Costly Commuting and the Job Ladder

Jean Flemming

2020-025

Costly Commuting and the Job Ladder

Jean Flemming[*]

March 13, 2020

Abstract

I show empirically that commuting time affects job acceptance, suggesting large indirect costs of traffic congestion. To quantify congestion's effect on the labor market, I build a model of frictional job search within a metropolitan area. Workers evaluate job offers based on productivity and commuting costs, taking congestion as given, but by accepting and commuting to distant jobs, affect others' labor market outcomes. Equilibrium employment transitions and wages are tightly linked to congestion. Calibrating the model to the London area, the congestion externality significantly decreases welfare and increases wage inequality. I show that the stronger are search frictions, the smaller are the welfare gains from a congestion tax. (JEL *E24, J32, J62, R13, R41*)

1   Introduction

In the 10 most congested cities in the world, commuters spend an average of 80 hours per year in rush hour traffic, equivalent to two work weeks of lost 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. Congestion in urban areas is an indication of a negative externality whereby the decision of one worker to commute affects the travel time of others. Taxes on commuters passing through congested areas can lead them to internalize their effect on others and improve welfare. These policies have been shown to successfully reduce congestion.[2] However, policies aimed at commuting and congestion can have large indirect effects because they potentially misallocate resources: more expensive commutes can stop workers from accepting productive job offers and lead them to accept less productive offers closer to home, or even to remain unemployed. Data from the Federal Reserve Bank of New York's Survey of Consumer Expectations shows that 13% of workers report the job's location as the main reason for rejecting an offer.[3] 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.

In order to fully understand the effects of congestion taxes and other urban policies, it is necessary to consider the key margins of the labor market that respond to these policies. To do so, I build a model in which the process of matching workers to jobs and commuting patterns are simultaneously determined. To the best of my knowledge, this is the first paper in which congestion arises endogenously within a frictional labor market. Workers' rankings of job offers depend on the productivity and commute required to get to a potential employer. By accepting a job offer that requires the worker to commute, she contributes to congestion which affects the labor market outcomes of others. The extent of congestion is closely tied to the frictions in the labor market, in particular how often workers receive job offers and how they rank them. In the frictionless limit, congestion will only arise if productivity and therefore wages differ by location to fully compensate workers for their commutes.

---

[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]For example, Singapore's electronic road pricing scheme and Stockholm's congestion tax.

[3]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.

With frictions, congestion arises because it takes time to find acceptable jobs, in terms of both location and productivity. The effect of a tax is therefore not only to change the locations of workers' jobs and their commutes, but also to affect unemployment, output, and wages. The model therefore shows that, even in the absence of a positive externality such as agglomeration, the welfare benefits of a congestion tax may be small due to the presence of search frictions.

This paper makes three contributions. First, I document new facts connecting commuting and job mobility using data from the British Household Panel Survey (BHPS, University of Essex, Institute for Social and Economic Research 2018). Focusing specifically on the London area, one of the most congested cities in the world, I show that wage changes and job mobility are strongly linked to commuting time. In particular, workers are more likely to change jobs via a job-to-job transition the higher is their current commuting time. In addition, workers are more likely to accept wage cuts if their commuting time falls relative to their previous job. Finally, using aggregate UK Census data, I show that commuting flows are positively correlated with proximity to the center of London's economic activity, the City of London.

My main contribution is to develop a novel spatial model of job search with an externality in commuting in order to understand the forces giving rise to these patterns and quantify the effects of congestion. 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 patterns 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. 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). 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, rent, and congestion are jointly determined in equilibrium. By allowing rent to adjust, the model is able to endogenously generate the pattern that the locations with the most productive jobs have both the most congestion and highest rent. By incorporating labor market frictions, the model generates the trade-offs evident in the micro data, and links these individual choices to aggregate congestion. 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

2

price exists in the steady state.[4]

I analytically show in a simple social planner's problem, that in the presence of labor market frictions, equilibrium congestion is not necessarily constrained inefficient. Because moving house or remaining unemployed is costly, given the moving costs and labor market frictions it can be optimal to commute to distant jobs, even after taking into account the effect of congestion on other workers' commuting costs. The higher is the marginal cost of congestion, the higher is the social value of reducing it, leading to an equilibrium in the decentralized economy in which congestion is inefficiently high. By making workers internalize their contribution to the externality, congestion taxes can lead some workers to move closer to their jobs, decreasing congestion for others and in turn improving welfare. Unlike models with frictionless labor markets, these taxes not only affect congestion but will also interact with job search, affecting unemployment, job mobility, and inequality.

My third contribution is quantitative. The model parameters are estimated to match features of the London labor market, taking into account productivity differentials across space in which the city center is most productive. Commuting costs are identified using two important moments. First, I use the share of job offers reported rejected because of the commute from the Federal Reserve Bank of New York's Survey of Consumer Expectations, discussed above. Second, to pin down the effect of congestion, I use data from Google Maps on commuting times during rush hour and in periods of low congestion. In this way, I disentangle the "no traffic" cost of commuting from the effect of the externality.

I use the model to estimate the effect of congestion on welfare and wage and utility dispersion across ex ante identical workers. Following Hall and Mueller (2018), inequality, or dispersion, is measured as the standard deviation of log wages or utility. 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 17% and 13% of wage and utility dispersion, respectively. With congestion, workers have higher average commuting costs, making jobs closer to home more attractive. Thus, it is more likely that workers with long commutes renegotiate, increasing their wages relative to new

---

[4]Halket and Vasudev (2014) use 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.

hires. However, this also implies that on average, workers with the highest wages have the highest commuting costs, resulting in relatively low dispersion in utility, defined as the difference between the wage and commuting and rent costs. Without congestion, commuting is less costly, leading to smaller average gains from renegotiation, and decreasing wage dispersion. At the same time, workers are more likely to make job-to-job transitions without congestion because the set of acceptable job offers is larger, resulting in higher wage growth with labor market experience. Because marginal commuting costs in the center are lower without congestion, high commuting costs offset high wages by less, resulting in a smaller differential in inequality in utils between the models with and without congestion. The model predicts that congestion decreases welfare, measured as the present value of workers at the time of entry into the labor market, by 3.8%.

Finally, I use the model to conduct a policy exercise related to congestion taxes. Because only productive jobs are accepted when the commuting cost is high, congestion and productivity are positively correlated, thus, the effect of the tax is to move workers from more to less productive firms. This leads to a less concentrated distribution of firms across space, causing rent to fall in the city center and rise closer to the periphery. By making job offers requiring a commute less attractive, the set of offers that workers would be willing to accept shrinks, causing the unemployment rate to increase. For the employed, the congestion tax affects wage growth and dispersion. By reducing the variance of commuting costs, the tax decreases wage growth as a function of labor market experience as well as wage and utility dispersion in the cross-section. In terms of magnitude, this tax is equivalent to roughly 0.7% of the monthly wage, and increases welfare by 0.5%. Due to the frictions in the labor market and substantial moving costs in my model, the welfare-maximizing tax rate and associated welfare gains are about one order of magnitude smaller than those in the baseline model of Brinkman (2016), discussed in the following section. The reason is that the benefit of lower congestion comes at the cost of decreasing the attractiveness of jobs located in the congested areas, which also have the highest productivity. With labor market frictions, the congestion benefits are smaller as fewer workers move away from jobs in the congested area, because it takes time to find acceptable jobs. I show that lower frictions, both in terms of higher job offer arrival rates and lower moving costs, lead to larger welfare gains in response to the tax.

4

Related Literature

This is the first paper to consider the interaction between commuting congestion and job search. I contribute to three main 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.

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.[5] 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. A related literature on spatial mismatch is concerned with the effect of distance and urban structure on labor market outcomes, mostly for the unemployed.[6] 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.

One of the few papers to consider on-the-job search and commuting is Van Vuuren (2018). In his model, 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 effects of rent on the labor market outcomes of recent entrants into the labor market. My model allows for both firms and workers to be located across the city and for workers' commutes to congest one another.[7] Another closely related paper is by Manning and Petrongolo (2017), who build a model of spatially directed job search by the unemployed without residential mobility. In their model, workers

---

[5]For comprehensive coverage of urban-labor models, see Zenou (2009). 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).

[6]e.g. Smith and Zenou 2003, Wasmer and Zenou 2006, Manning and Petrongolo 2017. For a review, see Gobillon et al. (2007).

[7]An earlier, more reduced-form model is developed by Van Ommeren et al. (1999), who study the joint determination of job and residential locations when labor and housing markets are frictional.

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 and congestion are much higher 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 a congestion tax.

Commuting has long been recognized as crucial to agents' location and work
decisions in urban economics.[8] Models in this literature typically assume frictionless
labor and housing markets; thus, workers are indifferent across all residential and job
locations in equilibrium.[9] Harari (2017) studies how city shape can impact welfare
and productivity. Recent papers by Tsivanidis (2018) and Heblich et al. (2018)
consider the effects of introducing public transportation infrastructure on welfare
and output. Like these papers, I build a model that captures the spatial variation
in employment and residential density, but focus on how these differences interact
with employment outcomes, taking the transportation technology, city shape, and
city limits as given.

A related literature has focused on 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.[10] Recent experimental results by Kreindler
(2018) show that congestion charges affect commuters' behavior but have only small
welfare benefits. Brinkman (2016) builds a theoretical model to jointly estimate con-
gestion and agglomeration externalities.[11] 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. Abstracting from the
benefits from endogenous agglomeration, I show that the welfare benefits from a
congestion tax are limited in the presence of search frictions and moving costs.

---

[8]e.g. Fujita 1989, and the seminal contributions of Alonso 1964, Mills 1967, and Muth 1969.

[9]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 em-
ployment status of workers. Conley and Topa (2002) and Glaeser et al. (2008) consider mechanisms
that give rise to spatial concentration of unemployment.

[10]e.g. Pigou (1932), Walters (1961), Vickrey (1969).

[11]Earlier papers that model the congestion externality include Cohen (1987), Anas and Kim
(1996), Arnott et al. (1993), and Arnott (2007). Parry and Bento (2001) examine the relationship
between congestion taxes and labor force participation.

The literature on non-wage job characteristics and compensating differentials aims to understand how wages reflect non-wage aspects of jobs.[12] 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.[13] This model is different from the literature studying compensating differentials through the lens of a search model because 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 scenic view from one's office, it is difficult to rationalize for some amenities, especially the commute.

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

---

[12]Frictionless models such as Rosen (1986) predict that wages should compensate for these values, equating utility across workers (see also Van Ommeren et al. 2000). 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 and Bonhomme and Jolivet 2009).

[13]Most models in this literature consider all non-wage amenities rather than focusing on one, with three exceptions being Dey and Flinn (2005), who focus on access to health care, and Pinheiro and Visschers (2015) and Jarosch (2016) who cosider job security.

evidence for the importance of congestion by comparing commuting time at different levels of congestion using data from Google Maps.

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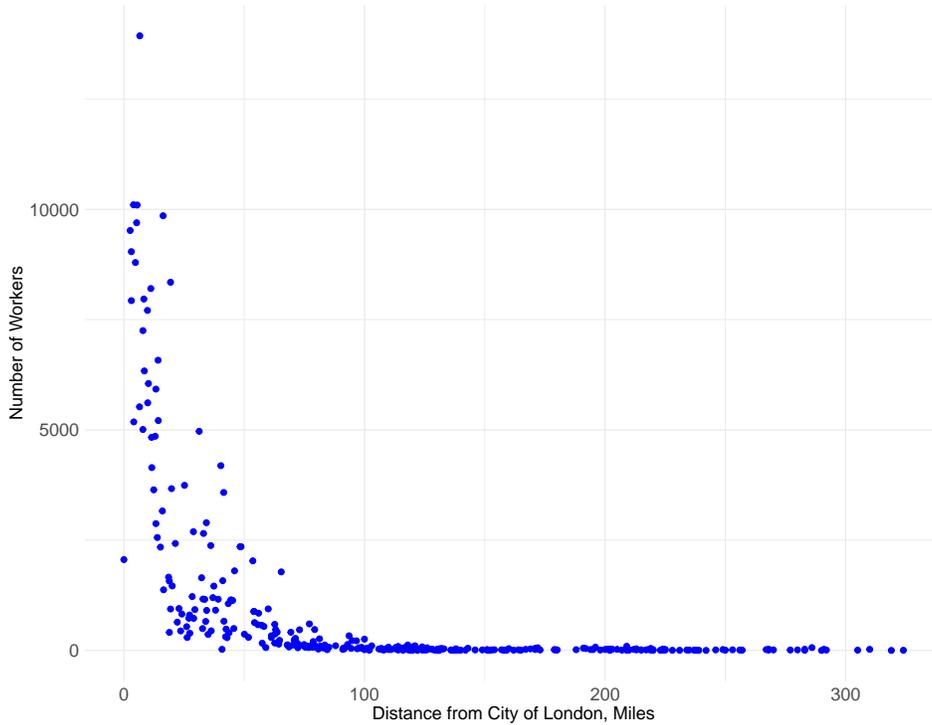
Data: UK Census Flow Data 2001, Special Workplace Statistics (Level 1). Number of workers commuting to the City of London as a function of driving distance in miles computed by Google Maps. Map data: Google, DigitalGlobe.

