Remote Work and City Structure*

Ferdinando Monte
Georgetown University

Charly Porcher
Georgetown University

Esteban Rossi-Hansberg
University of Chicago

July 25, 2023

Abstract

We study the adoption of remote work within cities and its effect on city structure and welfare. We develop a dynamic model of a city in which workers can decide to work in the central business district (CBD) or partly at home. Working in the CBD allows them to interact with other commuters, which enhances their productivity through a standard production externality, but entails commuting costs. Switching between modes of labor delivery is costly, and workers face idiosyncratic preference shocks for remote work. We characterize the parameter set in which the city exhibits multiple stationary equilibria. Within this set, a coordination mechanism can lead to stationary equilibria in which most workers commute or most of them work partially from home. In these cases, large shocks in the number of commuters, like the recent lockdowns and self-isolation generated by the COVID-19 pandemic, can result in dynamic paths that make cities converge to a stationary equilibrium with large fractions of remote workers. Using cell-phone-based mobility data for the U.S., we document that although most cities experienced similar reductions in CBD trips during the pandemic, trips in the largest cities have stabilized at levels that are only about 60% of pre-pandemic levels. In contrast, smaller cities have, on average, returned to pre-pandemic levels. House price panel data by city show consistent changes in house price CBD-distance gradients. We estimate the model for 274 U.S. cities and show that cities that have stabilized at a large fraction of remote work are much more likely to have parameters that result in multiple stationary equilibria. Our results imply welfare losses in these cities that average 2.7%.

*Monte: ferdinando.monte@georgetown.edu, Porcher: charly.porcher@georgetown.edu, Rossi-Hansberg: rossi-hansberg@uchicago.edu. We thank Shijian Yang for excellent research assistance and Treb Allen, Cecile Gaubert, Maria Hansen, Maurice Kugler, Maisy Wong, and numerous seminar participants for comments.

1 Introduction

Many of us work at least partially from home these days. A phenomenon that was only marginal a few years ago has become widespread and has challenged, in potentially profound ways, the basic organization of work in firms and cities, and across regions and countries. Information technology has improved to a point where workers can talk, chat, and send information in documents, databases, audio, and video without delays and at extremely low costs. However, although information technology clearly complements remote work, advances in information and communication technologies (ICT) increased remote work only marginally between 1980 and 2019. In contrast, since the COVID-19 pandemic, remote work has boomed. Suddenly, commuting has declined by half in many of our largest cities and seems to have stabilized at those low levels.[1] What generated this abrupt and seemingly permanent increase in work from home since the pandemic? What are its consequences for the structure of cities? Will the change be uniform across them? Importantly, what are the welfare consequences of permanent shifts in commuting and agglomeration patterns? In this paper, we aim to address these questions by proposing, and estimating for most U.S. cities, a dynamic theory of city structure with a labor delivery choice. We also contrast and verify its implications with data on individual cell phone mobility and zip code-level house prices.

At the core of our theory is the idea that the value of working at the CBD, in short "office work", depends on proximity to other workers. Workers' productivity is enhanced by interactions with others in their office, but also across businesses, and more generally with other workers at city centers. This is an old idea from Marshall, of course. Remote work, however, offers workers and employers the possibility to work in the same city but without interacting in person. Since workers do not internalize many of the benefits they convey to others, this alternative possibility leads to a novel coordination problem: Workers prefer to work in the CBD only if others do too, but prefer to work from home otherwise. In a dynamic economy, where workers face idiosyncratic preferences for remote work (e.g. having a baby or remodeling) and fixed costs from switching between labor delivery modes (e.g. setting up a home office or buying a car), this mechanism can lead to multiple stationary equilibria with different permanent shares of commuters. In these cases, abrupt changes to the number of commuters can make a city converge to a stationary equilibrium that is very different from the initial one. For many places, we postulate, this abrupt change was the COVID-19 pandemic and the lockdowns and self-isolation it generated. This one-time temporary change has led, for several U.S. cities, to a permanent change in one of the most enduring characteristics of human organization: work at city centers.

Of course, not all cities are equal. Some cities specialize in industries where interactions between workers lead to large spillovers and correspondingly large populations. Other cities specialize in industries with low spillovers and are small. Commuting costs can vary dramatically depending on transport infrastructure and geographic characteristics. Furthermore, the ability to work from

[1] These trends have been reported from survey data (Brynjolfsson et al., 2020; Aksoy et al., 2022; Brynjolfsson et al., 2023), administrative data (Dalton et al., 2022), changes in job advertisements for remote positions (Hansen et al., 2023), or changes in internet usage of firms (Matthies and Kwan, 2022).

home varies sharply with the worker's occupation, and occupations are, of course, not uniformly distributed in space. We all know about the concentration of bankers in New York or programmers in Palo Alto. All these characteristics of cities determine the extent to which the coordination problem we identify leads to multiplicity in stationary equilibria. We characterize in detail these conditions and, quite naturally, prove that the intensity of the coordination problems, and therefore the likelihood of multiple stationary equilibria, increases with the strength of agglomeration forces minus that of congestion forces and in the relative productivity of remote work to office work. We show that, in the U.S., the particular configuration of urban characteristics that leads to multiplicity in stationary equilibria was prevalent, before the pandemic, in many of the larger cities. That is, we estimate parameter configurations for large cities that, in general, although not always, put them in the "multiplicity cone": the parameter set where cities exhibit multiplicity in stationary equilibria.

