Abstract

Even though households in the UK spent over £1,000 per worker on commuting in 2017, the economic costs may be far higher because commuting and the associated congestion affect workers’ incentives for job search and acceptance. At the same time, workers’ job acceptance decisions determine commuting patterns and therefore aggregate congestion. Using UK data on commuting and employment outcomes, I find a strong positive relationship between commuting time and future job mobility. To understand the empirical patterns and quantify the aggregate implications, I build a novel model featuring a frictional labor market in which commuting gives rise to congestion as workers travel similar paths to work. By taking the frictions in the labor market and the congestion arising from commuting seriously, the model can be used to evaluate the spillovers from policies aimed at improving congestion or the functioning of the labor market. Since it takes time to find a close and productive job, and because moving house is costly, many workers commute to distant jobs. By contributing to congestion, they affect the incentives of other workers to accept job offers. The quantitative model is calibrated to match features of the London area and suggests that 25% of wage dispersion and 11% of utility dispersion is a result of congestion. The model predicts that targeted remote working policies can have large welfare benefits even with a small productivity loss from working at home. Focusing only on the congestion effects of policies may be misleading about their welfare implications due to their effects on workers’ progress up the job ladder.

*I am greatly indebted to Guido Menzio, Facundo Pigüillem, and Margaret Stevens for their advice and guidance. I’d also like to thank Isaac Baley, Paul Beaudry, James Best, Jake Bradley, Marco Casiraghi, Keshav Dogra, Cynthia Doniger, Gilles Duranton, Jan Eeckhout, Martin Ellison, Andrew Haughwout, Fatih Karahan, Francesco Lippi, Javier Fernández-Blanco, Ole Jann, Michael McMahon, Claudio Michelacci, Laura Pilossoph, Facundo Pigüillem, Diego Puga, Dan Quigley, Petr Sédlacek, Coen Teulings, Giorgio Topa, and numerous conference and seminar participants for helpful comments.
1 Introduction

In the 10 cities in the world with the most congestion, commuters spend an average of 80 hours per year in rush hour traffic, equivalent to two work weeks of lost productive time.\(^1\) During rush hour, many individuals are commuting to the same place at the same time, leading to significantly longer travel times for everyone. In addition to the direct costs of lost time and output, congestion might have even larger indirect costs because it potentially misallocates resources: longer commutes can stop workers from accepting job offers, halting their progress up the “job ladder”. Data from the New York Federal Reserve’s Survey of Consumer Expectations shows that 13% of workers report the job’s location as the main reason for rejecting an offer.\(^2\) This suggests that workers’ rankings of jobs depends on the commuting costs required to get to work, and therefore varies across workers depending on where they live and work. One way for workers to avoid congestion is by working remotely. As cities have developed and technology has improved, the number of remote workers in the UK alone has risen from 2.9 million in 1998 to 4.2 million in 2014.\(^3\) However, little is known about how remote working policies affect congestion, labor market outcomes, and welfare within metropolitan areas.

To better understand the link between congestion and the functioning of the labor market, I study how the process of matching workers to jobs and commuting patterns are jointly determined. To the best of my knowledge, this is the first paper in which congestion arises endogenously within a frictional labor market. Unlike other job amenities, such as flexible working hours, commuting costs depend on congestion, which is determined by the equilibrium choices of workers. This is especially important when jobs are spatially concentrated, as is the case in urban areas. Because it takes time to find a close and productive match, and because moving house is costly, many workers commute to distant jobs. The nature of the externality studied in this paper is that by accepting a job that requires the worker to commute, she affects the commuting costs of others by contributing to congestion. Workers take commuting costs as given, but changes in some workers’ commuting patterns affect congestion and therefore impact many other workers’ decisions about which jobs to accept and which to reject.

Figure 1 depicts how congestion affects job acceptance: if workers 1 and 2 are commuting to their jobs, indicated by arrows, they congest one another because their commuting paths cross. If worker 3 receives a job offer, she may reject it, indicated by the dashed line, if the congestion caused by worker 1 is high enough. Intuitively, workers reject otherwise acceptable matches if congestion is severe, and all workers value their matches differently depending on the level of congestion in the economy, which affects their wages, utility, and future employment outcomes. If instead worker 1 worked remotely, her own commuting costs would be eliminated, and because of her effect on congestion, worker 2’s match would improve and worker 3 would accept her job offer. Without

---

1 Source: INRIX Global Traffic Scorecard. Rush hour is defined as the hours 6:00-9:00am and 4:00-7:00pm, Monday-Friday. Traffic congestion is defined as speeds 65% or below the free flow speed.

2 I consider the location-related reasons for rejection to be that the commute time would be too long or the job would require a relocation. See the SCE Job Search Survey codebook for further details.

3 Source: Office for National Statistics.
the externality, the fact that worker 1 works remotely has no effect on other workers’ job values or acceptance. In addition to the static interaction across workers, the dynamic nature of job search will lead to further effects of remote working through all workers’ continuation values.

This paper makes three contributions. First, I document new facts connecting commuting and job mobility using data from the UK. In particular, workers are more likely to change jobs via a job-to-job transition the higher is their current commuting time. In addition, relative to their last job, workers are more likely to accept wage cuts if they also have a shorter commute. To understand these facts and evaluate the effects of policies, my second contribution is to develop a novel spatial model of job search with an externality in commuting. By combining the spatial features of an urban area with the bargaining process of Cahuc et al. (2006), the model remains tractable while matching key features in the data. The model is innovative in several ways, most importantly by incorporating endogenous congestion and commuting costs that affect all workers’ decisions both for job and residential mobility. Third, I calibrate the model using data from the London area and quantify the effect of congestion on wage and utility dispersion. I then perform counterfactual exercises to study how remote working policies affect labor market outcomes, congestion, and the allocation of workers and firms across space.

I begin by documenting new facts about job search within local labor markets, focusing specifically on the London area, one of the most congested cities in the world. Data from the British Household Panel Survey (BHPS) shows that wage changes and job mobility are strongly linked to commuting time. In the aggregate, commuting flows are positively correlated with proximity to the center of London’s economic activity, the City of London. At the individual level, I document evidence on compensating differentials by showing that the majority of job-to-job transitions

---

4I use the Cahuc et al. (2006) model to allow for wages to be functions of commuting costs for workers, which is crucial when thinking about non-wage amenities (e.g. Rosen 1986, Sorkin 2018). The model is more tractable than a wage posting model because it significantly simplifies the distributional dependence of wages as they are set after the worker and firm meet rather than ex-ante.
involving wage cuts are accompanied by a simultaneous decrease in commuting time.

To assess how congestion arises given labor market frictions in a dynamic setting, I build a spatial on-the-job search model with an externality in commuting. The labor market is made up of a number of locations that differ in the cost of living, given by the housing rent (henceforth, rent). Workers’ utility is increasing in wages, and decreasing in commuting costs and rent, thus, identical workers may value the same location differently depending on the location of their job, and value the same job differently depending on the location of their home. Commuting costs are a function of two components: a “no congestion” cost per unit of distance and a congestion cost. I identify this cost using differences in commuting time during rush hour and during times of little traffic using data from Google Maps. The value of a job reflects its productivity, rent, and commuting costs, leading to some highly productive matches being rejected when such costs are high. Commuting patterns, output, and congestion are jointly determined in equilibrium. Because wages are determined by the match surplus, which itself is a function of productivity and locations, the model is not observationally equivalent to one in which only wages net of commuting costs matter for workers. Unlike models with a frictionless labor market, wages depend on commuting and congestion.

Because the locations of workers and firms determine the demand for land, I allow rents to adjust in equilibrium. The model therefore is able to endogenously generate the pattern that the locations with the most productive jobs have both the most congestion and highest rent. Without allowing rents to change, the incentive to move to the most congested location would lead many workers and firms to locate in the same place. By considering endogenous rent, when more workers and firms move to the congested area, rent increases, making it less attractive to locate there. Using tools introduced in the economic theory literature on existence of equilibria with indivisible goods, most importantly that by Kaneko and Yamamoto (1986), I prove that an equilibrium rent price exists in the steady state. Recent work by Halket and Vasudev (2014) uses similar techniques to prove existence of a spatial equilibrium, but their focus is on homeownership and the choice of housing quantity, while abstracting from search frictions in the labor market. Conversely, I am interested in the effects of commuting to work and the associated externality, and thus model a much simpler housing market and a richer labor market.

To evaluate the model’s predictions, I calibrate it to match features of the London labor market, taking into account productivity differentials across space by including (exogenous) agglomeration whereby the city center is most productive. The model matches many features of urban labor markets, including commuting patterns by worker residence, with congestion and jobs concentrated in the city center. I use the model to estimate the effect of congestion on wage and utility dispersion across otherwise identical workers. Frictionless models with non-wage amenities such as Rosen (1986) predict that earnings inequality reflects differences in amenity values, but after taking into account the utility from both the wage and non-wage components of the job, utility is equalized across workers. Differently, Hwang et al. (1998) show that when labor markets are frictional, both wages and utility can exhibit dispersion. To understand how the equilibrium interaction of
the commuting amenity affects these measures, I compare the model’s predictions for dispersion with and without congestion, and find that congestion accounts for roughly 25% and 11% of wage and utility dispersion, respectively. Wages reflect the level of congestion through two channels: first, workers who have long commutes have higher costs due to congestion and therefore have low flow utility today, leading them to require more compensation for their commuting cost. Second, congestion affects workers’ future wages through the set of job offers they are willing to accept when searching on the job. Utility is defined as the difference between the wage and commuting and rent costs. Because workers with long commutes tend to live in areas with lower rent, commuting and rent costs to offset one another and may lead to less dispersion in utility. At the same time, workers climb the job ladder in both the spatial dimension and the productivity dimension, thus workers with short (long) commutes also tend to have high (low) wages, increasing dispersion. With congestion, commuting is costliest through areas with the most productive and highest paying jobs, leading to shorter commutes and less dispersion in commuting costs and utility. The model predicts that the latter effect dominates: congestion accounts for less of utility dispersion than wage dispersion.

Finally, I evaluate counterfactuals related to remote working. Without congestion, commuting costs are a feature of jobs that affect individual decisions, but do not interact across workers. Thus, allowing workers with long commutes to work remotely will only affect welfare directly, through workers’ own incentives to accept such jobs and the benefits they receive from reducing their commuting costs. Differently, workers with the longest commutes contribute greatly to congestion, thus, allowing these workers to work from home can have large indirect effects on welfare through changes in other workers’ commuting costs.

There are several key results. First, policies should take into account that congestion has heterogeneous effects across space and therefore should be spatially targeted. When output losses are zero, imposing remote working for the highest productivity jobs employing workers who would otherwise have to commute over 100 miles increases welfare by 1.5%. This policy increases the share of the highest productivity jobs held in equilibrium, causing output and welfare to rise and unemployment to fall. Because these workers would otherwise have to move house in order to accept an offer, rather than commute such a long distance, welfare for these workers increases. However, with more employed workers commuting at least part-time, congestion also increases. It is better in terms of welfare to target the workers who commute the farthest because they are likely to commute through locations where congestion is highest, closest to the center. If only firms in the center are targeted, regardless of the worker’s commute, welfare gains are lower because the policy is applicable only to a small share of job offers, and workers still must commute at least part time into the congested area. Targeting the policy only to firms in the center has little effect on workers commuting through the center, and therefore leads to smaller welfare gains. When remote working is associated with a 1% productivity loss, welfare gains are smaller, but can still be positive (about 1% higher) if the policy is targeted to the workers with the longest commutes.

In general, the more jobs that allow long distance workers to work from home, the lower is
congestion but also the lower is the job-to-job transition rate. This is because remote working can make workers unwilling to accept job offers that require a commute, leading to less job mobility and lower welfare. Policies that reduce congestion, for instance allowing all workers to work remotely 4 days per week regardless of location or commuting distance, can have large negative welfare effects because they make workers less likely to climb the job ladder into more productive jobs, and instead remain in low rent, low productivity jobs that do not require a commute.

Related Literature

This is the first paper to consider the interaction between commuting congestion and workers’ progress up the job ladder. I contribute to four broad strands of the literature. First, I add to the work on frictional labor markets and commuting, explicitly modeling a congestion externality across space, a key feature affecting commuting time in urban areas. Second, I contribute to the urban economics literature by studying the implications of this externality in a model with labor market frictions. Third, I contribute to the literature on compensating differentials by considering interactions of amenity values across workers. Finally, my model is linked to the literature on spatial mismatch by explicitly considering distance but endogenizing commuting time through workers’ ability to move house.

The connection between jobs and locations has been acknowledged since the island model of Lucas and Prescott (1974). A large literature has followed, explicitly modeling the interaction of commuting and labor markets. Commuting costs have been shown to be crucial determinants for a wide range of labor market outcomes, including labor supply (e.g. Wales 1978, Cogan 1981), work effort (e.g. Van Ommeren and Gutiérrez-i Puigarnau 2011), and wages (e.g. Van Der Berg 1992, Van Den Berg and Gorter 1997, Manning 2003).

Most theoretical models that consider commuting only allow for the unemployed to search for jobs and thus do not consider commuting’s effect on the job ladder. Studying the trade-off between wages and commutes in a model of on-the-job search is important for two reasons. First, job-to-job transitions account for a large share of total hires as well as wage growth (e.g. Topel and Ward 1992, Karahan et al. 2017, Moscarini and Postel-Vinay 2017). Additionally, as highlighted by Hornstein et al. (2011), it is essential to model on-the-job search to better match the frictional wage dispersion seen in the data.

An early paper that incorporates on-the-job search and commuting is Van Ommeren et al. (1999), who study the joint determination of job and residential locations when labor and housing markets are frictional. Using a reduced-form model which they estimate using data from the Netherlands, their analysis highlights the dynamic nature of mobility choices for workers’ lifetime utility. More recently, Van Vuuren (2018) introduces commuting in an on-the-job search model in which all firms are located in the city center and in which the arrival rate of job offers is a function of distance between a worker and vacancy. His focus is on the implications for new entrants to the labor market. The externality in his model is the limited amount of space for workers to live;

\footnote{For comprehensive coverage of urban-labor models, see Zenou (2009).}
by locating close to the center some workers push others to live farther from the location of jobs, increasing their commutes and decreasing their job-finding probabilities. I allow for both firms and workers to be located across the city and for workers’ commutes to congest one another. My interest lies in how the congestion aspect of commuting affects labor market outcomes and the implications of remote working policies.

Commuting has long been recognized as crucial to agents’ location and work decisions in urban economics (e.g. Fujita 1989). Models in this literature typically assume frictionless labor and housing markets; thus, workers are indifferent across all residential and job locations in equilibrium (e.g. Alonso 1964, Mills 1967, Muth 1969). More recent work considers the effect of frictions in labor and housing markets. Wasmer and Zenou (2002) study how the spatial configuration of an urban area affects the allocation and employment status of workers. Conley and Topa (2002) and Glaeser et al. (2008) consider mechanisms that give rise to spatial concentration of unemployment. Like these papers, I build a model that captures the spatial variation in employment and residential density, but focus on how these differences can interact across space via commuting and employment outcomes.

Since Pigou (1932), economists have acknowledged the importance of traffic density for travel speed and a large literature has developed to estimate the congestion externality and the tools necessary to achieve efficiency. The congestion externality explored in the literature measures the increase in time required to travel a given distance as the number of vehicles per unit of road increases (e.g. Walters 1961, Vickrey 1969). Brinkman (2016) builds an urban model in the spirit of Lucas and Rossi-Hansberg (2002) to jointly estimate congestion and agglomeration externalities. I model congestion similarly to his paper, as the mass of workers traveling through a location less the number of residents in that location (the “no-traffic” benchmark). I show how this congestion arises in a frictional labor market and affects the process of job search.

The literature on non-wage job characteristics and compensating differentials aims to understand how wages reflect non-wage aspects of jobs. Frictionless models such as Rosen (1986) predict that wages should compensate for these values, equating utility across workers. By highlighting the dynamic and frictional nature of the labor market, search theory has provided a framework through which the predictions for compensating differentials of these early models may be overturned (for instance, Hwang et al. 1998). Recent work by Sorkin (2018) and Hall and Mueller (2018) take two new approaches to understand specifically how dispersion in wages reflects dispersion in non-wage values, and how dispersion in utility may differ from dispersion in wages after taking into account non-wage amenities, respectively. So far, the literature has focused on these amenities at the match level and has not considered how the interaction of amenities across matches may affect dispersion in wages and utility.

Similar to the models in Taber and Vejlin (2016) and Jarosch (2016), I use the sequential auctions framework of Cahuc et al. (2006) to allow for a richer, yet more tractable, wage setting mechanism.\footnote{Becker (2011) builds a similar model of non-wage amenities with wages determined following Postel-Vinay and Robin (2002) and Nunn (2013) uses a bargaining model without renegotiation in the spirit of Mortensen and Pissarides (1994).} This model is different from the literature studying compensating differentials through the lens of a search model in several ways. Most importantly, I study an amenity that gives rise to a clear externality, traffic congestion. Although the assumption of independence of amenities across jobs may be reasonable for flexible working time or having a nice view from the office, it is difficult to rationalize for some amenities, especially the commute. Most models consider all non-wage amenities rather than focusing on one.\footnote{Two exceptions being Dey and Flinn (2005), who focus on access to health care, and Jarosch (2016) who considers job security.} Unlike Bonhomme and Jolivet (2009), who study commuting in addition to several other non-wage characteristics, the data I use in this paper contains information on the intensive margin of commuting, time, rather than an ordinal satisfaction measure of the worker’s commute.

Following the seminal paper of Kain (1968), many researchers have analyzed the prevalence of spatial mismatch.\footnote{For a review, see Gobillon et al. (2007).} Much of this literature is concerned with the effect of distance and urban structure on search effort or efficiency (e.g. Wasmer and Zenou 2006, Manning and Petrongolo 2017, Smith and Zenou 2003). The literature has shown that spatial mismatch cannot explain the increase in unemployment during the Great Recession (Şahin et al. 2014), nor in the level of unemployment (Marinescu and Rathelot 2018). Even if unemployed workers are willing to take jobs far from home and there is little spatial mismatch, the locations of workers and jobs and the congestion arising from commuting may have large effects on workers’ progress up the job ladder.

Manning and Petrongolo (2017) build a model of spatially directed job search by the unemployed without residential mobility. In their model, workers direct their search to jobs based on distance, taking into account how other workers’ search strategies affect their job finding probability. They estimate a high cost of distance using data that covers the entire UK. This paper considers the London area, where commuting time is longer than the UK average, and therefore locally targeted policies may have larger effects. The counterfactuals considered here focus on the consequences for the local labor market due to changes in commuting and congestion in response to remote working arrangements.