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 1 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.[14,15] The distance is computed using Google Maps,

---

[14]Data used in Figure 1 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.

[15]LADs 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.

with the location automatically selected by Google for each LAD, where distance is defined as the shortest driving distance between the two locations.[16,17]

On-the-Job Search

To provide evidence on the relationship between job search 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.[18] 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".[19,20] 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-

[16]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.

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

[18]About 5% of LADs in the sample have fewer than 50 observations, results when including these LADs are similar and are available upon request. For the full sample, there are 12,510 individuals; when restricting the sample to individuals living within 100 miles from the City, this falls to 4,276.

[19]I 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.

[20]The 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.

Table 1: Marginal Effects of Lagged Commute and Wage on Job-to-Job Transition Probability

	J2J$_t$	J2J$_t$
Commute$_{t-1}$ × 30	0.015***	0.011***
	(0.003)	(0.003)
Real log Wage$_{t-1}$	-0.035***	-0.041***
	(0.006)	(0.008)
Individual Characteristics		✓
Region, Time, Commute Method Industry & Occ FE	✓	✓
Pseudo R^2	.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 *** $p < .01$.

reported employment history.[21] 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 coefficient in the first row of column 2 can 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.

---

[21]I 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.

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 perform regressions using a linear probability model with individual fixed effects. The results are similar and can be found in Appendix A.

Table 2: After Job-to-Job Transition

		Wage		
		Down	Same	Up
	Down	0.50	0.44	0.34
Commute	Same	0.19	0.18	0.25
	Up	0.31	0.38	0.41
	N	157	103	437

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.

Table 2 shows that decreases in commuting time explain the prevalence of wage

11

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%.[22] 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 workers taking commuting increases and decreases would be equal for each column.

Differently, 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

---

[22]The 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).

Figure 2: Commuting Time to City of London: Effect of Congestion

Maps' recommended travel mode option. Figure 2 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 (dashed) 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 65% 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 ρ 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 ℓ and may be employed or unemployed. Unemployed (employed) workers draw job offers at rate λ_0 (λ_1). A job offer is a pair (ℓ_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 ℓ_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, ℓ and ℓ_F, respectively. Travel time is affected by the extent of the congestion externality, which is determined by the distribution of workers Ω 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 φ, which trigger a draw of the worker's location from

exogenous distribution $\Pi(\ell)$, where $\pi(\ell)$ denotes the probability that the worker draws $\ell \in \mathcal{L}$.[23]

When a worker is employed, she consumes her wage net of commuting and rental costs. Matches are destroyed exogenously at rate δ. When a match is destroyed, the worker becomes unemployed and the firm exits. Unemployed workers consume the flow value of home production less rental costs. All workers have the outside option of leaving the labor market and consuming $\overline{u}$, which is normalized to zero.

3.2 Wage Determination

Employment contracts consist of a wage and worker location, (w, ℓ_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 β. 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.[24] This assumption is further discussed at the end of the next section.

All value functions depend on the endogenous distributions due to the effect of congestion on commuting costs and through job search. As I will be solving the model in steady state, I omit dependence on these distributions for notational convenience. For a given location ℓ, define the value of unemployment as $U(\ell)$. For a given firm location and productivity (ℓ_F, y), worker location ℓ, and wage w, denote the value of an employed worker at the time of the match as $W(w, \ell_F, y, \ell)$. This value takes into account the worker's contracted location ℓ_e:

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

where variables with a tilde are indexed by the worker's location at the next instant. Similarly, the value of a filled job at the time of the match is

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

---

[23]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.

[24]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.

The match surplus is defined as

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

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 (ℓ_F, y, ℓ).[25] In the case of moving shocks, the worker cannot immediately return to her surplus-maximizing location. Define the match surplus after drawing ℓ' 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 in this case, and the match productivity and firm location are fixed, the worker will only move for job-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

$$W(\phi_0, \ell_F, y, \ell) - U(\ell) = \beta S(\ell_F, y, \ell) \tag{4}$$

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', ℓ_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 $\tilde{S}(\ell_F, y, \ell)$ and associated worker value is $\tilde{W}(w, \ell_F, y, \ell)$. 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 $w =$

<hr>

[25] In the quantitative model I allow for indifference by splitting workers evenly across the locations to which they are indifferent.

$\phi_1(y', \ell'_F, y, \ell_F, \ell)$ satisfying:

$$W(\phi_1, \ell'_F, y', \ell) - U(\ell) = \tilde{S}(\ell_F, y, \ell) + \beta\big(S(\ell'_F, y', \ell) - \tilde{S}(\ell_F, y, \ell)\big) \qquad (5)$$

In this case, the firms enter into Bertrand competition, and the worker extracts the full surplus of her previous match and a share β 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:

$$\tilde{W}(\phi_2, \ell_F, y, \ell) - U(\ell) = S(\ell'_F, y', \ell) \qquad (6)$$

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 (ℓ_F, y), who was previously living at ℓ and draws ℓ' 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 the latter three cases, the match continues. 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

$$\tilde{W}(\psi_1, \ell_F, y, \ell') - U(\ell') = 0 \tag{7}$$

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

$$\tilde{W}(\psi_2, \ell_F, y, \ell') - U(\ell') = \tilde{S}(\ell_F, y, \ell') \tag{8}$$

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 λ_0 is

$$(\rho + \chi)U(\ell) = b - r(\ell) + \lambda_0 \sum_{y', \ell'_F} \max\{0, \beta S(\ell'_F, y', \ell)\} g(\ell'_F, y') \tag{9}$$

At rate λ_0 the worker meets a vacancy drawn from distribution G. If the job is acceptable, the worker receives a share β of the surplus.

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

earning wage w, is given by:

$$
\begin{aligned}
(\rho + \chi)\tilde{W}(w, \ell_F, y, \ell) = & \; w - r(\ell) - T(\ell, \ell_F, \Omega) + (\delta + \varphi)\big(U(\ell) - \tilde{W}(w, \ell_F, y, \ell)\big) \\
+ \;& \varphi \sum_{\ell' \in \mathcal{L}} \big[\max\{\min\{\tilde{W}(w, \ell_F, y, \ell') - U(\ell'), \tilde{S}(\ell_F, y, \ell')\}, 0\} + U(\ell') - U(\ell)\big]\pi(\ell') \\
+ \;& \lambda_1 \sum_{(y', \ell'_F) \in B_1(y, \ell_F, \ell)} \big(\beta S(\ell'_F, y', \ell) + (1-\beta)\tilde{S}(\ell_F, y, \ell) - \tilde{W}(w, \ell_F, y, \ell) + U(\ell)\big)g(\ell'_F, y') \\
+ \;& \lambda_1 \sum_{(y', \ell'_F) \in B_2(w, y, \ell_F, \ell)} \big(S(\ell'_F, y', \ell) - \tilde{W}(w, \ell_F, y, \ell) + U(\ell)\big)g(\ell'_F, y') \tag{10}
\end{aligned}
$$

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 δ. At rate φ 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

$$
\begin{aligned}
(\rho + \chi)\tilde{J}(w, \ell_F, y, \ell) = & \; y - w - r(\ell_F) - (\delta + \varphi)\tilde{J}(w, \ell_F, y, \ell) \\
+ \;& \varphi \sum_{\ell' \in \mathcal{L}} \max\{\min\{\tilde{J}(w, \ell_F, y, \ell'), \tilde{S}(\ell_F, y, \ell')\}, 0\}\pi(\ell') - \lambda_1 \tilde{J}(w, \ell_F, y, \ell) \sum_{(y', \ell'_F) \in B_1(y, \ell_F, \ell)} g(\ell'_F, y') \\
+ \;& \lambda_1 \sum_{(y', \ell'_F) \in B_2(w, y, \ell_F, \ell)} \big(\tilde{S}(\ell_F, y, \ell) - S(\ell'_F, y', \ell) - \tilde{J}(w, \ell_F, y, \ell)\big)g(\ell'_F, y') \tag{11}
\end{aligned}
$$

where the first three terms are the match output, the wage that the firm pays its worker, and the firm's rent. At rate δ 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 φ, the worker receives a moving shock and 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 firm's payoff if the worker and firm renegotiate.

19

Combining these expressions gives the match value:

$$M(\ell_F, y, \ell) = \max_{\ell' \in \mathcal{L}} \left\{ \tilde{M}(\ell_F, y, \ell') - k_M \mathbb{1}\{\ell \neq \ell'\} \right\} \tag{12}$$

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

$$
\begin{aligned}
(\rho + \chi)\tilde{M}(\ell_F, y, \ell) &= y - r(\ell) - r(\ell_F) - T(\ell, \ell_F, \Omega) - (\delta + \varphi)\tilde{S}(\ell_F, y, \ell) \\
&+ \varphi \sum_{\ell' \in \mathcal{L}} (\max\{0, \tilde{S}(\ell_F, y, \ell')\} + U(\ell') - U(\ell))\pi(\ell') \\
&+ \lambda_1 \beta \sum \max\{0, S(\ell'_F, y', \ell) - \tilde{S}(\ell_F, y, \ell)\}g(\ell'_F, y')
\end{aligned} \tag{13}
$$

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 $\tilde{S}(\ell_F, y, \ell) = \tilde{M}(\ell_F, y, \ell) - \tilde{U}(\ell)$.

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.

20

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 (ℓ_F, y) living at ℓ as $e(\ell_F, y, \ell)$ and the mass of unemployed workers living at ℓ as $u(\ell)$. The distribution of workers is summarized by Ω.

The flow into employment for workers living at ℓ employed in matches located at ℓ_F with firm productivity y is made up of three groups. The first is the unemployed accepting an offer from a firm (ℓ_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 ℓ after receiving a moving shock, but remain employed with their firm (y, ℓ_F). In sum, the flow into such matches is given by:

$$
\begin{aligned}
e^+(\ell_F, y, \ell) &= g(\ell_F, y) \sum_{\ell' \in \mathcal{L}} \big[\lambda_0 u(\ell') \mathbb{1}\{ S(\ell_F, y, \ell') > 0 \} \\
&+ \lambda_1 \sum_{\ell'_F \in \mathcal{L}} \sum_{y' \in Y} \mathbb{1}\{ \tilde{S}(\ell'_F, y', \ell') < S(\ell_F, y, \ell') \} e(\ell'_F, y', \ell') \big] \mathbb{1}\{ \ell_e(\ell_F, y, \ell') = \ell \} \\
&+ \varphi \pi(\ell) \mathbb{1}\{ \tilde{S}(\ell_F, y, \ell) > 0 \} \sum_{\ell' \neq \ell} e(\ell_F, y, \ell')
\end{aligned}
\tag{14}
$$

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

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

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

$$
u^+(\ell) = \chi \mu(\ell) + \delta \sum_{\ell_F \in \mathcal{L}} \sum_{y \in Y} e(\ell_F, y, \ell) + \varphi \pi(\ell) \sum_{\ell'} \sum_{\ell_F \in \mathcal{L}} \sum_{y \in Y} e(\ell_F, y, \ell') \mathbb{1}\{ \tilde{S}(\ell_F, y, \ell) \leq 0 \}
\tag{16}
$$

The flow out of unemployment for workers living at ℓ is simply those unemployed

workers who successfully match or exit the economy:

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

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:

$$\sum_{y \in Y} \left(\sum_{\ell \in \mathcal{L}} e^+(\tilde{\ell}, y, \ell) + \sum_{\ell_F \in \mathcal{L}} e^+(\ell_F, y, \tilde{\ell}) - g(\ell_F, y)\left[\lambda_0 u(\tilde{\ell})\mathbb{1}\{S(\ell_F, y, \tilde{\ell}) > 0\}\right. \right.$$

$$+ \quad \lambda_1 \sum_{\tilde{\ell}_F \in \mathcal{L}} \sum_{\tilde{y} \in Y} \mathbb{1}\{\tilde{S}(\tilde{\ell}_F, \tilde{y}, \tilde{\ell}) < S(\ell_F, y, \tilde{\ell})\} e(\tilde{\ell}_F, \tilde{y}, \tilde{\ell}) \left. \right] \mathbb{1}\{\ell_e(\ell_F, y, \tilde{\ell}) = \tilde{\ell}\} \left. \right)$$

$$+ \quad u^+(\tilde{\ell}) - \delta \sum_{\ell_F \in \mathcal{L}} \sum_{y \in Y} e(\ell_F, y, \tilde{\ell}) \tag{18}$$

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:

$$\sum_{y \in Y} \left(\sum_{\ell \in \mathcal{L}} e^-(\tilde{\ell}, y, \ell) + \sum_{\ell_F \in \mathcal{L}} \left(e^-(\ell_F, y, \tilde{\ell}) \right. \right.$$

$$- \quad e(\ell_F, y, \tilde{\ell})\left[\delta + \lambda_1 \sum_{\ell'_F \in \mathcal{L}} \sum_{y' \in Y} \mathbb{1}\{S(\ell'_F, y', \tilde{\ell}) > \tilde{S}(\ell_F, y, \tilde{\ell})\}\mathbb{1}\{\ell_e(\ell'_F, y', \tilde{\ell}) = \tilde{\ell}\} g(\ell'_F, y')\right] \left. \right) \left. \right)$$

$$+ \quad u^-(\tilde{\ell}) - \lambda_0 u(\tilde{\ell}) \sum_{\ell_F \in \mathcal{L}} \sum_{y \in Y} \mathbb{1}\{S(\ell_F, y, \tilde{\ell}) > 0\}\mathbb{1}\{\ell_e(\ell_F, y, \tilde{\ell}) = \tilde{\ell}\} g(\ell_F, y) \tag{19}$$

3.5 Congestion

Recall that $d(\ell, \ell_F)$ is the distance in miles between a worker located in ℓ and a firm located in ℓ_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, $d(\ell, \ell) > 0$.

