},$$

where the second equality follows from substituting the optimal vacancy posting behaviour of p firms. This expression implies that the CDF of employment at quantile x depends on ζ_1^i, given $n(p)$, the per-vacancy profits $\pi_j^i(p)$ and the CDF of firms' value added.

When profits per vacancy increase more than proportionally with p, as in this class of model with monopsony power from search frictions, a higher ζ_1^i implies that the relative number of vacancies posted by productive and unproductive firms shrinks, and larger share of workers is employed at low-productivity firms. Lower ζ_1^i imply that a larger share of workers is employed at high-productivity firms.

To identify $\zeta_{1,j}^i$, I target six different quantiles of the distribution of employment: the 5th decile, and the difference between this and the successive 6th, 7th, 8th, 9th, and 10th deciles.[51]

[51] This is equivalent to targeting the shares of employment of the bottom 50% of firms, and the one of each additional 10% in the size distribution. I do not target the whole distribution in order to avoid putting too much weight on small employment shares at the bottom quantiles.

E.1.3 Search frictions.

Identifying the search costs requires pinning down ε_1, a^i and v. I identify the remaining parameters as follows.

Taking logs of Equation 18 and taking the derivative relatively to individual wage, I obtain

$$\frac{\partial \ln(S^i_{hj}(w))}{\partial \ln(w)} = \frac{\varepsilon_0}{\varepsilon_1 - \varepsilon_0} \frac{\partial \ln \left[\sum_{j'} \left(a^i_{hjj'} \theta^{1-\chi}_{ij'} \overline{V}^{E,i}_{hjj'}(w) \right)^{-\frac{1}{\varepsilon_0}} \right]}{\partial \ln(w)}.$$

Given a^i, market tightness and the value function, matching this derivative allows to pin down ε_1. While in the model all workers search on-the-job, in practice only a few workers search very intensively but sporadically. Thus, using the Labour Force Survey, I approximate $\frac{\partial \ln(S^i_{hj}(w))}{\partial \ln(w)}$ by regressing a dummy for searching on-the-job on log-wages and individual characteristics. Then, the coefficient on log-wages represents an estimate for the elasticity of effort to wages. The identifying assumption is that the derivative of the share of individuals searching on-the-job to income has the same slope as the derivative of effort $S_{hj}(w)$ to income. That is, the effort put into searching, conditional on being searching, is the same across the income distribution. Using the UK Time Use Survey 2011, I confirm that the amount of time spent searching by employed individual is large (55% of the unemployed's effort) and, conditioning on spending at least some search effort, approximately constant across the income distribution. From this regression, I recover $\hat{\beta}_w$, the coefficient on log-wage.

Given ε_1 and θ, it is then possible to pin down a^i. Consider the rate of moving from job to job within the same location j where the previous job was held,[52]. This is given by:

$$(\text{Move rate})^i_{hj} = a^i \theta^{1-\chi}_{ij} \int \frac{\left(s^i_{hj}(w)\right)^{1+\varepsilon_0}}{1+\varepsilon_0} \left[\int P^{i,E}_{hjj}(w,w') dF^i_j(w') \right] dE^i_{hj}(w).$$

Plugging in the expression for S^i and taking logs, we obtain:

$$\ln(\text{Move rate})^i_{hj} = \frac{1+\varepsilon_1}{\varepsilon_1 - \varepsilon_0} \ln(a^i) - \ln(1+\varepsilon_0) - \frac{1}{\varepsilon_0} \ln(\theta^{1-\chi}_{ij}) + \int \Gamma^i_{zj}(w) \int P^{i,E}_{hjj}(w,w') dF_j(w') dE_{hj}(w),$$

where $\Gamma^i_{zj}(w)$ is a term depending only on c^d, market tightness and value functions across

[52] This is guaranteed to be > 0 due to the individual preference shocks and the curvature $1+\varepsilon_0$ of the within-location effort efficiency.

locations. Thus, each skill-specific a^i can be written as

$$a^i = \exp\left[\frac{\varepsilon_1-\varepsilon_0}{1+\varepsilon_1}\left(\ln(\text{Move rate})^i_{hj} + \ln(1+\varepsilon_0) + \frac{\ln(\theta_{ij}^{1-\chi})}{\varepsilon_0} - \int \Gamma^i_{hj}(w) \int P^{i,E}_{hjj}(w,w')dF_j(w')dE_{hj}(w)\right)\right].$$

Since a^i depends on ε_1, in the MCMC estimation algorithm I draw $(a^i)^{\frac{1+\varepsilon_1}{\varepsilon_1-\varepsilon_0}}$ so to improve the efficiency of the algorithm.

E.1.4 Offer taste shock.

The main role of σ^i is to model how workers accept jobs with worse wages than their current job due to idiosyncratic preferences to how they like a location or an employer. Thus, I identify σ by matching the the share of workers with a negative wage gains after a job move.

$$P^i_j(\Delta w < 0) = \frac{\int_{\underline{w}_j}^{\overline{w}_j}\int_{\underline{w}_j}^{w} \mu^{E,i}_j(w,w')dF^i_j(w')dE^i_j(w)}{\int_{\underline{w}_j}^{\overline{w}_j}\int_{\underline{w}_j}^{\overline{w}_j} \mu^{E,i}_j(w,w')dF^i_j(w')dE^i_j(w)}.$$

This expression depends on the relative acceptance rates $\mu^{E,i}_j(w,w')$, the job distribution $E^i_j(w)$, and the job offer distribution $dF^i_j(w')$. Higher σ^i increase the acceptance rate - everything else equal - for offers with $w' < w$, and lowers it for $w' > w$. Hence, the expression is monotonically increasing in σ^i, given the distribution of offers, employment, and the value function.[53]

In the estimation algorithm, I do not estimate σ_i itself, but $\sigma_i(\alpha_i\underline{w}_i)^{-1}$. That is, the parameter scaled by the minimum wage received by a worker of skill i in the economy $\alpha_i\underline{w}_i = \alpha_i\iota\underline{p}_i \approx 0.565$.

E.1.5 Residence taste shock.

To identify the residence taste shock, I employ a similar strategy to σ^i, adapted to account for how we do not know - without knowing the value function - what would be the "preferred" moves.

[53]Since high σ^i make it more likely to accept $(w' < w)$ offers, and less likely to accept $(w' > w)$ offers, it must necessarily increase the proportion of $(w' < w)$ offers over the total, even if it affects the denominator in unclear ways, given the distribution of employment and offers.

If $\sigma_m^i = 0$, we should observe identical workers performing identical residence moves, everything else equal. For example, consider a worker residing in h, working in j at wage w and being observed accepting a job in j' at wage w'. Then, their residence move choice should be identical, as only one location h' (potentially identical to z) can maximise the value of accepting the offer, given that the employer-specific idiosyncratic shock is perfectly correlated within the residence cluster of the logit model. However, for $\sigma_m^i > 0$ this cannot be the case, and there would be increasing dispersion (up to the probability moving to each location being uniform, for $\sigma_m^i \to \infty$). Thus, one can pin down σ_m^i by matching the degree of dispersion of location residence moves. More straightforwardly, I directly match the whole set of residence moves probabilities .[54]

Since the variance of the preference shock necessary to generate a given degree of dispersion in workers' choices depends on its relative value to the value of employment, I first estimate a normalised parameter $\tilde{\sigma}_m^i = \frac{\sigma_m^i}{W_4^E(\underline{w})}$. That is: σ_m^i scaled by the value function at the minimum wage in the poorest location. Then, I recover σ_m^i by multiplying $\tilde{\sigma}_m^i$ for the equilibrium $W_{\text{Hull}}^E(\underline{w})$.

E.1.6 Commuting costs and amenities.

Finally, commuting costs and amenities are identified by matching the spatial distribution of residence, work, commutes, and wage gains of new commuters (job movers who were not commuting before the move).