Our finding that large cities tend to be in the multiplicity cone implies that the abrupt reduction in commuters caused by the COVID-19 pandemic made many of them switch to a path that seems to be converging to the low commuting stationary equilibrium. Of course, even in the stationary equilibrium with little commuting, some people commute; their idiosyncratic preference for office work and the fixed cost of switching make it the better choice for them. Our interpretation is that the pandemic reduced the share of commuters below the share of commuters in the low commuting stationary equilibrium. Hence, commuting has increased in some of these cities, but only slightly, and has stabilized way below pre-pandemic levels. In contrast, small cities tend to have parameter configurations outside the multiplicity cone, so their stationary equilibrium is unique. These cities experienced the same abrupt reduction in commuters as large cities did during the pandemic, but their dynamic evolution simply brings them back to the pre-pandemic stationary equilibrium. Hence, they exhibit no permanent change in the number of commuters. In sum, we argue that the pandemic has acted as an equilibrium selection device for cities where multiple stationary equilibria existed ex-ante. In a rare – in our view – empirical occurrence, the spatially-uniform temporary push of most workers into a remote state during the pandemic has revealed the set of cities featuring multiple stationary equilibria.

We contrast these predictions of our model with cell-phone-based data on worker mobility from SafeGraph, available from before the pandemic until mid-2022. We first identify CBDs in all cities by finding the group of census tracks that receive the largest numbers of commuters pre-pandemic.[2] We then track the number of trips to these CBDs over time. The results are revealing. In New York, NY, for example, relative to pre-pandemic levels, the pandemic reduced the number of trips to the CBD by more than 80%. After the pandemic, the number of trips increased slightly, but it has now stabilized at about 40% of the pre-pandemic level. San Francisco, CA, shows a similar pattern. In contrast, Madison, WI, experienced a drop in the pandemic as large as New York but has recovered fully to pre-pandemic levels. This apparent multiplicity of stationary equilibria

[2]In some cases, we identify multiple CBDs when there are several areas that have more than 80% of the commuters to the main CBD. One example is Midtown and Downtown in Manhattan.

is present more broadly in the data. The largest U.S. cities, with employment above 1.5 million workers, experienced a drop of 80% in trips to the CBD during the pandemic but now have stabilized at only 60% of pre-pandemic levels. In contrast, the smallest CBSAs, with employment below 150 thousand workers, experienced a comparable drop in CBD trips during the pandemic, but have now gone, on average, fully back to pre-pandemic levels.

The available housing data is also consistent with the mechanism in our theory. As others have documented, the pandemic generated a flattening of house price distance gradients in cities. Namely, prices at the CBD declined relative to prices further away from the city. This is a natural consequence of the expansion of work from home since the advantage of living close to work is diminished if workers do not commute to the office or do so only a couple of days a week. What has not been documented, but we document here using Zillow zip code-level data, is that the flattening in the rent gradient has continued in the largest cities but has reversed and disappeared in the smaller ones. The cities where trips to the CBD have gone back to pre-pandemic levels, like Madison, or Durham, NC, exhibit housing price gradients that are not statistically different from pre-pandemic ones. In contrast, the larger cities where CBD trips have stabilized at much lower levels, like New York and San Francisco, exhibit a large and significant flattening of this gradient, which has been exacerbated since the pandemic.

What are the welfare implications of our findings? For small cities outside the multiplicity cone, the long-run welfare effect is zero. These cities have simply reverted to the pre-pandemic equilibrium. The transition lasted a couple of years, but the transition costs are probably only marginal. In contrast, for cities that have permanently switched to the work-from-home intensive stationary equilibrium, the welfare implications depend on the welfare ranking of both stable stationary equilibria. A direct comparison of these stationary equilibria in the 155 cities that we estimate are inside the multiplicity cone reveals that the welfare costs are positive but relatively modest. The costs are slightly larger than 2.5% for New York and San Francisco, less than 2% for Phoenix, and around 3% for Los Angeles and San Jose. A stationary equilibrium with office work is indeed better than the remote work intensive stationary equilibrium, but the benefits from the lack of commuting compensate for most of the losses. Of course, these numbers depend on the details of the estimation strategy of our model and the many sources of data, including the ACS and the NLSY, that we use to implement it. We discuss the details of the estimation of our dynamic discrete choice model in the body of the paper.

Our paper is related to several strands of the literature. First, we contribute to the emerging literature studying the effects of remote work on urban environments. Liu and Su (2023) document a decrease in the urban wage premium of occupations with higher levels of remote work since the COVID-19 pandemic, consistent with a decline in the agglomeration externality for these occupations. Davis et al. (2021) develop a model where remote work productivity increases via adoption externalities and study the response of remote work employment to improvements in its technology. Parkhomenko and Delventhal (2023) develop a quantitative spatial equilibrium model with work from home and commuting that predicts the optimal frequency of remote work. They

study the spatial redistribution of economic activity induced by an increase in the amenity value of remote work. We propose an alternative interpretation to the large and persistent increase in remote work that does not rely on exogenous productivity or amenity shocks. Duranton and Handbury (2023) show that exogenous changes in which workers are allowed to work remotely in a monocentric city model can replicate observed changes in commuting and demand for housing. Instead, we show how the coordination of agents in a city can lead to multiple stationary equilibria, and how temporary shocks to the number of commuters can permanently affect the structure of cities by selecting a different stationary equilibrium.