## 2 Motivating Evidence

This section discusses motivating evidence pointing to the importance of commuting for on-the-job search using two data sources from the UK. First, the UK Census Flows contain aggregate information on workers’ commuting patterns, which show the importance of the distance between workers and firms for observed job outcomes in the cross-section. Second, the British Household Panel Survey (BHPS) is an individual-level panel with information on individuals’ commute times,
employment histories, housing, and wages, which I use to document within-worker patterns relating to commuting, wages, and on-the-job search.

I present several facts. First, in the Census data, there is a decreasing pattern between commuting flows and distance between workers and firms. Second, commuting time positively affects individual workers’ probabilities of future job-to-job transitions as well as the share of workers accepting wage cuts, showing the importance of commuting for outcomes related to on-the-job search. Finally, I provide evidence for the importance of congestion by comparing commuting time at different levels of congestion using data from Google Maps.

Figure 2: Number of Commuters to City of London by Distance

![Diagram showing the number of workers commuting to the City of London as a function of driving distance in miles. Data: UK Census Flow Data 2001, Special Workplace Statistics (Level 1). Number of workers commuting to the City of London as a function of driving distance in miles computed by Google Maps. Map data: Google, DigitalGlobe.]

Census Flows

The model presented in the next section considers a “local labor market” that I will define for the quantitative analysis as the London metropolitan area. Using the 2001 Census data, Figure 2 shows the number of residents in each Local Authority District (LAD) working in the City of
London as a function of distance between the residence and the City in miles. The distance is computed using Google Maps, with the location automatically selected by Google for each LAD, where distance is defined as the shortest driving distance between the two locations.

**Compensating Differentials**

To provide evidence on the relationship between wages and commutes, I use data from the BHPS between 1992 and 2008. The BHPS is an annual panel survey with information on individuals' commute times, employment histories, housing, and wages. I restrict the sample in several ways. First, I drop all LADs for which there are fewer than 50 observations or that are more than 100 miles driving distance from the City to focus on daily commuters. I restrict attention to full-time workers between 24 and 55 (inclusive), and consider the period 1992-2008 to exclude any effect of the Global Financial Crisis on the commuting-wage trade-off. I include workers reporting net monthly labor income within 3 standard deviations of the mean and a one-way commute up to 90 minutes, to exclude individuals with so-called “mega-commutes”. Reported one-way commutes are multiplied by 2 to determine the daily round-trip commute. Labor income is deflated by annual CPI in the UK from the Office for National Statistics. I exclude workers who are self-employed. The results presented in this section correspond to the unweighted sample; estimates using longitudinal weights are quantitatively similar and can be found in Appendix A.

Table 1 reports the marginal effects from a probit regression in which the dependent variable is equal to 1 if the individual made at least one job-to-job transition in the past year and 0 if she remained employed in the same job according to her self-reported employment history. The first column regresses the job-to-job transition probability on the commute and wage in the previous year and region, time, commute method, industry, and occupation fixed effects. The second column repeats the regression including a large set of individual characteristics. The estimates suggest that the commute in year \( t - 1 \) has a positive relationship with the probability of a job-to-job transition between years \( t - 1 \) and \( t \) even after controlling for the wage in \( t - 1 \). The results in column 2 can

---

9Data used in Figure 2 is the 2001 Census: Special Workplace Statistics (Level 1), available through the UK Data Service’s WICID. For each pair of locations \( i \) and \( j \), the flow data contains the number of individuals whose usual residence is in \( i \) and usual place of work is in \( j \).

10LADs are a unit of regional classification, with populations ranging from 2,300 to 1 million, and land area between 1 and 1,936 square miles.

11For instance, a Google Maps search for directions from Kensington to the City of London returns directions leaving from 112 Kensington High St, Kensington, London and arriving at 21 Bloomberg Arcade, London.

12For similar figures with distance in terms of commuting time and the commuters as a share of the LAD population, see Appendix A.

13About 5% of LADs in the sample have fewer than 50 observations, results when including these LADs are similar and are available upon request.

14I include workers reporting a commute of 0 minutes to take into consideration people working from home. Restricting the sample to those reporting a strictly positive commute or allowing for workers with wages outside 3 standard deviations of the mean does not significantly change the results.

15The Census Bureau defines a “mega-commuter” as a worker traveling more than 90 minutes to work, one-way, see Rapino and Fields (2013). In the BHPS data, these individuals represent less than 0.6% of the sample.

16I define a job-to-job transition as one with no more than 30 days of nonemployment between two employment spells. Results are similar if the nonemployment duration is limited to 0 days between the two spells.
Table 1: Marginal Effects of Lagged Commute and Wage on Job-to-Job Transition Probability

<table>
<thead>
<tr>
<th></th>
<th>J2J_t</th>
<th>J2J_t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commute_{t-1} × 30</td>
<td>0.015***</td>
<td>0.011***</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Real log Wage_{t-1}</td>
<td>-0.035***</td>
<td>-0.041***</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.008)</td>
</tr>
</tbody>
</table>

Individual Characteristics

| Region, Time, Commute Method | ✓ |
| Industry & Occ FE            | ✓ |
| Pseudo R²                    | .067 | .124 |
| N obs                        | 7,310 | 6,528 |

Notes: BHPS Sample 1992-2008, annual. Universe: respondents living in LADs within 100 miles of the City of London aged 24-55 working full-time in year t. Estimated marginal effects from a probit regression of J2J_t, which is a dummy equal to one if a worker made a job-to-job transition in the past year and 0 if she remained in the same job, on commute time and wages in the previous year. Individual characteristics include the annual regional house price index, a quadratic term in labor market experience, age, education, marital status, and number of children, 1-year lagged tenure, 1-year lagged dummies for outright homeownership, mortgage holding, whether the individual moved in the past year, real housing expenditures, whether the worker was unemployed in the past year, the number of employment spells, whether the spouse or partner was employed last year, a government job dummy, and union status. All regressions include region, month, year, 1-digit industry and occupation fixed effects. Robust standard errors are reported in parentheses. * denotes p < .1, ** p < .05, and *** p < .01.

be interpreted as follows: a 30 minute increase in last year’s commute is associated with about a 1 percentage point increase in the probability of making a job-to-job transition in the current year.

The results in Table 1 potentially suffer from endogeneity as they may be driven by unobservable heterogeneity in workers: those who commute more today could be more diligent workers, leading them to be more likely to switch jobs. Conversely, it is possible that the workers who commute more and earn lower wages are less motivated and therefore may switch jobs more often because they are more likely to be unsuitable for any job. To consider this possibility, I run similar regressions using a linear probability model with individual fixed effects. The results are similar and can be found in Appendix A.

Table 2 shows that decreases in commuting time explain the prevalence of wage cuts following job-to-job transitions. To construct the table, I compare wages and commutes in the years before and after a job-to-job transition took place. “Up” and “Down” indicate differences from the last reported wage or commute of more than 5%, and “Same” indicates differences less than 5%. Each column of the table shows, for a given accepted wage which was higher, lower, or the same as her previous wage, the share of workers whose commute was higher, lower, or the same as her previous commute. N indicates the total number of wage cuts, no change, or wage increases. If commuting were irrelevant in workers’ decisions to make job-to-job transitions, we would expect the share of

\[ \text{median wage cut in the sample is 6.3\% of the real monthly wage (£116) and the median commuting decrease in the sample is 33\% of commuting time (20 minutes).} \]
workers taking commuting increases and decreases would be equal for each column.

<table>
<thead>
<tr>
<th>Commute</th>
<th>Wage Down</th>
<th>Wage Same</th>
<th>Wage Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Down</td>
<td>0.50</td>
<td>0.44</td>
<td>0.34</td>
</tr>
<tr>
<td>Same</td>
<td>0.19</td>
<td>0.18</td>
<td>0.25</td>
</tr>
<tr>
<td>Up</td>
<td>0.31</td>
<td>0.38</td>
<td>0.41</td>
</tr>
<tr>
<td>N</td>
<td>157</td>
<td>103</td>
<td>437</td>
</tr>
</tbody>
</table>

Notes: BHPS Sample 1992-2008, annual. Universe: respondents living in LADs within 100 miles of the City of London aged 24-55 working full-time, who changed jobs since year \( t - 1 \). “Up” and “Down” indicate differences from the last reported wage or commute of more than 5%, and “Same” indicates differences less than 5%.

Models with renegotiation such as Cahuc et al. (2006) generate wage cuts when workers move to more productive jobs with a high future value. Workers are willing to take wage cuts because they expect to extract more of the surplus through future renegotiation of their wages. Other models explain wage cuts as the result of an unobserved advance notice of a separation: when workers know that their job will be destroyed, they are forced to accept any offer that is better than unemployment. This paper follows the literature on non-wage amenities and argues that wage cuts can be due to increases in the non-wage value of the job. In general, these amenities are difficult to measure or are subjective (e.g. measures of job satisfaction). One benefit of studying commuting using the BHPS is the availability of time use data on workers’ daily commutes. Assuming that the amenity value is monotonic in the time spent commuting, the data provides an objective measure of the commuting amenity.

The table shows that half of all workers reporting wage cuts in the year in which a job-to-job transition takes place also report decreases in their commuting time relative to the previous year. The relationship is roughly symmetric: a larger share of wage increases are accompanied by increases in commuting time. Thus, the wage-commuting trade-off can help to explain the existence of wage cuts in the data. Appendix A contains similar tables showing that this pattern is robust: the share of workers taking wage cuts and commuting cuts is stable across genders, occupation switchers, residential movers, and marital status. The results are also robust to defining a commuting change in terms of minutes rather than percent and restricting the definition of a job-to-job transition to have zero days of nonemployment between employment spells.

**Congestion**

Finally, I show that congestion is important for commuting time by comparing travel time at rush hour and on weekends. Though I do not consider mode of transportation here, which clearly affects travel time, the regressions using BHPS data presented above and in Appendix A do control for workers’ commuting method as well as time. Because the BHPS does not contain information on
firms’ locations, I use Google Maps to estimate the underlying relationship between travel time from each LAD to the City of London, which I will consider to be the city center when estimating commuting costs in the model. To do so, I compute the minimum commuting time required to arrive by 9am on an arbitrary weekday and Saturday from the location selected by Google for each LAD as described above, using Google Maps’ recommended travel mode option. Figure 3 shows results for each of the LADs in the Census Flow data for which a positive number of residents commute to the City, with a distance of at most 100 miles. The difference between the observation and the 45 degree line is the additional time required to arrive in the City if traveling on a weekday relative to traveling at the same time of day on the weekend. The data shows that on average, it takes 50% more time to travel into the City during periods of high congestion.

3 Model

I build a spatial random search model where commuting costs are an important determinant of match surplus. The goal of the model presented here is to consider how workers’ employment patterns and wages are affected by the spatial configuration of the labor market and how externalities in commuting shape equilibrium outcomes. I take rent in each location as given and assume that land supply is perfectly elastic. In Section 4 I assume land is scarce and introduce a land market.
through which rent is endogenously determined.

### 3.1 Set-up

Time is continuous and agents are infinitely lived. The economy is populated by a continuum of workers of measure one and a continuum of firms of a positive measure. A firm in the model is simply a job, and I will use these terms interchangeably. Workers exit the labor market at rate $\chi > 0$, and new workers enter at the same rate. All agents share a common discount rate $\rho$ and are risk neutral.

There is a single labor market within which all workers and firms are located. The labor market is defined by $2N + 1$ locations, $\ell \in \mathcal{L} \equiv \{1, \ldots, N + 1, \ldots 2N + 1\}$, with location $N + 1$ defining the city center. Each worker and each filled job is located in a “residence” of size 1. For simplicity, all residences are assumed to be rented. Both workers and filled jobs pay rent in their location, $r(\ell)$, which is exogenous.

Workers differ in their location $\ell$ and may be employed or unemployed. Unemployed (employed) workers draw job offers at rate $\lambda_0$ ($\lambda_1$). A job offer is a pair $(\ell_F, y)$ drawn from exogenous distribution $G$, where $\ell_F \in \mathcal{L}$ is the location of the job and $y$ is the match-specific productivity drawn from a finite set $Y$. Thus, search is random in both the spatial and productivity dimensions. Both $\ell_F$ and $y$ are fixed for the duration of the match. Firms have access to a constant returns to scale technology where employed workers supply their labor inelastically and match output is given by $y$. Employed workers receive a wage $w$ and unemployed workers receive their value of home production, $b$.

In order to earn her wage, an employed worker must pay an out-of-pocket flow commuting cost $T(\ell, \ell_F, \Omega)$, increasing in the commuting distance and travel time. Commuting does not affect the output of a match; it is assumed that firms are rigid about the hours that workers must be present at the firm (full-time), and that commuting therefore provides disutility to the worker but does not directly affect output. The worker’s commuting distance $d(\ell, \ell_F)$ is determined by the location of the worker’s residence and of her workplace, $\ell$ and $\ell_F$, respectively. Travel time is affected by the extent of the congestion externality, which is determined by the distribution of workers $\Omega$ and is discussed in detail in Section 3.5.

A worker may choose a new location in which to live by paying a fixed moving cost $k_M$. To allow workers to move for reasons other than the job, employed workers face moving shocks at rate $\varphi$, which trigger a draw of the worker’s location from exogenous distribution $\Pi(\ell)$, where $\pi(\ell)$ denotes the probability that the worker draws $\ell \in \mathcal{L}$.\(^{18}\)

When a worker is employed, she consumes her wage net of commuting and rental costs. Matches are destroyed exogenously at rate $\delta$. When a match is destroyed, the worker becomes unemployed and the firm exits. Unemployed workers consume the flow value of home production less rental

\(^{18}\)For simplicity I assume that unemployed workers cannot change their location. It is straightforward to allow for this extension; with moving costs of the magnitude calibrated below this assumption is insignificant for the quantitative results.
costs. All workers have the outside option of leaving the labor market and consuming $\bar{u}$, which is normalized to zero.

### 3.2 Wage Determination

Employment contracts consist of a wage and worker location, $(w, \ell_e)$. Wages are renegotiated with the agreement of both parties, following Dey and Flinn (2005) and Cahuc et al. (2006). The outcome of this game is identical to the generalized Nash-bargaining solution, where the worker’s bargaining power is given by $\beta$. Similar to search intensity in Bagger and Lentz (Forthcoming), I assume that the worker’s location is contractible at the time of matching, to ensure that the equilibrium can be characterized independently of the split of the surplus between firm and worker.\(^{19}\) This assumption is further discussed at the end of the next section.

For a given location $\ell$, define the value of unemployment as $U(\ell)$. For a given firm location and productivity $(\ell_F, y)$, worker location $\ell$, and wage $w$, denote the value of an employed worker as $W(w, \ell_F, y, \ell)$. This value takes into account the worker’s contracted location $\ell_e$:

$$W(w, \ell_F, y, \ell) = \tilde{W}(w, \ell_F, y, \ell_e) - k_M \mathbb{1}\{\ell \neq \ell_e\} \quad (1)$$

where variables with a tilde denote those taking the worker’s location as given. Similarly, the value of a filled job is

$$J(w, \ell_F, y, \ell) = \tilde{J}(w, \ell_F, y, \ell_e) \quad (2)$$

The match surplus is defined as

$$S(\ell_F, y, \ell) = \max_{\ell' \in \mathcal{L}} \{ \tilde{M}(\ell_F, y, \ell') - k_M \mathbb{1}\{\ell' \neq \ell\} \} - U(\ell) \quad (3)$$

where $\tilde{M}(\ell_F, y, \ell) = \tilde{W}(w, \ell_F, y, \ell) + \tilde{J}(w, \ell_F, y, \ell)$, with $\tilde{W}$ and $\tilde{J}$ defined below. The location in the worker’s contract is given by

$$\ell_e(\ell_F, y, \ell) = \arg \max_{\ell' \in \mathcal{L}} \{ \tilde{M}(\ell_F, y, \ell') - k_M \mathbb{1}\{\ell' \neq \ell\} \}.$$  

I assume that the surplus-maximizing location is unique for each $(\ell_F, y, \ell)$.\(^{20}\) In the case of moving shocks, the worker cannot immediately return to her surplus-maximizing location. Define the match surplus after drawing $\ell'$ following a moving shock as $\tilde{S}(\ell_F, y, \ell') = \tilde{M}(\ell_F, y, \ell') - U(\ell')$. Following a moving shock, the worker’s location may change only after making a job-to-job transition or receiving another moving shock. Because the worker cannot immediately move following a moving shock, and the match productivity and firm location are fixed, the worker will only move for job-

---

\(^{19}\)The lack of tractability when workers are able to choose their location independently comes from the fact that the worker’s location choice affects the probability of renegotiation, whereas the location choice maximizing the match value does not.

\(^{20}\)In the quantitative model I allow for indifference by splitting workers evenly across the locations to which they are indifferent.
related reasons when she accepts a new job, thus, if a worker is continuing in a match I denote the surplus as \( \tilde{S}(\ell_F, y, \ell) \), and the workers’ value as \( \tilde{W}(w, \ell_F, y, \ell) \).

When a firm meets an unemployed worker, her wage \( \phi_0(y, \ell_F, \ell) \) is set to satisfy
\[
W(\phi_0, \ell_F, y, \ell) - U(\ell) = \beta S(\ell_F, y, \ell)
\]
(4)

Henceforth, I will drop the arguments of the wage functions where possible. I use a prime to denote the values of a job offer (e.g. \((y', \ell'_F)\)). When a worker is employed, she may renegotiate her wage upon the arrival of an outside offer. Suppose a worker’s current surplus is given by \( S(\ell_F, y, \ell) \) and associated worker value is \( W(w, \ell_F, y, \ell) \). If she is contacted by a firm with which her surplus would be \( S(\ell'_F, y', \ell) \), there are three possibilities:

1. \( \tilde{S}(\ell_F, y, \ell) < S(\ell'_F, y', \ell) \)
2. \( \tilde{W}(w, \ell_F, y, \ell) - U(\ell) < S(\ell'_F, y', \ell) < \tilde{S}(\ell_F, y, \ell) \)
3. \( S(\ell'_F, y', \ell) < \tilde{W}(w, \ell_F, y, \ell) - U(\ell) < \tilde{S}(\ell_F, y, \ell) \)

In case (1), the worker will be poached by the new firm, with wage \( w = \phi_1(y', \ell'_F, y, \ell_F, \ell) \) satisfying:
\[
W(\phi_1, \ell'_F, y', \ell) - U(\ell) = \tilde{S}(\ell_F, y, \ell) + \beta S(\ell'_F, y', \ell) - \tilde{S}(\ell_F, y, \ell)
\]
(5)

In this case, the firms enter into Bertrand competition, and the worker extracts the full surplus of her previous match and a share \( \beta \) of the gains in surplus between the poaching firm and her previous employer. Note that the new job may require the worker to move and therefore has no tilde. Following the terminology of Postel-Vinay and Turon (2010), the surplus \( \tilde{S}(\ell_F, y, \ell) \) becomes the worker’s new “negotiation baseline.” In case (2), the worker remains at her current firm, but renegotiates her wage to \( w = \phi_2(y, \ell_F, y', \ell'_F, \ell) \) with the outside offer becoming her negotiation baseline:
\[
\tilde{W}(\phi_2, \ell_F, y, \ell) - U(\ell) = S(\ell'_F, y', \ell)
\]
(6)

Finally, in case (3), the outside offer is too low to warrant renegotiation between the worker and her current firm, therefore she remains at the same wage in her current match.