E.2 Second block Parameters

Having identified the parameters of the first block, I can proceed with the identification of the parameters of the second block. These involve: i) the productivity of the local housing sector, $\tilde{A}_{H,j}$; ii) the productivity of the local non-tradeable sector, A_j^L; iii) vacancy costs by skill and location $\zeta_{0,j}^i$; and iv) the distribution of floorspace use $\phi_j^i(p)$ and tradeable efficiency $\beta_j^i(p)$, by

[54]Notice that a non-independent set of moments was already used to pin down i κ^i and one κ_d. Using 5 locations ensures that there is a rich enough set of hh' moves to pin down the 5 parameters. Increasing the number of locations can help further.

i. In total, this involves identifying 20 parameters and 2 distributions.[55]

Vacancy posting costs. I identify $\zeta_{0,j}^i$ in order to perfectly fit the aggregate mass of vacancies v_j^i within each location and skill by inverting Equation 28:

$$\zeta_{0,j}^i = \frac{M_j^i}{v_j^i} \int \left(\frac{\pi_j^i(p)}{\overline{\pi}_j^i} \right)^{\frac{1}{\zeta_1^i}} dF_j^{i,f}(p).$$

Tradeables. I fit non-parametrically the distribution of employment across firms, by firm size. Using the NESS survey, I have data on what firms declare to mostly sell non-tradeables, and the share of employees being high or low-skilled. Merging it with the ABS census data on all firms and plants, I obtain a CDF of employment across firms, by location and skill. Then, for each point on my productivity grid, I fit the corresponding share.

Housing sector productivity. $\tilde{A}_{H,j}$ can be identified through Equation 13. Given the demand of floorspace by workers and firms, whose identification is described above, we can express $A_{H,j}$ as a function of known parameters and equilibrium objects, yielding a closed-form solution:

$$\tilde{A}_{H,j} = \left[\frac{r}{(1-b_j)P_{H,j}} \left[H_{w,j}^D + H_{f,j}^D \right]^{b_j} \right]^{\frac{1}{1-b_j}}.$$

Floorspace per employee. I obtain the expenditure on rents as a share of wage bills for each decile of firms' employment. So that, for $q = \{0, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$,[56] I have, for each location j and skill i

$$\tilde{\phi}_{j,q}^i = \frac{P_{H,j}\phi_j^i(p_{j,q}^i)p_{j,q}^i}{\alpha^i w_j^i(p_{j,q}^i)}.$$

Then, I fit back this distribution to my firm-location-skill grid using the model's employment equilibrium CDF within each location. I rescale the overall expenditure so that the mean cost-to-wages ratio fits the national average of 9.4%. Finally, I recover $\hat{\phi}_j^i(\tilde{p}) = \phi_j^i(p_j^i)p_j^i(\tilde{p})$ by dividing by the local price index and multipliying by local wages at each firm on the grid.

[55] Technically, there are 20 distributions to be estimated: the distribution of β and ϕ in each location and skill, along the produtivity margin.

[56] I start skip the first and second decile

Non-tradeables productivity and p. In Equilibrium, the non-tradeable market must clear in each location. That is, A_j^N must satisfy

$$A_j^N \sum_{i \in I} M_j^i \int_p s(1-\beta_j^i(s)) n_{i,j}^*(s) v_{i,j}^*(s) dF_j^{i,f}(s) = \eta_1 \frac{Y_j^r}{P_{l,j}}. \tag{29}$$

All terms on the right-hand side of the equation are known. However, on the left-hand side all the terms but M_j^i are unknown, as they are either A_j^N, p (indexed as s within the integral), and functions that depend on p. Assume $\tilde{p}_j^i(p)$ is a one-to-one function of p at current prices $P_{l,j}$ and $P_{H,j}$. Call $\hat{\beta}_j^i(\tilde{p}) = \beta_j^i(p(\tilde{p})$, and similarly so $\hat{n}, \hat{v}, \hat{\gamma}$. Then,

$$p_j^i(\tilde{p}) = \frac{\left(\tilde{p} + \hat{\phi}_j^i(\tilde{p})\right)}{\hat{\beta}_j^i(\tilde{p}) + P_{l,j} A_j^N (1-\hat{\beta}_j^i(\tilde{p}))}. \tag{30}$$

In this equation, we know all terms but A_j^N. Performing a change of variable from p to $\tilde{p}$ in Equation 29 yields

$$A_j^N \sum_{i \in I} M_j^i \int_{\tilde{p}} p_j^i(s)(1-\hat{\beta}_j^i(s)) \hat{n}_{i,j}^*(s) \hat{v}_{i,j}^*(s) dF_j^{i,f}(s) = \eta_1 \frac{Y_j^r}{P_{l,j}}. \tag{31}$$

If $\sum_{i \in I} M_j^i > 0, \forall j \in \mathscr{L}$ (there is a positive mass of firms in each location) and $\forall j \in \mathscr{L}, \int \beta_j^i(p) dF_j^i(p)) > 0$ (the production of tradeables is non-zero in each location), then the left-hand side of Equation 31 is increasing in $A_{l,j}$.[57] Hence, for $\eta_1 \frac{Y_j^r}{P_{l,j}} \neq 0$,[58] there must exist one and only one $A_j^N > 0$ that satisfies this equation. This pins down the parameter A_j^N, as well as the functional form of $p_j^i(\tilde{p})$ from Equations 30 and. Finally, this also allows to recover $\phi_j^i(p) = \frac{\hat{\phi}_j^i(\tilde{p}(p))}{p}$.

I now describe how to recover p, A^N and the distribution of γ. Recall that A_j^N was identified by iterating over different guesses of A_j^N until convergence to a value that guaranteed market clearing in the local non-tradeable good market, where supply depends on A_j^N both directly and through the distribution of productivities p implied by $\beta(p)$, $\phi(p)$, and A_j^N.

Identifying A_j^N, p, and γ follows a similar procedure. Call $\hat{\beta}_j^i(\tilde{p})$ the value of β^* for firms of type i, in location j, with value added $(\tilde{p})$, as fitted to the CDF of employment using NESS

[57] Simply notice how, expanding $p(s)$ and assuming $A_j^N \neq 0$, Equation 31 can be written as $\sum_{i \in I} M_j^i \int_{\tilde{p}} \frac{\left(\tilde{p} + P_{H,j}\hat{\phi}_j^i(\tilde{p})\right)}{\frac{\hat{\beta}_j^i(\tilde{p})}{A_{l,j}} + P_{l,j}(1-\hat{\beta}_j^i(\tilde{p}))} (1-\hat{\beta}_j^i(s)) \hat{n}_{i,j}^*(s) \hat{v}_{i,j}^*(s) \hat{\gamma}_j^i(s) ds = \eta_1 \frac{Y_j^r}{P_{l,j}}$. As long as $P_{H,j}, P_{l,j} \neq \infty$, the term within the integral is increasing in A_j^N.

[58] Notice that this must the case as long as $\exists i \in I : \sigma_m^i > 0$, and the price is not infinite.

data.

Given these results, I show how to recover A^N_j, p and ϕ. I apply the following algorithm:

1. Start from a guess of A^N_j.

2. Calculate $p^i_j(\tilde{p})$ by computing the following expression:
$$p^i_j(\tilde{p}) = \frac{\left(\tilde{p} + \hat{\phi}^i_j(\tilde{p})\right)}{\hat{\beta}^i_j(\tilde{p}) + P_{l,j}A^N_j(1-\hat{\beta}^i_j(\tilde{p}))}.$$

3. Check whether the local non-tradeable markets clear at the given A^N_j by verifying if
$$A_{l,j}\sum_{i\in I}M^i_j\int_{\tilde{p}} p^i_j(s)(1-\hat{\beta}^i_j(s))^{1-\tilde{\gamma}^i_j(\tilde{p})}\hat{n}^*_{i,j}(s)\hat{v}^*_{i,j}(s)dF^{i,f}_j(s) = \eta_1\frac{Y^r_j}{P_{l,j}}.$$

4. If markets do not clear, take a new guess of A^N_j (larger if excess demand is positive, smaller if negative) and repeat points 2-4 until convergence.

5. Once A^N_j and $p^i_j(\tilde{p})$ are such that the market clearing condition is satisfied, obtain $\phi^i_j(p) = \frac{\hat{\phi}^i_j(\tilde{p}(p))}{p}$.

F Details on targeted moments

F.1 Workers by location

I estimate the relative share of full-time employees, by occupation $\times$ residence location, from the cleaned ASHE cross-sectional dataset.

Table F.1: Share of residents (workers only), by skill

	Location			
Skill	London	Productive	Fringe	RoC
High	0.167	0.244	0.202	0.388
Low	0.121	0.239	0.200	0.439

F.2 Unemployment

Since all the flows of workers across firms are estimated using full-time workers only, I consider "unemployed" all individuals that are: i) unemployed, according to the ILO definition; 2) in involuntary part-time employment. Adding workers in unvoluntary part-time employent is important in order to correctly capture the labour supply in the economy relative to the current amount of full-time employed.