Remote work has been linked to changes in residential and commercial housing prices within and across cities (Mondragon and Wieland, 2022; Gupta et al., 2022; Brueckner et al., 2021; Liu and Su, 2021; Ramani and Bloom, 2021; Althoff et al., 2021). Our finding that residential housing rent gradients with respect to distance to the CBD are diverging between larger and smaller cities is, to our knowledge, a new fact. We also provide an interpretation for this fact, since these housing rent gradients map directly into structural parameters of our model.[3] We show how the divergence between larger and smaller cities is predicted by our model, given that larger cities are more likely than smaller cities to converge to the remote-work-intensive stationary equilibrium.

Understanding the productivity of remote work relative to in-person work remains an active area of research. Building on earlier studies (Oettinger, 2011; Cappelli and Keller, 2013), recent work indicates that remote workers may perform better than in-person workers in some contexts (Choudhury et al., 2022; Bloom et al., 2022; Barrero et al., 2021; Etheridge et al., 2020; Bloom et al., 2015). However, workers employed in occupations involving a significant amount of problem-solving tend to experience lower performance when working remotely (Emanuel et al., 2022; Künn et al., 2020). Gibbs et al. (2023) study employee productivity before and during the COVID-19 pandemic among skilled professionals at an Indian information technology services company. They find that employee productivity fell by 8% to 19%, and point to higher communication costs and longer time spent on coordination activities and meetings as sources for this decline. They also find that employees networked with fewer individuals and business units, both inside and outside the firm.[4] Other studies emphasize that the higher earnings of remote workers can be partly explained by their selection into remote work based on unobservable characteristics (Harrington and Emanuel, 2022; Atkin et al., 2022). Our model recognizes workers' heterogeneous preferences for remote work (Mas and Pallais, 2017; Lipowska et al., 2022). Our analysis of the NLSY79 provides new estimates of the remote work premium from non-experimental data that account for unobserved persistent heterogeneity in worker performance. Our findings indicate that the technology of remote work has been improving steadily over the past decades while remaining less productive than in-person work on average across occupations.

Agglomeration economies are a central tenet of urban economics (Henderson, 2003; Combes

[3] Duranton and Puga (2019) also link housing rent gradients to the elasticity of urban costs with respect to city size.

[4] See also Yang et al. (2022). They find that firm-wide remote work resulted in a less collaborative network among workers, with fewer connections between different parts.

et al., 2012). When cities successfully bring people together, positive spillovers abound. People interact, learn from each other, and foster entrepreneurship, leading to greater productivity within the city (Moretti, 2004; Gennaioli et al., 2012; Glaeser and Maré, 2001; Baum-Snow and Pavan, 2011; De La Roca and Puga, 2016).[5] We contribute to the literature examining urban agglomeration economies by studying how remote work alters their workings. Canonical theories of cities predict that more central locations command higher house prices because of better accessibility to the locus of production (Alonso, 1964; Muth, 1969). By removing the need to travel across locations to reach the workplace, remote work deprives cities of one of their most fundamental organizational principles. The main insight from our analysis is that cities with strong agglomeration externalities, typically viewed as successful, are also the most at risk of dramatic shifts in their structure when people choose to stay home. Similar to us, Owens et al. (2020) uses coordination problems generated by residential externalities to rationalize the existence of depopulated and successful neighborhoods with similar underlying characteristics in Detroit.

The rest of the paper is organized as follows. Section 2 presents stylized facts on remote work that emphasize the sudden increase in remote work since 2020, despite no large increase in the earnings premium for remote work. Section 3 presents our dynamic monocentric model with a choice of labor delivery mode. Section 4 quantifies the model, including our estimation of the relative productivity of remote work and externality parameters. Section 5 presents additional evidence on the divergence of large and smaller cities in terms of commuting patterns and housing price gradients. It also discusses how our estimated model rationalizes this divergence and the resulting welfare implications. We relegate many of the model's and quantification's details, additional robustness exercises, and counterfactuals to the Appendix.

2 Stylized Facts on Remote Work

We begin by documenting changes in the shares of workers delivering work remotely and changes in the earnings per hour premium of remote work between 1980 and 2021. We rely on two data sources. First, we use the decennial census waves of 1980, 1990, and 2000, and the annual American Community Survey waves between 2005 and 2021. From this source, we obtain large repeated cross-sectional information on the individual earnings per hour of workers who work in person or remotely. Second, we use the National Longitudinal Survey of Youth 1979 (NLSY79), a panel of 4,147 individuals from which we collect information on the number of hours of remote work, hourly pay, and occupation every two years from 1998 to 2018. The repeated observations of the same workers allow us to control for permanent unobserved heterogeneity between workers and obtain estimates of the remote work premium that are not biased by selection on unobservable characteristics (Harrington and Emanuel, 2022; Atkin et al., 2022). We provide details about the construction of our samples in Appendix A.2.

[5]The productivity premium in larger cities has also been linked to the combination of more efficient firms and workers selecting these areas (Combes et al., 2008; D'Costa and Overman, 2014; Gaubert, 2018).