Similarly to Postel-Vinay and Turon (2010), the arrival of a moving shock may trigger a renegotiation due to a change in the worker or firm’s values. Consider an employed worker in a match \((\ell_F, y)\), who was previously living at \( \ell \) and draws \( \ell' \) upon the arrival of the shock. The worker cannot immediately return to her previous location following the shock, thus, the worker and firm will remain matched if \( \tilde{S}(\ell_F, y, \ell') > 0 \). She can renegotiate using her new location, thus, her value of employment is given by \( \tilde{W}(w, \ell_F, y, \ell') \).

There are four possibilities:

(a) \( \tilde{S}(\ell_F, y, \ell') < 0 \)
(b) \( \tilde{S}(\ell_F, y, \ell') > 0 \), and \( \tilde{W}(w, \ell_F, y, \ell') - U(\ell') < 0 \)
(c) \( \tilde{S}(\ell_F, y, \ell') > 0 \), and \( \tilde{W}(w, \ell_F, y, \ell') - U(\ell') > \tilde{S}(\ell_F, y, \ell') \)

(d) \( \tilde{S}(\ell_F, y, \ell') > 0 \), \( \tilde{W}(w, \ell_F, y, \ell') - U(\ell') > 0 \) and \( \tilde{W}(w, \ell_F, y, \ell') - U(\ell') \leq \tilde{S}(\ell_F, y, \ell') \)

In case (a), the match is endogenously destroyed due to a negative surplus at the worker’s new location. In case (b), the worker’s surplus is negative under the old wage and is reset to make the worker indifferent between remaining in the match or quitting into unemployment, \( w = \psi_1(\ell_F, y, \ell') \) such that

\[
\tilde{W}(\psi_1, \ell_F, y, \ell') - U(\ell') = 0
\]  

(7)

In this case, the firm gets the full surplus. Conversely, in case (c), the wage is reset to \( w = \psi_2(\ell_F, y, \ell') \) to make the firm indifferent between continuing in the match or exiting. Equivalently, the worker gets the full surplus

\[
\tilde{W}(\psi_2, \ell_F, y, \ell') - U(\ell') = \tilde{S}(\ell_F, y, \ell')
\]  

(8)

Finally, in case (d), both the worker and firm have positive surplus under the new realization of the worker’s location, and therefore the match will continue with no change in the wage.

### 3.3 Value Functions

The flow value for an unemployed worker who consumes home production \( b \) and receives job offers at rate \( \lambda_0 \) is

\[
(\rho + \chi)U(\ell) = b - r(\ell) + \lambda_0 \sum_{y', \ell_F'} \max\{0, \beta S(\ell_F', y', \ell')\} g(\ell_F', y')
\]  

(9)

At rate \( \lambda_0 \) the worker meets a vacancy drawn from distribution \( G \). If the job is acceptable, the worker receives a share \( \beta \) of the surplus.

Next, consider an employed worker living in \( \ell \) in current firm \( (y, \ell_F) \) earning wage \( w \). Define the sets for which a worker makes a job-to-job transition and renegotiates, respectively, as \( B_1(\ell_F, y, \ell) = \{(\ell_F', y') \in \mathcal{L} \times \mathcal{Y} : S(\ell_F', y', \ell) > \tilde{S}(\ell_F, y, \ell)\} \), and \( B_2(w, \ell_F, y, \ell) = \{(\ell_F', y') \in \mathcal{L} \times \mathcal{Y} : \tilde{W}(w, \ell_F, y, \ell) - U(\ell) < S(\ell_F', y', \ell) \leq \tilde{S}(\ell_F, y, \ell)\} \). The flow value for a worker living in a location \( \ell \), in current firm \( (\ell_F, y) \) earning wage \( w \), is given by:

\[
(\rho + \chi)\tilde{W}(w, \ell_F, y, \ell) = w - r(\ell) - T(\ell, \ell_F, \Omega) + (\delta + \varphi)(U(\ell) - \tilde{W}(w, \ell_F, y, \ell)) \\
+ \varphi \sum_{\ell' \in \mathcal{L}} \left[ \max\{\tilde{W}(w, \ell_F, y, \ell') - U(\ell'), \tilde{S}(\ell_F, y, \ell')\}, 0\right] + U(\ell') - U(\ell) \right] \pi(\ell') \\
+ \lambda_1 \sum_{(y', \ell_F') \in B_1(y, \ell_F, \ell)} (\beta S(\ell_F', y', \ell) + (1 - \beta) \tilde{S}(\ell_F, y, \ell) - \tilde{W}(w, \ell_F, y, \ell) + U(\ell)) g(\ell_F', y') \\
+ \lambda_1 \sum_{(y', \ell_F') \in B_2(w, y, \ell_F, \ell)} (S(\ell_F', y', \ell) - \tilde{W}(w, \ell_F, y, \ell) + U(\ell)) g(\ell_F', y')
\]  

(10)

Where the value function \( W \), taking into account the location for this worker is given by (1). I verify below that this location is independent of the worker’s current wage.
The first three terms on the right hand side are the wage, rent, and commuting cost. The worker exogenously separates from her match at rate $\delta$. At rate $\varphi$ the worker experiences a moving shock and may separate into unemployment, renegotiate her wage, or remain at her old wage. The last two lines report the payoffs to the worker when an outside offer arrives. The worker separates if the surplus of the new offer is strictly larger than her current surplus, that is, $(\ell_F', y') \in B_1(\ell_F, y, \ell)$. The worker and firm renegotiate if the surplus of the outside offer exceeds the worker’s current negotiation baseline but does not exceed her current surplus, that is, $(\ell_F', y') \in B_2(w, \ell_F, y, \ell)$.

The value of a continuing filled job is given by

\[
(\rho + \chi)\bar{J}(w, \ell_F, y, \ell) = y - w - r(\ell_F) - (\delta + \varphi)\bar{J}(w, \ell_F, y, \ell) + \varphi \sum_{\ell' \in \mathcal{L}} \max\{\min\{\bar{J}(w, \ell_F, y, \ell'), \bar{S}(\ell_F, y, \ell')\}, 0\} \pi(\ell') - \lambda_1 \bar{J}(w, \ell_F, y, \ell) \sum_{(y', \ell_F') \in B_1(y, \ell_F, \ell)} g(\ell_F', y') + \lambda_1 \sum_{(y', \ell_F') \in B_2(w, y, \ell_F, \ell)} \left(\bar{S}(\ell_F, y, \ell) - S(\ell_F', y', \ell) - \bar{J}(w, \ell_F, y, \ell)\right) g(\ell_F', y')
\]

where the first three terms are the match output is $y$, the wage that the firm pays its worker, and the firm’s rent $\ell_F$. At rate $\delta$ the match is destroyed exogenously, and the value of the job if the worker separates is zero, since the job is destroyed when the match ends. At rate $\varphi$, the worker receives a moving shock and may may separate into unemployment, renegotiate her wage, or remain at her old wage. The last term on the second line is the loss to the firm if the match is destroyed when the worker makes a job-to-job transition, and the third line shows the payoff if the worker and firm renegotiate.

Combining these expressions gives the match value:

\[
M(\ell_F, y, \ell) = \max_{\ell' \in \mathcal{L}} \{\bar{M}(\ell_F, y, \ell') - k_M 1\{\ell \neq \ell'\}\}
\]

where $\bar{M}(\ell_F, y, \ell) = \bar{J}(w, \ell_F, y, \ell) + \bar{W}(w, \ell_F, y, \ell)$, with:

\[
(\rho + \chi)\bar{M}(\ell_F, y, \ell) = y - r(\ell) - r(\ell_F) - T(\ell, \ell_F, \Omega) - (\delta + \varphi)\bar{S}(\ell_F, y, \ell) + \varphi \sum_{\ell' \in \mathcal{L}} \left(\max\{0, \bar{S}(\ell_F, y, \ell')\} + U(\ell') - U(\ell)\right) \pi(\ell') + \lambda_1 \beta \sum \max\{0, S(\ell_F', y', \ell) - \bar{S}(\ell_F, y, \ell)\} g(\ell_F', y')
\]

where the associated policy function for the worker’s location is denoted $\ell_e(\ell_F, y, \ell)$. When an outside offer arrives and the worker is able to change location, the match surplus is given by $S(\ell_F, y, \ell) = M(\ell_F, y, \ell) - U(\ell)$. After a moving shock, when the worker cannot move, the surplus that is used to determine whether to renegotiate or continue in the match is given by $\bar{S}(\ell_F, y, \ell) = \bar{M}(\ell_F, y, \ell) - \bar{W}(\ell_F, y, \ell)$. 

18
Discussion on Contractability of Location

Wages in this model are determined by the current and next-best match surplus. Thus, otherwise identical workers are compensated for their commutes through wages, a seemingly strong assumption. The main reason for using this wage-setting mechanism is to facilitate the analysis with externalities and endogenous rent prices below. Although one may argue that wages offered to new hires should not vary by commuting cost, in this model with renegotiation firms will compensate their worker in order to avoid her being poached by an outside offer. Thus, observed wages after at least one renegotiation will vary for otherwise identical workers due to differences in threat points reflecting heterogeneous commuting costs.

In addition, many contracts include a “mobility clause” defining the maximum distance the worker is expected to live from the workplace. The reason for this assumption here is technical, since the worker’s location affects the match value through commuting and rent costs, but the worker’s choice of location will generally not coincide with the match value-maximizing choice. This is because the worker’s location affects the set of offers resulting in wage renegotiation. It is beyond the scope of this paper to consider the incentive problems related to the interaction between wage contracts and workers’ location choices.

3.4 Worker Flows

This section defines the flows into and out of the worker distributions across employment states and space. Denote the mass of employed workers currently in a match \((\ell_F, y)\) living at \(\ell\) as \(e(\ell_F, y, \ell)\) and the mass of unemployed workers living at \(\ell\) as \(u(\ell)\). The distribution of workers is summarized by \(\Omega\).

The flow into employment for workers living at \(\ell\) employed in matches located at \(\ell_F\) with firm productivity \(y\) is made up of three groups. The first is the unemployed who match with a firm \((\ell_F, y)\). The second is the employed making job-to-job transitions, who were previously in a match with \(y' \neq y\) and/or \(\ell'_F \neq \ell_F\). The final group are those employed workers who arrive in \(\ell\) after receiving a moving shock, but remain employed with their firm \((y, \ell_F)\). In sum, the flow into such matches is given by:

\[
e^+(\ell_F, y, \ell) = g(\ell_F, y) \sum_{\ell' \in \mathcal{L}} \lambda_0 u(\ell') \mathbb{1}\{S(\ell_F, y, \ell') > 0\} \\
+ \lambda_1 \sum_{\ell'_F \in \mathcal{L}} \sum_{y' \in \mathcal{Y}} \mathbb{1}\{S(\ell'_F, y', \ell') < S(\ell_F, y, \ell')\} e(\ell'_F, y', \ell') \mathbb{1}\{\ell_e(\ell_F, y, \ell') = \ell\} \\
+ \varphi \pi(\ell) \mathbb{1}\{S(\ell_F, y, \ell) > 0\} \sum_{\ell' \neq \ell} e(\ell_F, y, \ell')
\]  

Similarly, the flow out of such matches is given by the employed previously in matches located at \(\ell_F\) with firm productivity \(y\), who separate either exogenously or endogenously by making a
job-to-job transition or after receiving a moving shock:

\[
e^{-\ell_F, y, \ell} = e(\ell_F, y, \ell) \left[ \chi + \delta + \varphi (1 - \pi(\ell)) + \lambda_1 \sum_{\ell' \in \mathcal{L}} \sum_{y' \in Y} 1 \left\{ S(\ell'_F, y', \ell) > \tilde{S}(\ell_F, y, \ell) \right\} g(\ell'_F, y') \right] \tag{15}
\]

The flow into of unemployment of workers living at \( \ell \) consists of those employed workers who separate exogenously following a separation shock \( \delta \) or endogenously following a moving shock \( \varphi \), and newborn workers who draw location \( \ell \) with probability \( \mu(\ell) \):

\[
u^+(\ell) = \chi \mu(\ell) + \delta \sum_{\ell' \in \mathcal{L}} \sum_{y \in Y} e(\ell_F, y, \ell) + \varphi \pi(\ell) \sum_{\ell' \in \mathcal{L}} \sum_{y \in Y} e(\ell_F, y, \ell') 1 \left\{ \tilde{S}(\ell_F, y, \ell) < 0 \right\} \tag{16}
\]

The flow out of unemployment for workers living at \( \ell \) is simply those unemployed workers who successfully match or exit the economy:

\[
u^-(\ell) = u(\ell) \left[ \chi + \lambda_0 \sum_{\ell' \in \mathcal{L}} \sum_{y \in Y} 1 \left\{ S(\ell_F, y, \ell) > 0 \right\} g(\ell_F, y) \right] \tag{17}
\]

Turning to the flows across space, the flow of workers and active firms into a given location \( \tilde{\ell} \) is comprised of the flow of newly filled jobs located at \( \tilde{\ell} \), those newly hired workers not previously living at \( \tilde{\ell} \) whose contract specifies that they live at \( \tilde{\ell} \), employed workers who draw \( \tilde{\ell} \) upon arrival of moving shock, and the newly unemployed who are located at \( \tilde{\ell} \). This inflow is given by:

\[
\sum_{y \in Y} \left( \sum_{\ell' \in \mathcal{L}} e^+(\tilde{\ell}, y, \ell) + \sum_{\ell' \in \mathcal{L}} e^+(\ell_F, y, \tilde{\ell}) - g(\ell_F, y) \right) \left[ \lambda_0 u(\tilde{\ell}) 1 \left\{ S(\ell_F, y, \tilde{\ell}) > 0 \right\} \right.
+ \lambda_1 \sum_{\ell' \in \mathcal{L}} \sum_{y' \in Y} 1 \left\{ \tilde{S}(\ell'_F, y', \tilde{\ell}) > S(\ell_F, y, \tilde{\ell}) \right\} e(\ell_F, y', \tilde{\ell}) \left. \right\} 1 \{ \ell_e(\ell_F, y, \tilde{\ell}) = \tilde{\ell} \} + u^+(\tilde{\ell}) - \delta \sum_{\ell' \in \mathcal{L}} \sum_{y \in Y} e(\ell_F, y, \tilde{\ell}) \right) \tag{18}
\]

Similarly, the flow of workers and firms out of location \( \tilde{\ell} \) is made up of the firms whose matches were located at \( \tilde{\ell} \) and are destroyed, plus the employed and unemployed who lived at \( \tilde{\ell} \) and moved following a change in their employment state. This flow is given by:

\[
\sum_{y \in Y} \left( \sum_{\ell \in \mathcal{L}} e^-(\tilde{\ell}, y, \ell) + \sum_{\ell' \in \mathcal{L}} e^-(\ell_F, y, \tilde{\ell}) \right.
- e(\ell_F, y, \tilde{\ell}) \left[ \delta + \lambda_1 \sum_{\ell' \in \mathcal{L}} \sum_{y' \in Y} 1 \left\{ S(\ell'_F, y', \tilde{\ell}) > \tilde{S}(\ell_F, y, \tilde{\ell}) \right\} 1 \{ \ell_e(\ell_F, y', \tilde{\ell}) = \tilde{\ell} \} g(\ell'_F, y') \right) \left. \right\} \right) + u^-(\tilde{\ell}) - \lambda_0 u(\tilde{\ell}) \sum_{\ell' \in \mathcal{L}} \sum_{y \in Y} 1 \left\{ S(\ell_F, y, \tilde{\ell}) > 0 \right\} 1 \{ \ell_e(\ell_F, y, \tilde{\ell}) = \tilde{\ell} \} g(\ell_F, y) \tag{19}
\]
3.5 Congestion

Recall that \(d(\ell, \ell_F)\) is the distance in miles between a worker located in \(\ell\) and a firm located in \(\ell_F\). Since locations are discrete, I assume that all workers experience a positive commute even if their firm is in the same location, that is, \(d(\ell, \ell) > 0\).

Similar to Brinkman (2016), I define congestion as the share of commuters passing through a location less the number of workers living in that location. I consider all of the pairs of locations \((\ell, \ell_F)\) and compute the “commuting path,” defined as the locations through which a worker living at \(\ell\) must travel in order to get to \(\ell_F\). For instance, to travel from location 2 to location N, a worker must travel through \(N - 1\) locations. Define the commuting path as \(cp(\ell_F, \ell)\) where I impose symmetry: \(cp(\ell_F, \ell) = cp(\ell, \ell_F)\), where \(cp(\ell, \ell_F) = \{\ell' \in \mathcal{L} : \ell \leq \ell' \leq \ell_F \text{ if } \ell \leq \ell_F \text{ and } \ell_F \leq \ell \text{ otherwise}\}\). Congestion in a given location \(\tilde{\ell}\) is given by the difference between commuters and residents:

\[
c(\tilde{\ell}) = \sum_y \sum_{(\ell_F, \ell) \in cp(\ell, \ell_F)} [c(\ell_F, y, \ell) + e(\ell, y, \ell_F)] - \left(\sum_{\ell_F} e(\ell_F, y, \ell) + u(\ell)\right) - \left(\sum_{\ell_F} e(\ell_F, y, \ell) + u(\ell)\right)
\]

(20)

Note that a worker who lives and works in the same location does not add to congestion; similarly, the unemployed decrease congestion in the location in which they live.