Similarly to Brinkman (2016), I define congestion as the share of commuters passing through a location less the number of employed workers living in that lo-

cation. I consider all of the pairs of locations (ℓ, ℓ_F) and compute the "commuting path," defined as the locations through which a worker living at ℓ must travel in order to get to ℓ_F. Letting $cp(\ell_F, \ell)$ denote this path and imposing symmetry[26] implies $cp(\ell, \ell_F) = \{\ell' \in \mathcal{L} : \min\{\ell, \ell_F\} \leq \ell' \leq \max\{\ell, \ell_F\}\}$. Congestion in a given location $\tilde{\ell}$ is given by the difference between commuters and residents:

$$c(\tilde{\ell}) = \sum_y \left(\sum_{(\ell_F, \ell): \tilde{\ell} \in cp(\ell, \ell_F)} [e(\ell_F, y, \ell) + e(\ell, y, \ell_F)] - \sum_{\ell_F} e(\ell_F, y, \tilde{\ell}) \right) \quad (20)$$

Note that a worker who lives and works in the same location does not add to congestion; similarly, the unemployed have no effect on congestion.

The total cost of commuting between ℓ and ℓ_F is defined as

$$T(\ell, \ell_F, \Omega) = \left(\tau d(\ell, \ell_F) + \kappa \sum_{\ell' \in cp(\ell, \ell_F)} c(\ell') \right) b$$

where τ is the "no traffic" cost of commuting an extra unit of distance and κ is the marginal congestion cost. The summation computes all of the congestion that a worker encounters by commuting between ℓ and ℓ_F; in this way commuting costs depend on the distribution Ω. In the case where there is no externality, $\kappa = 0$. The cost of time to the worker is proportional to the opportunity cost of time, given by the value of home production b.

3.6   Efficiency

I use the quantitative model below to evaluate the welfare implications of congestion taxes, 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 I show in a simple static framework that the decentralized equilibrium need not align with that of a social planner when moving is costly. I consider workers who receive job offers from two locations, and upon receiving an offer from a location different from their own, must decide whether to commute or pay a moving cost. The social value of some workers moving to the location of their jobs includes the net benefits for all remaining commuters from less congestion plus the output gain from any additional match creation. Because moving is costly, individual matches may find it privately optimal to commute, failing

---

[26]i.e. $cp(\ell_F, \ell) = cp(\ell, \ell_F)$.

to take into account their effect on congestion. I show that for any sufficiently large moving cost, there exists an interval of values for the marginal cost of congestion such that it is socially beneficial but privately suboptimal to have workers move to the location of their jobs rather than commute. When rent is exogenous, I derive the optimal tax rate to align the incentives of the planner and workers in the decentralized economy, which holds under a sufficient condition for the difference in rent between the two locations. The optimal tax is proportional to the marginal congestion cost, κ, and equal to the share of workers who benefit from the reduction in congestion, which workers in the decentralized economy do not take into account when making their moving decisions. The tax leads workers to internalize the positive externality they create from reducing congestion. Using this intuition, I consider linear congestion taxes in Section 5.

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 $\ell \in \mathcal{L}$ contains total land L.[27] 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:

$$\hat{L}(\ell) = \underbrace{L(\ell)}_{\text{total land}} - \sum_y \sum_{\tilde{\ell}} \underbrace{\left(e(\ell, y, \tilde{\ell}) - e^-(\ell, y, \tilde{\ell})\right)}_{\text{existing filled jobs}}$$

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

$$\max_{\zeta(\ell) \in [0, \hat{L}(\ell)]} \nu\big(\zeta(\ell), r(\ell)[L - \zeta(\ell)]\big) \tag{21}$$

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 ν 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, remaining land demand is determined by the total mass of employed and unemployed workers residing in ℓ and newly filled jobs located in ℓ. Equilibrium in the land market requires the supply of land be equal to the demand of for land:

$$L - \zeta(\ell) = u(\ell) + \sum_{y \in Y} \sum_{\tilde{\ell} \in \mathcal{L}} e(\tilde{\ell}, y, \ell) + e(\ell, y, \tilde{\ell}) \tag{22}$$

I now define a steady state equilibrium.

Definition 1. *A steady state equilibrium consists of value functions $M : \mathcal{L} \times Y \times \mathcal{L} \to \mathbb{R}$ 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 $\ell_e(y, \ell_F, \ell)$ 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 congestion 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.*

The proof is shown in Appendix B. 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 both the demand correspondence for land as well as the mass of workers to whom that demand correspondence belongs are both endogenous functions of the price. The aggregate 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.

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

26

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 3. As discussed in Section 2, I calibrate the model to match the metropolican 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 a large share of workers commute within the zone's boundaries, is discussed in Appendix D. The reason for using the 100 mile radius is quantitative: in the alternative definition, many observations from the BHPS data are lost, and the empirical moments used in the calibration are much more noisy. Given two locations ℓ and ℓ_F, the distance in miles is $d(\ell, \ell_F) = |\ell - \ell_F| + 10\mathbb{1}\{\ell = \ell_F\}$. In words, this functional form for distance implies that the round trip distance between the worker and firm in the same location ℓ is 10 miles. The quantitative results below are robust to changes in this value.

Figure 3: Locations

| -200 | -133 | -67 | 0 | 67 | 133 | 200 |
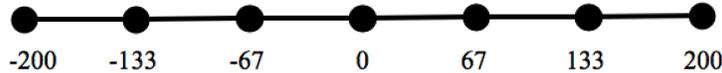
Regarding the other distributional assumptions, the location and productivity of vacancies is drawn from the joint distribution G. The location distribution is assumed to be symmetric around the center, where the probability of drawing a location, from left to right, is given by $[g(1), g(2), g(3), g(4), g(3), g(2), g(1)]$, with $g(4) + 2\sum_{i=1}^{3} g(i) = 1$. The distribution for y is conditional on the draw of the firm location ℓ_F and drawn from a Normal distribution with mean $1 + A(\ell_F)$ and standard deviation σ_y, discretized using the Tauchen method to span two standard deviations from the mean in each location. Finally, the location distribution for newborn workers (μ) is equal to the distribution after moving shocks (π) 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 ρ is set to match an annual interest rate of 5% and the death rate χ is chosen to match an average working lifetime[28] 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, standard errors are omitted. The standard labor market parameter values and their targets are shown in Tables 3 and 4, respectively, and are discussed below.

The parameters related to transitions are δ, φ, λ_0, λ_1, and β. The separation rate into unemployment (EU rate) is driven by exogenous separations, occurring at rate δ, and endogenous separations which may occur after the arrival of a moving shock φ. The unemployment to employment (UE) rate is mainly driven by the arrival rate of offers for the unemployed, λ_0. The job-to-job transition (EE) rate is primarily driven by λ_1. Together with the variance of match-specific productivities, bargaining power β determines the average wage growth between jobs. Because φ is a spatial parameter, I leave its discussion to the next section.

The standard deviation of productivity, σ_y, is chosen to match the variance of the observed distribution of year-on-year wage growth.[29] Targeting the variance of wage growth rather than the variance of wages controls for individual fixed effects possibly present in the data. 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.[30]

---

[28]The working lifetime corresponds to ages 24-55, the same range considered in the empirical results of Section 2.

[29]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.

[30]The benefits of unemployment include the average commuting costs saved from not working,

Table 3: Standard Parameter Values

Parameter		Value
Discount Rate	ρ	.0041
Exit rate	χ	.0027
Bargaining Power	β	.21
Home Production	b	.68
Separation Rate	δ	.004
Arrival Rate, Unemployed	λ_0	.128
Arrival Rate, Employed	λ_1	.100
Standard Deviation, Productivity	σ_y	.045

Table 4: Targets: Standard Parameters

Moment	Data	Model
Annual interest rate	.05	.05
Years in LF	31	31
Average wage growth, J2J	.065	.065
Replacement rate (Hall and Milgrom 2008)	.71	.70
EU rate (Q, Gomes 2012)	.014	.015
UE rate (Q, Gomes 2012)	.26	.28
EE rate (Q, Gomes 2012)	.027	.025
Variance, wage growth	.039	.039

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 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. The remaining values for A are symmetric around the center and 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 5 miles of the City. Lower rent in the

---

for a discussion of a similar point see Bils et al. (2012).

periphery offsets the effect of lower productivity, 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 27% at the same distance. Here, workers have an option value of remaining unemployed, therefore there is a trade-off between low A and high unemployment.

The choice of φ 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 φ implies a non-job related move once every 22 years, or 1.4 times 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, which is about 20% larger than the present value of the average surplus.[31]

Landlords' preferences are assumed to be CES:

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

where $\omega_L \in (0, 1)$ is the weight on the landlord's consumption and σ_L is the elasticity of substitution between consumption and rental income. The parameter ω_L is chosen to match an average monthly rent to wage income ratio and σ_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, 2.[32]

Recall that the commuting cost function is

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

The parameter τ determines the marginal disutility from commuting in the absence of the externality (the "no traffic" cost) and κ is the marginal 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

---

[31]This large value is qualitatively similar to the literature: Kennan and Walker (2011) estimate that the cost of moving house in the US is as high as $300,000.

[32]As long as L is constant, a change in L would be equivalent to a rescaling of the parameter ω_L.

Table 5: Spatial Parameter Values

Parameter		Value
Moving cost	k_M	23.9
No-traffic commuting cost	τ	.004
Congestion Cost	κ	.79
Moving rate	φ	.004
Landlords' preferences	ω_L	.195
Landlords' preferences	σ_L	.97
Location-specific productivity, 5-33 mi	$A(3)$	.63
Location-specific productivity, 34-66 mi	$A(2)$	.55
Location-specific productivity, 67-100 mi	$A(1)$	.45
Vacancy location probability, 5-33 mi	$g(3)$	.081
Vacancy location probability, 34-66 mi	$g(2)$	.174
Vacancy location probability, 67-100 mi	$g(1)$	.173

Table 6: Targets: Spatial Parameter Values

Moment	Data	Model
Moving probability (A)	.045	.041
Share of offers rejected for commute	.13	.15
Relative increase in commuting cost with congestion	.65	.65
Share of job-related moves	.129	.123
Avg rent/income	.365	.372
Elasticity of supply (Barker, 2003)	.3	.3
Relative wage, residence 5-33 mi	.88	.93
Relative wage, residence 33-66 mi	.87	.88
Relative wage, residence 66-100 mi	.80	.85
Jobs density, 5-33 mi from center	.83	.87
Jobs density, 33-66 mi from center	.80	.86
Jobs density, 66-100 mi from center	.78	.76

share of job offers rejected because of the commute in the model is the share of offers with higher productivity that a worker rejects; in the data this corresponds to the question in the Federal Reserve Bank of New York's Survey of Consumer Expectations regarding the main reason for rejecting a job offer. To pin down κ, I use the average commuting time with and without congestion from Google Maps shown in Figure 2. 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, κ. The value of κ can be interpreted as follows: when 10% of workers in the economy congest a location, commuting costs increase by $0.1\kappa b = 0.05$ units of output, or roughly 5% of the average wage.[33]

4.2 Properties of the Calibrated Economy

In this section I evaluate the model's performance against the data. Table 7 compares several untargeted moments. 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 channels 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, leading to high expected wage growth, 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. Each of these channels contributes roughly half of the wage cuts in the model. The share of movers in the model who remain in the same job is due to the moving shock φ. This moment is lower than in the data, possibly due to other voluntary reasons for moving house not present in the model. The next two lines show average wages for new hires out of unemployment relative to all workers and wage growth within jobs. The model captures these magnitudes well.

In the data, the share of workers residing in the periphery (locations more than 66 miles one-way from the City of London) making job-to-job transitions is less than for workers living in the center (within 5 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 center are more likely to work in the most productive jobs. Because there are

---

[33]The average wage in the calibrated model is 1.13.