Using the Labour Force Survey, I calculate unemployment as the share of individuals between 25 and 65 years old such that: i) are classified as unemployed; or work part-time but i) declare they would accept more hours at the same pay rate, or ii) declare they are actively searching for a job with more hours. This gives an average unemployment rate of 10.6%.

Since LFS are geocoded at the regional level, I can distinguish only between the location I classify as "London" and the rest of the country. I impute the share of unemployed in each model location as the weighted average of the regional unemployment levels, where the weights are given by the sum of the population weights of all Local Authority Units that are part of the (model) location j and of each administrative region. The results are presented in Table F.2

Table F.2: Unemployment rate

	Location				
Estimation	London	Bristol	Liverpool	Hull	South-East
LFS	11.6%		10.31%		
Imputation	11.6%	10.7%	10.7%	11.5%	9.5%

F.3 Commuting

I estimate the relative share of full-time employees, by occupation $\times$ residence location $\times$ workplace, from ASHE.

F.4 Wage gains

To estimate the wage gains, I compute $\ln(y_{t+1}) - \ln(y_t)$ across all workers. I subset the observations in order to look only at job movers who: i) were not commuting at time t, ii) do not change residence location between t and $t+1$, and iii) are not commuting in $t+1$. Then, for each group of initial occupations at time t, I regress the gains over a sex dummy and a set of age dummies. The excluded category are male individuals aged 30-34 with high skill. The parameters of interest are the location-specific constant, reported with their standard errors in Table F.3.

Table F.3: Wage gains of within-location movers

	Location			
Skill	London	Productive	Fringe	RoC
High	7.64%	6.67%	6.48%	6.18%
	(0.5)	(0.6)	(0.7)	(0.5)
Low	6.88%	5.57%	5.54%	4.99%
	(1.5)	(1.5)	(1.6)	(1.4)

I compute the wage gains of commuters by subsetting the same cleaned wages by all workers who: i) were not commuting before a move, ii) do so after the move, and iii) do not move residence. I regress the wage gains of this subset of movers over a sex dummy and a set of age dummies. The estimates with their standard errors are reported in Table F.3. In the estimation of the model, I match the moments provided in Table F.3, and the differences between these and those reported Table F.4, so to give more weight to the difference between commuting and local gains.

Table F.4: Wage gains of new commuters

	Location			
Skill	London	Productive	Fringe	RoC
High	4.95%	8.22%	8.88%	8.75%
	(0.9)	(0.7)	(0.7)	(0.6)
Low	7.12%	9.05%	6.51%	8.75%
	(2.1)	(1.7)	(1.7)	(0.6)

F.5 Job movers

I obtain job movers from ASHE. I consider a "job mover" any individual who, between year t and $t+1$, has moved to either: i) a different firm , or ii) the same firm, but in a different Local Authority Unit.

F.6 Residence movers

I obtain residence movers from ASHE. While I observe every residence move, as long as it happens between different LSOAs (similar to Census tracts) and worker do not drop out of the dataset, my model has two limitations. First, workers do not move unless they change job. Hence, I focus only on residence moves of job movers. Second, my model doesn't make any distinction between residence movers and stayers within the same model location. Hence, I do not discriminate between workers who did not move at all, and those who moved to a different house within the same model location.

F.7 Firm size distribution

From the cleaned ABS data, I obtain the employment distribution across firms. Since I do not model "sectors", I obtain the log-size deviation of each firm from its 3-digits sector mean. Then, I obtain a distribution of firm size by summing this deviation to the log-mean of the data, and taking the exponential. Then, I compute the CDF of total employment over the distribution of firms, ranked by size, for each location. Since the bottom quantile have very small shares of firms, thus making the fit sensitive to small deviations from these moments, I target 6 part of the distribution: the share of employment at firms up to the 50th decile of size, and the share between each additional decile.

F.8 Relative GVA

Using ONS's regional GVA data, I compute the mean GVA per worker of each location. Then, I rescale each absolute quantity into relative terms to London, which is normalised to 1.

Table F.5: Relative GVA

London	Productive	Fringe	RoC	South-East
1.000	0.761	0.635	0.611	0.850
(-)	(0.011)	(0.012)	(0.009)	(0.011)

F.9 Vacancies

I estimate the mass of skill-location mass of vacancies as follows. Using the NESS dataset, I obtain vacancies, by 1-digit occupations. Using the definition of high- and low-skilled I adopt throughout the paper, I obtain vacancies by skill $i \in I$ for each firm n in j (v^i_{jn}). Then, I calculate the ratio of vacancies to employment in each location by dividing the weighted sum of vacancies by the weighted sum of employment:

$$\tilde{v}^i_j = \frac{\sum\limits_{n \in N_j} w_j v^i_{jn}}{\sum\limits_{n \in N_j} w_j e^i_{jn}}.$$

Since v^i_j represents the mass of vacancies, where the mass of the *population* is normalised to one, I recover v^i_j from $\tilde{v}^i_j$ by dividing by the share of employees over the total of the labour force, given by 1 minus the unemployment rate u:

$$v^i_j = \frac{\tilde{v}^i_j}{1-u}.$$

G Estimation Algorithm

The objective is to find a set of parameter $\phi^* \in \mathbb{F}$, for $\mathbb{F}$ being the hyperset of admissible parameters, which minimises a loss function $\mathscr{L}(\phi)$:

$$\phi^* = \arg\min_{\phi \in \mathbb{F}} \mathscr{L}(\phi) = \arg\min_{\phi \in \mathbb{F}} \sum_{n \in N} \omega_n \left(\frac{\hat{X}_n(\phi) - X_n}{SE_{X_n}} \right)^2,$$

where $\hat{X}_n(\phi)$ represent the simulated moment n from solving the parameter using ϕ, X_n is the data moments, and SE_{X_n} is the standard error of the empirical moment. ω_n represents weights applied to moment n, chosen so that: i) the sum of weights of moments within one of the families of moments described in Table 4, ii) a few of these families are given half-weights, to take into account how their very tight standard errors would force the model to overfit a few moments.

I choose as estimation algorithm an MCMC approach. First, I pick 100 strings of parameters, each 1000 elements long. I choose as initial points 100 random point from my parameter space. Then, I follow the following procedure for each of these strings. For each iteration, I pick a set of parameters ϕ_k using the Metropolis-Hasting sampling procedure, starting from the previously stored point. Then, I follow the following procedure: I solve the model using those parameters, and simulate a dataset of 1,500,000 observations to recover the simulated moments $\hat{X}(\phi_k) = \{\hat{X}_1(\phi_k), \ldots, \hat{X}_n(\phi_k)\}$. Then, I calculate the loss function of iteration k $\mathscr{L}(\phi_k)$.

Then, I keep the draw of parameters ϕ_k with probability

$$P(\text{keep } \phi_k | \phi_{k-1}) = \min\{1, T_k \exp\left[\ln(\mathscr{L}(\phi_{k-1})) - \ln(\mathscr{L}(\phi_k))\right]\},$$

where T_k is a temperature parameter used to target a rejection rate of 0.18. The temperature is update every 200 draws.

Then, I use the best estimates from each of these strings as starting points for 1000 further draws, each 1000 elements long. For each of these strings, I discard the first 800 draws, and keep the last 200. The final estimate is given by the set of parameters providing the best fit among the 1,000,000 parameter draws thus obtained. The 200 observations kept at the end of each string (for a total of 200,000) will be used to calculate the standard errors of the estimates.

The procedure is thus a mix of the approaches of Heise and Porzio (2023) and Jarosch (2023), which adjusts for the finite domain of the parameters. Like Heise and Porzio (2023) and Engbom and Moser (2022), I choose the draw that delivers the minimum $\mathscr{L}(\phi)$, rather than an average of all the draws.