The worker’s marginal commuting cost is assumed to take the form

\[
\tau + \kappa \left(\sum_{i \in cp(\ell, \ell_F)} c(i)\right)
\]

where \(\tau\) is the “no traffic” cost of commuting an extra unit of distance and \(\kappa\) measures the congestion cost. The summation computes all of the congestion that a worker encounters by commuting between \(\ell\) and \(\ell_F\); in this way commuting costs depend on the distribution \(\Omega\). In the case where there is no externality, \(\kappa = 0\).

The total cost of commuting between \(\ell\) and \(\ell_F\) is given by

\[
T(\ell, \ell_F, \Omega) = \left(\tau + \kappa \left(\sum_{i \in cp(\ell, \ell_F)} c(i)\right)\right) d(\ell, \ell_F) b
\]

where the cost of time to the worker is proportional to the opportunity cost of time, given by the value of home production \(b\). Before defining the equilibrium, in the next section I extend the model to be used for the quantitative results.

3.6 Efficiency

I use the quantitative model below to evaluate the welfare implications of congestion targeted policies, including working remotely, while taking into account the dynamic nature of job search, workers’ location decisions, and congestion. To develop intuition for cases when the congestion externality is inefficient, in Appendix C.1 I show in a simple static framework that the decentralized
equilibrium need not align with that of a social planner when reducing the commute is costly. The social value of reducing the commute for a large mass of workers includes the net benefits for all remaining commuters from less congestion plus the output gain from any additional match creation. If working remotely implies a small output loss, individual matches may find it privately optimal to commute, failing to take into account their effect on congestion. I show conditions under which is is socially beneficial but privately suboptimal to use the remote working technology.

4 Quantitative Model

I incorporate a land market with endogenous supply to take into account how rent prices reflect congestion and labor market frictions. Each location $\ell \in L$ contains total land $L$. Both workers and firms with filled jobs pay rent, $r(\ell)$, which is determined endogenously. Land is endowed to a continuum of homogeneous landlords in each location. Since workers can move at any time but the locations of all filled jobs are fixed, it is useful to consider the stock of available land, which I denote $\hat{L}(\ell)$, equal to the total land less that occupied by the existing filled jobs:

$$\hat{L}(\ell) = \frac{L(\ell)}{\text{total land}} - \sum_{y} \sum_{\tilde{\ell}} \left( e(\ell, y, \tilde{\ell}) - e^-(\ell, y, \tilde{\ell}) \right)$$

The representative landlord maximizes her utility by choosing how much of the available land to consume, denoted $\zeta(\ell)$. Her maximization problem is written

$$\max_{\zeta(\ell) \in [0, \hat{L}(\ell)]} \left\{ \nu(\zeta(\ell), r(\ell)[L - \zeta(\ell)]) \right\}$$

where the first argument is the amount of land the landlord consumes, and the second argument is the amount of resources she gets from renting the remaining land $L - \zeta(\ell)$ at price $r(\ell)$. I assume that $\nu$ is twice continuously differentiable and strictly increasing in both of its arguments. The strictly positive derivative with respect to the first argument insures that the landlord will not rent any of her land at a price of zero.

Given the land occupied by existing jobs, land demand is determined by the total mass of employed and unemployed workers residing in $\ell$ and newly filled jobs located in $\ell$. Equilibrium in the land market requires the supply of land be equal to the demand of for land:

$$L - \zeta(\ell) = u(\ell) + \sum_{y \in Y} \sum_{\ell \in L} e(\ell, y, \ell) + e(\ell, \ell, \ell)$$

I now define a steady state equilibrium.

**Definition 1.** A steady state equilibrium consists of value functions $M : L \times Y \times L \rightarrow \mathbb{R}$ for
each \( \ell_F \in \mathcal{L} \) and \( U : \mathcal{L} \to \mathbb{R} \), a policy function \( \ell_e(\ell_F, y, \ell) \in \mathcal{L} \), wage functions when employed \( \phi_1 : \mathcal{L} \times Y \times \mathcal{L} \times Y \times \mathcal{L} \to \mathbb{R}_+ \), \( \phi_2 : \mathcal{L} \times Y \times \mathcal{L} \times Y \times \mathcal{L} \to \mathbb{R}_+ \), \( \psi_1 : \mathcal{L} \times Y \times \mathcal{L} \to \mathbb{R}_+ \) and \( \psi_2 : \mathcal{L} \times Y \times \mathcal{L} \to \mathbb{R}_+ \), a wage function when unemployed, \( \phi_0 : \mathcal{L} \times Y \times \mathcal{L} \to \mathbb{R}_+ \), a rent function \( r : \mathcal{L} \to \mathbb{R}^{2N+1}_+ \), landlords’ consumption \( \zeta : \mathcal{L} \to \mathbb{R}^{2N+1}_+ \), and distributions of workers across employment states, and of workers and firms across space such that given the commuting externality \( c : \mathcal{L} \to \mathbb{R}^{2N+1}_+ \),

(i) For each \((\ell_F, y, \ell) \in Y \times \mathcal{L} \times \mathcal{L}\), \( M(\ell_F, y, \ell) \) satisfies (13) and \( \ell_e(\ell_F, y, \ell) \) is the associated policy function. For each \( \ell \in \mathcal{L} \), \( U(\ell) \) satisfies (9).

(ii) When an outside offer arrives, wages \( \phi_1(y', \ell'_F, y, \ell_F, \ell) \) and \( \phi_2(y, \ell_F, y', \ell'_F, \ell) \) are determined by the surplus splitting equations (5) and (6). When an offer arrives to the unemployed worker, the wage \( \phi_0(y, \ell_F, \ell) \) is determined by (4). When a moving shock is realized, wages \( \psi_1(\ell_F, y, \ell') \) and \( \psi_2(\ell_F, y, \ell') \) are determined by (7) and (8).

(iii) For each \((\ell_F, y, \ell) \in Y \times \mathcal{L} \times \mathcal{L}\), the distributions across employment states satisfy \( e^+(\ell_F, y, \ell) = e^- (\ell_F, y, \ell) \) and \( u^+(\ell) = u^- (\ell) \), given by (14)-(17). The distributions across space equate (18) and (19).

(iv) The commuting externality \( c \) is consistent with the distributions and given by (20).

(v) For each \( \ell \in \mathcal{L} \), \( \zeta(\ell) \) satisfies (21) and rent \( r(\ell) \) adjusts such that (22) holds.

I restrict attention to the steady state for tractability. In the steady state, equating the flows of employed and unemployed workers determines \( e(\ell_F, y, \ell) \) and \( u(\ell) \). Under the assumption that vacant firms do not occupy space and thus do not enter into the flows across space, imposing steady state in the labor market implies the spatial steady state, that is, the equality of (18) and (19).

Existence

The steady state distributions \( \Omega = (e, u) \) are pinned down by the surplus value function \( S \) and policy functions determining workers’ locations. Proposition 1 states that there exists a bounded price vector \( r \) satisfying the conditions of a steady state equilibrium.

Proposition 1. A steady state price vector \( r \) exists.

Proving existence of the equilibrium for the model with endogenous rent is nontrivial due to the dependence of the distributions of workers and firms on the rent price. The complication arises here because the rent price function \( r \) affects steady state distributions of workers across firms and locations, and these distributions must also be consistent with the land market clearing condition. The demand correspondence that takes prices and returns a vector of demands corresponding to the right hand side of (22) is clearly nonempty. However, it is not clear that the correspondence is upper hemicontinuous nor convex-valued, since the distributions of vacancies and workers may not
respond continuously to changes in price and agents can choose only one discrete location in which to reside.

For simplicity, I assume that there is no free entry of vacancies, but extending the model to allow for this is straightforward as long as firms draw, rather than choose, their location. If vacancies were able to choose their location freely, all would be posted in the location with the highest surplus, violating continuity. Regarding convexity, I use results from Kaneko and Yamamoto (1986), who show that an equilibrium price vector exists in an economy where agents may choose only one of many indivisible goods by first solving a convex analog to the problem, and then showing that any solution to such a problem is also a solution to the original problem.

The proof follows and extends the arguments presented in Kaneko and Yamamoto (1986). It differs from the previous literature because not only do the demands for land change by worker type, but also the mass of each type of worker and of filled jobs, that is, the worker’s employment status and current match, responds to changes in rent prices. Intuitively, when types are fixed, workers and firms are permanently matched. An increase in rent near the most productive jobs will make the surplus of these matches fall, and workers will respond by trading off their commuting and rent costs. When workers and firms can choose who to match with, an increase in rent will not only cause changes in residential patterns of workers in existing firms, but also change the set of acceptable matches and the level of employment. This causes the distributions of workers and firms, match surplus, and wages all to change in response to rent prices. The proof thus hinges on showing that the steady state distributions respond continuously to the price. This mechanism by which spillovers exist between labor and rental markets is a fundamental feature of the model as it is tightly linked to the level of congestion.

4.1 Calibration

The model is calibrated at a monthly frequency. The number of match-specific productivities is set to 5. The local labor market is divided into 7 locations defined by the round trip commuting distance from the center, shown in Figure 4. As discussed in Section 2, I calibrate the model to match the metropolitan area defined by those LADs within 100 miles (one-way) of the City of London. An alternative definition of the local labor market defined by the “commuting zone” around the City, which consists of those locations for which over half of all workers commute within the zone’s boundaries, is discussed in Appendix D. Given two locations \( \ell \) and \( \ell_F \), the distance in miles is \( d(\ell, \ell_F) = |\ell - \ell_F| + 101 \{\ell = \ell_F\} \). In words, this functional form for distance implies that the round trip distance between the worker and firm in the same location \( \ell \) is 10 miles.

Figure 4: Locations
Regarding the other distributional assumptions, the location and productivity of vacancies is drawn from the joint distribution $G$. The location distribution is assumed to be symmetric around the center, where the probability of drawing a location, from left to right, is given by $[g(1), g(2), g(3), g(4), g(3), g(2), g(1)]$, with $g(4) + 2 \sum_{i=1}^{3} g(i) = 1$. The distribution for $y$ is conditional on the draw of the firm location $\ell_F$ and drawn from a Normal distribution with mean $1 + A(\ell_F)$ and standard deviation $\sigma_y$. Finally, the location distribution for newborn workers ($\mu$) is equal to the distribution after moving shocks ($\pi$) and is chosen to match the spatial distribution in the BHPS for workers 24 years old with less than 2 years of labor market experience.

Appendix E describes the numerical solution algorithm. I first discuss the standard parameters of the model and then turn to those particular to the spatial dimension and commuting.

4.1.1 Standard Parameters

Two parameters are calibrated exogenously. The discount rate $\rho$ is set to match an annual interest rate of 5% and the death rate $\chi$ is chosen to match an average working lifetime\textsuperscript{22} of 31 years. The remaining parameters, including those in the next section, are calibrated jointly using the Simulated Method of Moments. Given a set of parameter values, I solve for the steady state of the model and simulate 3,000 worker histories over 3,000 months, discarding the first 300 months. Because solving the model is computationally intensive, I do not compute standard errors. The standard labor market parameter values and their targets are shown in Tables 3 and 4, respectively. I next discuss the relationship between the targeted moments and structural parameters.

The parameters related to transitions are $\delta$, $\varphi$, $\lambda_0$, $\lambda_1$, and $\beta$. The separation rate into unemployment (EU rate) is driven by exogenous separations, occurring at rate $\delta$, and endogenous separations which may occur after the arrival of a moving shock $\varphi$. The unemployment to employment (UE) rate is mainly driven by the arrival rate of offers for the unemployed, $\lambda_0$. The job-to-job transition (EE) rate is primarily driven by $\lambda_1$. Bargaining power $\beta$ is chosen to match the ratio of the average wage for new entrants relative to the average wage for all employed workers. Because $\varphi$ is a spatial parameter, I leave its discussion to the next section.

The standard deviation of productivity, $\sigma_y$, is chosen to match the variance of the observed log wage distribution.\textsuperscript{23} Since the mean of $y$ is location-specific, it is discussed in the next section. The home production value $b$ is chosen to match a replacement rate of 0.71 following Hall and Milgrom (2008), defined as the flow benefits of unemployment relative to the average wage.\textsuperscript{24}

4.1.2 Spatial Parameters

The probabilities $g(1), g(2),$ and $g(3)$ are chosen to match the average jobs density for LADs 5-33, 33-66, and 66-100 miles (one way) from the City of London as defined by Google Maps and discussed

\textsuperscript{22}The working lifetime corresponds to ages 24-55, the same range considered in the empirical results of Section 2.

\textsuperscript{23}Wages used to compute the calibration moments are in real terms and in logs, after subtracting the year, month, education level, and commuting method effects, see Appendix E.

\textsuperscript{24}The benefits of unemployment include the commuting costs saved from not working, for a discussion of a similar point see Bils et al. (2012).
Table 3: Standard Parameter Values

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discount Rate ( \rho )</td>
<td>.0041</td>
</tr>
<tr>
<td>Exit rate ( \chi )</td>
<td>.0027</td>
</tr>
<tr>
<td>Bargaining Power ( \beta )</td>
<td>.42</td>
</tr>
<tr>
<td>Home Production ( b )</td>
<td>.83</td>
</tr>
<tr>
<td>Separation Rate ( \delta )</td>
<td>.003</td>
</tr>
<tr>
<td>Arrival Rate, Unemployed ( \lambda_0 )</td>
<td>.11</td>
</tr>
<tr>
<td>Arrival Rate, Employed ( \lambda_1 )</td>
<td>.09</td>
</tr>
<tr>
<td>Standard Deviation, Productivity ( \sigma_y )</td>
<td>.02</td>
</tr>
</tbody>
</table>

Table 4: Targets: Standard Parameters

<table>
<thead>
<tr>
<th>Moment</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual interest rate</td>
<td>.05</td>
<td>.05</td>
</tr>
<tr>
<td>Years in LF</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>Relative wage, new hires from U</td>
<td>.87</td>
<td>.92</td>
</tr>
<tr>
<td>Replacement rate (Hall and Milgrom 2008)</td>
<td>.71</td>
<td>.68</td>
</tr>
<tr>
<td>EU rate (Q, Gomes 2012)</td>
<td>.014</td>
<td>.013</td>
</tr>
<tr>
<td>UE rate (Q, Gomes 2012)</td>
<td>.26</td>
<td>.23</td>
</tr>
<tr>
<td>EE rate (Q, Gomes 2012)</td>
<td>.027</td>
<td>.025</td>
</tr>
<tr>
<td>Variance, ( \ln(w) )</td>
<td>.191</td>
<td>.197</td>
</tr>
</tbody>
</table>

in Section 2. Jobs density is defined as the ratio of the number of jobs to resident working age population. The probability of drawing a vacancy in the city center, \( g(4) \) is given by \( 1 - 2 \sum_{i=1}^{3} g(i) \).

Returning to the conditional mean for match-specific productivity \( y \), the value of \( A \) in the center is normalized to one, and is symmetric around the center. The remaining values for \( A \) are chosen to match the average wage for workers residing in LADs 5-33, 33-66, and 66-100 miles from the City of London relative to the average wage for workers residing within 10 miles of the City. The parameter values imply that output is much higher on average in the center: the level of \( A \) in the center is almost 50% higher than in the periphery. Lower rent in the periphery offsets the effect of lower production, increasing wages and creating the need for large productivity differences across locations to match relative wages as a function of distance. Brinkman (2016) estimates agglomeration effects implying a 40% productivity loss at a radius 100 miles from the city center, here the estimate for \( A(1) \) implies an average productivity loss of 15%. In my model, workers have an option value of remaining unemployed, there is a trade-off between low \( A \) and high unemployment.

The choice of \( \varphi \) targets the share of job-related moves: in the data, workers state that the main reason for moving was for employment reasons just 12.9% of the time. The remaining share of moves is attributed to moving shocks in the model. The calibrated value for \( \varphi \) implies a non-job related move once every 30 years, or about one time between the ages of 24 and 55. Given the
rate of exogenous moves, the annual average moving rate pins down the moving cost $k_M$. This moving cost is equivalent to roughly 8 years of average wages, in line with the literature: Kennan and Walker (2011) estimate that the cost of moving house in the US is as high as $300,000.

Landlords’ preferences are assumed to be CES:

$$
\nu(\zeta(\ell), r(\ell)(L - \zeta(\ell))) = \left[ \omega_L \zeta(\ell)^{\frac{\sigma_L-1}{\sigma_L}} + (1 - \omega_L)(r(\ell)(L - \zeta(\ell)))^{\frac{\sigma_L-1}{\sigma_L}} \right]^{\frac{\sigma_L}{\sigma_L-1}} 
$$

where $\omega_L \in (0, 1)$ is the weight on the landlord’s consumption and $\sigma_L$ is the elasticity of substitution between consumption and rental income. The parameter $\omega_L$ is chosen to match an average monthly rent to wage income ratio and $\sigma_L$ is chosen to match the price elasticity of housing supply in the UK (Barker 2003). The endowment of land $L$ is constant and chosen such that there is enough land in the economy for the maximum mass of workers and filled jobs.\(^{25}\)

Table 5: Spatial Parameter Values

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moving cost</td>
<td>$k_M$</td>
</tr>
<tr>
<td>No-traffic commuting cost</td>
<td>$\tau$</td>
</tr>
<tr>
<td>Congestion Cost</td>
<td>$\kappa$</td>
</tr>
<tr>
<td>Moving rate</td>
<td>$\varphi$</td>
</tr>
<tr>
<td>Landlords’ preferences</td>
<td>$\omega_L$</td>
</tr>
<tr>
<td>Landlords’ preferences</td>
<td>$\sigma_L$</td>
</tr>
<tr>
<td>Location-specific productivity, 10-33 mi</td>
<td>$A(3)$</td>
</tr>
<tr>
<td>Location-specific productivity, 34-66 mi</td>
<td>$A(2)$</td>
</tr>
<tr>
<td>Location-specific productivity, 67-100 mi</td>
<td>$A(1)$</td>
</tr>
<tr>
<td>Vacancy location probability, 10-33 mi</td>
<td>$g(3)$</td>
</tr>
<tr>
<td>Vacancy location probability, 34-66 mi</td>
<td>$g(2)$</td>
</tr>
<tr>
<td>Vacancy location probability, 67-100 mi</td>
<td>$g(1)$</td>
</tr>
</tbody>
</table>

Recall that the commuting cost function is

$$
T(\ell, \ell_F, \Omega) = \left( \tau + \kappa \left( \sum_{i \in \mathcal{O}(\ell, \ell_F)} c(i) \right) \right) d(\ell, \ell_F)b
$$

The parameter $\tau$ determines the marginal disutility from commuting in the absence of the externality (the “no traffic” cost) and $\kappa$ is the congestion cost. Together, these two parameters determine the share of job offers rejected because of the commute and the average increase in commuting costs with congestion. The share of job offers rejected because of the commute is the share of offers with higher productivity that a worker rejects; in the data this corresponds to the

\(^{25}\)As long as $L$ is constant, a change in $L$ would be equivalent to a rescaling of the parameter $\omega_L$.\)
question in the New York Federal Reserve’s Survey of Consumer Expectations regarding the main reason for rejecting a job offer. To pin down $\kappa$, I use the average commuting time with and without congestion from Google Maps shown in Figure 3. For all LADs considered, the average increase in time to commute to the City of London on a weekday morning is 65% higher than on a weekend. Assuming that commuting costs are proportional to time, this moment identifies the congestion cost, $\kappa$.