Table 7: Untargeted Moments

Moment	Data	Model
Share of wage cuts	.16	.25
Share of movers who stay in the same job	.018	.011
Relative wage, new hires from U	.87	.83
Average wage growth, within job	.016	.021
Annual J2J transition rate, residence in center	.147	.068
Annual J2J transition rate, residence in periphery	.105	.153
Relative Moving Distance, J2J to All Movers	1.39	1.38

a finite number of productivity levels, these workers reach the top of the productivity "ladder" quickly, after which they no longer change jobs. 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 4: Commuting Distance by Place of Residence
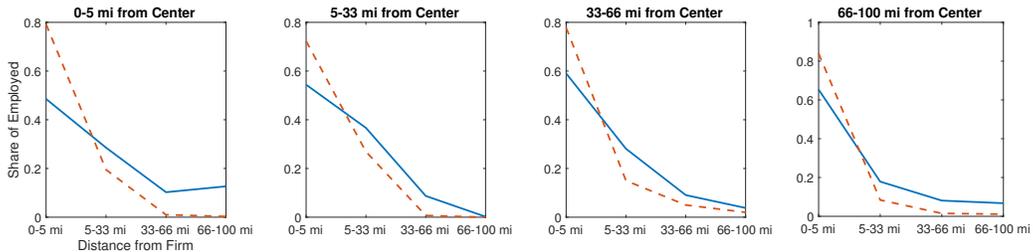
Figure 4 shows the commuting patterns by place of residence in the model and data. The horizontal axis corresponds to the distance from the place of residence, and the vertical axis shows the share of commuters living in each location and commuting a given distance. In all four locations, the model matches the fact that the farther the commuting distance, the lower is the share of commuters. However, the

model underestimates the share of workers living and working in the same location, due to the random arrival of job offers across space and the fact that commuting costs capture only the cost of time, and not the monetary and psychological costs.[34]

Figure 5 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 14% 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 gradient is similar, though slightly steeper, than the rent gradient in the data.

Figure 5: Workers and Firms

(a) Residential Locations

(b) Firm Productivity

---
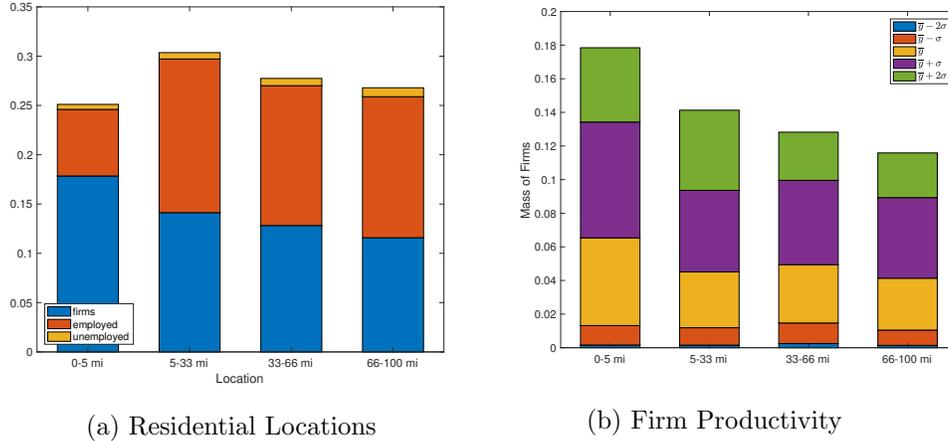
[34]One feature that would improve the model's performance here would be a location choice for new entrants, rather than the assumed exogenous distribution of newborn workers. Because welfare is on average higher in the periphery, this would increase the share of workers living and working in the center.

4.3 Effect of Congestion

This section examines the effect of commuting and congestion on labor market outcomes. Figure 6 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.

Figure 6: On-the-Job Search Outcomes
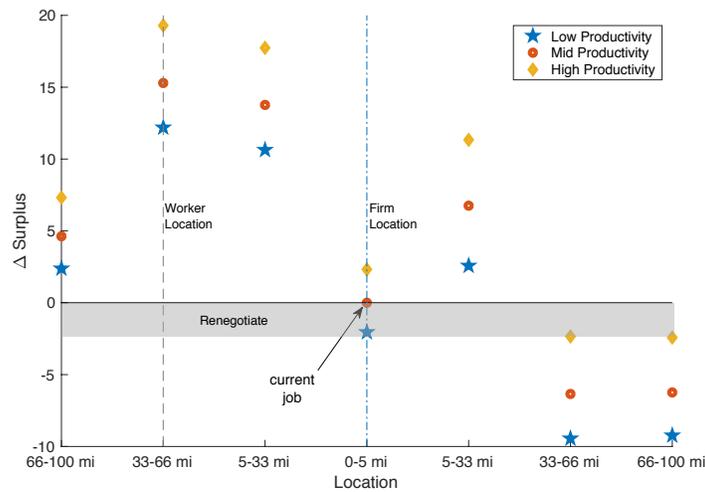
Figure 7 compares aggregate outcomes in the full model to the model without the externality, setting $\kappa = 0$. The top row of the figure shows the changes in average commuting costs, average wages, and welfare by worker location. The top left

panel shows that commuting costs are lower without the externality. The difference between the two lines is the effect of congestion on commuting costs. 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 about 50% higher without congestion, but this varies widely depending on the worker's location. Because congestion is highest near the city center, wages increase to compensate workers for their more costly commutes. Across all locations, welfare, measured as the expected value at entry into the labor market, is 3.8% higher in the no externality model. The increase in welfare is highest for workers in the periphery, since the set of acceptable job offers was decreased the most by the high congestion closer to the city center, due to the long commuting distance.

Figure 7: Aggregate Patterns Across Space
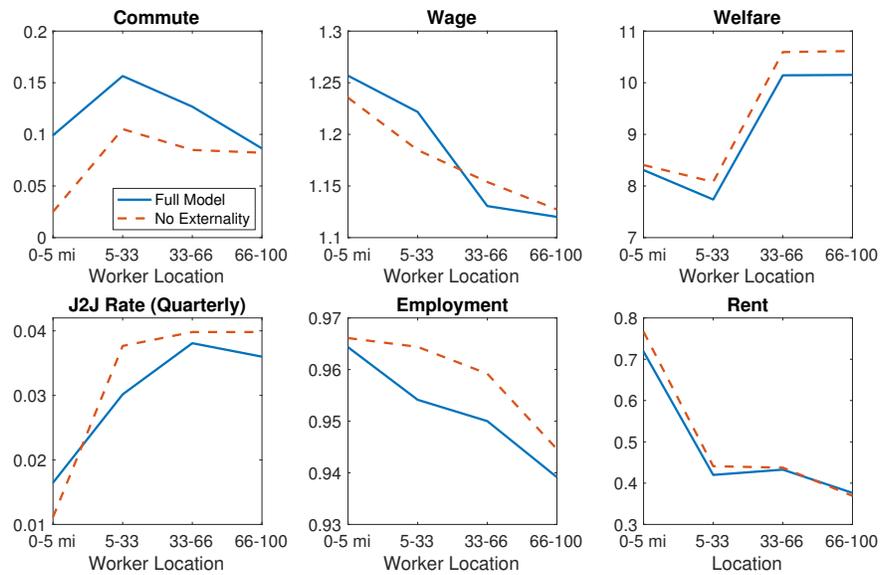
In the second row of Figure 7, I present the job-to-job transition rate, employment, and rent. The externality causes the job-to-job transition rate to decrease because workers are less likely to accept distant offers. With congestion, unemployed workers reject many more job offers and therefore are less likely to accept offers once employed, making the job-to-job transition rate for workers living in most location lowers. This rate is lowest closest to the center because workers living in the center are also more likely to work in the center (over 50% in steady state,

36

relative to just 6% of workers living 100 miles from the center). Since these jobs are the most productive, workers are less likely to accept outside offers. Employment increases in all locations in the absence of congestion, as the flow value of all matches increases relative to the value of unemployment. The change in employment is smallest for workers living in the center, but the gain from decreased congestion is largest, making jobs in the center relatively more attractive and decreasing the job-to-job transition rate for workers residing there. Finally, rent increases in locations within 33 miles of the center due to less congestion and more demand for land from both workers and firms. Indeed, in the full model congestion and rent both make the city center less attractive: without congestion, equilibrium rent would increase far more in the center in order to clear the land market. With congestion, the center is less attractive as a workplace from the perspective of commuting workers and thus fewer jobs are created in the center, decreasing firms' demand for land and therefore rent.

Figure 8 plots average wage growth with experience in the models with and without the externality. Without congestion, wage growth is twice as large as in the full model. By increasing the share of acceptable job offers, as shown in Figure 7, wage growth increases as workers make more job-to-job transitions for which they receive significant wage gains on average. In addition, the average wage level falls, especially for workers living in the city center, leading to higher wage growth. After roughly 8 years of labor market experience, wages become flat in both models, as the model has no skill accumulation or returns to experience and the job ladder has finitely many steps. The next section shows that although wage growth with experience is higher in the model without the congestion externality, cross-sectional dispersion is lower without congestion.

4.4   Wage and Utility Dispersion

As highlighted by Hornstein et al. (2011), there is a large amount of wage dispersion across otherwise identical workers in the data. My model highlights the role of commuting costs and congestion for job search, giving rise to a non-wage amenity that depends on a negative externality. This externality is largest in the most productive areas, a channel which can give rise to high frictional wage dispersion.

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 ($\kappa = 0$) and the model with

Figure 8: Wage Growth with Labor Market Experience
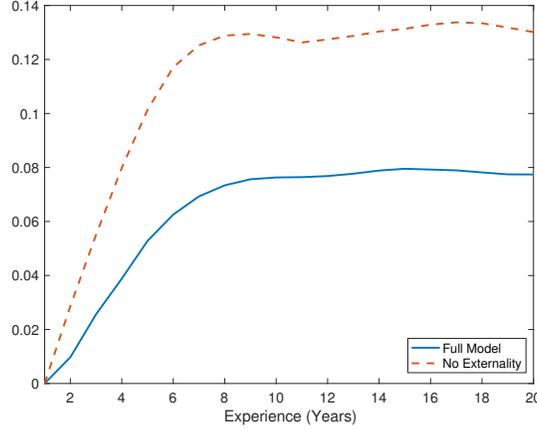
zero commuting costs ($\kappa = 0$ and $\tau = 0$).

Comparing the first two rows in the first column, wage dispersion is 17% 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.

Table 8: Standard Deviation: Wage and Utility

Model	$\ln w$	$\ln w - \ln T(\ell, \ell_F)$	$\ln w - \ln(T(\ell, \ell_F) + r(\ell))$
Full	.197	.806	.257
No Congestion	.164	1.04	.224
No Commuting Cost	.038	-	.109

Before taking into account differences in housing rent, utility dispersion is much larger than wage dispersion, and is higher in the no congestion model. This is especially large because the model fails to capture the positive correlation between wages and commuting times.[35] The reason for this is that the dynamic sorting from

----

[35]This correlation is .14 in the data, and -.01 in the calibrated model.

the job ladder is strong enough to reverse the static trade-off between wages and commutes: for newly hired workers out of unemployment, the correlation between wages and commutes is positive, but as workers climb the job ladder, they move toward jobs with short commutes and high productivity, causing the cross-sectional correlation to turn (slightly) negative.

Without the externality, the correlation between wages and commuting times becomes more negative, leading to a larger dispersion in wages net of commuting costs relative to the full model. After taking rent into account utility dispersion is 13% 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 compensated for long commutes with low rent prices (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. However, the model predicts that the increase in dispersion as measured by utility relative to wages is lower in the full model (31%) than in the model without congestion (36%), since congestion discourages workers from accepting jobs with the longest commutes, compressing differences in utility.

5 Counterfactual Exercise: Congestion Taxes

This section examines the effects of a linear congestion tax on welfare. This model could be used for many policy experiments, for instance the effects of infrastructure investment, remote working arrangements, or flexible working time.[36] I consider a congestion tax to understand the labor market implications of the simplest policy directly targeting the externality, as well as to follow the intuition of the simple

---

[36]Regarding infrastructure, Duranton and Turner (2011) show that improvements in roads have little effect on congestion, because better infrastructure makes it easier to commute, causing an increase in commuting distance and in congestion. Bloom et al. (2014) empirically evaluate the benefits of remote working for call center employees. Arnott et al. (1993) consider workers' choice of the time of departure to avoid peak-period congestion.

planner's problem discussed in Section 3.6 and Appendix C. The welfare measure corresponds to the expected value of unemployment at entry, defined as the weighted average of the value of unemployment, with weights corresponding to the location distribution for newborn workers. I compare the results for the baseline model to two parameterizations highlighting the additional frictions present in the model: one with lower labor market frictions and the other with lower moving costs. These models respectively correspond to doubling the arrival rates of job offers λ_0 and λ_1, and decreasing the moving cost k_M by half. The congestion tax, t increases commuting costs through congestion:

$$T(\ell, \ell_F, \Omega) = \left(\tau d(\ell, \ell_F) + (1 + t)\kappa \sum_{i \in cp(\ell, \ell_F)} c(i) \right) b$$

Revenues from the tax are rebated lump sum to workers, and are computed as a fixed point consistent with the distribution of workers in response to the tax.

The introduction of a tax on congestion and lump sum transfer redistributes resources from workers in matches with more congested commutes to unemployed workers and those with less congested commutes. The tax leads workers to accept less productive jobs with shorter commutes on average, since productivity is highest in the most congested location, the city center. Because moving is costly, most workers find jobs closer to home rather than moving closer to their workplace, resulting in the steady state distribution of firms becoming flatter as the tax increases. By decreasing the probability that a worker accepts a job that requires a commute, the congestion tax changes the shape of the job ladder: in the limit, workers accept only jobs in their own location and evaluate offers based only on their productivity (net of moving costs), decreasing the job-to-job transition rate and increasing unemployment. By the definition of congestion, tax revenues in this limiting case are equal to zero.