G.1 Model Fit

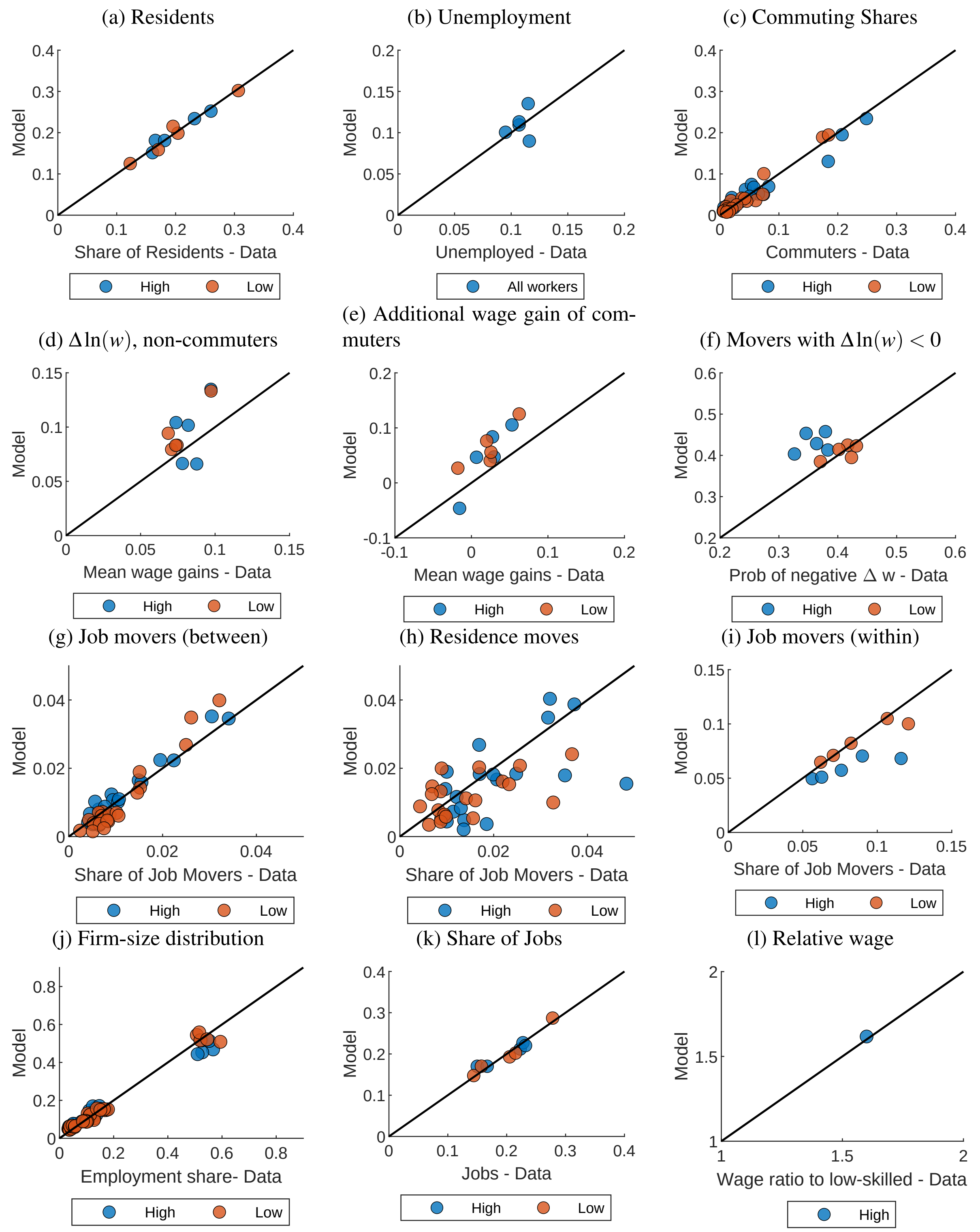
In this section, I present the results from the MCMC estimation. Table G.1 shows how the moments from the model (vertical axis) match with the moments from the data (horizontal axis). Figure G.1 plots the data moments against the model's simulated moments, showing how most of the moments have a very good fit, apart from a few outliers. The model is able to jointly capture the distribution of residents (both high and low-skilled), unemployment, the different commuting shares between locations, the number of job movers within each location, the share of job movers between locations, and the distribution of employment across firms. Moreover, it performs sufficiently well is all other dimensions. The model fits well most parameters. The best fits are on the low-to-high skilled wage ratio, the relative wages of locations, the number of residents, and the one of individuals who live in he same location where they work. The worst fit is obtained on the search effort derivative (32.1% deviation from the data), although the sign is correctly captured. The wage gains are well fitted, apart from the one of new commuters outside of London, estimated to be particularly negative.

Table G.1: Targeted Moments

	Mean error (pp.)
Resident share	4.1
Job share	5.9
Unemployment	10.8
Commuting shares (within)	2.1
Commuting shares (between)	21.3†
Search effort	32.1
Job moves (within)	14.9
Job moves (between)	19.9†
Residence moves	4.3†
Local wages	1.6
Wage gains (within)	1.4‡
Wage gains (new commuters)	4.1‡
Share negative wage gains	12.8
Employment CDF	12.0
GVA	5.0
High-to-low wage ratio	0.3

Notes: The table shows the mean percentage deviation (so that 1 = 1% deviation) between simulated and data moments. Exceptions are momenta that present very small moments, and hence whose percentage deviation is very sensitive. For these, I provide two remedies: i) † implies that this is the mean moment deviation, divided by the mean moment; ii) ‡ represents the deviation in pure percentage points, used for wages.

Figure G.1: Model Fit - Targeted Moments



Note: data's moments (x-axis) and model's moments (y-axis) for each of the targeted groups of moments, generated using the estimated parameters from the MCMC procedure.

G.1.1 Non-targeted moments

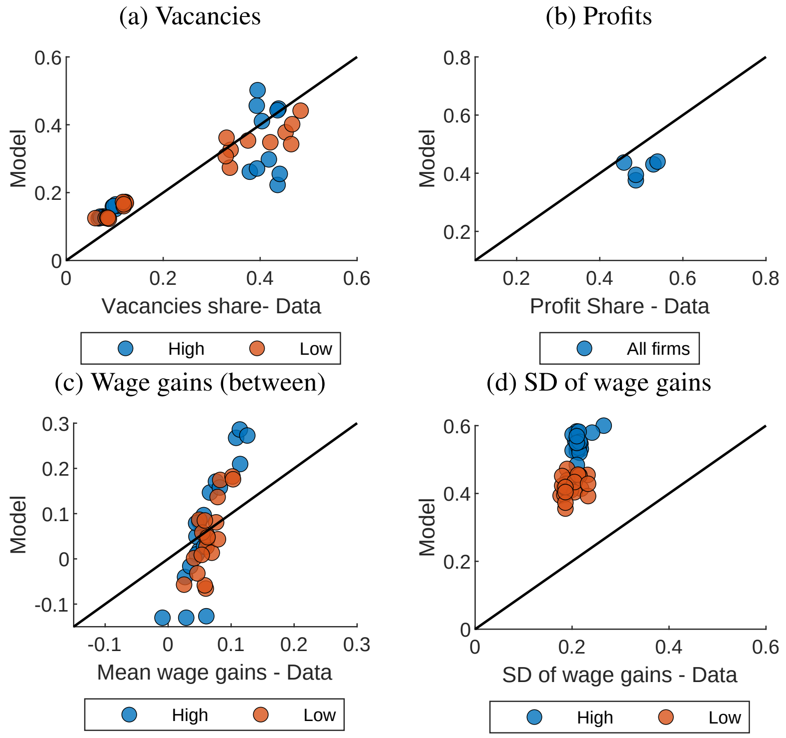
To validate the model, I show how it is able to capture features of the data not used for the calibration. With regards to the chosen functional form of firms' vacancy posting, the model fits well the distribution of vacancies along the employment distribution, as well as the profit share of total value added for four out of five locations. Turning to wage gains across space, the model is well centred around the mean values, although the slope of the curve appears steeper, overestimating the largest gains and losses, related to moving in/out of London. However, the model fails to fit the standard deviation of all wage gains, overestimating it across all locations and skills. Both these latter results may be due to the use of only two skills, which does not fully capture skill sorting and the multiple job ladders and the lack of a "home" location for workers (as in Heise and Porzio (2023)), implying larger compensating differentials for moving out of productive locations.

Table G.2: Non-targeted Moments

	Mean error (pp.)
Vacancies CDF	25.3†
Profits	13.8
Wage Gains, by (j,j)	7.2‡
SD of wage gains, by (j,j)	128.2

Notes: The table shows the mean percentage deviation (so that 1 = 1% deviation) between simulated and data moments. Exceptions are momenta that present very small moments, and hence whose percentage deviation is very sensitive. For these, I provide two remedies: i) † implies that this is the mean moment deviation, divided by the mean moment; ii) ‡ represents the deviation in pure percentage points, used for wages.