4.2 Properties of the Calibrated Economy

In this section I evaluate the model’s performance against a set of untargeted moments in the data. Table 7 compares several untargeted moments in the model to the data. The share of wage cuts is equal to the proportion of job to job transitions resulting in a wage at least 5% lower than the worker’s previous wage. The model predicts this share to be far higher than in the data, because there are two mechanisms giving rise to wage cuts. First, workers are willing to take lower wages today when the continuation value of the match is much higher than their previous job, similarly to Cahuc et al. (2006). Second, workers face a trade-off between productivity and commuting, and will take wage cuts when job offers are located at a shorter commuting distance. In the model, the share of movers who remain in the same job is purely due to the moving shock $\varphi$. This moment is lower in the model than in the data, possibly because respondents to the survey mis-report the share of job-related moves. The next two lines show wage growth between and within jobs in the data and in the model. The model captures within-job wage growth well, but wage growth between jobs is lower than in the data because of the large share of wage cuts that the model predicts.

In the data, the share of workers residing in the periphery (locations more than 66 miles one-way
Table 7: Untargeted Moments

<table>
<thead>
<tr>
<th>Moment</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share of wage cuts</td>
<td>.16</td>
<td>.34</td>
</tr>
<tr>
<td>Share of movers who stay in the same job</td>
<td>.018</td>
<td>.009</td>
</tr>
<tr>
<td>Average wage growth, J2J</td>
<td>.065</td>
<td>.018</td>
</tr>
<tr>
<td>Average wage growth, within job</td>
<td>.016</td>
<td>.015</td>
</tr>
<tr>
<td>Annual J2J transition rate, residence in center</td>
<td>.147</td>
<td>.063</td>
</tr>
<tr>
<td>Annual J2J transition rate, residence in periphery</td>
<td>.105</td>
<td>.110</td>
</tr>
<tr>
<td>Relative Moving Distance, J2J to All Movers</td>
<td>1.39</td>
<td>1.41</td>
</tr>
</tbody>
</table>

from the City of London) making job-to-job transitions is less than for workers living in the center (within 10 miles from the City). In the model, job-to-job transitions in the center are lower than in the periphery because workers who live in the periphery receive relatively many job offers close to home where congestion is low, leading to more frequently make transitions. The last line in Table 7 shows the moving distance for workers making job-to-job transitions relative to all movers in the model and data. Movers in the data are workers moving within the local labor market (those who appear in the sample the year prior to the move). Although the model predicts much larger distances due to the discrete number of locations, it reproduces the longer moving distance for job related reasons because workers will move for a job when the net gain in surplus is larger than the moving cost. Since this cost is fixed, workers will tend to move farther for job related reasons.

Figure 5: Commuting Distance by Place of Residence

Figure 5 shows the commuting patterns by place of residence in the model and data. The horizontal axis corresponds to the distance from the place of residence, and the vertical axis shows the share of commuters living in each location and commuting a given distance. In all four locations, the model matches the data well, both in terms of the share of workers living and working in the same location as well as the farther the commuting distance, the lower is the share of commuters.

Figure 6 shows the spatial distribution of workers and firms and of firm productivity. The model successfully captures the declining density of jobs in distance from the center, and the larger share
of the highest productivity jobs in the center. The exogenous distribution of vacancies has just 10% of offers in the center; more filled jobs are located in the center because of the higher average output through the location-specific productivity distribution. The increase in productivity more than offsets the higher congestion that workers must face if they need to commute to these jobs. The rent price by location is shown in Figure 12 in Appendix F. In the calibration, I target the average rent to income, and the average wage by location, but these do not determine the slope of the rent price as a function of distance (the “rent gradient”). As seen in the figure, the model successfully captures the gradient.

Figure 6: Workers and Firms

4.3 Effect of Congestion

This section examines the effect of commuting and congestion on labor market outcomes. Figure 7 shows the possible outside offers for a worker living in a location 33 miles from the center working for a firm with average productivity in the center. The vertical axis shows the change in surplus from an offer in each location with low, medium, and high productivity. The worker accepts all offers above the horizontal line, which have a higher surplus than her current job. When productivity is the only match-specific factor driving job mobility, workers accept all offers with higher productivity. The figure shows the trade-off between the commute and productivity: when the firm location is closer to her own, the worker accepts offers with productivity that is lower than that in her current match. Conversely, when the firm is sufficiently far away, offers with higher productivity give little gain in terms of surplus. The worker uses some offers that have a lower commuting cost to renegotiate her wage, shown by the shaded region in the figure. Moving decisions are shown for workers living in various locations in Appendix F.

I compare the quantitative model to one without an externality by setting $\kappa = 0$. Commuting costs in the no externality model are therefore given by $T(\ell, \ell_F) = \tau d(\ell, \ell_F)b$. Figure 8 compares aggregate outcomes in the model with and without the externality. The top row of the figure
shows the changes in average commuting costs and average wages by worker location, and total output by firm location. In the absence of congestion, commuting costs increase for all workers as the marginal cost of commuting is constant. For workers in living 33 miles from the center, who face the most congestion in the full model, commuting costs rise the most without the externality. Because congestion and productivity are highest near the center, in the full model workers receiving jobs far from home must decide between commuting through the congested area or paying the large moving cost to relocate to the location of their job. Without congestion, the moving rate falls, and workers commute farther distances. Aggregate output produced in the center increases and falls in the periphery due to a shift in the allocation of workers to firms in the absence of congestion towards firms in the center.

In the second row of Figure 8, I present the job-to-job transition rate, employment, and rent. In the full model, the job-to-job transition rate near the center is lower because workers are less willing to accept low productivity matches when congestion is high. In the periphery, the job-to-job transition rate falls in both models because this is the location with the lowest rent, increasing the value of the match despite the lower productivity. This is reflected in the higher share of workers living and working in the same location in the periphery relative to other locations in the city. Employment slightly increases for firms 33 miles from the center, where congestion is high in the full model, and slightly decreases near the periphery, where congestion was lowest. Without the congestion effect, more jobs are created in locations through which many workers commute, and it is less attractive to create jobs far from the center because there is no trade-off between congestion and productivity. Finally, rent is almost identical in the two models, because it is driven by location-specific productivity, which is left unchanged.
Figure 8: Aggregate Patterns Across Space

Figure 9: Equilibrium Congestion

Figure 9 shows the average commute for workers living in each location, decomposed into congestion and distance costs. The distance cost simply sets the cost of congestion $\kappa = 0$. Workers who travel through more than one location experience the sum of all congestion in the path between their residence and the location of their firm. On average, commuting costs are 8% higher with congestion, but this varies widely depending on the worker’s location. The average commuting cost
for workers living in the center is lower than workers living just outside the center because a smaller share of workers in the center commute outside of the center to work. Conversely, more workers living just outside the center commute into the center to work, traveling a longer distance and thus having higher commuting costs.

### 4.4 Wage and Utility Dispersion

To understand how wage and utility dispersion are affected by the congestion externality, I compute wage dispersion, defined as the standard deviation of the log of the monthly wage, and utility dispersion, defined as log wages net of commuting costs, or as log wages net of commuting and rent costs. Table 8 contains the results, comparing the model with and without the externality and the model with zero commuting costs. As in the previous section, the model without the externality corresponds to setting $\kappa = 0$.

Comparing the first two rows, wage dispersion is 25% higher in the model with the externality. Wages reflect the level of congestion through two channels: first, workers who have long commutes have higher costs due to congestion and therefore have low flow utility today. This decreases the worker’s surplus as well as that for the match, but increases the worker’s wage because of the out-of-pocket cost she must pay to commute. Second, congestion affects workers’ future job outcomes through the set of offers they are willing to accept through on-the-job search.

Before taking into account differences in housing rent, utility dispersion is more than twice wage dispersion, and is higher in the no congestion model. After taking rent into account utility dispersion is 11% higher in the full model relative to the no congestion model. It is important to take into account differences in rent for workers' utility. A standard prediction of models in urban economics is that workers will trade-off rent and commuting costs and therefore will be indifferent across locations in the absence of frictions (e.g. Fujita 1989), I therefore focus on the results in the last column. Because workers with lower rent live in the periphery and tend to commute farther, rent and commuting costs have offsetting effects. With the congestion externality, the strength of this relationship differs depending on the worker’s location: workers in the center have high rents, but face more congestion to commute a short distance, raising their marginal cost of commuting. Since dispersion in both wages and commuting costs is higher because of heterogeneity in the commuting cost per mile, utility dispersion may be larger in the presence of congestion. The model predicts that the increase in dispersion as measured by utility relative to wages is lower in the full model (13%) than in the model without congestion (27%).

<table>
<thead>
<tr>
<th>Model</th>
<th>$\ln w$</th>
<th>$\ln w - \ln T(\ell, \ell_F)$</th>
<th>$\ln w - \ln (T(\ell, \ell_F) + r(\ell))$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full</td>
<td>.446</td>
<td>.988</td>
<td>.504</td>
</tr>
<tr>
<td>No Congestion</td>
<td>.358</td>
<td>1.08</td>
<td>.455</td>
</tr>
<tr>
<td>No Commuting Cost</td>
<td>.081</td>
<td>-</td>
<td>.171</td>
</tr>
</tbody>
</table>
5 Counterfactual Exercises: Remote Working

I perform several counterfactuals to study the effects of remote working on congestion and labor market outcomes. I impose that workers in jobs targeted by the policy work from home a share $h$ of the time. When working remotely, workers incur only a share $h$ of their commuting costs and their contribution to congestion decreases proportionally. In each exercise, I compute the welfare gains, and changes in employment, congestion, and wage and utility dispersion. I consider several ways of targeting the remote working policy. The next section presents the results when the policy is available to the firms with the highest productivity in each location. The following two sections examine the effects of the policy when it is targeted to firms located in the city center, or to matches in which the worker lives at least 100 miles from the firm.

Empirical measures of productivity for remote workers relative to office workers are rare. Bloom et al. (2014) find that remote workers from a call center in China saw their productivity, measured as the number of calls made, increase by 13%. White (Forthcoming) finds that the wages of remote workers in the US in 2014 were 5% higher than those of office workers, and exhibited the same amount of dispersion. In all of these studies, remote working is voluntary; here, I impose the policy on all jobs to compute the extent of the possible welfare gains. I assume that by working remotely, a match produces $(1 - hp)y_H$, where $p \in [0, 1]$ is the productivity loss. I vary the productivity loss in the following sections, comparing the cases between $p = 0\%$ and $p = 10\%$.

5.1 All High Productivity Firms

In this section the remote working policy is used by firms in each location that have the highest productivity, corresponding to two standard deviations above the location-specific mean. Table 9 contains the results. The first row contains the results for the full model with no remote working (baseline). The next three rows show the results when workers are remote 1 day per week ($h = 0.2$). When there is no productivity loss, a larger share of the highest productivity jobs are created relative to the baseline, and the inflow to unemployment falls because moving shocks no longer cause the highest productivity matches to separate. However, the outflow rate also falls because the unemployed are more efficient at job search and the option value of waiting for a remote working job increases. Wage dispersion and utility dispersion remain roughly constant, and average wages increase because fewer workers move house in order to accept a job, raising the continuation value for all matches which is reflected in the wage.

Congestion falls when more workers accept the remote working technology, but because these workers represent a small share of the population, other workers in lower productivity matches...
commute more, offsetting much of the effect of remote working on congestion. Average annual wage growth is only mildly affected, since most jobs offers do not have the highest productivity, and the indirect effect on the surplus of other offers and therefore on wages is small. Finally, welfare gains are highest when there is no productivity loss associated with remote working: for remote working 1 day per week, the welfare gains are 1.7%, and for 4 days per week the gains increase to 7.7%. As long as the productivity gains are small (less than 10%), more remote working leads to higher welfare when all high productivity firms have access to the technology. When productivity losses are large, fewer high productivity matches are created, leading to lower welfare than in the baseline. When this loss is 10% and workers must work remotely 4 days per week, the model predicts that these losses can be as large as 7.5%.

Table 9: Effect of Remote Working, All High Productivity Firms

<table>
<thead>
<tr>
<th>Experiment</th>
<th>UE</th>
<th>EE</th>
<th>σ(Wage)</th>
<th>σ(Utility)</th>
<th>Average Congestion</th>
<th>Wage Growth</th>
<th>PDV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>.232</td>
<td>.025</td>
<td>.446</td>
<td>.504</td>
<td>.055</td>
<td>1.41%</td>
<td>48.1</td>
</tr>
<tr>
<td>1 day/wk, p = 0</td>
<td>.237</td>
<td>.025</td>
<td>.445</td>
<td>.508</td>
<td>.051</td>
<td>1.52%</td>
<td>48.4</td>
</tr>
<tr>
<td>1 day/wk, p = 0.01</td>
<td>.237</td>
<td>.025</td>
<td>.433</td>
<td>.496</td>
<td>.051</td>
<td>1.51%</td>
<td>48.3</td>
</tr>
<tr>
<td>1 day/wk, p = 0.1</td>
<td>.236</td>
<td>.025</td>
<td>.442</td>
<td>.501</td>
<td>.053</td>
<td>1.45%</td>
<td>47.6</td>
</tr>
<tr>
<td>4 days/wk, p = 0</td>
<td>.351</td>
<td>.011</td>
<td>.142</td>
<td>.315</td>
<td>-.042</td>
<td>1.17%</td>
<td>36.2</td>
</tr>
<tr>
<td>4 days/wk, p = 0.01</td>
<td>.352</td>
<td>.012</td>
<td>.130</td>
<td>.307</td>
<td>-.041</td>
<td>1.12%</td>
<td>33.9</td>
</tr>
<tr>
<td>4 days/wk, p = 0.1</td>
<td>.232</td>
<td>.029</td>
<td>.405</td>
<td>.487</td>
<td>.029</td>
<td>1.56%</td>
<td>43.2</td>
</tr>
</tbody>
</table>

5.2 High Productivity Firms, Center Only

In this section the remote working policy is used by firms in the city center that have the highest productivity. In this way, the remote working policy is targeted to the most productive jobs that also require workers to commute through the most congested area, and therefore can have large effects relative to the size of the population that is targeted.

Relative to when remote working is used by all high productivity firms, the policy targeted toward firms in the center generally has positive, but generally lower, welfare gains. When remote working is only 1 day per week, a large productivity loss can still lead to welfare gains because workers’ rankings of jobs are less affected. In the previous section, if the productivity loss is large and all high productivity firms allow for remote working, it has an effect on workers’ rankings of jobs across all locations. To minimize commuting costs without working remotely, it is most attractive to live and work in the same location, therefore many workers will choose lower productivity matches in their own location rather than the highest productivity firms in any location. Instead, when only firms in the center lose productivity with remote working, workers still prefer to work in their own location.
location, but the most productive firms in all locations except the center have no productivity loss, resulting in higher welfare. In this way, productivity losses for some firms can lead to higher welfare through workers’ dynamic sorting into other firms. Welfare is higher than in the model without remote working because even with a productivity loss, there are fewer workers commuting into the city center because some jobs do not require it. This decreases congestion in the center relative to the periphery, and makes workers commuting to less productive jobs in the center better off.

When only the firms in the center are targeted, there is a welfare loss of 1.3% if the productivity loss and share of remote working time is large, but otherwise the policy results in welfare gains, ranging from 0.4% to 3.3%. Though these gains are less than the largest gain from applying the policy to firms regardless of location discussed in the previous section, the welfare losses in the case of a large productivity loss are far smaller.

Table 10: Effect of Remote Working, High Productivity Firms in Center

<table>
<thead>
<tr>
<th>Experiment</th>
<th>UE</th>
<th>EE</th>
<th>σ(Wage)</th>
<th>σ(Utility)</th>
<th>Average Utility</th>
<th>Wage Growth</th>
<th>PDV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>.232</td>
<td>.025</td>
<td>.446</td>
<td>.504</td>
<td>.055</td>
<td>1.41%</td>
<td>48.1</td>
</tr>
<tr>
<td>1 day/wk, p = 0</td>
<td>.232</td>
<td>.026</td>
<td>.450</td>
<td>.510</td>
<td>.058</td>
<td>1.51%</td>
<td>48.1</td>
</tr>
<tr>
<td>1 day/wk, p = 0.01</td>
<td>.232</td>
<td>.026</td>
<td>.440</td>
<td>.501</td>
<td>.056</td>
<td>1.47%</td>
<td>48.1</td>
</tr>
<tr>
<td>1 day/wk, p = 0.1</td>
<td>.232</td>
<td>.025</td>
<td>.445</td>
<td>.504</td>
<td>.058</td>
<td>1.43%</td>
<td>48.1</td>
</tr>
<tr>
<td>4 days/wk, p = 0</td>
<td>.235</td>
<td>.023</td>
<td>.447</td>
<td>.528</td>
<td>.041</td>
<td>1.67%</td>
<td>48.5</td>
</tr>
<tr>
<td>4 days/wk, p = 0.01</td>
<td>.235</td>
<td>.023</td>
<td>.447</td>
<td>.528</td>
<td>.043</td>
<td>1.65%</td>
<td>48.4</td>
</tr>
<tr>
<td>4 days/wk, p = 0.1</td>
<td>.234</td>
<td>.026</td>
<td>.447</td>
<td>.516</td>
<td>.052</td>
<td>1.49%</td>
<td>47.7</td>
</tr>
</tbody>
</table>

5.3 High Productivity Firms, Workers With Long Commutes

In the final remote working experiment, I consider the welfare implications when specific matches, rather than firms, are targeted. In particular, the policy is implemented when a worker meets a vacancy that is at least 100 miles (3 locations) away. The welfare gains from this policy are almost always as large as those for all firms, and larger than those targeting only firms in the center. This is because by allowing workers living far from their firms to work remotely part of the time, workers are more likely to accept high productivity offers without having to pay the large moving cost, which significantly increases match surplus and therefore welfare. The benefits to working from home a larger share of the time are largest in this case, because the cost of commuting over 100 miles is significant, especially if the workers can work remotely only one day per week.