Because solving a dynamic planner's problem is beyond the scope of this paper, I determine the optimal tax by varying the tax rate and computing welfare. Figure 9 graphically shows the results for the three models. I first describe the welfare effect in the benchmark model, shown by the solid line. An increase in the tax causes the average commuting distance to fall. In addition, the unemployment rate rises as workers reject more offers requiring a long commute. Together, these two effects decrease the share of workers paying the tax: at high levels of the tax, this "tax base" effect becomes large enough that the welfare gains begin to decrease in

Figure 9: Change in Welfare, Congestion Tax
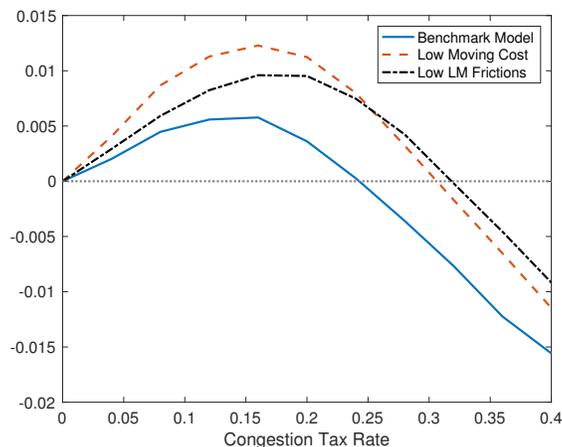
the tax rate. This is due to the fact that as the tax increases the job acceptance threshold in increases for all workers: the normality of the match-specific productivity distribution in each location implies that for a large enough tax, the share of offers rejected will rise faster than the tax rate, causing tax revenues to fall. Higher taxes have the desired effect of decreasing congestion, due to fewer commuters both because of higher unemployment and a larger share of workers living and working in the same location. The decrease in congestion benefits those workers continuing to commute, but does not fully offset the negative effect of higher unemployment, leading to a decrease in average welfare for higher levels of the tax.

As labor market frictions decrease (i.e. the arrival rates λ_0 and λ_1 increase), the unemployment rate falls and workers more easily find jobs that are highly productive and have short commutes. Therefore, the level of congestion is lower and output is higher than in the baseline model. For a given tax rate, revenues are lower in the model with low labor market frictions than in the baseline model, since the tax base is larger but congestion is significantly lower. In the model with lower labor market frictions, shown by the dash-dotted line, the welfare-maximizing tax rate is slightly higher because the workers are more able to adjust to the tax, and the unemployment rate increases by less for a given tax rate since job offers arrive more frequently.

Lastly, the welfare effects in the model with low moving costs are shown by the dashed line in the figure. With lower moving costs, an increase in the congestion

41

tax induces more workers to move to the location of their job, quickly decreasing congestion but also the tax base. When it is easier to move to the location of any job offer, congestion taxes induce more workers to move to the most productive jobs rather than rejecting offers located in congested areas. Overall, the welfare gains are largest in the model with low moving costs, followed by the model without labor market frictions and the benchmark model.

Table 9: Steady State Effects of a 16% Congestion Tax

	Baseline	Low Labor Market Frictions	Low Moving Cost
Increase in Commute, Congestion	-4.3 pp	-4.0 pp	-4.7 pp
Average y	-3.2%	-1.9%	-0.5%
Share of Jobs, Center	-1.0 pp	-0.8 pp	-0.5 pp
Share of job-related moves	-0.3 pp	-1.9 pp	+0.8 pp
EE Rate	+0.2 pp	+0 pp	-0.1pp
UE Rate	-3.3 pp	-0.5 pp	-0.6pp
EU Rate	+0 pp	+0 pp	+0 pp
Unemployment Rate	+0.4 pp	+0.1 pp	+0.1 pp
Average Wage	+0.3%	+0.1%	-0.6%
Wage Growth, J2J	-0.8 pp	-0.6 pp	-0.8 pp
$\sigma \ln w$	-0.8%	-0.7%	-2.9%
$\sigma \ \ln w - \ln(T(\ell, \ell_F) + r(\ell))$	-2.4%	-1.5%	-3.0%
% Change Welfare	0.53%	0.92%	1.21%
Tax Revenues/Total Output	.008	.006	.008
Average Tax/Wage	.007	.008	.011

Welfare is measured as the expected value of unemployment at entry. Columns correspond to the changes in observed variables for the baseline calibration, the model with a 50% lower moving cost, and the model with 2 times the level of job offer arrival rates, relative to the corresponding models with tax rate equal to zero. The tax is equal to the welfare-maximizing tax rate in the benchmark model, shown in Figure 9. "pp" denotes the percentage point change.

Table 9 shows the effects of the congestion tax on steady state labor market outcomes. The first column compares the baseline model under the welfare-maximizing tax rate of 16% to the corresponding moments under no tax, where the parameter values are set to those in Tables 3 and 5. The second and third columns compare

the same moments in the model with lower labor market frictions and with lower moving costs under a 16% tax to the corresponding parameterizations under no tax.

The first row of the table shows the direct effect of the tax: the average increase in commuting costs due to congestion, the moment targeted by the congestion cost parameter κ in Section 4.1, falls by nearly 5 percentage points under the tax. Most of this decrease comes from the city center, which has the highest congestion and therefore the largest response in terms of fewer workers traveling through that location. However, this comes at a cost of lower average output as fewer workers accept jobs in and near the city center. The fall in average output is the result of two offsetting effects: first, the share of jobs in the center falls, leading to a productivity loss. Second, the higher the tax, the more workers prefer to work close to home, and therefore more jobs with the highest location-specific productivity are created, causing average output in the periphery to increase. Because the city center is far more productive than the periphery, this leads to a decline in output.

The largest effect on the transition rates comes through a lower unemployment to employment rate, which falls by over 3 percentage points, causing the unemployment rate to increase. Job-to-job transitions (EE) increase slightly as workers accept more jobs in the periphery, where vacancies are more likely to be located. The separation rate into unemployment is unaffected because very few endogenous separations occur due to the low arrival rate of moving shocks. Average wages slightly increase because fewer workers accept wage cuts when making job-to-job transitions. Wage growth following job-to-job transitions falls due to lower average productivity gains when changing jobs: because workers are more likely to live and work in the same location, the gain that they get in match surplus when switching jobs is small because of the small variance of the productivity distribution. Finally, dispersion in both wages and utility fall because the congestion tax compresses the distribution of match surplus by taxing the most productive jobs relatively more due to location specific productivity differences. Dispersion in utility falls by more than dispersion in wages because the tax decreases the accepted wage variance as well as leads to a flatter rent gradient across space by decreasing the desirability of the city center. Together, these lead to significantly lower dispersion in utility.

With lower labor market frictions, workers can more quickly climb the job ladder, leading to lower unemployment and congestion relative to the baseline model. The effect of the tax is therefore to decrease average congestion, albeit from a lower level, by less than in the baseline model, because workers strictly prefer high pro-

43

ductivity jobs in the center and a small tax rate does not significantly change the ranking of accepted job offers. Because workers are less sensitive to the tax, average output falls by less, and the unemployment to employment rate falls by less, causing only a small increase in the unemployment rate. Average wages slightly increase because more workers pay higher commuting costs with the tax, leading to larger out of pocket costs and higher wages. Finally, dispersion in wages falls by a similar magnitude as in the baseline model, while the effect of higher commuting costs through congested locations causes utility dispersion to fall by less than in the baseline. The welfare gains with lower labor market frictions are larger than in the baseline model because average output and unemployment fall by less, despite the smaller gains from reducing congestion.

The last column shows the results for the model with lower moving costs. Here, workers are more able to move for job-related reasons, indeed, the moving rate and the share of workers living and working in the same location are almost double those in the baseline model. The introduction of a congestion tax leads even more workers to move to the location of their job, causing congestion to significantly fall, by even more than in the baseline model. However, average output does not fall because more workers are able to move to the city center when they get a highly productive job offer. The resulting increase in rent is offset by fewer low productivity firms and commuters locating in the center. Because workers are more willing to move, the unemployment rate is not significantly affected by the tax. Since many more workers move to the location of their job with a tax, wages and wage growth fall because the moving cost decreases match surplus and workers are unlikely to enjoy wage growth to compensate for costly commutes. In the cross section, wage dispersion and utility dispersion both fall by more than in the baseline model because workers moving to their job locations compresses the distribution of match values as well as that of commuting costs. However, the large gains from reducing congestion but relatively small increase in unemployment and average output lead to far larger gains in welfare relative to the baseline. Because the share of commuters is low when moving costs are low, the benefits to increasing the tax further are smaller, while the threshold for accepting and moving to a job offer near the center increases, therefore above a tax rate of 16% welfare in this case also begins to decrease.

By using a frictional model of the labor market, the model is able to quantify the monetary gains from commuting without imposing that the value of time is equal to the wage at the margin. Instead, workers who commute to full-time jobs

without flexible working hours must commute during the time they would otherwise be enjoying leisure, valued at the flow value of home production, b. Because the value of time is less than the average wage, commuting costs are smaller than in a model without labor market frictions. In the model with endogenous rent, the tax bill is almost 0.8% of total output, and the average tax paid by workers is 0.7% of wages.

Comparing this result to the literature, Brinkman (2016) estimates that with exogenous, location-specific productivity, welfare gains are in the order of 5% with the optimal congestion tax, but can turn negative if agglomeration externalities are strong enough.[37] The estimated gains here are about one order of magnitude lower, because this model has two additional sources of frictions that prevent larger welfare gains: moving costs and labor market frictions. Both of these decrease the welfare gains because many workers still find it optimal to commute and not to change their residence even after the tax is implemented. Importantly, with a frictionless labor market, employment increases in response to the tax for nearly all of his specifications. Here employment falls slightly due to the increase in the share of job offers rejected by the unemployed and resulting decrease in the UE transition rate. The novel channel in this model through which labor market frictions and moving costs contribute to congestion has the largest effect on unemployment and inequality, both of which are absent in the previous literature.

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. Without congestion, welfare is almost 4% higher, due

---

[37]Welfare gains correspond to changes in aggregate production reported in Table 4 of Brinkman (2016).

to the lower time cost of commuting as well as the positive effect on workers' job acceptance decisions.

The model can directly speak to crucial questions faced by policymakers and urban planners, and introduces important frictions necessary to properly price the value of time and therefore of congestion taxes. I find that the optimal congestion tax increases welfare by 0.5%, increasing the cost of commuting through congested areas by 16% at a cost to workers of 0.7% of average wages. More interestingly, congestion taxes raise the unemployment rate due to a lower rate of job acceptance for the unemployed. In addition, congestion taxes cause wage growth following job-to-job transitions to fall significantly. Regarding inequality, the model predicts that the welfare-maximizing tax on congestion leads to a 0.8% decline in wage dispersion and a 2.4% decline in utility dispersion. Lowering labor market frictions leads to a smaller response of unemployment to the tax, and therefore higher welfare gains. Reducing moving costs leads to the largest welfare gains from the tax due to a smaller negative effect of disincentivizing workers from accepting jobs in the most productive areas. Even absent the benefits of agglomeration, urban policies may not achieve their full welfare benefits as long as the labor market is affected by search frictions.

An important feature absent from this model is the labor-demand response to spatial frictions, which I rule out by assuming an exogenous 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 associated externality. 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 the link between spatial mismatch and congestion when search is directed or when wages are posted rather than the result of bargaining. Relatedly, incorporating search effort (e.g. Bagger and Lentz, Forthcoming) into the model would allow for workers to respond to congestion taxes in between job offers, with potentially ambiguous effects on the efficacy of the policy. The model could also be used to study alternative urban policies such as infrastructure investment or spatially-targeted employment subsidies and their effects on the local labor market.

References

Alonso, W. (1964). Location and land use. toward a general theory of land rent.

Anas, A. and I. Kim (1996). General equilibrium models of polycentric urban land use with endogenous congestion and job agglomeration. *Journal of Urban Economics 40*(2), 232–256.

Arnott, R. (2007). Congestion tolling with agglomeration externalities. *Journal of Urban Economics 62*(2), 187 – 203.

Arnott, R., A. De Palma, and R. Lindsey (1993). A structural model of peak-period congestion: A traffic bottleneck with elastic demand. *The American Economic Review*, 161–179.

Bagger, J. and R. Lentz (Forthcoming). An Equilibrium Model of Wage Dispersion and Sorting. *Review of Economic Studies*.

Barker, K. (2003). *Review of Housing Supply: Securing Our Future Housing Needs: Interim Report: Analysis*. HM Stationery Office.

Bils, M., Y. Chang, and S.-B. Kim (2012). Comparative advantage and unemployment. *Journal of Monetary Economics 59*(2), 150–165.

Bloom, N., J. Liang, J. Roberts, and Z. J. Ying (2014). Does working from home work? evidence from a chinese experiment. *The Quarterly Journal of Economics 130*(1), 165–218.

Bonhomme, S. and G. Jolivet (2009). The pervasive absence of compensating differentials. *Journal of Applied Econometrics 24*(5), 763–795.

Brinkman, J. C. (2016). Congestion, agglomeration, and the structure of cities. *Journal of Urban Economics 94*, 13–31.

Cahuc, P., F. Postel-Vinay, and J.-M. Robin (2006). Wage Bargaining with On-the-Job Search: Theory and Evidence. *Econometrica 74*(2), 323–364.