Figure G.2: Model Fit - Non-Targeted Moments



Note: data's moments (x-axis) and model's moments (y-axis) for each of the targeted groups of moments, generated using the estimated parameters from the MCMC procedure.

G.1.2 Jacobian Matrix

While the previous analysis points at a good fit of the model to the data, it is important to ask whether the parameters are indeed identified according to the theoretical identificaiton strategy, loading more on those moments that should be able to pin them down. To do so, I estimate an approximated Jacobian matrix of the fit of each moment, given a variation in the parameter value. From the final estimate, for each parameter θ I compute $\overline{\theta} = \theta \times 1.05$ and $\underline{\theta} = \theta \times 0.95$. Call θ^* all other parameters apart from θ itself. Then, I compute two new likelihood function of moment x, $L_x(\theta^*, \overline{\theta})$ and $L_x(\theta^*, \underline{\theta})$, once for each parameter, in order to obtain

$$\Delta_{x,\theta} = L_x(\theta^*, \overline{\theta}) - L_x(\theta^*, \underline{\theta}).$$

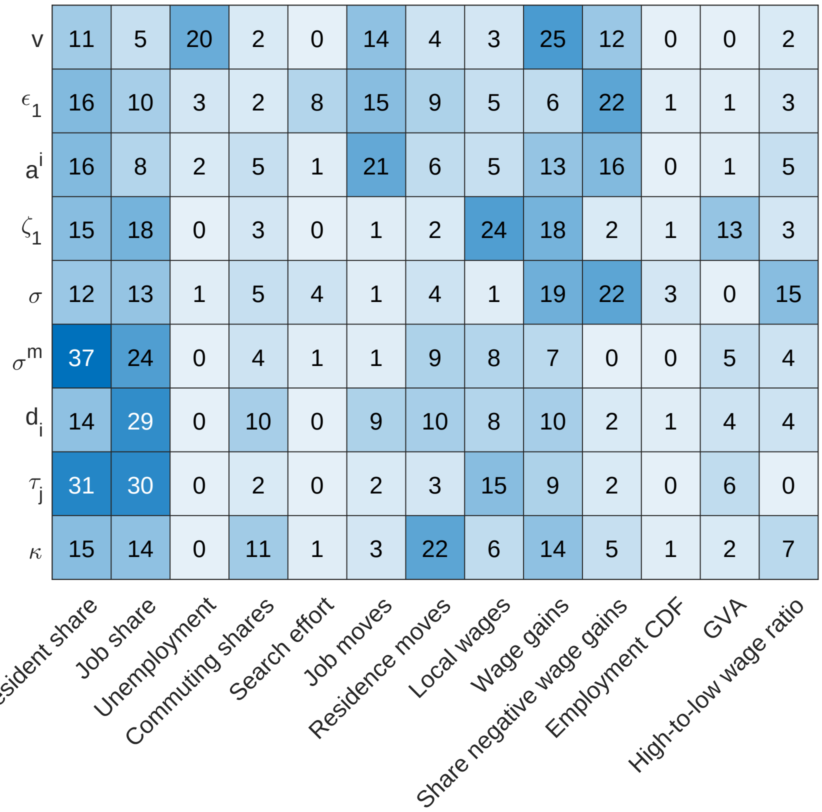
In Figure G.3 I report these deltas, averaged by parameter type (to improve readibility) and normalised so that each rows sums to 100, in order to highlight what moments' fit is most affected by a given parameter.[59] The maximum colour scale is set to 37, the highest value on the chart. Overall, the jacobian suggests that the identification strategy works well to identify the parameters of interest. I now discuss this in more detail.

Some moments appear to be more important than others. These are the resident and job shares, and wage gains. This reflects two things. First, the fact that some moments are more precisely estimated than others, hence small changes in the fit have a larger effect on the aggregate likelihood relatively to other moments. Second, the fact that most parameters are identified up to knowing the value function, for which wage gains (and hence continuation values) are fundamental.

Most parameters affect the total likelihood mainly through the predicted moments described in Appendix E.1. The parameters related to spatial frictions, d_i (commuting costs) and κ (residence moving costs), τ_j (amenities), all heavily load on the respective moments (commuter and movers shares, residence moving shares and origin/destination, residents and worker shares). v, the relative search effectiveness of the unemployed, is the only parameter heavily loading on unemployment. ε_1 and σ are the only ones loading on the derivative of search effort. a^i, the amount of applications generated per unit of effort in the home location, loads on the amount of job moves. σ, the variance of the job preference shock, loads on the share of negative wage gains. σ_m, the variance of the residence shock, loads on residence moves. Only ζ_1 does not depend on the predicted parameters, but on an affine one (the relative GVA by locations), possibly because the 10 parameters affect only a few parameters in the associated distribution of employment CDF (6 each out of 60).

[59] This is needed to account for how a 10% change in a parameter could have very different effects on the total likelihood.

Figure G.3: Jacobian matrix of likelihood function



Note: each row represents the mean change, across all parameter sets, in the fit of each set of moments, with the sum normalised to 100.

H Additional counterfactual results

H.1 Agglomeration effects

I relax the initial assumption of no agglomeration effects. In particular, I allow agglomeration to change with the population density in a location. I assume A_j^T has functional form

$$\ln(A_j^T) = 0 + A\ln\left(\frac{(\text{density})_{jt}}{(\text{density})_{j0}}\right), \tag{32}$$

where A captures the "proportional" effect of changes in density on tradeable productivity. For example, for A=0.01 a small increase in density would increase all firms' productivity in the tradeable sector by 1%. $(\text{density})_{jt}$ captures the population density of a location, calculated as $(\text{density})_{jt} = \frac{(\text{population mass})_{jt}}{(\text{Area})_{jt}}$. $(\text{density})_{j0}$ is the population density calculated directly from the data.

Equation 32 specifies the 0 to highlight that, in the calibration of my model, I implicitly assume that the baseline A_j^T is equal to 1 in all locations.[60] Hence, all baseline productivity effects of agglomeration will go into the estimated distribution of productivity p.

Notice that while this is a proportional effect in A^T, high-productivity firms are more likely to allocate their labour force to the production of tradeables. Hence, accounting for localness heterogeneity implies that changes in local agglomeration effects do not increase total productivity, nor value added, proportionally across firms. This is reminiscent of the results of Combes et al. (2012), who show that agglomeration economies have both "shift" and "dilation" effects on the productivity distribution. Notice that, when ignoring localness heterogeneity, "dilation" effects disappear, as the production of tradeable goods has the same weight on value added across all firms.

Calibration. Agglomeration effects have been estimated to be around 0.02-0.06 log-points for each 1-log point increase in density. Accounting for the fact that tradeable sales represent, on average, 61% of all firms' revenues, I calibrate the model to $A = \frac{0.06}{0.61} \approx 0.098$. Meaning that doubling a region's density would increase the region's tradeable productivity by approximately

[60] Notice that, in the baseline calibration, $(\text{density})_{jt} = (\text{density})_{j0}$. Hence, $A_j^T = \exp(0 + A\ln(1)) = \exp(0) = 1$.

10%.

Calibrating A to an high value also allows to show that accounting for realistic agglomeration effects is mostly inconsequential for the qualitative results of my analysis. While - once I account for localness heterogeneity - large businesses are more affected by changes in agglomeration of tradeables than smaller ones, there are two things to consider:

1. Having productive businesses benefiting more from agglomeration works towards making the additional congestion in productive ($\approx$ high-amenities, in the UK) areas less costly. Hence, this is expected to (slightly) strengthen the baseline results.

2. The number of migrants between regions is relatively small. Hence, the quantitative importance of this channel is likely to be small.

I calculate

Results. Indeed, I find that the results of the agglomeration exercise mostly follow the baseline counterfactual. Table H.1 show that the total costs of spatial frictions are slightly underestimated when not accounting for agglomeration, regardless of the stance taken of localness heterogeneity. The difference between the two counterfactuals with and without localness heterogeneity are very similar to those in the baseline result without agglomeration, although slightly larger.