Remote work has become an increasingly common form of labor delivery over the past few decades. This pattern is typically reported by looking at the share of all working hours performed at home over time (Oettinger, 2011; Cappelli and Keller, 2013; Liu and Su, 2023). Part of this increase is likely due to improvements in communication technologies, while another part may be due to gradual changes in the composition of the labor force, both in terms of educational attainment and employment by occupation and industry. Our first exercise aims to isolate the role played by technological improvement by controlling for individual characteristics.

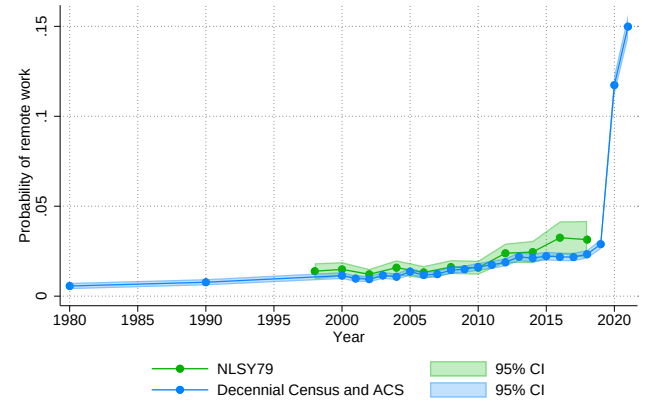
In Figure 1a, we report our estimates of the likelihood of being a remote worker over time, controlling for a number of observable covariates. Using Census/ACS and NLSY data separately, we estimate a logit model for the probability of being a remote worker. In both datasets, we control for workers' education, gender, age, marital status, race, experience on the job, and broad industry and occupation fixed effects. In the Census/ACS data, we define as "remote workers" those who report working from home when asked about their means of transportation to work. In the NLSY data, we define as "remote workers" those who work for more than 24 hours at home in a week (the equivalent of three days per week). We then predict the shares of remote workers, varying the year and setting the values of all covariates to their mean value over the sample. Both samples report a gradual increase in the share of remote workers over time. The share increases from less than 1 percent in 1980 to 2.6 percent in 2019. In 2020, this share jumps to 12.0 percent and reaches 15.0 percent in 2021.[6] The estimated share from the NLSY features a similar gradual increase from 1.4 percent in 1998 to 3.7 percent in 2018. These estimates can be interpreted as the probability of being a remote worker for a worker with typical and time-invariant characteristics. That is, the gradual increase from 1980 to 2019 is not explained by changes in the composition of employment by occupation or industry over time, nor is it explained by the changes in worker individual characteristics such as educational attainment. Rather, they likely stem from the increased attractiveness of remote work.

In Figure 1b, we estimate the hourly wage premium of remote work relative to in-person work over time. In the Census/ACS data, we regress log hourly wages on year-specific indicators of remote work, the same controls as the ones used in the logit model described above, allowing in addition for year-specific effects for occupation, industry, education, region, and demographics. In the NLSY data, we complement the specification by adding individual fixed effects. Both exercises indicate a steady increase in the premium for remote work. In 1980, remote work implied a discount of 44.5 percent in hourly rates compared to in-person work, according to the Census/ACS estimates. In 2021, workers with similar observable characteristics who were working remotely earned 4.1 percent higher earnings. According to estimates from the NLSY sample, the discount for remote work went from 18.5 percent in 2000 to only 4.9 percent in 2018. The lower values for the remote work premium in the NLSY are consistent with remote workers having characteristics

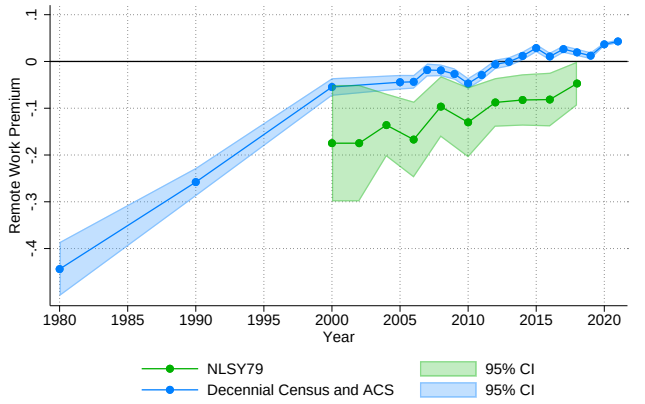
[6]Note that the value of 15.0 percent for the predicted share of remote workers does not represent the actual average share of remote workers in 2021. This predicted value is obtained by setting the value of all controls to their mean over the entire 1980-2021 period and does not represent the employment composition in 2021. In Table A1 in Appendix A.2, we report that the actual share of remote workers measured in the ACS in 2021 is 19.2 percent.

Figure 1: Prevalence and Earnings Premium of Remote Work

(a) Predicted Share of Remote Work



(b) Hourly Wage Premium of Remote Work



Panel (a) reports the predicted share of remote workers from a logit model with year fixed effects and controlling for workers' education, gender, age, marital status, race, experience on the job, and broad industry and occupation fixed effects. The estimates report the varying years, setting the values of all other covariates to their mean value over the sample. Panel (b) reports the coefficients of yearly remote work indicators from the regression of log hourly wages, including the same controls as in Panel (a) plus year-specific effects for occupation, industry, education, region, and demographics. In the Census/ACS data, "remote workers" are defined as those reporting working from home when asked about their means of transportation to work. In the NLSY data, "remote workers" are defined as those who work for more than 24 hours at home per week. Each point represents an estimate for a year. The shaded areas represent the 95 confidence interval around the estimates.

that make them likely to obtain higher earnings, beyond the ones observed in the Census/ACS. Remarkably, the remote work premium did not increase dramatically in 2020 and 2021. This is consistent with the large increase in remote work being driven by other mechanisms rather than by a large productivity increase of remote work. The model and evidence below describe such a mechanism.