Cogan, J. F. (1981). Fixed costs and labor supply. *Econometrica*, 945–963.

Cohen, Y. (1987). Commuter welfare under peak-period congestion tolls: who gains and who loses? *International Journal of Transport Economics*, 239–266.

Conley, T. G. and G. Topa (2002). Socio-economic distance and spatial patterns in unemployment. *Journal of Applied Econometrics 17*(4), 303–327.

Dey, M. S. and C. J. Flinn (2005). An equilibrium model of health insurance provision and wage determination. *Econometrica 73*(2), 571–627.

Duranton, G. and M. Turner (2011). The fundamental law of road congestion: Evidence from us cities. *The American Economic Review 101*(6), 2616–2652.

Fujita, M. (1989). *Urban Economic Theory: Land Use and City Size*. Cambridge university press.

Glaeser, E. L., M. E. Kahn, and J. Rappaport (2008). Why do the poor live in cities? the role of public transportation. *Journal of urban Economics 63*(1), 1–24.

Gobillon, L., H. Selod, and Y. Zenou (2007). The mechanisms of spatial mismatch. *Urban studies 44*(12), 2401–2427.

Gomes, P. (2012). Labour market flows: Facts from the united kingdom. *Labour Economics 19*(2), 165 – 175.

Halket, J. and S. Vasudev (2014). Saving up or settling down: Home ownership over the life cycle. *Review of Economic Dynamics 17*(2), 345–366.

Hall, R. E. and P. R. Milgrom (2008). The limited influence of unemployment on the wage bargain. *American Economic Review 98*(4), 1653–74.

Hall, R. E. and A. I. Mueller (2018). Wage dispersion and search behavior: The importance of nonwage job values. *Journal of Political Economy 126*(4), 1594–1637.

Harari, M. (2017). Cities in bad shape: Urban geometry in india.

Heblich, S., S. J. Redding, and D. M. Sturm (2018). The making of the modern metropolis: Evidence from london. Technical report, National Bureau of Economic Research.

Hornstein, A., P. Krusell, and G. L. Violante (2011). Frictional wage dispersion in search models: A quantitative assessment. *American Economic Review 101*(7), 2873–98.

Hwang, H.-s., D. T. Mortensen, and W. R. Reed (1998). Hedonic wages and labor market search. *Journal of Labor Economics 16*(4), 815–847.

Jarosch, G. (2016). Searching for Job Security and the Consequences of Job Loss.

Kaneko, M. and Y. Yamamoto (1986). The existence and computation of competitive equilibria in markets with an indivisible commodity. *Journal of Economic Theory 38*(1), 118 – 136.

Kennan, J. and J. R. Walker (2011). The effect of expected income on individual migration decisions. *Econometrica 79*(1), 211–251.

Kreindler, G. E. (2018). The welfare effect of road congestion pricing: Experimental evidence and equilibrium implications.

Lucas, R. E. and E. C. Prescott (1974). Equilibrium Search and Unemployment. *Journal of Economic theory 7*(2), 188–209.

Manning, A. (2003). The real thin theory: Monopsony in modern labour markets. *Labour Economics 10*(2), 105–131.

Manning, A. and B. Petrongolo (2017). How local are labor markets? evidence from a spatial job search model. *American Economic Review 107*(10), 2877–2907.

Marinescu, I. and R. Rathelot (2018, July). Mismatch unemployment and the geography of job search. *American Economic Journal: Macroeconomics 10*(3), 42–70.

Mills, E. S. (1967). An aggregative model of resource allocation in a metropolitan area. *The American Economic Review 57*(2), 197–210.

Muth, R. (1969). Cities and housing: The spatial patterns of urban residential land use.

Parry, I. W. H. and A. Bento (2001). Revenue recycling and the welfare effects of road pricing. *The Scandinavian Journal of Economics 103*(4), 645–671.

Pigou, A. C. (1932). *The Economics of Welfare* (4 ed.). MacMillan, London.

Pinheiro, R. and L. Visschers (2015). Unemployment risk and wage differentials. *Journal of Economic Theory 157*, 397–424.

Postel-Vinay, F. and H. Turon (2010). On-the-Job Search, Productivity Shocks and the Individual Earnings Process. *International Economic Review 51*(3), 599–629.

Rapino, M. and A. Fields (2013). Mega Commuters in the U.S.: Time and Distance in Defining the Long Commute using the American Community Survey. US Census Bureau Working Paper.

Rosen, S. (1986). The theory of equalizing differences. *Handbook of labor economics 1*, 641–692.

Şahin, A., J. Song, G. Topa, and G. L. Violante (2014). Mismatch unemployment. *American Economic Review 104*(11), 3529–64.

Smith, T. E. and Y. Zenou (2003). Spatial mismatch, search effort, and urban spatial structure. *Journal of Urban Economics 54*(1), 129–156.

Sorkin, I. (2018). Ranking firms using revealed preference. *The Quarterly Journal of Economics 133*(3), 1331–1393.

Tsivanidis, N. (2018). The aggregate and distributional effects of urban transit infrastructure: Evidence from bogotá‚Äôs transmilenio. Technical report.

University of Essex, Institute for Social and Economic Research (2018). British Household Panel Survey: Waves 1-18, 1991-2009. Special License Access, Local Authority Districts. *8th Edition. http://doi.org/10.5255/UKDA-SN-5151-2.*

Van Den Berg, G. J. and C. Gorter (1997). Job search and commuting time. *Journal of Business & Economic Statistics 15*(2), 269–281.

Van Der Berg, G. J. (1992). A structural dynamic analysis of job turnover and the costs associated with moving to another job. *The Economic Journal 102*(414), 1116–1133.

Van Ommeren, J., P. Rietveld, and P. Nijkamp (1999). Job moving, residential moving, and commuting: a search perspective. *Journal of Urban Economics 46*(2), 230–253.

Van Ommeren, J. N. and E. Gutiérrez-i Puigarnau (2011). Are workers with a long commute less productive? an empirical analysis of absenteeism. *Regional Science and Urban Economics 41*(1), 1–8.

Van Ommeren, J. N., G. J. Van Den Berg, and C. Gorter (2000). Estimating the marginal willingness to pay for commuting. *Journal of Regional Science 40*, 541–563.

Van Vuuren, A. (2018). City structure and the location of young college graduates. *Journal of Urban Economics 104*, 1–15.

Vickrey, W. S. (1969). Congestion theory and transport investment. *The American Economic Review 59*(2), 251–260.

Wales, T. J. (1978). Labour supply and commuting time: an empirical study. *Journal of Econometrics 8*(2), 215–226.

Walters, A. A. (1961). The theory and measurement of private and social cost of highway congestion. *Econometrica 29*(4), 676–699.

Wasmer, E. and Y. Zenou (2002). Does city structure affect job search and welfare? *Journal of urban Economics 51*(3), 515–541.

Wasmer, E. and Y. Zenou (2006). Equilibrium search unemployment with explicit spatial frictions. *Labour Economics 13*(2), 143–165.

Zenou, Y. (2009). Urban search models under high-relocation costs. theory and application to spatial mismatch. *Labour Economics 16*(5), 534–546.

Online Appendix

FOR ONLINE PUBLICATION

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 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 uses 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 (*payn*). 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.

Figure 10: Number of Commuters to City of London by Time
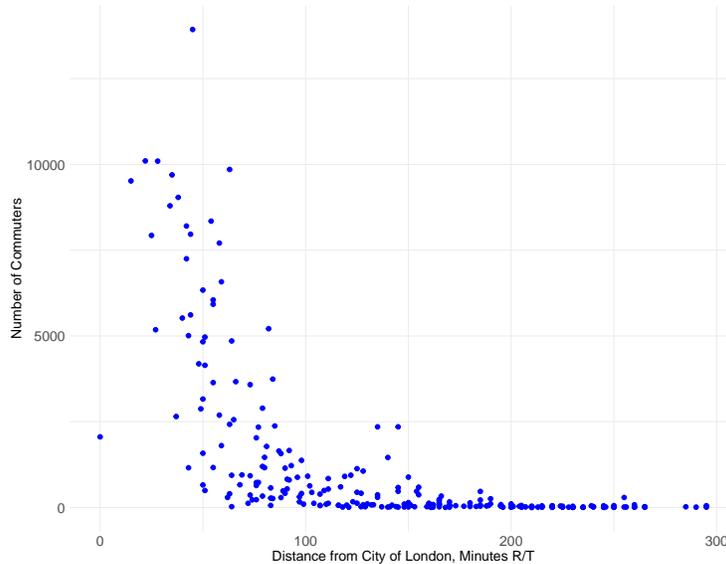
Data: UK Census Flow Data 2001, Special Workplace Statistics (Level 1). Number of workers commuting to the City of London as a function of commuting time in minutes, round trip, computed by Google Maps.

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 1. 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.

Figure 11: Share of Commuters to City of London by Distance
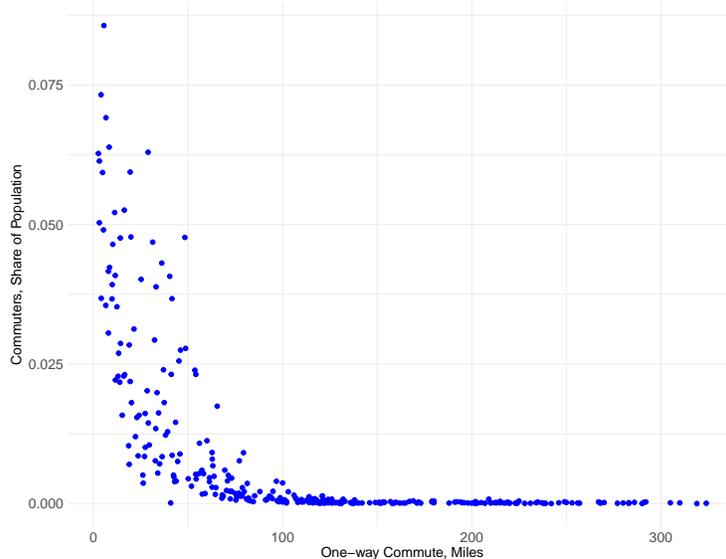
Data: UK Census Flow Data 2001, Special Workplace Statistics (Level 1). Share of workers commuting to the City of London relative to population of LAD of residence, as a function of commuting distance in miles computed by Google Maps. Map data: Google, DigitalGlobe.

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).[38] 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.[39] Figures 10 and 11

---

[38]To ensure to be firmly within Google's terms of service, I hand-collected the data by searching for each LAD.

[39]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

show plots identical to Figure 1 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.

A.2 EE and UE Transitions: Robustness

Table 10 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 11 to 20 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 21 and 22 respectively define commuting increases and decreases as $+/-$ 5 or 10 minutes of commuting time, rather than by percent changes in the commute. Table 23 shows the table after allowing for job-to-job transitions with zero days of nonemployment between the two employment spells.

Bloomberg Arcade, London.

Table 10: Effect of Lagged Commute and Wage on Job-to-Job Transition Probability, Linear Probability Model

	J2J_t	J2J_t
$\text{Commute}_{t-1} \times 30$	0.018***	0.022***
	(0.005)	(0.005)
Real log Wage_{t-1}	-0.071***	-0.069***
	(0.013)	(0.014)
Individual Characteristics		✓
Individual FE	✓	✓
Region, Time, Commute Method Industry & Occ FE	✓	✓
R^2	.051	.059
N ind	1,259	1,122
N obs	7,023	6,338

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 coefficients from a linear probability model 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 individual fixed effects, and region, month, year, 1-digit industry and occupation fixed effects. Robust standard errors are reported in parentheses. * denotes $p < .1$, ** $p < .05$, and *** $p < .01$.