Table H.1: Agglomeration counterfactual

	Het. VA (%)	Fixed VA (%)	Difference
GDP	5.57	4.55	1.02
VA of firms	1.48	0.07	1.41
VA of firms: per worker	-1.2	-2.6	1.41
Wages	1.89	0.13	1.76
Wages: per worker	-0.79	-2.54	1.75
Welfare	31.12	29.81	1.3

Notes: All variables are in real terms, adjusted by the average price index of consumption of all agents. "GDP" is the total value of production, including housing. "VA of firms" is the total value added of firms. "Wages" are the total wage bill, and "Wages: per workers" are wages by the mass of employed individuals. "Welfare" is the mean value of the value function of all individuals. The "Het. VA" and "Fixed VA" column shows the differences between the "Heterogeneous VA effect" and the "Fixed VA effect" counterfactuals and the baseline calibration, in percentage points. The last column shows the difference, in percentage points, between the two counterfactuals.

I Additional results for planning policy

I.1 First stage

In this section, I present the first stage results for the IV specifications reported in Table 8.

Table I.1: Estimation of regulation impact on housing price elasticity - First Stage

	ln(y) × Refusal		
	(1)	(2)	(3)
ln(income) × Delay change	-0.080***	-0.169***	
	(0.019)	(0.025)	
ln(income) × Labour share 1983	-1.99***		-2.10***
	(0.194)		(0.204)
ln(income)	0.524	0.019	0.617
	(0.426)	(0.492)	(0.464)
Delta elevation	0.238	0.347	0.217
	(0.166)	(0.226)	(0.168)
Share built 1990	4.42*	0.157	5.11*
	(2.26)	(5.58)	(2.64)
ln(income) × Delta elevation	-0.032	-0.050	-0.029
	(0.023)	(0.031)	(0.023)
ln(y) × Share built	-0.615**	-0.240	-0.701*
	(0.312)	(0.744)	(0.363)
Observations	3,444	3,444	3,444
R^2	0.456	0.189	0.443
Within R^2	0.456	0.189	0.443
Year fixed effects	✓	✓	✓

Notes: p-values: *** ≤ 0.01, ** ≤ 0.05, * ≤ 0.10. Standard errors are Conley spatial standard errors, with a cut-off of 100km. Observations are Local Authority Unit × year, for the period 1998-2019. The table shows the first stage results for the IV specifications whose second stage is reported in Table 8 in the main text.

I.2 Robustness: Expenditure

Since my model predicts interactions between the role of local expenditure on housing and business floorspace, a more precise independent variable to estimate local demand would be one that takes into account: i) the local expenditure on housing, and ii) the local expenditure on floorspace. Due to differences in data availability, I perform the regression between 2001 and 2019 (instead of 1995 and 2019 as in the main text), but report the results of the standard specification for the new time frame as a guide. I report the findings in Table I.2. I find that accounting for floorspace costs increases the size of β^{reg} by approximately 15% in both the OLS and the IV specification.

Table I.2: Estimation of regulation impact on housing price elasticity - Expenditure

	ln(house price)			
	OLS		IV	
	(1)	(2)	(3)	(4)
ln(income)	0.046		0.109	
	(0.029)		(0.142)	
Delta elevation	0.001	0.0006	0.035	0.034
	(0.029)	(0.026)	(0.096)	(0.089)
Share built 1990	-2.33**	-2.11**	0.232	0.280
	(1.03)	(0.903)	(4.22)	(3.87)
ln(income) × Delta elevation	-0.001		-0.005	
	(0.004)		(0.012)	
ln(income) × Refusal share	0.111***		0.177***	
	(0.014)		(0.015)	
ln(income) × Share built 1990	0.313**		0.004	
	(0.135)		(0.532)	
ln(expenditure)		0.051*		0.123
		(0.030)		(0.153)
ln(expenditure) × Delta elevation		-0.001		-0.006
		(0.004)		(0.013)
ln(expenditure) × Refusal share		0.127***		0.202***
		(0.016)		(0.018)
ln(expenditure) × Share built 1990		0.327**		-0.003
		(0.137)		(0.560)
Observations	2,296	2,296	2,296	2,296
R^2	0.692	0.695	0.566	0.567
Within Adjusted R^2	0.577	0.581	0.403	0.405
Kleibergen-Paap			1.34	1.37
Year fixed effects	✓	✓	✓	✓
Instruments			All	All
Endogenous			Refusal, Area	Refusal, Area

Notes: p-values: $^{***} \leq 0.01$, $^{**} \leq 0.05$, $^{*} \leq 0.10$. Standard errors are Conley spatial standard errors, with a cut-off of 100km. Observations are Local Authority Unit × year, for the period 1998-2019. The table shows the results of a regression of ln(house prices) at the Local Authority Unit level over two sets of covariates. The first one uses ln(income), which captures the income of individuals residing in the LAU. The second uses ln(expenditure), which adds the business rates of local businesses as a proxy for the firms' expenditure on "space".

I.3 Robustness: Bartik instrument for income shock

Another possible worry is that changes in local demand are endogenous. For example, Diamond (2016) considers how local changes in construction costs may be correlated to local changes in demand growth. When construction costs are not observed by the econometrician, their contribution ends up in the residual. However, if construction cost shocks are correlated to local income shocks (for example, due to local construction workers' wages ending up in both measure), this would mean that the error term and the independent variables are correlated to each other, introducing bias in the fitted coefficient $\hat{\beta}^{\text{reg}}$.

To take care of this possible endogeneity bias, I instrument local income with a Bartik instrument constructed from national industry-level value added growth, and local industry shares. I find that the estimates of the key parameter, "ln(income) $\times$ Refusal share" are almost identical to the standard IV method.

Table I.3: Estimation of regulation impact on housing price elasticity - Expenditure

	ln(house price)		
	OLS	IV	
	(1)	(2)	(3)
ln(income)	0.046	0.041	0.096
	(0.029)	(0.029)	(0.138)
Delta elevation	0.001	0.001	0.030
	(0.029)	(0.028)	(0.095)
Share built 1990	-2.33**	-2.12**	0.254
	(1.03)	(1.02)	(4.28)
ln(income) × Refusal share	0.111***	0.111***	0.177***
	(0.014)	(0.014)	(0.014)
ln(income) × Share built 1990	0.313**	0.289**	0.003
	(0.135)	(0.134)	(0.537)
ln(income) × Delta elevation	-0.001	-0.001	-0.004
	(0.004)	(0.004)	(0.012)
Observations	2,296	2,296	2,296
R^2	0.692	0.692	0.564
Within Adjusted R^2	0.577	0.576	0.401
Kleibergen-Paap		21.2	0.218
Year fixed effects	✓	✓	✓
Instruments		Bartik	Bartik, Delay, Labour votes
Endogenous		Income	Income, Refusal, Area

Notes: p-values: *** ≤ 0.01, ** ≤ 0.05, * ≤ 0.10. Standard errors are Conley spatial standard errors, with a cut-off of 100km. Observations are Local Authority Unit × year, for the period 1998-2019. The table shows the results of a regression of ln(house prices) at the Local Authority Unit level over a ln(income), ln(income) interacted by the refusal rate of major project applications in the LAU, and a set of controls. The first column reports OLS results. In the second column, ln(income) is instrumented using a Bartik-style shift-share instrument. In the third column, the Barkit instrument is added on top of the baseline IV specification reported in Table 8 in the main text.

J Other Policies Details and Results

J.1 Work from home

Although eliminating commuting costs would be unfeasible, recent advancements in technologies and changes in workplace culture have led many high-skilled workers to work remotely. As of early 2023, the UK was the English-speaking country with the largest share of job ads including at least some degree of hybrid or remote work (Hansen et al., 2023). In this section, I study the effects of a 30% reduction in commuting costs for high-skilled workers only, consistent with the current evidence about work-from-home trends.[61] I interpret this reduction in commuting costs as a shift toward hybrid work, rather than full remote work for a few workers, as the former has not been associated (unlike the latter) with a fall in productivity or mentoring (Barrero et al., 2023).

Results. I illustrate the results in Table J.1. Introducing a work-from-home policy for high-skilled workers leads to lower GDP, but higher value added of firms. Employment increases slightly, but real wages per worker fall, as employers can lower wages and still be able to recruit more workers. However, welfare increases as high-skilled workers pay considerably less utility and pecuniary costs when commuting. Accounting for localness heterogeneity leads to larger gains, but also to slightly exacerbated inequalities between the gains of high- and low-skilled workers. For example, the welfare gains of high-skilled workers are 9% larger (1.11 pp. vs 1.02 pp.), while the welfare gains of low-skilled workers are even lower (0.06 pp. vs 0.07 pp.). This is due to high-skilled workers' location choices causing additional congestion on low-skilled firms, reducing their value added gains, and thus real wages paid to their workers.