When there is little productivity loss and workers work remotely 4 days per week, average congestion significantly increases because of the large decline in the moving rate, leading to longer commutes for non-remote workers through the city center. Because only workers commuting over
100 miles have access to remote working, and there is little productivity loss, workers are willing to take long commutes 1 day per week in order to work from home the remaining 4 days. Commuting costs across the city increase by up to 20%, despite the fact that workers are commuting only 20% of the time.

Table 11: Effect of Remote Working, High Productivity Firms with Commute ≥ 100 mi

<table>
<thead>
<tr>
<th>Experiment</th>
<th>UE</th>
<th>EE</th>
<th>σ(Wage)</th>
<th>σ(Utility)</th>
<th>Average Congestion</th>
<th>Wage Growth</th>
<th>PDV</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>.232</td>
<td>.025</td>
<td>.446</td>
<td>.504</td>
<td>.055</td>
<td>1.41%</td>
<td>48.1</td>
<td></td>
</tr>
<tr>
<td>1 day/wk, p = 0</td>
<td>.244</td>
<td>.025</td>
<td>.448</td>
<td>.506</td>
<td>.058</td>
<td>1.51%</td>
<td>48.8</td>
<td></td>
</tr>
<tr>
<td>1 day/wk, p = 0.01</td>
<td>.244</td>
<td>.025</td>
<td>.449</td>
<td>.511</td>
<td>.058</td>
<td>1.58%</td>
<td>48.6</td>
<td></td>
</tr>
<tr>
<td>1 day/wk, p = 0.1</td>
<td>.236</td>
<td>.025</td>
<td>.438</td>
<td>.499</td>
<td>.056</td>
<td>1.52%</td>
<td>47.6</td>
<td></td>
</tr>
<tr>
<td>4 days/wk, p = 0</td>
<td>.238</td>
<td>.018</td>
<td>.390</td>
<td>.474</td>
<td>.087</td>
<td>1.76%</td>
<td>51.9</td>
<td></td>
</tr>
<tr>
<td>4 days/wk, p = 0.01</td>
<td>.238</td>
<td>.019</td>
<td>.354</td>
<td>.475</td>
<td>.084</td>
<td>1.67%</td>
<td>50.1</td>
<td></td>
</tr>
<tr>
<td>4 days/wk, p = 0.1</td>
<td>.232</td>
<td>.027</td>
<td>.437</td>
<td>.497</td>
<td>.058</td>
<td>1.52%</td>
<td>45.5</td>
<td></td>
</tr>
</tbody>
</table>

5.4 Summary of Results

The remote working policy can greatly increase welfare if productivity losses are small. Assuming these losses are known ex ante, the welfare gains are largest when there is no welfare loss or only a small loss (1%) by allowing those workers with the longest commutes to work from home as much as possible. However, to achieve the same gains, congestion on average increases as long as these workers must commute at least some of the time. If policymakers’ objective is to decrease congestion, they may therefore reject policies with the largest welfare gains, and instead prefer to apply the policy to all high productivity firms regardless of the commutes of their workers or the firms’ locations. When productivity losses are small, applying remote working to all firms can achieve substantial welfare gains, as well as lower unemployment, wage and utility dispersion, and congestion.

When the loss is large (10%), there can be welfare gains from assigning the highest productivity firms in the city center to have remote work, because of the way in which lower congestion in the city center affects workers’ job acceptance decisions. Applying a blanket remote working policy regardless of the location of firm or worker always has smaller welfare gains than a targeted policy, either to the location with the most congestion or the workers with the longest commutes.

Regarding the implementation of the remote working policy, with no productivity loss, remote working increases match surplus and therefore is always optimal. When there is some productivity loss, a subsidy could be given to firms according to their location and output, or according to the distance between the worker and firm.
6 Conclusion

In this paper I argue that commuting congestion and labor market outcomes are inherently linked in urban areas. I develop a model that can address this link and that allows for endogenous wages and housing prices, both of which are key in the determination of the spatial configuration of workers and firms. Workers’ commuting paths across space contribute to congestion, which has important effects on individual workers’ search strategies, and on the dispersion in wages and utility across workers. Adopting a frequently used bargaining protocol, the model remains tractable but rich in its ability to replicate features of urban labor markets. I show several empirical patterns linking the commute to current and future labor market outcomes for individuals. The model can directly speak to crucial questions faced by policymakers and urban planners, and introduces important trade-offs to evaluate policies related to infrastructure and remote working arrangements. If the objective is to maximize employment and reduce congestion, untargeted remote working policies to all of the highest productivity firms throughout the city achieve the largest welfare gains when productivity losses from remote working are small.

One important feature absent from this model is the labor-demand response to spatial frictions, which I rule out by assuming an exogenous location-productivity distribution for vacancies. Allowing firms to choose their locations is an important avenue for future research. Relatedly, I do not explore how entry is affected by the presence of the amenity or the externality associated with it. Another important topic that I leave to future work is to understand how the conclusions of this model are affected by alternative mechanisms for job search and bargaining. Spatially-directed job search has been explored for the unemployed by Marinescu and Rathelot (2018) and Manning and Petrongolo (2017) among others. It is important to understand how spatial mismatch affects workers’ progress up the job ladder when search is directed or when wages are posted rather than the result of bargaining. Finally, I do not consider rationing of land; an interesting exercise would be to examine the effects of policies such as rent control in the present model.

References


White, D. R. (Forthcoming). Agency theory and work from home. *LABOUR*. 

41
A Data Description and Micro-Level Regressions

In this section I first describe the data and main variables used in the quantitative analysis. I then document the robustness of the results in Section 2, and then run a series of regressions to show the importance of commuting for job-to-job transitions.

A.1 Data

The data used in Section 2 and the remainder of this section come from three sources (1) the British Household Panel Survey (BHPS), (2) the Census flow data, and (3) Google Maps. This section discusses details and use of each source.

A.1.1 BHPS

For each wave of the survey, I merge data from the household questionnaire (hhresp), the full and proxy questionnaires (indresp), employment history (jobhist), relationship between household members (egoalt), residential locations at the local authority district level (oslaua_protect), and House price statistics for small areas in England (HPSSA) from the ONS. The records are linked using the household and person ID (hid and pid), and then years are linked using the person ID, keeping individuals appearing in at least 2 years of the survey. All nominal variables are deflated using annual CPI in the UK from the ONS.

Transitions across labor market states are identified from the employment history. The employment history contains information from September 1 of the previous year through the date of the interview. A job-to-job transition in year $t$ is defined as a worker who was employed at the date of the interviews in year $t$ and $t-1$, who reports a different job in the job history between the two interviews with no more than 30 days of nonemployment between the previous and current job. Unemployment-to-employment transitions are identified by workers who are employed in year $t$ in the first job following an unemployment spell reported at the date of the interview in $t-1$.

Wages are given by the CPI-deflated usual net pay per month ($paym$). Commuting time ($jbttwt$) is defined as the answer to “About how much time does it usually take for you to get to work each day, door to door (in minutes)?” Commuting method is the answer to “And what usually is your main means of travel to work?”, with possible responses: British rail, train; Underground, tube, metro; Bus, minibus or coach (public or private); Motor cycle, scooter, moped; Driving a car or van; Passenger in car or van; Pedal cycle; On foot (walks all the way); Other.

A.1.2 Census Flow Data

Responses from the decennial Census on usual place of residence and usual place of work are aggregated to compute the flow tables. For each pair of locations (A,B), the Flow tables contain the number of individuals originating in A and finishing in B, either for residential moves or daily commutes. The data used are the Special Workplace Statistics (Level 1), which contain commuting flows for each LAD. For consistency with the time span of the BHPS sample, the 2001 Census is
used to construct Figure 2. I consider all LADs of residence for which a positive number of workers commute to the City of London LAD ("the City"), plotted on the vertical axis.

A.1.3 Google Maps

Using Google Maps, I collect data on the shortest driving distance from each LAD to the City, as well as the recommended travel mode to arrive at the destination by 9am on an arbitrary weekday (July 19, 2018). I select those LADs for which the Census Flow Data reports a positive number of individuals whose usual place of work is the City (see above). The recommended travel mode is the fastest way to get from the origin to destination according to Google’s algorithm, taking into account traffic and other delays. The location of the LAD corresponds to the latitude and longitude automatically selected by Google for each district. Figures 10 and 11 show plots identical to Figure 2 using distance as the fastest commuting time rather than shortest driving distance and commuters as a share of the LAD population going to the City, respectively.

---

28 To ensure to be firmly within Google’s terms of service, I hand-collected the data by searching for each LAD.
29 For instance, a Google Maps search for directions from Kensington to the City of London returns directions leaving from 112 Kensington High St, Kensington, London and arriving at 21 Bloomberg Arcade, London.
A.2 EE and UE Transitions: Robustness

Table 12 contains estimates from a linear probability model similar to the marginal effects in the probit regression of Table 1. To test whether unobservable worker fixed effects are driving the results, I include individual fixed effects in both columns of the table. The regressions therefore rely on individuals that I observe employed for at least 3 years of the sample.

Tables 13 to 22 show the wage cut tables similar to Table 2 for the subsamples of men, women, occupation switchers and stayers, residential movers and stayers, married and single, and with and without children. Tables 23 and 24 respectively define commuting increases and decreases as +/- 5 or 10 minutes of commuting time, rather than by percent changes in the commute. Table 25 shows the table after allowing for job-to-job transitions with zero days of nonemployment between the two employment spells.

Reservation Wages

Following the arguments of Hall and Mueller (2018), I consider workers who made an UE transition between years \( t - 1 \) and \( t \), for whom both the wage and commute in year \( t \) are available and who reported their reservation wage in year \( t - 1 \). I compute the log difference between the realized wage and reservation wage, and subtract the average realized to reservation wage for all workers making
Table 12: Effect of Lagged Commute and Wage on Job-to-Job Transition Probability, Linear Probability Model

<table>
<thead>
<tr>
<th></th>
<th>J2J&lt;sub&gt;t&lt;/sub&gt;</th>
<th>J2J&lt;sub&gt;t&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commute&lt;sub&gt;t-1&lt;/sub&gt; × 30</td>
<td>0.018***</td>
<td>0.022***</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Real log Wage&lt;sub&gt;t-1&lt;/sub&gt;</td>
<td>-0.071***</td>
<td>-0.069***</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>Individual Characteristics</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Individual FE</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Region, Time, Commute Method</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Industry &amp; Occ FE</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>R&lt;sup&gt;2&lt;/sup&gt;</td>
<td>.051</td>
<td>.059</td>
</tr>
<tr>
<td>N ind</td>
<td>1,259</td>
<td>1,122</td>
</tr>
<tr>
<td>N obs</td>
<td>7,023</td>
<td>6,338</td>
</tr>
</tbody>
</table>

Notes: BHPS Sample 1992-2008, annual. Universe: respondents living in LADs within 100 miles of the City of London aged 24-55 working full-time in year <i>t</i>. Estimated coefficients from a linear probability model of J2J<sub>t</sub>, which is a dummy equal to one if a worker made a job-to-job transition in the past year and 0 if she remained in the same job, on commute time and wages in the previous year. Individual characteristics include the annual regional house price index, a quadratic term in labor market experience, age, education, marital status, and number of children, 1-year lagged tenure, 1-year lagged dummies for outright homeownership, mortgage holding, whether the individual moved in the past year, real housing expenditures, whether the worker was unemployed in the past year, the number of employment spells, whether the spouse or partner was employed last year, a government job dummy, and union status. All regressions include individual fixed effects, and region, month, year, 1-digit industry and occupation fixed effects. Robust standard errors are reported in parentheses. * denotes <i>p < .1</i>, ** <i>p < .05</i>, and *** <i>p < .01</i>.

Table 13: After Job-to-Job Transition, Men

<table>
<thead>
<tr>
<th>Commute</th>
<th>Wage</th>
<th>Down</th>
<th>Same</th>
<th>Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Down</td>
<td></td>
<td>0.43</td>
<td>0.50</td>
<td>0.36</td>
</tr>
<tr>
<td>Same</td>
<td></td>
<td>0.24</td>
<td>0.17</td>
<td>0.26</td>
</tr>
<tr>
<td>Up</td>
<td></td>
<td>0.32</td>
<td>0.33</td>
<td>0.38</td>
</tr>
<tr>
<td>N</td>
<td></td>
<td>78</td>
<td>54</td>
<td>234</td>
</tr>
</tbody>
</table>

Notes: BHPS Sample 1992-2008, annual. Universe: respondents living in LADs within 100 miles of the City of London aged 24-55 working full-time, who changed jobs since year <i>t</i>−1. "Up" and "Down" indicate differences from the last reported wage or commute of more than 5%, and "Same" indicates differences less than 5%.

UE transitions. Table 26 shows this difference for workers who accepted jobs with commutes below (center column) and above (right column) the mean commute of 53 minutes round trip. The first row of the table can be interpreted in the following way: for workers commuting less than average,
their realized to reservation wage is 10% below average, while the realized to reservation wage ratio for workers commuting more than average is 14% above average. Further splitting the sample shows that this difference in means, though not statistically significant, is qualitatively robust across
Table 17: After Job-to-Job Transition, Residential Movers

<table>
<thead>
<tr>
<th>Commute</th>
<th>Down</th>
<th>Same</th>
<th>Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Down</td>
<td>0.33</td>
<td>0.42</td>
<td>0.35</td>
</tr>
<tr>
<td>Same</td>
<td>0.19</td>
<td>0.05</td>
<td>0.19</td>
</tr>
<tr>
<td>Up</td>
<td>0.48</td>
<td>0.53</td>
<td>0.46</td>
</tr>
</tbody>
</table>

Notes: BHPS Sample 1992-2008, annual. Universe: respondents living in LADs within 100 miles of the City of London aged 24-55 working full-time, who changed jobs and residence since year $t - 1$. "Up" and "Down" indicate differences from the last reported wage or commute of more than 5%, and "Same" indicates differences less than 5%.

Table 18: After Job-to-Job Transition, Residential Stayers

<table>
<thead>
<tr>
<th>Commute</th>
<th>Down</th>
<th>Same</th>
<th>Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Down</td>
<td>0.52</td>
<td>0.44</td>
<td>0.34</td>
</tr>
<tr>
<td>Same</td>
<td>0.19</td>
<td>0.21</td>
<td>0.27</td>
</tr>
<tr>
<td>Up</td>
<td>0.29</td>
<td>0.35</td>
<td>0.40</td>
</tr>
</tbody>
</table>

Notes: BHPS Sample 1992-2008, annual. Universe: respondents living in LADs within 100 miles of the City of London aged 24-55 working full-time, who changed jobs but not residence since year $t - 1$. "Up" and "Down" indicate differences from the last reported wage or commute of more than 5%, and "Same" indicates differences less than 5%.

Table 19: After Job-to-Job Transition, Married

<table>
<thead>
<tr>
<th>Commute</th>
<th>Down</th>
<th>Same</th>
<th>Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Down</td>
<td>0.48</td>
<td>0.44</td>
<td>0.31</td>
</tr>
<tr>
<td>Same</td>
<td>0.17</td>
<td>0.22</td>
<td>0.28</td>
</tr>
<tr>
<td>Up</td>
<td>0.35</td>
<td>0.34</td>
<td>0.41</td>
</tr>
</tbody>
</table>

Notes: BHPS Sample 1992-2008, annual. Universe: respondents living in LADs within 100 miles of the City of London aged 24-55 working full-time, who changed jobs since year $t - 1$. "Up" and "Down" indicate differences from the last reported wage or commute of more than 5%, and "Same" indicates differences less than 5%.
Table 20: After Job-to-Job Transition, Single

<table>
<thead>
<tr>
<th>Commute</th>
<th>Wage</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Down</td>
<td>Same</td>
<td>Up</td>
<td></td>
</tr>
<tr>
<td>Down</td>
<td>0.51</td>
<td>0.43</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td>Same</td>
<td>0.20</td>
<td>0.15</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td>Up</td>
<td>0.28</td>
<td>0.42</td>
<td>0.40</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>82</td>
<td>53</td>
<td>228</td>
<td></td>
</tr>
</tbody>
</table>

Notes: BHPS Sample 1992-2008, annual. Universe: respondents living in LADs within 100 miles of the City of London aged 24-55 working full-time, who changed jobs since year $t-1$. "Up" and "Down" indicate differences from the last reported wage or commute of more than 5%, and "Same" indicates differences less than 5%.

Table 21: After Job-to-Job Transition, Children

<table>
<thead>
<tr>
<th>Commute</th>
<th>Wage</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Down</td>
<td>Same</td>
<td>Up</td>
<td></td>
</tr>
<tr>
<td>Down</td>
<td>0.46</td>
<td>0.38</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>Same</td>
<td>0.17</td>
<td>0.24</td>
<td>0.28</td>
<td></td>
</tr>
<tr>
<td>Up</td>
<td>0.37</td>
<td>0.38</td>
<td>0.48</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>59</td>
<td>21</td>
<td>138</td>
<td></td>
</tr>
</tbody>
</table>

Notes: BHPS Sample 1992-2008, annual. Universe: respondents living in LADs within 100 miles of the City of London aged 24-55 working full-time, who changed jobs since year $t-1$. "Up" and "Down" indicate differences from the last reported wage or commute of more than 5%, and "Same" indicates differences less than 5%.

Table 22: After Job-to-Job Transition, No Children

<table>
<thead>
<tr>
<th>Commute</th>
<th>Wage</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Down</td>
<td>Same</td>
<td>Up</td>
<td></td>
</tr>
<tr>
<td>Down</td>
<td>0.52</td>
<td>0.45</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td>Same</td>
<td>0.20</td>
<td>0.17</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>Up</td>
<td>0.28</td>
<td>0.38</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>98</td>
<td>82</td>
<td>299</td>
<td></td>
</tr>
</tbody>
</table>

Notes: BHPS Sample 1992-2008, annual. Universe: respondents living in LADs within 100 miles of the City of London aged 24-55 working full-time, who changed jobs since year $t-1$. "Up" and "Down" indicate differences from the last reported wage or commute of more than 5%, and "Same" indicates differences less than 5%.
Table 23: After Job-to-Job Transition, Commute ± 5 min

<table>
<thead>
<tr>
<th>Commute</th>
<th>Wage Down</th>
<th>Wage Same</th>
<th>Wage Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Down</td>
<td>0.50</td>
<td>0.44</td>
<td>0.33</td>
</tr>
<tr>
<td>Same</td>
<td>0.20</td>
<td>0.21</td>
<td>0.28</td>
</tr>
<tr>
<td>Up</td>
<td>0.30</td>
<td>0.35</td>
<td>0.39</td>
</tr>
<tr>
<td>N</td>
<td>157</td>
<td>103</td>
<td>437</td>
</tr>
</tbody>
</table>

Notes: BHPS Sample 1992-2008, annual. Universe: respondents living in LADs within 100 miles of the City of London aged 24-55 working full-time, who changed jobs since year \( t - 1 \). Wage “Up” and “Down” indicate differences from the last reported wage of more than 5%, and “Same” indicates differences less than 5%. Commute “Up” and “Down” indicate differences from the last reported commute of more than 5 minutes.