Table 11: After Job-to-Job Transition, Men

		Wage		
		Down	Same	Up
	Down	0.43	0.50	0.36
Commute	Same	0.24	0.17	0.26
	Up	0.32	0.33	0.38
	N	78	54	234

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 12: After Job-to-Job Transition, Women

		Wage		
		Down	Same	Up
	Down	0.56	0.37	0.31
Commute	Same	0.14	0.20	0.25
	Up	0.30	0.43	0.44
	N	79	49	203

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 13: After Job-to-Job Transition, Occupation Switchers

		Wage		
		Down	Same	Up
	Down	0.53	0.48	0.29
Commute	Same	0.17	0.14	0.24
	Up	0.30	0.38	0.46
	N	96	56	234

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 2-digit occupation 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 14: After Job-to-Job Transition, Occupation Stayers

		Wage		
		Down	Same	Up
	Down	0.44	0.38	0.39
Commute	Same	0.23	0.23	0.27
	Up	0.33	0.38	0.34
	N	61	47	203

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 2-digit occupation 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 15: After Job-to-Job Transition, Residential Movers

		Wage		
		Down	Same	Up
	Down	0.33	0.42	0.35
Commute	Same	0.19	.05	0.19
	Up	0.48	0.53	0.46
	N	21	19	74

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 16: After Job-to-Job Transition, Residential Stayers

		Wage		
		Down	Same	Up
	Down	0.52	0.44	0.34
Commute	Same	0.19	0.21	0.27
	Up	0.29	0.35	0.40
	N	136	84	363

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 17: After Job-to-Job Transition, Married

		Wage		
		Down	Same	Up
	Down	0.48	0.44	0.31
Commute	Same	0.17	0.22	0.28
	Up	0.35	0.34	0.41
	N	75	50	209

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 18: After Job-to-Job Transition, Single

		Wage		
		Down	Same	Up
	Down	0.51	0.43	0.37
Commute	Same	0.20	0.15	0.23
	Up	0.28	0.42	0.40
	N	82	53	228

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 19: After Job-to-Job Transition, Children

		Wage		
		Down	Same	Up
	Down	0.46	0.38	0.24
Commute	Same	0.17	0.24	0.28
	Up	0.37	0.38	0.48
	N	59	21	138

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, No Children

		Wage		
		Down	Same	Up
Commute	Down	0.52	0.45	0.38
	Same	0.20	0.17	0.24
	Up	0.28	0.38	0.37
	N	98	82	299

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, Commute $\pm$ 5 min

		Wage		
		Down	Same	Up
Commute	Down	0.50	0.44	0.33
	Same	0.20	0.21	0.28
	Up	0.30	0.35	0.39
	N	157	103	437

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 22: After Job-to-Job Transition, Wage ± 10 min

| | | Wage | | |
		Down	Same	Up
	Down	0.44	0.31	0.25
Commute	Same	0.35	0.42	0.46
	Up	0.21	0.27	0.29
	N	157	103	437

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 23: After Job-to-Job Transition, 0 Days Nonemployment

| | | Wage | | |
		Down	Same	Up
	Down	0.50	0.42	0.34
Commute	Same	0.19	0.20	0.25
	Up	0.30	0.38	0.41
	N	149	97	417

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 24: Log Realized to Reservation Wage Relative to Average

Group	Commute Below Median	Commute Above Median
All	-0.10	0.15
Some College or Less	-0.02	0.03
College Graduates	-0.21	0.23
Men	-0.20	0.25
Women	-0.01	0.01

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.

Reservation Wages

Following the arguments of Hall and Mueller (2018), I consider workers who made a 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 UE transitions. Table 24 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 all groups.

Results Using Longitudinal Weights

In this section I weigh 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

Table 25: After Job-to-Job Transition, Weighted Sample

		Wage		
		Down	Same	Up
	Down	0.51	0.45	0.31
Commute	Same	0.19	0.17	0.26
	Up	0.20	0.38	0.43
	N	101.44	61.50	277.1

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.

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 26: Log Realized to Reservation Wage Relative to Average, Weighted Sample

Group	Commute Below Median	Commute Above Median
All	-0.07	0.13

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. $U(\ell; r)$, omitting the distributional dependence on Ω. I assume that the distributions contained in Ω, $u(\ell; r)$ and $e(\ell_F, y, \ell; r)$ are continuous in r.

Part 1a - Demand Correspondence

Define the demand correspondence for newly matched vacancies for each $\ell \in \mathcal{L}$ and $y \in Y$ as

$$
\begin{aligned}
D_V(\ell, y; r) \;=\;& e^{\ell+1} \sum_{\tilde{\ell}} \Big[\lambda_0 u(\tilde{\ell}; r) \mathbb{1}\{S(\ell, y, \tilde{\ell}; r) > 0\} \\
&+ \; \lambda_1 \sum_{\hat{\ell}} \sum_{\hat{y}} \mathbb{1}\{S(\ell, y, \tilde{\ell}; r) > \tilde{S}(\hat{\ell}, \hat{y}, \tilde{\ell}; r)\} e(\hat{\ell}, \hat{y}, \tilde{\ell}; r) \Big] \Big\}
\end{aligned}
$$

where e^{ℓ} is the ℓ-th unit vector of dimension $2N + 2$, and the second term is the probability a meeting occurs and the match is accepted. The unit vector's dimension is the sum of the outside option of leaving the city and the number of locations $2N + 1$, where e^1 is the demand for leaving the city, and $e^{\ell+1}$ is the demand for land in location $\ell = 1, \ldots, 2N + 1$.

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

$$
\begin{aligned}
D_E(\ell_F, y, \hat{\ell}; r) \;=\;& \big\{ x \in \{e^1, e^2, \ldots, e^{2N+2}\} : \beta S(\ell_F, y, \hat{\ell}; r) > \bar{u} \text{ and} \\
& x(\ell) \in \arg\max M(\ell_F, y, \hat{\ell}; r) \big\}
\end{aligned}
$$

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

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

for each $\hat{\ell} \in \mathcal{L}$. I denote the location choice implied by x as $x(\ell) \in 0 \bigcup \mathcal{L}$. 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 $U(\ell; r)$ 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, and J are all continuous in r. Given these results, I now prove each element of $D(r)$ is upper hemicontinuous (UHC). Below I assume that the unemployed can move at fixed cost k_M. The case in which the unemployed cannot move, assumed in the model, is a special case.

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

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

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

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

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

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

But U is continuous in r, so

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

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, $D_U(\hat{\ell}; r)u(\hat{\ell}; r)$ 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 that $x^0 \notin D_E(\ell_F, y, \hat{\ell}; r^0)$. Then there exists $x' \in D_E(\ell_F, y, \hat{\ell}; r^0)$. Then $x'(\ell) \in \arg\max M(\ell_F, y, \hat{\ell}; r^0)$. Since M is continuous $\exists N$ such that for $n > N$, $x^n(\ell) \notin \arg\max M(\ell_F, y, \hat{\ell}; r^n)$, a contradiction.

Since (ℓ_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: $\sum_{\hat{y}} D_V(\hat{\ell}, \hat{y}; r)g(\hat{\ell}, \hat{y})$

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

$$
\sum_\ell \left[\lambda_0 u(\ell; r) \mathbb{1}\{S(\ell_i, y, \ell; r) > 0\} \right.
$$
$$
+ \quad \lambda_1 \sum_{\hat{\ell}} \sum_{\hat{y}} \mathbb{1}\{S(\ell_i, y, \ell; r) > S(\hat{\ell}, \hat{y}, \ell; r)\}e(\hat{\ell}, \hat{y}, \ell; r) \left. \right]\}
$$

is continuous since S is continuous under the assumption that u and e are continuous. Thus, D_V is UHC for all $(\hat{\ell}, \hat{y}) \in \mathcal{L} \times Y$, 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 $\hat{\ell} \in \mathcal{L}$, 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 ℓ available to allocate to each group making up the demand correspondence above is given by the land choice of the landlord, $\zeta(\ell)$, less the land occupied by continuing jobs, for whom the location is fixed. In steady

state, the amount of land occupied by continuing jobs is constant and equal to the mass of jobs located in ℓ less the mass of endogenously or exogenously destroyed matches. In each location $\ell \in \mathcal{L}$, the supply is given by

$$\Sigma(\ell, r) \;=\; \zeta(\ell, r) - \sum_y \sum_{\hat{\ell}} (e(\ell, y, \hat{\ell}; r) - e^-(\ell, y, \hat{\ell}; r))$$

where $\Sigma(\ell, r) \geq 0$ if $r \geq 0$ and 0 otherwise since landlords will not rent any land at a negative price. 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

$$E(r) = [\Sigma(0, r), \ldots, \Sigma(2N + 1, r)]$$

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 $P = [0, P_{max}]^{2N+2}$, where

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

Defining 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 $C = [\underline{E}, \overline{E}]^{2N+2}$ 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 $covE(r)$ inherits the nonemptiness and upper hemicontinuity of $E(r)$. By Lemma 5 in KY, $covE(r)$ is convex-valued. Further, denoting x as an element of $E(r)$, the following

two results hold:

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

2. If $r(\ell) = S_{max}$ for some $\ell \in \mathcal{L}$ and $x \in covE(r)$, 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 ℓ must be weakly positive. To prove 2, since for $r(\ell) = S_{max}$, supply is weakly positive and the surplus of all matches is weakly negative. Thus, no vacancies that draw ℓ will match, and no employed or unemployed workers will demand to locate in ℓ 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 $0 \in covE(r)$.

Denote this solution by x_{ch} and σ_{ch}, where x_{ch}^i, $i = 0, \ldots, 2N+1$ is given by the sum of each type of worker's demand for land in location i. Including the outside option of leaving the city, there are $2N + 2$ "types" of unemployed workers, and $(2N + 2)^2 Y$ "types" of employed workers, thus, the maximum number of types in each location is $M \equiv 2N + 2 + (2N + 2)^2 Y$. I therefore denote $x_{ch}^{j,i}$ as the demand for land in location i by type $j = 1, \ldots, M$. Similarly, σ_{ch}^i is the supply of land in each location i, after taking into account the demand for new vacancies. The market clearing condition for land can be written

$$\sum_{j=1}^{M} x_{ch}^{i,j} + \sigma_{ch}^i = L \quad \text{for all } i = 0, \ldots, 2N+1$$

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

$$\sum_{i=0}^{2N+1} x_{ch}^{i,j} \leq 1 \quad \text{for all } j = 1, \ldots, M$$

Since x_{ch} and σ_{ch} may not be integers, consider the system

$$\sum_{j=1}^{M} \hat{x}^{i,j} + \hat{\sigma}^i = L \quad \text{for all } i = 0, \ldots, 2N+1$$

$$\sum_{i=0}^{2N+1} \hat{x}^{i,j} \leq 1 \quad \text{for all } j = 1, \ldots, M$$

where $\hat{x}^{i,j} \in \mathbb{R}_+$ and $\hat{\sigma}^i \in \mathbb{R}_+$ for all $i = 0, \ldots, 2N+1$ and $j \in M$.

Let $\boldsymbol{x} = [x^{0,1}, x^{1,1}, \ldots, x^{2N+1,1}, \ldots, x^{0,M}, \ldots, x^{2N+1,M}]$ be an $1 \times (2N+2)M$ vector and $\boldsymbol{\sigma} = [\sigma^0, \ldots, \sigma^{2N+1}]$ is a $1 \times (2N+2)$ vector. The system above can be written as

$$A(\boldsymbol{x}, \boldsymbol{\sigma})' \begin{matrix} \leq \\ \vdots \\ \leq \\ = \\ \vdots \\ = \end{matrix} \begin{bmatrix} 1 \\ \vdots \\ 1 \\ L \\ \vdots \\ L \end{bmatrix}$$

where A is an $(2N+2)M \times (2N+2)(M+1)$ matrix

$$A = \begin{bmatrix} \begin{array}{cccc} 1 \ldots 1 & & & \\ & \ddots & & \\ & & 1 \ldots 1 \\ \hline 1 & 1 & 1 & \\ \ddots & \ddots & & \ddots \\ 1 & 1 & & 1 \end{array} & \begin{array}{c} 0_{M \times (2N+2)} \\ \\ \\ \hline \\ I_{2N+2} \\ \end{array} \end{bmatrix}$$

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 where A is a (totally) unimodular matrix,[40] 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. 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 ℓ is given by $r(\ell)$, paid by both workers and firms located in ℓ. Land is assumed to be perfectly elastic. Workers are born unemployed, with their location ℓ determined by a draw from exogenous distribution F. Workers can change their

---

[40]Total unimodularity requires that every square submatrix has a determinant of -1, 0 or 1.

location by paying a moving cost, k_M.

Workers are born at the beginning of the period and draw a vacancy with probability λ. A vacancy is a draw of the location of the match, ℓ_F from distribution G, with the probability of drawing a vacancy in location ℓ 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 move, 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 β of the match surplus, where ℓ 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 Ω 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 Ω, and their beliefs are assumed to be consistent with the realized value of Ω in equilibrium.

Suppose that in the decentralized equilibrium all workers receiving offers from another location prefer to commute rather than move. If it is optimal for both of these groups to commute, the following conditions must hold:

$$1 - r(\ell) - T(e(\ell, \ell') + e(\ell', \ell)) > b, \quad \ell = 1, 2 \tag{23}$$

$$T(e(\ell, \ell') + e(\ell', \ell)) + r(\ell) - r(\ell') < k_M, \quad \ell = 1, 2, \ell' \neq \ell \tag{24}$$

where $e(\ell, \ell') = \lambda f(\ell) g(\ell')$ is the mass of workers born in location $\ell = 1, 2$ receiving an offer from $\ell' \neq \ell$. 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 parameters are assumed to satisfy (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 moves 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.

$$(1 - 2r(1))e(1,1) + (1 - 2r(1) - k_M)\pi(1)e(1,2) + (1 - 2r(2) - k_M)\pi(2)e(2,1)$$

$$+ \left(1 - r(1) - r(2) - T\big((1 - \pi(1))e(1,2) + (1 - \pi(2))e(2,1)\big)\right)(1 - \pi(1))e(1,2) + (1 - 2r(2))e(2,2)$$

$$+ \left(1 - r(1) - r(2) - T\big((1 - \pi(1))e(1,2) + (1 - \pi(2))e(2,1)\big)\right)(1 - \pi(2))e(2,1)$$

$$+ (b - r(1))u(1) + (b - r(2))u(2) \quad (25)$$

Subject to the feasibility constraints:

$$e(\ell, \ell_F) \leq \lambda f(\ell)g(\ell_F), \quad for \ \ell, \ell_F = 1, 2$$

$$e(\ell, \ell_F) \geq 0 \ \forall(\ell, \ell_F)$$

$$u(\ell) \geq (1 - \lambda)f(\ell) \ \forall \ell$$

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

$$\pi(\ell) \in [0, 1] \ \forall \ell$$

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 moving and commuting to be a share between 0 and 1 of the mass of workers with offers in a location different from their own.