[61] Bick et al. (2023) find, using survey data and a structural model, that the equilibrium work-from-home level for all workers in the U.S. will be approximately 1-in-5 days, or a 20% fall.

Table J.1: Work-from-home counterfactual

	Het. VA (%)	Fixed VA (%)	Difference
GDP	-0.14	-0.31	0.17
VA of firms	0.14	0.09	0.05
- VA: High Skilled	0.07	0	0.07
- VA: Low Skilled	0.39	0.44	-0.05
Wages	-0.29	-0.41	0.12
Wages: per worker	-0.37	-0.48	0.11
- Wages: High Skilled	-0.6	-0.75	0.15
- Wages: Low Skilled	0.37	0.39	-0.02
Welfare	0.88	0.8	0.07
- Welfare: High Skilled	1.11	1.02	0.1
- Welfare: Low Skilled	0.06	0.07	-0.01

Notes: All variables are in real terms, adjusted by the average price index of consumption of all agents. "GDP" is the total value of production, including housing. "VA of firms" is the total value added of firms. "Wages" are the total wage bill, and "Wages: per workers" are wages by the mass of employed individuals. "Welfare" is the mean value of the value function of all individuals. The "Het. VA" and "Fixed VA" column shows the differences between the "Heterogeneous VA effect" and the "Fixed VA effect" counterfactuals and the baseline calibration, in percentage points. The last column shows the difference, in percentage points, between the two counterfactuals.

K Extensions

K.1 Microfundation for $\beta(p)$

In the main text, I have assumed that the share of labour dedicated to the production of tradeables and non-tradeables is exogenous. In this section, I relax this assumption.

Assume that the production function of a firm is

$$\text{Tradeables: } T_j^i(p,n) = A_j^T \beta^{\gamma^i(p)} pn,$$

$$\text{Non-Tradeables: } N_j^i(p,n) = A_j^N (1-\beta)^{1-\gamma^i(p)} pn,$$

where β is the share of labour assigned to the production of tradeables, $\gamma^i(p)$ is an exogenous parameter that regulates how much a firm's productivity is decreasing in allocating additional units of labour to the production of tradeables, and $1-\gamma^i(p)$ is the same for the production of non-tradeables.

Consider the problem of labour allocation of a firm that wants to maximise revenues per unit of labour and productivity pn:

$$\beta_{i,j}^* = \arg\max_{\beta} A_j^T \beta^{\gamma^i(p)} + P_{l,j} A_j^N (1-\beta)^{1-\gamma^i(p)}.$$

Then, assuming an internal solution, the optimal $\beta_{i,j}^*$ solves the equation

$$\frac{(1-\beta_{i,j}^*)^{\gamma^i(p)}}{(\beta_{i,j}^*)^{1-\gamma^i(p)}} \frac{\gamma^i(p)}{1-\gamma^i(p)} = \frac{P_{l,j} A_j^N}{A_j^T}, \tag{33}$$

and value added per worker is

$$\tilde{p}_j^i = p\left[A_j^T (\beta_{i,j}^*)^{\gamma^i(p)} + P_{l,j} A_j^N (1-\beta_{i,j}^*)^{1-\gamma^i(p)} - P_{H,j}\phi_j^i(p)\right]. \tag{34}$$

K.1.1 Estimation

Assuming this microfoundation does not change the estimation of the first block of parameters. However, it affects the estimation procedure of the second block. In this section, I provide details of the assumptions needed to identify $\gamma(p)$ and details of the new algorithm.

Proposition 2. *Assume* $0 < \frac{P_{l,j} A_j^N}{A_j^T} < \infty$. *Then,* $\forall \beta_{i,j}^* \in (0,1)$, $\exists\, \gamma_i^*(p) \in (0,1)$ *such that, for* $\gamma^i(p) = \gamma_i^*(p)$, $\gamma_i^*(p)$ *is a solution to Equation 33.*

Proof. Call, to simplify notation, $\gamma = \gamma_i^*(p)$. Assume $\beta_{i,j}^* \in (0,1)$ and $\gamma \in (0,1)$. Then, taking logs in Equation 33,

$$\gamma\left(\ln(1-\beta_{i,j}^*) + \ln(\beta_{i,j}^*)\right) + \ln\left(\frac{\gamma}{1-\gamma}\right) = \ln\left(\frac{P_{l,j}A_j^N}{A_j^T}\right) + \ln(\beta_{i,j}^*),$$

where all the logs are defined since $\frac{P_{l,j}A_j^N}{A_j^T} > 0$, $\beta_{i,j}^* > 0$ and $\gamma > 0$. Consider how

$$\lim_{\gamma \to 0}\left[\gamma\left(\ln(1-\beta_{i,j}^*) + \ln(\beta_{i,j}^*)\right) + \ln\left(\frac{\gamma}{1-\gamma}\right)\right] = -\infty.$$

$$\lim_{\gamma \to 1}\left[\gamma\left(\ln(1-\beta_{i,j}^*) + \ln(\beta_{i,j}^*)\right) + \ln\left(\frac{\gamma}{1-\gamma}\right)\right] = \infty.$$

Since $f(\gamma) = \gamma\Big(\ln(1-\beta_{i,j}^*) + \ln(\beta_{i,j}^*)\Big) + \ln\left(\frac{\gamma}{1-\gamma}\right)$ is a continuous function in γ, for $\gamma \in (0,1)$, for the Intermediate Value Theorem, the function must attain $c = \frac{P_{l,j}A_j^N}{A_j^T}$ for some value of γ. However, applying the theorem requires a closed interval. I construct this as follows. Consider the closed interval $[a,b] \subset R_+$. Choose $a = 0$.[62] Choose $b = M$, where $\frac{P_{l,j}A_j^N}{A_j^T} < M < \infty$. Such a number must exist, as $\frac{P_{l,j}A_j^N}{A_j^T} < \infty$ by assumption. Since $\frac{P_{l,j}A_j^N}{A_j^T} > 0$ by assumption, $c \in [0,M]$ and, due to the intermediate value theorem, $\exists \gamma^* \in (0,1): \ f(\gamma^*) = c$.

That is, given any value of $\beta \in (0,1)$, then $\exists \gamma \in (0,1)$ such that it solves Equation 33. $\square$

Notice that this proof does not guarantee uniqueness, but just existence of a solution $\gamma \in (0,1)$. That is, a parameter such that the Inada conditions are satisfied for a given β. To take care of this issue, I introduce Assumption 3 and Proposition 3. In practice, this never happens within the domain of my empirical $\hat{\beta}$, as the solution to 3 implies that the solution is not unique only for β being larger than 0.981 or smaller than 0.019.

Assumption 3. *If $\exists \gamma, \gamma' \in (0,1)$, $\gamma \neq \gamma'$ such that $\gamma^i(p) = \{\gamma, \gamma'\}$ are solutions to Equation 33, assume that $\gamma_i^*(p) = \min\{\gamma, \gamma'\}$*

Proposition 3. *There exists one and only one $\gamma \in (0,1)$ such that $\gamma^i(p) = \gamma$ is the unique solution to Equation 33 whenever*

$$\beta^2 - \beta + \exp(-4) < 0.$$

[62] The fact that $a = 0$ is in the codomain of $f(\gamma)$ can be proven by applying the intermediate value theorem on the interval $[-1,1]$.