3 Model

3.1 Overview

We model a linear city with an exogenous population of size L. Each city location houses one resident. Time is discrete. A worker starts period t in a labor delivery mode $\ell \in \{c, m\}$ (commuting or remote) that has been determined the previous period, and is thus fixed within the period. The worker chooses a residence location, housing, and consumption, supplying labor according to mode ℓ. Workers then see their idiosyncratic shocks $\varepsilon_t = \{\varepsilon_{c,t}, \varepsilon_{m,t}\}$ for each choice, make a labor delivery mode decision for next period, and enjoy the associated idiosyncratic benefit. These benefits are drawn from a Gumbel distribution with location parameter $\tilde{\gamma}$ (the Euler constant) and

scale parameter s, $G(\varepsilon) = \exp\left\{-e^{\varepsilon/s-\tilde{\gamma}}\right\}$. The current period ends, and in the next one the new composition of labor delivery modes is again pre-determined.

Locations in the city are ordered in a line and indexed by $d \in [0, +\infty)$. Production occurs in the central business district (CBD), at $d = 0$. Land is owned by absentee landlords. A worker living in location d pays a rent $r(d)$ per unit of housing. Land has an alternative use that yields an exogenous return equal to r_a and landowners can enhance it to supply the units of housing demanded by the household.

In what follows, we describe a simple production environment and then formalize the problem of a worker. The within-period part determines indirect flow utilities and the city rent function. The dynamic part determines the transition shares across labor delivery modes, and hence the aggregate evolution of the city. We conclude with a characterization of the equilibrium of a city.

3.2 Production

There are two goods in the economy: final consumption and housing. Final consumption is a homogeneous good produced in the CBD by profit-maximizing firms under constant returns to scale: one efficiency unit of labor produces one unit of good, at a constant marginal cost of $\bar{w}$. The number of efficiency units of a worker, and hence her total wage w_ℓ, both depend on the chosen labor delivery mode ℓ. In particular, define

$$\tilde{L}_c \equiv L_c^\mu L^{1-\mu} \tag{1}$$

as the measure of the economic activity that takes place in the city center, a function of the current number of commuters L_c, the total city size L, and the fraction of days per week a remote worker spends at home, μ. Then, we postulate

$$w_\ell \equiv \begin{cases} w_c \equiv \bar{w} \cdot A\tilde{L}_c^\delta & \text{if } \ell = c, \\ w_m \equiv \bar{w} \cdot z^\mu (A\tilde{L}_c^\delta)^{1-\mu} & \text{if } \ell = m. \end{cases} \tag{2}$$

In particular, the efficiency units offered by a commuting worker depend on the in-office productivity $A > 0$, and on how many workers are on average in the CBD, $\tilde{L}_c$, with an elasticity $\delta > 0$. The efficiency units of a remote worker are a combination of the efficiency of remote work $z > 0$, and the efficiency of commuting work, weighted by the fraction of time in each mode. Free entry in production implies $\bar{p} = \bar{w} = 1$, after a normalization.

Absentee landlords invest their effort to produce housing on their unit-size plot and receive $r(d)$ per unit of housing. Land has an alternative use that yields r_a.

3.3 Problem of the Worker

A worker starts period t in state $\ell \in \{c, m\}$, solves a within-period problem delivering a flow utility u_ℓ, observes idiosyncratic shocks associated with the next period's labor delivery mode options,

and makes a new choice ℓ'.

3.3.1 Within-Period Problem

We write the within-period problem of a worker in state ℓ as

$$u_\ell(L_c) = \max_{Q_\ell, H_\ell, d} \quad \log T^{\mu_\ell} \left(\frac{B(\tilde{L}_c)}{\tau(\tilde{L}_c) d^\gamma} \right)^{1-\mu_\ell} Q_\ell^\alpha H_\ell^{1-\alpha}, \tag{3}$$

$$s.t. \quad \mu_\ell \equiv \begin{cases} 0 & if\ \ell = c \\ \mu \in (0,1] & if\ \ell = m \end{cases}, \tag{4}$$

$$w_\ell = \bar{p} Q_\ell + r(d) H_\ell. \tag{5}$$

The worker chooses consumption of a homogeneous good, Q_ℓ, housing H_ℓ, and housing location d. μ_ℓ is the fraction of days per week spent at home in mode ℓ. Amenity and disutility terms include $T > 0$, a benefit of working from home; $B(\tilde{L}_c)$, a benefit of visiting the CBD, increasing in $\tilde{L}_c$; $\tau(\tilde{L}_c)$, the general level of congestion in commuting, also increasing in $\tilde{L}_c$, and commuting costs d^γ, where $\gamma > 0$ captures the elasticity of commuting costs to distance. The exponent $\alpha \in (0,1)$ is the weight of consumption in utility. Workers obtain a wage w_ℓ described in 2, and spend it on consumption, with unit price $\bar{p}$, and housing with unit price $r(d)$.