Table 24: After Job-to-Job Transition, Wage ± 10 min

<table>
<thead>
<tr>
<th>Commute</th>
<th>Wage Down</th>
<th>Wage Same</th>
<th>Wage Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Down</td>
<td>0.44</td>
<td>0.31</td>
<td>0.25</td>
</tr>
<tr>
<td>Same</td>
<td>0.35</td>
<td>0.42</td>
<td>0.46</td>
</tr>
<tr>
<td>Up</td>
<td>0.21</td>
<td>0.27</td>
<td>0.29</td>
</tr>
<tr>
<td>N</td>
<td>157</td>
<td>103</td>
<td>437</td>
</tr>
</tbody>
</table>

Notes: BHPS Sample 1992-2008, annual. Universe: respondents living in LADs within 100 miles of the City of London aged 24-55 working full-time, who changed jobs since year \( t - 1 \). Wage “Up” and “Down” indicate differences from the last reported wage of more than 5%, and “Same” indicates differences less than 5%. Commute “Up” and “Down” indicate differences from the last reported commute of more than 10 minutes.

Table 25: After Job-to-Job Transition, 0 Days Nonemployment

<table>
<thead>
<tr>
<th>Commute</th>
<th>Wage Down</th>
<th>Wage Same</th>
<th>Wage Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Down</td>
<td>0.50</td>
<td>0.42</td>
<td>0.34</td>
</tr>
<tr>
<td>Same</td>
<td>0.19</td>
<td>0.20</td>
<td>0.25</td>
</tr>
<tr>
<td>Up</td>
<td>0.30</td>
<td>0.38</td>
<td>0.41</td>
</tr>
<tr>
<td>N</td>
<td>149</td>
<td>97</td>
<td>417</td>
</tr>
</tbody>
</table>

Notes: BHPS Sample 1992-2008, annual. Universe: respondents living in LADs within 100 miles of the City of London aged 24-55 working full-time, who changed jobs since year \( t - 1 \) with 0 days nonemployment between employment spells. “Up” and “Down” indicate differences from the last reported wage or commute of more than 5%, and “Same” indicates differences less than 5%.
Table 26: Log Realized to Reservation Wage Relative to Average

<table>
<thead>
<tr>
<th>Group</th>
<th>Commute Below Median</th>
<th>Commute Above Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>-0.10</td>
<td>0.15</td>
</tr>
<tr>
<td>Some College or Less</td>
<td>-0.02</td>
<td>0.03</td>
</tr>
<tr>
<td>College Graduates</td>
<td>-0.21</td>
<td>0.23</td>
</tr>
<tr>
<td>Men</td>
<td>-0.20</td>
<td>0.25</td>
</tr>
<tr>
<td>Women</td>
<td>-0.01</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Notes: BHPS Sample 1993-2008, annual. Universe: respondents living in LADs within 100 miles of the City of London aged 24-55 employed in year $t$ and reporting a minimum weekly wage ("Reservation wage") while unemployed in year $t-1$. Columns report the log difference between the realized and reservation wage relative to the average across all commutes, conditional on the new job requiring an above or below-mean commuting time. In the sample, the mean round trip commute is 53 minutes. All: N=103, Men: N=57, Women: N=46, College Graduates: N=47, Some College or Less: N=56.

Table 27: After Job-to-Job Transition, Weighted Sample

<table>
<thead>
<tr>
<th>Commute</th>
<th>Wage</th>
<th>Down</th>
<th>Same</th>
<th>Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Down</td>
<td></td>
<td>0.51</td>
<td>0.45</td>
<td>0.31</td>
</tr>
<tr>
<td>Same</td>
<td></td>
<td>0.19</td>
<td>0.17</td>
<td>0.26</td>
</tr>
<tr>
<td>Up</td>
<td></td>
<td>0.20</td>
<td>0.38</td>
<td>0.43</td>
</tr>
<tr>
<td>N</td>
<td></td>
<td>101.44</td>
<td>61.50</td>
<td>277.1</td>
</tr>
</tbody>
</table>

Notes: BHPS Sample 1992-2008, annual. Universe: respondents living in LADs within 100 miles of the City of London aged 24-55 working full-time, who changed jobs since year $t-1$. "Up" and "Down" indicate differences from the last reported wage or commute of more than 5%, and "Same" indicates differences less than 5%. Observations are weighted by the Longitudinal individual respondent weights.

almost all groups.

Results Using Longitudinal Weights

In this section I weight the observations by their longitudinal sample weights ($lrwght$) and reproduce the empirical analysis above. Longitudinal weights are nonzero only for individuals in year $t$ who gave full interview starting with the first wave up to and including year $t$ (or children who were under 16 in families present in the first wave, who at the age of 16 are allocated the minimum of the longitudinal weights of their parents). Because using these weights further reduces the already small sample size for workers making UE transitions, I do not report results for the subsamples by education or gender.
Table 28: Log Realized to Reservation Wage Relative to Average, Weighted Sample

<table>
<thead>
<tr>
<th>Group</th>
<th>Commute Below Median</th>
<th>Commute Above Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>-0.07</td>
<td>0.13</td>
</tr>
</tbody>
</table>

Notes: BHPS Sample 1993-2008, annual. Universe: respondents living in LADs within 100 miles of the City of London aged 24-55 employed in year $t$ and reporting a minimum weekly wage (“Reservation wage”) while unemployed in year $t-1$. Columns report the log difference between the realized and reservation wage relative to the average across all commutes, conditional on the new job requiring an above or below-median commuting time. In the sample, the median round trip commute is 40 minutes (standard deviation = 36.7 minutes). Observations are weighted by the Longitudinal individual respondent weights. All: N=73.
B Proof of Proposition 1

For the purposes of the proof, I define value functions, policy functions, and distributions as functions of the price vector \( r \), e.g., \( U(\ell; r) \). I assume that the distributions of workers, \( u(\ell; r) \) and \( e(\ell_F, y, \hat{\ell}; r) \) are continuous in \( r \).

Part 1a - Demand Correspondence

Define the demand correspondence for newly matched vacancies for each \( i = 0, \ldots, N \) and \( y \in Y \) as

\[
D_V(\ell_i, y; r) = e^{i+1} \sum_{\ell} \{ \lambda_0 u(\ell; r) \mathbb{1}\{S(\ell, y, \ell; r) > 0\} + \lambda_1 \sum_{\hat{\ell}} \sum_{\hat{y}} \mathbb{1}\{S(\ell, y, \ell; r) > S(\hat{\ell}, \hat{y}, \ell; r)\} e(\hat{\ell}, \hat{y}, \ell; r) \}
\]

where \( e^i \) is the \( i \)-th unit vector of dimension \( N + 1 \), and the second term is the probability a meeting occurs and the match is accepted. This demand correspondence defines the number of vacancies created in location \( i - 1 = 0, \ldots, N \).

Similarly, define the demand correspondence for the employed and unemployed, respectively, as

\[
D_E(\ell_F, y, \hat{\ell}; r) = \{ x \in \{0, e^1, e^2, \ldots, e^{N+1}\} : \beta S(\ell_F, y, \hat{\ell}; r) > \overline{u} \text{ and } x(\ell) \in \arg \max M(\ell_F, y, \hat{\ell}; r) \}
\]

for each \( (\ell_F, y) \in \mathcal{L} \times Y \) and \( \hat{\ell} \in \mathcal{L} \) and

\[
D_U(\hat{\ell}; r) = \{ x \in \{0, e^1, e^2, \ldots, e^{N+1}\} : U(\hat{\ell}; r) \geq \overline{u} \text{ and } x(\ell) \in \arg \max U(\hat{\ell}; r) \}
\]

for each \( \hat{\ell} \in \mathcal{L} \). I denote the location choice implied by \( x \) as \( x(\ell) \in \mathcal{L} \). Clearly, since all of the sets include 0, all three sets are nonempty. Define the aggregate demand correspondence \( D(r) \) as

\[
D(r) = \sum_{\hat{\ell}} \sum_{\hat{y}} D_V(\hat{\ell}, \hat{y}; r) g(\hat{\ell}, \hat{y}) + D_U(\hat{\ell}; r) u(\hat{\ell}; r) + \sum_{\ell_F} \sum_y D_E(\ell_F, y, \hat{\ell}; r) e(\ell_F, y, \hat{\ell}; r)
\]

The first term is the demand correspondence for land for each vacancy which is created, multiplied by the mass of vacancies of each type \((\hat{\ell}, \hat{y})\). The second and third terms are the demand correspondences, respectively, of the unemployed and employed workers living in \( \hat{\ell} \) times their mass. Importantly, the mass of workers and firms in each location and of workers in each employment state depends on the price vector \( r \).

It is clear that \( M(\ell_F, y, \hat{\ell}; r) \) and \( U(\hat{\ell}; r) \) are continuous functions of \( r \). By assumption, \( u(\hat{\ell}; r) \) and \( e(\ell_F, y, \hat{\ell}; r) \) are continuous in \( r \). Thus, \( U, W, \text{ and } J \) are all continuous in \( r \). Given these results, I now prove each element of \( D(r) \) is upper hemicontinuous (UHC).

(i) Unemployed: \( D_U(\hat{\ell}; r) u(\hat{\ell}; r) \)
Take an arbitrary $\hat{\ell} \in \mathcal{L}$. Take a sequence $r^n \to r^0$, $x^n \to x^0$ with $x^n \in D_U(\hat{\ell}; r^n)$ for all $n$.

Since $S$ is continuous and $x^n \in \arg \max U(\hat{\ell}; r^n)$ for all $n$, it follows that $x^0$ must be feasible. Suppose $x^0 \notin \arg \max U(\hat{\ell}; r^0)$. Then since $D_U(\hat{\ell}; r^n)$ is nonempty, there exists $x' \in \{0, e^1, e^2, \ldots, e^{N+1}\}$ with $x' \in \arg \max U(\hat{\ell}; r^0)$. Thus,

$$\tilde{U}(x'(\ell); r^0) - k_M 1\{x'(\ell) \neq \hat{\ell}\} \geq \hat{U}(x^0(\ell); r^0) - k_M 1\{x^0(\ell) \neq \hat{\ell}\}$$

Since $x' \neq x^0$, it must be the case that $x'(\ell) \neq \hat{\ell}$ or $x^0(\ell) \neq \hat{\ell}$ or both. Then

$$\tilde{U}(x'(\ell); r^0) - k_M 1\{x'(\ell) \neq \hat{\ell}\} - \tilde{U}(x^n(\ell); r^n) + k_M 1\{x^n(\ell) \neq \hat{\ell}\}$$
$$> \tilde{U}(x^n(\ell); r^n) - k_M 1\{x^n(\ell) \neq \hat{\ell}\} - \tilde{U}(x^n(\ell); r^n) + k_M 1\{x^n(\ell) \neq \hat{\ell}\}$$

For $n$ large enough, the right hand side can be made arbitrarily close to zero, thus

$$\tilde{U}(x'(\ell); r^0) - k_M 1\{x'(\ell) \neq \hat{\ell}\} > \hat{U}(x^n(\ell); r^n) - k_M 1\{x^n(\ell) \neq \hat{\ell}\}$$

But $U$ is continuous in $r$, so

$$\tilde{U}(x'(\ell); r^n) - k_M 1\{x'(\ell) \neq \hat{\ell}\} > \hat{U}(x^n(\ell); r^n) - k_M 1\{x^n(\ell) \neq \hat{\ell}\}$$

which is a contradiction since $x^n \in D_U(\hat{\ell}; r^n)$. Since $\hat{\ell} \in \mathcal{L}$ was arbitrary, it follows that $D_U(\ell; r)$ is UHC for all $\ell \in \mathcal{L}$, and given the guess that $u(\ell; r)$ is continuous, $D_U(\ell; r)u(\ell; r)$ is UHC.

(ii) Employed: $\sum_{\ell_F} \sum_y D_E(\ell_F, y, \hat{\ell}; r)e(\ell_F, y, \hat{\ell}; r)$

Similarly to the argument in part (i), consider an arbitrary $(\ell_F, y) \in \mathcal{L} \times Y$ and $\hat{\ell} \in \mathcal{L}$, and let $r^n \to r^0$, $x^n \to x^0$ with $x^n \in D_E(\ell_F, y, \hat{\ell}; r^n)$ for all $n$. Clearly $x^0$ is feasible, but suppose there exists $x' \in D_E(\ell_F, y, \hat{\ell}; r^n)$. Then $x'(\ell) \in \arg \max M(\ell_F, y, \hat{\ell}; r^0)$. Since $M$ is continuous for $\exists N$ such that for $n > N$, $x^n(\ell) \in \arg \max M(\ell_F, y, \hat{\ell}; r^n)$, a contradiction.

Since $(\ell_F, y)$ and $\hat{\ell}$ were arbitrary, and using our guess $e(\ell_F, y, \hat{\ell}; r)$ is continuous in $r$, the above argument holds for all $(\ell_F, y) \in \mathcal{L} \times Y$ and $\hat{\ell} \in \mathcal{L}$, and thus $\sum_{\ell_F} \sum_y D_E(\ell_F, y, \hat{\ell}; r)e(\ell_F, y, \hat{\ell}; r)$ is upper hemicontinuous.

(iii) Vacancies: $\sum_y D_V(\hat{\ell}, \hat{y}; r)g(\hat{\ell}, \hat{y})$

Consider an arbitrary $(\hat{\ell}, \hat{y}) \in \mathcal{L} \times Y$. Each

$$\sum_{\ell} [\lambda_0 u(\ell; r)1\{S(\ell, y, \ell; r) > 0\}]$$
$$+ \lambda_1 \sum_{\ell} \sum_{\hat{y}} 1\{S(\ell, y, \ell; r) > S(\hat{\ell}, \hat{y}, \ell; r)e(\hat{\ell}, \hat{y}, \ell; r)\}$$
is continuous since \( S \) is continuous under the assumption that \( u \) and \( e \) are continuous. Thus, \( D_V \) is UHC for all \((\hat{\ell}, \hat{y}) \in L \times Y\), and \( \sum_{\hat{y}} D_V(\hat{\ell}, \hat{y}; r) g(\hat{\ell}, \hat{y}) \) is UHC.

Given the continuity of \( S \), it follows from the steady state conditions equating the flows into and out of employment and unemployment that \( u(\hat{\ell}; r) \) and \( e(\ell_F, y, \hat{\ell}; r) \) are continuous in \( r \) for each \( \hat{\ell} \in L \), verifying the guess. Thus, the aggregate demand correspondence \( D(r) \) is nonempty and UHC.

**Part 1b - Supply Correspondence**
The supply of land at each location \( \ell \) available to allocate to each group making up the demand correspondence above is given by the land choice of the landlord, \( \psi(\ell) \), less the land occupied by continuing jobs, for whom the location is a state. In steady state, the amount of land occupied by continuing jobs is constant and equal to the mass of jobs located in \( \ell \) less the mass of endogenously or exogenously destroyed matches. In each location \( \ell \in L \), the supply is given by

\[
\Sigma(\ell, r) = \psi(\ell, r) - \sum_y \sum_{\hat{\ell}} \left( e(\ell, y, \hat{\ell}; r) - e^-(\ell, y, \hat{\ell}; r) \right)
\]

where \( \Sigma(\ell, r) \geq 0 \) if \( r \geq 0 \) and 0 otherwise since landlords will not rent at a negative price by assumption. The term \( e^-(\ell, y, \hat{\ell}; r) \) corresponds to (15). The supply function in each location is clearly continuous in \( r \) given landlords’ preferences and given that \( e \) is continuous in \( r \). Thus, the aggregate supply correspondence is the vector

\[
E(r) = [\Sigma(0, r), \ldots, \Sigma(N, r)]
\]

which is nonempty and upper hemicontinuous.

**Part 2 - Excess Demand**
Define the excess demand correspondence as

\[
E(r) = D(r) - \Sigma(r)
\]

It follows from Part 1 that the excess demand correspondence maintains the nonemptiness and upper hemicontinuity of the supply and demand correspondences. Define the set of prices as \( P = [0, P_{max}]^{N+1} \), where

\[
P_{max} = \max_{\ell_F, y, \hat{\ell}} \left\{ \left( y + \delta U(\ell) + \lambda_1 \beta \sum_{\hat{\ell}_F, \hat{y}} \max \left\{ 0, M(\hat{\ell}_F, \hat{y}, \ell; r) - M(\ell_F, y, \ell; r) \right\} g(\hat{\ell}_F, \hat{y}) \right) / (\rho + \chi + \delta) \right\}
\]

Letting the minimum and maximum values of \( E(r) \) given the set \( P \) as \([\underline{E}, \overline{E}]\), the excess demand correspondence maps the convex set \( P \) into the convex set \( C = [\underline{E}, \overline{E}]^{N+1} \) where \(-\infty < \underline{E} < \overline{E} < \infty\).

**Part 3 - Existence of a Fixed Point**
This part of the proof uses results presented in Kaneko and Yamamoto (1986) (henceforth KY). By Lemma 4 in KY, the convex hull of \( E(r) \), denoted \( \text{cov}E(r) \) inherits the nonemptiness and upper hemicontinuity of \( E(r) \). By Lemma 5 in KY, \( \text{cov}E(r) \) is convex-valued. Further, denoting \( x \) as an element of \( D(r) \), the following two results hold:

1. If \( r(\ell) = 0 \) for some \( \ell \in \mathcal{L} \) and \( x \in \text{cov}E(r) \), then \( x(\ell) \geq 0 \)

2. If \( r(\ell) = S_{\text{max}} \) for some \( \ell \in \mathcal{L} \) and \( x \in \text{cov}E(r) \), then \( x(\ell) \leq 0 \)

To prove 1, notice that if \( r(\ell) = 0 \), \( \Sigma(r, \ell) = 0 \). Since \( D(r) \) is nonempty, aggregate demand at \( \ell \) must be weakly positive. To prove 2, since for \( r(\ell) \geq 0 \), supply is weakly positive and the surplus of all matches is weakly negative. Thus, no vacancies that draw \( \ell \) will match, and no employed or unemployed workers will demand to locate in \( \ell \) since the employed can at most extract their total match surplus and the unemployed have a lower value than the maximum match surplus. By Kakutani’s fixed point theorem, there exists a price vector \( r \in P \) such that \( 0 \in \text{cov}E(r) \).