Proof. Taking the derivative of the log-version of 33, I obtain

$$\frac{\partial f(\gamma)}{\partial \gamma} = \ln((1-\beta^*_{i,j})\beta^*_{i,j}) + \frac{1}{\gamma} + \frac{1}{1-\gamma}.$$

Consider how $\frac{\partial f(\gamma)}{\partial \gamma} > 0 \, \forall \gamma \in (0,1)$ whenever

$$\frac{1}{\gamma} + \frac{1}{1-\gamma} > -\ln((1-\beta^*_{i,j})\beta^*_{i,j}).$$

In particular, we want to check whether

$$\min_{\gamma \in (0,1)} \left(\frac{1}{\gamma} + \frac{1}{1-\gamma} \right) = 4 > -\ln((1-\beta^*_{i,j})\beta^*_{i,j}),$$

where the equality is given by recognising how the left is minimised at $\gamma = \frac{1}{2}$. Taking the exponentials, the function must satisfy

$$\exp(4) > \frac{1}{(1-\beta^*_{i,j})\beta^*_{i,j}}.$$

$$-\beta^2 + \beta - \exp(-4) > 0.$$

The case for $\frac{\partial f(\gamma)}{\partial \gamma} < 0 \, \forall \gamma \in (0,1)$ is not relevant as $\sup_{\gamma \in (0,1)} \left(\frac{1}{\gamma} + \frac{1}{1-\gamma} \right) = \infty$. Hence, there is no β such that the function can be negative for any γ. □

Then, for each location, skill, and $\tilde{p}$ we can pin down $\gamma^i_j(p)$ such that the share of employment allocated to the production of tradeables for a firm with value added $\tilde{p}$, skill i in location j is $\tilde{\beta}^*_{ij}(\tilde{p}) = \hat{\beta}^i_j(\tilde{p})$, where $\hat{\beta}^i_j(\tilde{p})$ is the corresponding value estimated from the NESS data at the corresponding CDF of employment of $E^i_j(\tilde{p})$.

Changes to second-block identification. Since γ is a component of value added, microfounding β does not change the identification of the first block of parameters. Hence, it only affects the procedure to identify the second block. The calibration of $\hat{\beta}^i_j(\tilde{p})$ and $\hat{\phi}^i_j(\tilde{p})$, the distribution of employment allocated to producing tradeables and the floorspace costs, is unchanged, as it requires to fit the empirical moments to the employment CDF of each (i,j) market. However, notice that $\hat{\beta}^i_j(\tilde{p})$ is not anymore a parameter in itself, but an equilibrium object (β^*) I observe and I am going to target to identify γ.

I now describe how to recover p, A^N and the distribution of γ. Recall that A^N_j was identified by iterating over different guesses of A^N_j until convergence to a value that guaranteed market clearing in the local non-tradeable good market, where supply depends on A^N_j both directly and through the distribution of productivities p implied by $\beta(p)$, $\phi(p)$, and A_j^N.

Identifying A_j^N, p, and γ follows a similar procedure. Call $\hat{\beta}_j^i(\tilde{p})$ the value of β^* for firms of type i, in location j, with value added $(\tilde{p})$, as fitted to the CDF of employment using NESS data. Then, apply the following algorithm:

1. Start from a guess of A_j^N.

2. Compute $\tilde{\gamma}_j^i(\tilde{p})$ such that it solves
$$\frac{(1-\hat{\beta}_j^i(\tilde{p}))^{\tilde{\gamma}_j^i(\tilde{p})}}{(\hat{\beta}_j^i(\tilde{p}))^{1-\tilde{\gamma}_j^i(\tilde{p})}}\frac{\tilde{\gamma}_j^i(\tilde{p})}{1-\tilde{\gamma}_j^i(\tilde{p})}=\frac{P_{l,j}A_j^N}{A_j^T}.$$

3. Calculate p by computing the following expression:
$$p_j^i(\tilde{p})=\frac{\left(\tilde{p}+\hat{\phi}_j^i(\tilde{p})\right)}{(\hat{\beta}_j^i(\tilde{p}))^{\tilde{\gamma}_j^i(\tilde{p})}+P_{l,j}A_{l,j}(1-\hat{\beta}_j^i(\tilde{p}))^{1-\tilde{\gamma}_j^i(\tilde{p})}}.$$

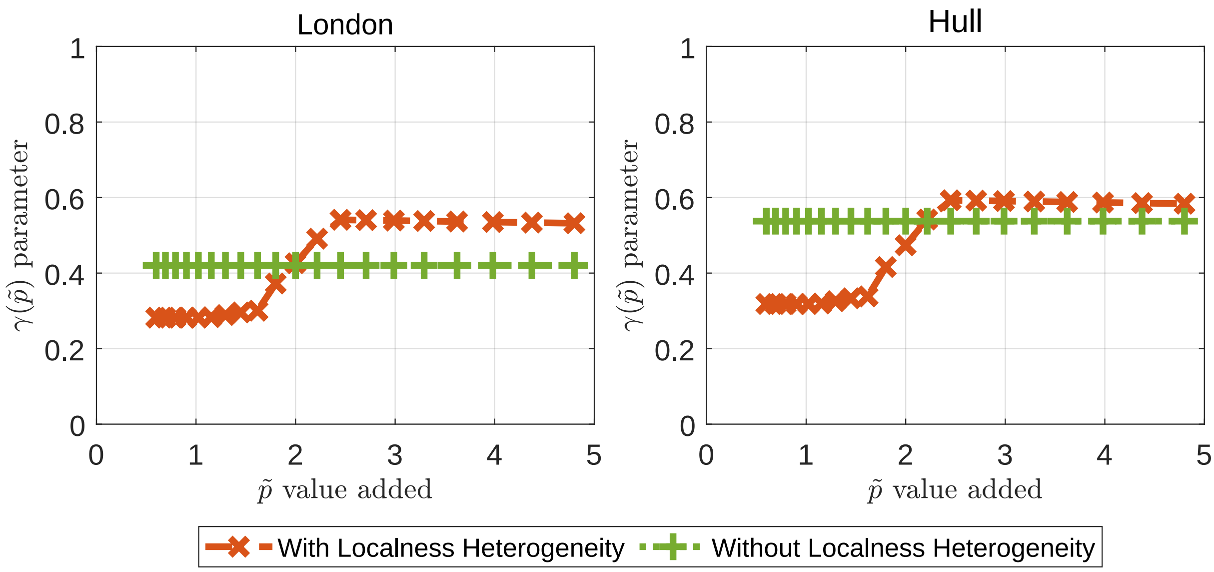
4. Check whether the local non-tradeable markets clear at the given A_j^N by checking if
$$A_{l,j}\sum_{i\in I}M_j^i\int_{\tilde{p}}p_j^i(s)(1-\hat{\beta}_j^i(s))^{1-\tilde{\gamma}_j^i(\tilde{p})}\hat{n}_{i,j}^*(s)\hat{v}_{i,j}^*(s)dF_j^{i,f}(s)=\eta_1\frac{Y_j^r}{P_{l,j}}.$$

5. If markets do not clear, take a new guess of A_j^N (larger if excess demand is positive, smaller if negative) and repeat points 2-5 until convergence.

6. Once A_j^N and $p_j^i(\tilde{p})$ are such that the market clearing condition is satisfied, obtain $\phi_j^i(p)=\frac{\hat{\phi}_j^i(\tilde{p}(p))}{p}$.

I provide the details of the estimated distribution of $\tilde{\gamma}(\tilde{p})$ for high-skilled firms in the richest and poorest locations, in Figure K.1.

Figure K.1: Value of $\gamma(\tilde{p})$, with vs without localness heterogeneity



K.1.2 Counterfactual

In Table K.1, I provide the results when estimating the model using the microfounded β. The results are approximately similar to those reported in the baseline counterfactual, although somewhat smaller (between 0.01 percentage points for VA of firms, and 0.26 percentage points for GDP). This is for two reasons:

1. The changes in the price of non-tradeables across counterfactuals is mostly small, meaning that the channel is of small quantitative importance.

2. Changes in house prices are larger, and firms are more heterogeneous in $\phi(p)$ than in $\beta(p)$. Most of the additional misallocation arising from this channel is unaffected by the microfoundation assumptions on $\beta(p)$.

Table K.1: Key variables, real change relatively to baseline (microfounded β model)

	Het. VA (%)	Fixed VA (%)	Difference
GDP	5.05	4.20	0.85
VA of firms	1.42	0.08	1.34
VA of firms: per worker	-1.28	-2.60	1.32
Wages	1.79	0.14	1.65
Wages: per worker	-0.92	-2.54	1.62
Welfare	30.95	29.76	1.20

Notes: All variables are expressed in real terms. Nominal GDP is adjusted to real terms using a price deflator at the product level. All other variables are adjusted using the average price index of consumption of all agents. "GDP" is the total value of production, including housing. "VA of firms" is the total value added of firms. "Wages" are the total wage bill, and "Wages: per workers" are wages by the mass of employed individuals. "Welfare" is the mean value of the value function of all individuals. The "Het. VA" and "Fixed VA" column shows the differences between the "Heterogeneous VA effect" and the "Fixed VA effect" counterfactuals and the baseline calibration, in percentage points. The last column shows the difference, in percentage points, between the two variations.