Our formulation captures the key trade-offs between commuting and remote work. On the amenity side, remote work allows an individual to enjoy the flexibility of working from home and to pay lower commuting and congestion costs, at the price of lower enjoyment of the social and consumption opportunities in the CBD, all for a fraction μ of the time. On the wage side, remote work trades off the positive externalities arising from in-office interactions with the efficiency of communication technology z, again for a fraction μ of the time. The scale of economic activity in the CBD, as measured by $\tilde{L}_c$, shapes the balance of these tradeoffs: when more agents visit the city center, individual productive efficiency is higher, the CBD is more lively, but congestion is heavier.

3.3.2 Dynamic Problem

The solution to the within-period problem described above delivers the within-period utility flow. The worker then observes two new idiosyncratic shocks for the next period's labor delivery mode, $\varepsilon_{\ell',t}$, and chooses ℓ' accounting for the evolution of the state of the economy. In particular, the dynamic decision of a worker who starts the period in state ℓ is

$$U_\ell \left(\{\varepsilon_{c,t}, \varepsilon_{m,t}\}; L_{c,t}, \omega^t \right) = u_\ell \left(L_{c,t} \right) + \max_{\ell' \in \{c,m\}} \left\{ \beta V_{\ell'} \left(L_{c,t+1}; \omega^{t+1} \right) - F_{\ell\ell',t} + \varepsilon_{\ell',t} \right\}, \tag{6}$$

where $\omega^t \equiv \{\omega_k\}_{k=t}^{+\infty}$ and $\omega_t \equiv \{A_t, z_t, F_{\ell\ell',t}, T_t, B_t(\cdot), \tau_t(\cdot)\}$ is a known exogenous sequence of (possibly) time-varying technology, fixed costs, and amenities, and

$$V_\ell\left(L_{c,t};\omega^t\right) = \int_{\varepsilon_m}\int_{\varepsilon_c} U_\ell\left(\{\varepsilon_c,\varepsilon_m\};L_{c,t},\omega^t\right) dG\left(\varepsilon_c\right) dG\left(\varepsilon_m\right) \tag{7}$$

is the expected utility of a worker starting in state ℓ if $L_{c,t}$ workers are commuting. In these expressions, $F_{\ell\ell',t}$ is the cost of transitioning from mode ℓ to mode ℓ' at time t. Although the theory can accommodate any non-negative values, we will set $F_{\ell\ell} = 0$ for $\ell \in \{c, m\}$. The term ε_ℓ is the idiosyncratic shock received by the worker in mode ℓ.

3.4 Equilibrium

We start by defining a within-period equilibrium condition that has to hold at each time, given the current distribution of commuters and remote workers. We then define a dynamic equilibrium, which describes the evolution of the economy when the within-period equilibrium holds at each time, and forward-looking workers make commuting and remote work choices based on—and consistent with—the aggregate evolution of the economy. Finally, we define a stationary equilibrium as a dynamic equilibrium where all endogenous variables are constant.

3.4.1 Within-Period Equilibrium

Definition 3.1 *Given L_c and ω, a within-period equilibrium is a rent function $r(d)$, $d \in [0, L]$ and an allocation $\{Q_c, Q_m, H_c, H_m\}$ such that i) individuals solve (3)-(5), ii) firms maximize profits and free entry drives profits to zero and iii) the housing market clears.*

We construct a within-period equilibrium by first solving for consumption and housing demand conditional on a distance, and then letting the agent pick the distance that delivers the highest (indirect) utility. From the Cobb-Douglas structure, we have

$$Q_\ell = \alpha w_\ell, \tag{8}$$

$$H_\ell(d) = \frac{(1-\alpha)w_\ell}{r(d)}, \tag{9}$$

which implies an indirect utility of

$$u_\ell = \log \bar{\alpha} T^{\mu_\ell} \left(\frac{B(\tilde{L}_c)}{\tau(\tilde{L}_c) d^\gamma}\right)^{1-\mu_\ell} \frac{w_\ell}{r(d)^{1-\alpha}}, \tag{10}$$

with $\bar{\alpha} \equiv \alpha^\alpha (1-\alpha)^{1-\alpha}$, for $\ell \in \{c, m\}$. Taking the first-order conditions for the optimal distance in this equation,

$$\frac{r'_\ell(d)d}{r_\ell(d)} = -\frac{\gamma(1-\mu_\ell)}{(1-\alpha)} \implies r_\ell(d) = k_\ell d^{-\frac{\gamma(1-\mu_\ell)}{(1-\alpha)}}, \tag{11}$$

where $r'_\ell(d)$ is the slope of the rent function in the segment relevant to workers with delivery mode ℓ, and k_ℓ is a constant of integration. At the optimal residence, the savings on rents of a marginally farther location exactly compensate for the increased disutility in travel time. Since remote workers suffer less for a given travel time increase (by eq. (4), $\mu_m \equiv \mu \in (0,1] > \mu_c = 0$), their willingness to pay for proximity to the CBD is lower. Hence, commuters live in $[0, L_c]$, remote workers live in $[L_c, L]$, and the equilibrium rent function has a steeper (constant) elasticity for commuters than for remote workers.