Denote this solution by \( x_{\text{ch}} \) and \( \sigma_{\text{ch}} \), where \( x_{\text{ch}}^i, i = 0, \ldots, N \) is given by the sum of each type of worker’s demand for land in location \( i \). There are \( N \) “types” of unemployed workers, and \( N^2Y \) “types” of employed workers, thus, the maximum number of types in each location is \( M \equiv N + N^2Y \). I therefore denote \( x_{\text{ch}}^{j,i} \) as the demand for land in location \( i \) by type \( j = 1, \ldots, M \). Similarly, \( \sigma_{\text{ch}}^i \) is the supply of land in each location \( i \), after taking into account the demand for new vacancies. Note that in steady state, the demand for land by newly filled vacancies is equal to the inflow to employment. Thus,

\[
\sigma_{\text{ch}}^i = \psi(i) - \sum_y \sum_{\hat{\ell}} e(i, y, \hat{\ell})
\]

The market clearing condition for land can be written

\[
\sum_{j=1}^{M} x_{\text{ch}}^{i,j} + \sigma_{\text{ch}}^i = L \quad \text{for all } i = 1, \ldots, N
\]

Since each worker can consume at most 1 unit of land:

\[
\sum_{i=0}^{N} x_{\text{ch}}^{i,j} \leq 1 \quad \text{for all } j = 1, \ldots, M
\]

Since \( x_{\text{ch}} \) and \( \sigma_{\text{ch}} \) may not be integers, consider the system

\[
\sum_{j=1}^{M} \hat{x}^{i,j} + \hat{\sigma}^i = L \quad \text{for all } i = 1, \ldots, N
\]

\[
\sum_{i=0}^{N} \hat{x}^{i,j} \leq 1 \quad \text{for all } j = 1, \ldots, M
\]

where \( \hat{x}^{i,j} \in \mathbb{R}_+ \) and \( \hat{\sigma}^i \in \mathbb{R}_+ \) for all \( i \in N \) and \( j \in M \).
Let $x = [x_0^1, x_1^1, \ldots, x_N^1, \ldots, x_0^M, \ldots, x_N^M]$ be an $1 \times (N+1)M$ vector and $\sigma = [\sigma^0, \ldots, \sigma^N]$ is a $1 \times (N + 1)$ vector. The system above can be written as

\[
A(x, \sigma) \leq \begin{bmatrix} 1 \\ \vdots \\ L \end{bmatrix}
\]

where $A$ is an $(N + 1)M \times (N + 1)(M + 1)$ matrix

\[
A = \begin{bmatrix}
1 & \ldots & 1 \\
\vdots & \ddots & \vdots \\
1 & \ldots & 1 \\
1 & 1 & 1 \\
\end{bmatrix}
\begin{bmatrix}
0_{M \times (N+1)} \\
I_{N+1}
\end{bmatrix}
\]

and it follows directly from the proof of the Theorem in KY that the fixed point is also in $E(r)$, since any fixed point of the convex hull can be written in matrix form and shown to satisfy the unimodular property, implying that the solution is also an integer solution. Thus, an equilibrium price vector $r$ exists.

### C Toy Model: A Static Planner’s Problem

To build intuition, I consider a static model analyzing the inefficiency introduced by the commuting externality. In the next section, I build a dynamic model with a rich set of features which I will use to study the welfare implications of remote working quantitatively. Time lasts for one period. The economy is defined by two locations $\{1, 2\}$ in which workers and firms may locate. The exogenous rent in location $\ell$ is given by $r(\ell)$, paid by both workers and firms located in $\ell$. Land is assumed to be perfectly elastic. Workers are born unemployed, with their location $\ell$ determined by a draw from exogenous distribution $F$.

Workers are born at the beginning of the period and draw a vacancy with probability $\lambda$. A vacancy is a draw of the location of the match, $\ell_F$ from distribution $G$, with the probability of drawing a vacancy in location $\ell$ denoted $g(\ell)$, with $g(1) = 1 - g(2)$.
The Decentralized Economy

Upon arrival of an offer, a worker has three options. She may (i) accept the job and remain in her current location, (ii) accept the job and work remotely, or (iii) reject the job and stay unemployed. If a worker accepts a job, she receives a wage conditional on her match and location: \( w(\ell_F, \ell) \) which is a share \( \beta \) of the match surplus, where \( \ell \) is her chosen location. To highlight the role of the inefficiency in location decisions, I assume that \( \beta = 1 \). If the worker commutes, that is \( \ell \neq \ell_F \), she pays commuting cost \( T(\Omega) \), where \( \Omega \) is the mass of commuters.

Employed workers consume their wages net of commuting and rent costs, and unemployed workers consume the value of home production \( b \) net of rent. Workers take expectations over the equilibrium value of \( \Omega \), and their beliefs are assumed to be consistent with the realized value of \( \Omega \) in equilibrium.

Suppose that in the decentralized equilibrium a mass \( e_1 + e_2 \) commute to their jobs, where \( e_\ell \) are workers born in location \( \ell = 1, 2 \). If it is optimal for both of these groups to commute, the following conditions must hold:

\[
1 - r(\ell) - T(e_1 + e_2) > b, \quad \ell = 1, 2
\]

(23)

\[
T(e_1 + e_2) < p_r
\]

(24)

Conditions (23) and (24) imply that all workers in the decentralized economy located in 1 (2) receiving an offer located in 2 (1) accept and commute. In this case, the commuting cost in the decentralized equilibrium is given by \( T(\lambda(f(2)g(1) + f(1)g(2))) \). The conditions that are assumed to hold in the decentralized equilibrium are (23), (24), and: \( 1 - r(\ell) > b \) for \( \ell = 1, 2 \), which states that workers who receive an offer in their own location always accept.

A Constrained Planner

The planner’s objective is to maximize output by choosing the mass of workers in each match and employment state, and whether a worker works remotely or commutes in the case of a job offer in a different location than the worker’s residence. The planner is subject to feasibility constraints due to the labor market frictions and a restriction on the total mass of workers in the economy.

Consider an alternative allocation in which the mass \( e_1 \) works remotely and \( e_2 \) continues to commute. All other workers’ contributions to output are unchanged. The planner prefers this allocation to the decentralized allocation if

\[
T(e_1 + e_2)(e_1 + e_2) - T(e_2)e_2 > p_re_1
\]

(25)

where \( e_1 + e_2 \) is the mass of commuters in the decentralized allocation and \( e_2 \) is the mass of commuters under the alternative allocation. The condition states that the benefit to all workers from lower commuting costs must be greater than the output loss. Assume that commuting costs
are of the form $T(\Omega) = \Omega^\gamma$, then we can write (24) and (25) as

$$(e_1 + e_2)^\gamma < p_r < \left(\frac{(e_1 + e_2)^{\gamma+1} - e_2^{\gamma+1}}{e_1}\right)$$

If the above condition holds, then the decentralized equilibrium is inefficient and the planner will choose for some workers to work remotely. In other words, it can be welfare improving to use the remote working technology but privately optimal to commute when the productivity loss is greater than the individual benefit from decreasing the commute but it is less than the social benefit of reducing congestion.

Proposition 2 describes the conditions under which the decentralized equilibrium is inefficient.

**Proposition 2.** Suppose $T(\Omega) = \Omega^\gamma$, $e_1, e_2 \in (0, 1)$ and $e_1 + e_2 < 1$. Then the following are true:

(i) For any $p_r \in (0, 1)$ there exists $\gamma > 0$ such that the decentralized equilibrium is inefficient.

(ii) As $p_r$ approaches zero, the value of $\gamma$ such that the decentralized equilibrium is inefficient approaches infinity.

The next section contains the proof of the results. In this static setting, the congestion externality arises when two conditions are met. First, the reduction in output is sufficiently small, and second, the effect of congestion on commuting costs is sufficiently convex. The inefficiency arises when a mass of workers finds it privately optimal to commute but would make all other workers better off by working remotely. The social value of remote working includes the net benefits for all remaining commuters from less congestion, net of the output loss for the remote workers. Part (i) of Proposition 2 states that when the commuting cost is sufficiently convex, by decreasing congestion by a small amount, the cost for all commuters significantly decreases. Because output is linear in the mass of workers, the output loss is smaller than the output gain due to the fall in congestion. Part (ii) states that the smaller is the output loss, the greater must be the benefits from reducing commuting for the decentralized economy to be inefficient. When output losses are zero, there is only a positive effect of remote working and it is privately and socially optimal for all workers who would otherwise commute to work remotely. The smaller the output loss, the more likely that the decentralized equilibrium is efficient.

**C.1 Proof of Proposition 2**

The planner’s objective is to maximize output by choosing the mass of workers in each match and employment state:

\begin{align*}
(1-2r(1))e(1, 1) + (1-r(1)-r(2)-T(e(1, 2)+e(2, 1)))(1-\pi(1, 2))e(1, 2) + ((1-p_r)-r(1)-r(2))\pi(1, 2)e(1, 2) \\
+(1-2r(2))e(2, 2) + (1-r(1)-r(2)-T(e(1, 2)+e(2, 1)))(1-\pi(2, 1))e(2, 1) + ((1-p_r)-r(1)-r(2))\pi(2, 1)e(2, 1) \\
+(b-r(1))u(1) + (b-r(2))u(2) \tag{26}
\end{align*}
Subject to the feasibility constraints:

\[
\sum_{\ell=1}^{2} e(\ell, i) \leq \lambda(\ell)g(i), \text{ for } i = H, L
\]

\[
e(\ell, \ell_f) \geq 0 \forall (\ell, \ell_F)
\]

\[
u(\ell) \geq (1 - \lambda)f(\ell) \forall \ell
\]

\[
\sum_{\ell=1}^{2} \left( u(\ell) + \sum_{\ell_F=1}^{2} e(\ell, \ell_F) \right) = 1
\]

\[
\pi(\ell, \ell_F) \in [0, 1] \text{ for } \ell \neq \ell_F
\]

The first condition restricts the mass of employed in each match to be at most the mass of workers who received that offer. The second and third conditions restrict the mass of employed workers of each type to be weakly positive and the mass of unemployed to be at least equal to the mass of workers receiving no job offer. The fourth condition requires that the total mass of workers in the economy is 1. The last condition restricts the planner’s choice between working remotely and commuting to be a share between 0 and 1 of the mass of workers with offers in a location different from their own.

I now compare aggregate output in the allocation of the decentralized economy to an allocation in which the planner assigns all workers receiving offers in location 2 and living in 1 to work remotely. Commuting costs in the first case are given by those in the decentralized economy: \( T(e_1 + e_2) \). In the second case, commuting costs are \( T(e_2) \), where \( e_1 = \lambda f(1)g(2) \) and \( e_2 = \lambda f(1)g(2) \).

In the decentralized allocation, the contribution to aggregate output of workers living in 1 and receiving offers on 2, and workers living in 2 receiving offers in 1 is given by

\[
(1 - r(2) - r(1) - T(e_1 + e_2))e_1 + (1 - r(2) - r(1) - T(e_1 + e_2))e_2 
\]  

(27)

All other terms are assumed to remain constant, since all workers accept offers and only commute if the offer comes from the location in which they do not live.

Now, consider the contribution to aggregate output of the same groups after the planner changes the allocation.

\[
((1 - p_r) - r(2) - r(1))e_1 + (1 - r(2) - r(1) - T(e_2))e_2 
\]  

(28)

Output is increased under the new allocation if (28)\( > \) (27), that is:

\[
(T(e_1 + e_2) - T(e_2))e_2 > (p_r - T(e_1 + e_2))e_1
\]

Rearranging the above condition, we have

\[
p_r e_1 < (e_1 + e_2)T(e_1 + e_2) - e_2T(e_2) 
\]  

(29)
Suppose $T(\Omega) = \Omega^\gamma$. Then

\[ p_r < \frac{(e_1 + e_2)^{\gamma+1} - e_2^{\gamma+1}}{e_1} \]  

(30)

Condition (24) can be rearranged to obtain

\[ p_r > (e_1 + e_2)^\gamma \]  

(31)

It is easy to show that for all $\gamma > 0$,

\[ \frac{(e_1 + e_2)^{\gamma+1} - e_2^{\gamma+1}}{e_1} > (e_1 + e_2)^\gamma \]

Further, for all $\gamma > 0$, $(e_1 + e_2)^{\gamma+1} - e_2^{\gamma+1} > 0$, and $(e_1 + e_2)^\gamma < 1$. Suppose $\exists p_r \in (0, 1)$ such that both (30) and (31) do not simultaneously hold for any $\gamma > 0$. First note that

\[ \frac{\partial}{\partial \gamma} (e_1 + e_2)^\gamma < 0 \]

and for $\gamma$ sufficiently large,

\[ \frac{\partial}{\partial \gamma} \frac{(e_1 + e_2)^{\gamma+1} - e_2^{\gamma+1}}{e_1} < 0 \]

As $\gamma \to \infty$, $(e_1 + e_2)^\gamma$ converges to zero, thus if $p_r > 0$, $\exists \gamma > 0$ such that (31) holds. As $\gamma \to 0$, the right hand side of (30) converges to 1. Thus, it must be the case that $\exists \gamma > 0$ such that $p_r$ satisfies both eqrefub and (31), which completes the proof of part (i) of the proposition. To prove part (ii), note that both right hand sides of (30) and (31) are strictly decreasing for large enough $\gamma$ and converge to zero as $\gamma \to \infty$, thus, as $p_r \to 0$, the value of $\gamma$ satisfying (30) and (31) converges to infinity.

D Alternative Definition of the Local Labor Market

I redefine the local labor market used in the empirical analysis and derivation of moments. In Section 4, the local labor market is defined as those LADs within 100 miles of the City of London, seen in Figure 2 as the point below which a significant number of workers commute to the City. An alternative definition is explored here, specifically, by defining the local labor market as a “commuting zone”. The commuting zone is comprised of those locations for which over half of all workers commute within its boundaries.

As in Section 2, I use the 2001 Census Flow Data to identify the set of regions for which a positive number of workers commute to the City of London, and consider all origin-destination pairs for which both LADs are in this set. For each origin, I compute the number of workers commuting to each destination as a share of the population. The LAD with the largest share is Chelmsford (0.76). The City of London’s share is .73. I define the commuting zone as those LADs with a share above half of the maximum, .38. Results for the empirical analysis in Section 2 and Appendix A
are similar and are available upon request.

E Numerical Solution Algorithm

I solve the model in the following steps:

1. Guess a rent function, initial distributions of workers, and initial value functions for $U$, $W$, and $S$. Set the initial externality to zero.

2. Given the rent function and distributions of workers, solve for the value functions $U$, $W$, and $S$ and optimal moving choices using (1), (3), (9), (10), and (13) by iterating on the value functions until convergence.

3. Given the value functions computed in step 2, compute the worker distributions given steady state flow conditions (14)-(17).

4. Given worker distributions, compute the level of congestion using (20).

5. Given worker distributions, compute the difference between the left and right hand sides of (22) for each $\ell$. Denoting the excess demand in (22) as $D_E(\ell)$, update the rent price as

$$r'(\ell) = r(\ell) + f(D_E(\ell))$$

where $f$ is a normalization to avoid large jumps in the price: $f(x) = x/50$.

6. Using solutions for $W$, $U$, $S$, $\tilde{W}$, and $\tilde{S}$, compute wages satisfying (4)-(8).

7. Compute the difference between the guessed and updated worker distributions and repeat steps 2 to 6 until convergence.

F Calibration

F.1 Moments for Calibration

All moments used in the calibration relating to wages correspond to the residuals of log real monthly wages on year, month, education level, and commuting method. Wage observations more than 3 standard deviations outside of the mean are omitted. Further, the moments correspond to the subset of workers considered in Section 2: age 24-55, in the labor force, with a commute up to 180 minutes round trip, living within 100 miles from the City of London as defined by Google Maps. Locations in the model correspond to LADs within 10 miles of the center (location 4), and 10-33 miles (locations 3 and 5), 34-66 (2 and 6), and 66-100 (1 and 7) miles from the center.
To calibrate the productivity process, I discretize the $y$ distribution to approximate a Normal $(1 + A(\ell_F), \sigma_y)$ using the Tauchen method, where the parameters $A(\ell_F)$ are calibrated to match relative wages by worker location.

The moving probability is the share of workers who report moving more than three miles (5km) from their residence in the last wave of the survey. The share of job-related moves is the share of movers who state that their main reason for moving was for their own job. Jobs density is available from the Office for National Statistics, defined as the ratio of the number of jobs to resident working age population (16-64), and corresponds to the year 2001, though at the aggregated level used here the figures are stable over time.

The value for the replacement rate is standard in the literature; Bils et al. (2012) point out that this value should include the saved costs due to not working which here include commuting costs. The UE, EU, EE rates come from Gomes (2012) and correspond to the quarterly rates for the UK between 1994 and 2009. The variance of the wage corresponds to the cross-sectional residual log wage variance. Average rent to income is the average net monthly housing cost for individuals who rent their residence relative to the net unresidualized wage. The elasticity of UK housing supply comes from Barker (2003). The newborn distribution across locations is given by $[0.1475, 0.1550, 0.1180, 0.1630, 0.1180, 0.1550, 0.1475]$.

### F.2 Figures

**Figure 12: Housing Rent**

![Housing Rent Graph](image_url)

Figure 13 shows the policy function for offers with the highest productivity, given the worker’s initial location. The left panel shows the surplus and location decisions of a match $(\ell_F, y_H, 1)$ as a function of $\ell_F$. When the firm is in or near the worker’s location, at the left side of the figure,
surplus is high and the worker does not move. As the job nears the center, the surplus drops because congestion faced by the worker and rent for the firm rise. When commuting costs are high enough, the worker moves to the firm’s location. The moving decision is different for a worker with initial location in the center (right panel). For this worker, again the surplus is higher when the firm and worker locations are the same, but decreases in both directions as the firm moves toward the periphery. If the firm is two or more locations away, the worker moves in order to decrease her rent and commute. Comparing the two figures, if an unemployed worker living in location 1 gets a job offer from location 3, she will accept it and commute, whereas if a worker living in 4 gets an offer from 2, he will move to the firm’s location. Both are equidistant in terms of miles, but the commuting time is very different due to differences in congestion in the center and periphery.