Consider an alternative allocation in which the mass $e(1,2)$ moves to location 2 and $e(2,1)$ continues to commute to location 1, that is $\pi(1) = 1, \pi(2) = 0$. Recall

that in the decentralized allocation, all workers receiving offers different from their own commute. All other workers' contributions to output are unchanged. The planner prefers this allocation to the decentralized allocation if

$$T(e(1,2) + e(2,1))(e(1,2) + e(2,1)) - T(e(2,1))e(2,1) > (r(2) - r(1) + k_M)e(1,2) \quad (26)$$

where $e(1,2) + e(2,1)$ is the mass of commuters in the decentralized allocation and $e(2,1)$ 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 moving cost. Assume that commuting costs are of the form $T(\Omega) = (\tau + \kappa\Omega)b$, then we can write (24) and (26) as

$$r(1) - r(2) + (\tau + \kappa(e(1,2) + e(2,1)))b < k_M < r(1) - r(2) + (\tau + \kappa(e(1,2) + 2e(2,1)))b$$

If the above condition holds, then the decentralized equilibrium is inefficient and the planner will choose for some workers to move. Rewriting the right hand side as the sum of the private benefit of reducing commuting and the social benefit: $r(1) - r(2) + (\tau + \kappa(e(1,2) + e(2,1)))b + \kappa e_2 b$ it is clear that if the social benefit, $\kappa e_2 b$, is large enough, it can be welfare-improving to have some workers move, but privately optimal to commute.

In order for the above allocation to be constrained efficient, the planner must prefer to have $e(1,2)$ move and $e(2,1)$ commute rather than $e(2,1)$ move and $e(1,2)$ commute, or for both $e(1,2)$ and $e(2,1)$ to move. First, note that the planner will never have both $e(1,2)$ and $e(2,1)$ move, since it is privately suboptimal for both to move, and there is no social benefit since no workers would commute in this case. Second, the planner will prefer to have $e(1,2)$ move and $e(2,1)$ commute rather than $e(2,1)$ move and $e(1,2)$ commute if

$$(1 - r(1) - r(2) - T(e(1,2)))e(1,2) + (1 - 2r(1) - k_M)e(2,1)$$
$$< (1 - r(1) - r(2) - T(e(2,1)))e(2,1) + (1 - 2r(2) - k_M)e(1,2) \quad (27)$$

Noticing that $r(1) - r(2) - T(e(2,1)) + k_M > 0$ by (24), rearranging this expression gives

$$\frac{r(2) - r(1) - T(e(1,2)) + k_M}{r(1) - r(2) - T(e(2,1)) + k_M} < \frac{e(2,1)}{e(1,2)}$$

The numerator and denominator on the left hand side are the output losses to

workers $e(1,2)$ and $e(2,1)$, respectively, if the worker moves rather than commutes. From the decentralized equilibrium, it is clear that if there were no benefits from reducing congestion, no workers will choose to move. The expression above compares the net gain for each worker from moving relative to commuting when the other type of worker moves. First, observe that if the left hand side is equal to the right hand side, the net gain from workers $e(1,2)$ moving and $e(2,1)$ commuting and vice versa are equal, and the planner is indifferent. This is the case if $e(1,2) = e(2,1)$. If rent in the two locations are equal, since commuting costs are increasing in the mass of workers, the left hand side is greater than one if $e(1,2) < e(2,1)$, that is, the mass of workers paying the moving cost must be less than the mass of workers benefiting from lower congestion. As rent in location 2 increases relative to rent in location 1, the loss from moving workers $e(1,2)$ increases, implying that it is only optimal to move these workers if their mass is small relative to $e(2,1)$.

The constrained efficient allocation is to have workers $e(1,2)$ move and $e(2,1)$ commute, and all other workers to make choices identical to the decentralized equilibrium, if (26) and (27) hold.

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

Proposition 2. *Suppose $T(\Omega) = (\tau + \kappa\Omega)b$. Then the following are true:*

(i) *For any $k_M > r(1) - r(2) + \tau b$ there exists $\underline{\kappa}$, $\overline{\kappa}$ such that the decentralized equilibrium is inefficient for all $\kappa \in (\underline{\kappa}, \overline{\kappa})$.*

(ii) *As k_M approaches τb from above, $\underline{\kappa}, \overline{\kappa} \to 0$. As k_M increases, the distance $\overline{\kappa} - \underline{\kappa}$ increases.*

(iii) *A tax implementing the constrained efficient allocation, t, satisfies $t = \frac{e(2,1)}{e(1,2)+e(2,1)}$, where $e(1,2)$ is the mass of movers in the constrained efficient equilibrium and $e(2,1)$ is the mass of commuters benefiting from a reduction in congestion, if $r(1)$ is sufficiently large relative to $r(2)$.*

Proof. Define $\underline{\kappa}$ as
$$\underline{\kappa} = \frac{k_M - r(1) + r(2) - \tau b}{b(e_1 + 2e_2)}$$

Similarly, define $\overline{\kappa}$ by
$$\overline{\kappa} = \frac{k_M - r(1) + r(2) - \tau b}{b(e_1 + e_2)}$$

73

For $\kappa \in (\underline{\kappa}, \overline{\kappa})$, both (24) and (26) hold, proving part (i).

For part (ii), notice that as k_M approaches τb, the values of $\underline{\kappa}$, $\overline{\kappa}$ approach zero. Similarly, we can define

$$\Delta \kappa = \overline{\kappa} - \underline{\kappa} = \frac{k_M - r(1) + r(2) - \tau b}{b} \left(\frac{1}{e_1 + e_2} - \frac{1}{e_1 + 2e_2} \right)$$

which is strictly increasing in k_M.

To prove part (iii), recall that the decentralized equilibrium is constrained inefficient if

$$r(1) - r(2) + (\tau + \kappa(e(1,2) + e(2,1)))b < k_M < r(1) - r(2) + (\tau + \kappa(e(1,2) + 2e(2,1)))b$$

Letting t be the tax on congestion, we can write the left hand side as

$$r(1) - r(2) + (\tau + (1+t)\kappa(e(1,2) + e(2,1)))b$$

Setting the above expression equal to $r(1) - r(2) + (\tau + \kappa(e(1,2) + 2e(2,1)))b$ and solving for t gives

$$t = \frac{e(2,1)}{e(1,2) + e(2,1)}$$

This tax rate achieves constrained efficiency if and only if

$$(\tau + (1+t)\kappa(e(1,2) + e(2,1)))b + r(1) - r(2) \geq k_M$$

$$(\tau + (1+t)\kappa(e(1,2) + e(2,1)))b + r(2) - r(1) < k_M$$

which hold if

$$(\tau + \kappa(e(1,2) + e(2,1)) + \kappa e(2,1))b + r(1) - r(2) \geq k_M$$

$$(\tau + \kappa(e(1,2) + e(2,1)) + \kappa e(2,1))b + r(2) - r(1) < k_M$$

A sufficient condition for these expressions to hold is

$$r(1) - r(2) > \frac{e(1,2) + 2e(2,1)}{3e(1,2) + 4e(2,1)}(k_M - \tau b)$$

$\square$

In this static setting, the congestion externality arises when two conditions are met. First, the moving cost must be sufficiently large, and second, the effect of congestion on commuting costs is bounded by two values. The first condition must hold in order for workers in the decentralized economy to prefer to commute rather than move. The second condition bounds the congestion cost so that the planner prefers that a group of workers moves rather than commutes. The social benefit of moving is the private benefit of lower commuting costs for the movers plus the marginal reduction in commuting costs for the commuters who benefit from less congestion, times their mass. Part (i) of Proposition 2 states that for any sufficiently large moving cost, there exist a range of values for the marginal cost of congestion such that the decentralized equilibrium is inefficient. The lower bound for κ insures that the social value of reducing congestion is high enough to overcome the cost of moving. Symmetrically, the upper bound restricts the congestion cost to be small enough such that workers in the decentralized economy do not always take the constrained efficient action. Part (ii) states that the smaller is the moving cost, the smaller is the range and the lower is the value for congestion costs such that the decentralized equilibrium is constrained inefficient. When moving costs are small, there is only a positive effect of moving and it is privately and socially optimal for all workers who would otherwise commute to move to the location of their firms. The smaller the moving cost, the more likely that the decentralized equilibrium is efficient. Part (iii) derives the optimal tax, which is equal to the share of workers who continue to commute after some workers move. These workers benefit from the reduction in congestion, which workers in the decentralized economy do not take into account when making their moving decision. By increasing their own commuting costs by taxing congestion, the optimal tax aligns their incentives with the planner's under some parameter values. In particular, if rent in the location where the movers originally live is high enough, the private benefits from moving are greater for this group for any given tax rate, relative to the group of workers for whom the planner finds it optimal to commute, and the tax achieves constrained efficiency.

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 1 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 total number of commuters. I define the commuting zone as those LADs with a share above 30%. Under this definition, the location farthest from the City of London is Colchester (65 miles). There are 1,655 individuals under this definition, equal to 13% of the sample size when using the 100-mile radius in the main model. Results for the empirical analysis in Section 2 and Appendix A are similar and are available upon request. Table 27 shows how the parameters used in the calibration and the untargeted moments differ in the two definitions. Because more than 80% of individual observations are lost, the estimates in the second column are much more imprecise than using the 100 mile radius, though overall they are quantitatively similar.

Table 27: Targets: Parameters

Moment	$\leq$ 100 miles	"Comm. Zone"
Relative wage, new hires from U	.87	.87
Variance, $ln(w)$	.186	.205
Moving probability (A)	.036	.034
Relative increase in commuting cost with congestion	.65	.73
Share of job-related moves	.129	.132
Avg rent/income	.365	.378
Relative wage, residence 5-33 (5-22) mi	.82	.84
Relative wage, residence 33-66 (22-44) mi	.79	.84
Relative wage, residence 66-100 (44-65) mi	.70	.89
Jobs density, 5-33 (5-22) mi from center	.83	.77
Jobs density, 33-66 (22-44) mi from center	.80	.85
Jobs density, 66-100 (44-65) mi from center	.78	.80
Share of wage cuts	.16	.15
Share of movers who stay in the same job	.033	.035
Average wage growth, J2J	.065	.083
Average wage growth, within job	.016	.018
Annual J2J transition rate, residence in center	.147	.147
Annual J2J transition rate, residence in periphery	.105	.101
Relative Moving Distance, J2J to All Movers	1.40	0.67

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 ℓ. 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 5 miles of the center (location 4), and 5-33 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
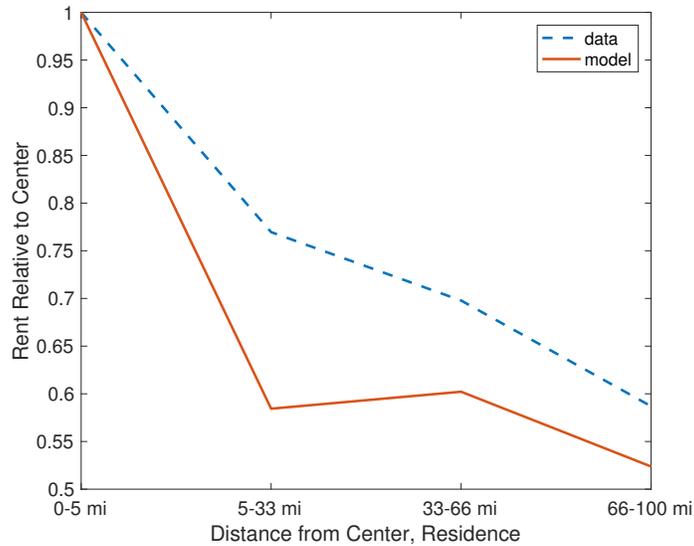
Figure 13 shows the policy function for offers with the highest productivity, given the worker's initial location. The left panel shows the surplus and location decisions of a match $(\ell_F, y_H, 1)$ as a function of ℓ_F. When the firm is in or near the worker's location, at the left side of the figure, 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

Figure 13: Moving Decisions and Match Surplus
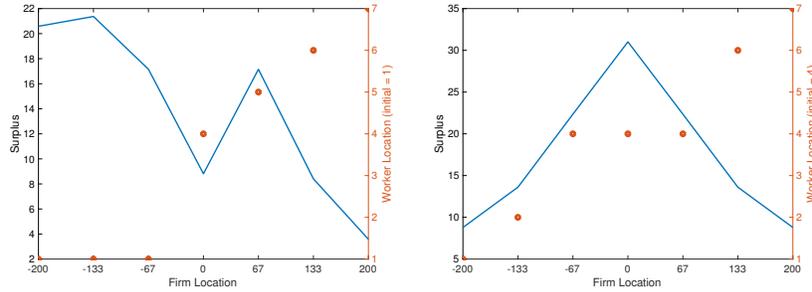
worker with initial location in the center (right panel). For this worker, again the surplus is higher when the firm and worker locations are the same, but decreases in both directions as the firm moves toward the periphery. If the firm is two or more locations away, the worker moves in order to decrease 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.