In our formulation, these conditions translate into a constant elasticity rent, with constants of integration k_c and k_m. These constants of integration are characterized by two no-arbitrage conditions. First, the rent per unit of housing in the last residence location has to coincide with the alternative land use r_a, so that $r_m(L) = r_a$. Second, land owners must be indifferent between renting to commuters and remote workers at $d = L_c$, so that $r_c(L_c) = r_m(L_c)$. Using these two conditions, the equilibrium rent function is

$$r(d) = \begin{cases} (L/d)^{\gamma/(1-\alpha)} (L_c/L)^{\mu\gamma/(1-\alpha)} r_a & \text{if } d < L_c \\ (L/d)^{(1-\mu)\gamma/(1-\alpha)} r_a & \text{if } d \geq L_c \end{cases}. \tag{12}$$

An implication of this result is that the average elasticity of rent to distance is given by:

$$\frac{\partial r(d)}{\partial d} \frac{d}{r(d)} = -\frac{\gamma}{1-\alpha} \left[(1-\mu) L + \mu L_c \right]. \tag{13}$$

Hence, the rent gradient reflects the strength of distance frictions, as measured by γ, and the average value of being close to the CBD, as proxied by the commuting population L_c.

A within-period equilibrium is then described by this rent function, and the allocation implied by equations (8)-(9). The indirect utility function (10) becomes,

$$u_\ell = \begin{cases} \log \left[r_a^{\alpha-1} \bar{\alpha} \left(\frac{B(\tilde{L}_c)}{\tau(\tilde{L}_c)} \right) w_c / \tilde{L}_c^{\gamma} \right] & \text{if } \ell = c \\ \log \left[r_a^{\alpha-1} \bar{\alpha} T^{\mu} \left(\frac{B(\tilde{L}_c)}{\tau(\tilde{L}_c)} \right)^{1-\mu} w_m / L^{\gamma(1-\mu)} \right] & \text{if } \ell = m \end{cases}. \tag{14}$$

These indirect utilities of each labor delivery mode are an element of the dynamic problem of a worker, to which we now turn.

3.4.2 Dynamic Equilibrium

A worker starts the period in state ℓ and enjoys the utility flow u_ℓ. The worker then observes two new idiosyncratic shocks $\varepsilon_{\ell',t}$ and decides the new state for the next period ℓ', accounting for the exogenous evolution ω^t and the aggregate measure of commuting workers $L_{c,t+1}$. With some

manipulation,[7] we can rewrite equation (7) as

$$V_\ell\left(L_{c,t};\omega^t\right)=u_\ell\left(L_{c,t};\omega_t\right)+\beta V_\ell\left(L_{c,t+1};\omega^{t+1}\right)-F_{\ell\ell,t}+\Omega\left(\bar{\varepsilon}_{\ell,t}\right), \tag{15}$$

where $\bar{\varepsilon}_{\ell,t}\equiv\{\bar{\varepsilon}_{\ell c,t},\bar{\varepsilon}_{\ell m,t}\}$, the term

$$\bar{\varepsilon}_{\ell\ell',t}\left(L_{c,t+1}\right)\equiv\beta\left[V_{\ell'}\left(L_{c,t+1};\omega^{t+1}\right)-V_\ell\left(L_{c,t+1};\omega^{t+1}\right)\right]-\left(F_{\ell\ell',t}-F_{\ell\ell,t}\right), \tag{16}$$

is the average value of switching from ℓ to ℓ', and $\Omega\left(\bar{\varepsilon}_t^\ell\right)$ is the value of the option of being able to move out of ℓ if circumstances require, or the "option value" of ℓ. Under our distributional assumptions, the option value takes the form

$$\Omega\left(\bar{\varepsilon}_{\ell,t}\right)=-s\ln\lambda_{\ell\ell,t}, \tag{17}$$

where the share of workers in state ℓ transitioning to ℓ' is given by

$$\lambda_{\ell\ell',t}=\frac{\exp\left(\bar{\varepsilon}_{\ell\ell',t}\left(L_{c,t+1}\right)/s\right)}{\exp\left(\bar{\varepsilon}_{\ell c,t}\left(L_{c,t+1}\right)/s\right)+\exp\left(\bar{\varepsilon}_{\ell m,t}\left(L_{c,t+1}\right)/s\right)}, \tag{18}$$

and the evolution of the measure of commuting workers is governed by the law of motion

$$L_{c,t+1}=L_{c,t}\lambda_{cc,t}+\left(L-L_{c,t}\right)\left(1-\lambda_{mm,t}\right). \tag{19}$$

Definition 3.2 *Given a known exogenous sequence of parameters $\{\omega^t\}_{t=0}^{+\infty}$ and an initial condition $0\leq\bar{L}_{c,0}\leq L$, a dynamic equilibrium is a pair of value functions V_ℓ, $\ell\in\{c,m\}$ and a matrix of transition probabilities $\lambda_t^{\ell\ell'}$, $\ell,\ell'\in\{c,m\}$ such that eq. (15)-(18) hold, and the state of the economy evolves according to (19) with $L_{c,0}=\bar{L}_{c,0}$.*

We are ready to define a dynamic equilibrium, as well as a stationary equilibrium.

Definition 3.3 *Given a time-invariant sequence of exogenous parameters $\omega\equiv\{\bar{\omega}\}_{t=0}^{+\infty}$, a stationary equilibrium is a dynamic equilibrium where all endogenous variables are constant over time.*

Throughout the text, we focus on multiplicity of stationary equilibria. In principle, the model might feature multiple equilibria, something we have not fully explored theoretically. In our numerical exercises, we have not found indications of the presence of such multiplicity.

3.4.3 Characterization of the Equilibrium

To discuss the existence and other properties of the stationary equilibria, we specify the shape of $B(\cdot)$ and $\tau(\cdot)$, and ensure that the model does not deliver corner solutions. In the rest of the paper, we maintain the following two assumptions.

[7] All calculations are reported in Appendix B.1.

Assumption 1 *$B(\cdot)$ and $\tau(\cdot)$ are isoelastic functions:*

$$B(\tilde{L}_c) \equiv \bar{B}\tilde{L}_c^{\xi} \text{ and } \tau(\tilde{L}_c) = \bar{\tau}\tilde{L}_c^{\theta}, \text{ with } \bar{B}, \bar{\tau}, \xi, \theta > 0.$$

Assumption 2 *The dispersion in the idiosyncratic shocks is large enough:*

$$s > \mu^2(\xi + \delta - \theta - \gamma/\mu).$$

Assumption 1 allows a characterization of the equilibria in terms of comparisons among (constant) elasticity values. Assumption 2 guarantees that as an economy approaches a purely remote equilibrium, the average value of switching from commuting to remote goes to $-\infty$.

Proposition 3.1 *There always exists at least one stationary equilibrium.*
Proof. *See Appendix B.2.* ■

The proof makes a formal case that the average benefit of switching to commuting approaches $+\infty$ when almost everybody is working remotely, which causes the number of commuters to stay above zero. Conversely, when almost all employees commute, the number of remote workers tends to stay above zero. Since the law of motion for L_c is continuous, there must be at least one number of current commuters for which the number of commuters in the following period remains constant.

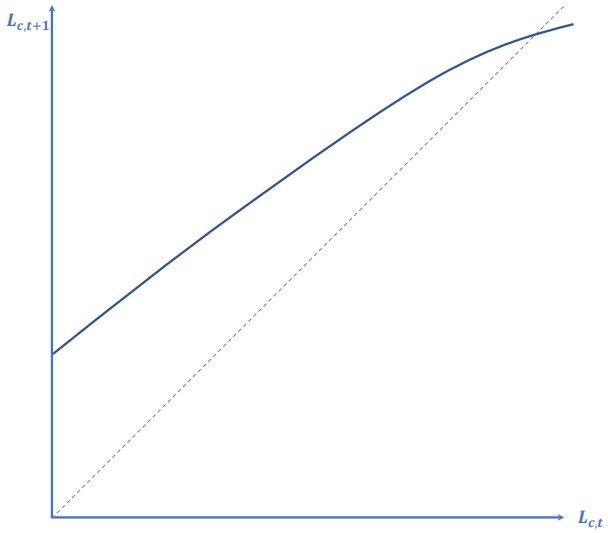
Having established existence, we now investigate the properties of the set of equilibria in this economy. Proposition 3.2 establishes that if agglomeration forces are weak enough, there is a unique stationary equilibrium.

Proposition 3.2 *Assume $\delta + \xi < \theta + \gamma/\mu$. Then there is a unique stationary equilibrium. The stationary equilibrium value of the commuting population $L_{c,ss}$ monotonically increases with A and $\bar{B}$, and monotonically falls with Z, T and $\bar{\tau}$.*
Proof. *See Appendix B.3.* ■

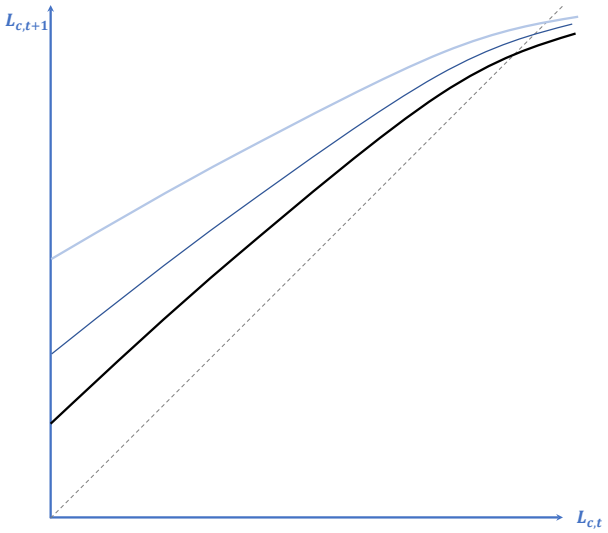
A unique stationary equilibrium requires that having more commuters increases the individual incentive to become a remote worker. We can break down this individual incentive into two parts: the current benefit of actually working remotely instead of commuting, and the option value of becoming a remote worker. If the measure of commuters increases, the option value of becoming a remote worker always grows - a dynamic congestion force. Under the assumptions of Proposition 3.2, static congestion forces also increase the net benefit $u_m - u_c$ when the number of commuters grows, and this is sufficient for uniqueness. The left panel of Figure 2 depicts a typical phase diagram for this case. The blue line indicates how the measure of commuters changes from one period t (horizontal axis) to the next $t+1$ (vertical axis). The city's unique stationary equilibrium occurs at the crossing with the dashed 45° line. A city below the stationary equilibrium value will experience an increase in the number of commuters over time. However, congestion forces will eventually slow this growth as the number of commuters rises.

Figure 2: Phase Diagram for a Unique Stationary Equilibrium

(a) Unique Stationary Equilibrium



(b) Increasing Remote Work Productivity



The left panel of this figure shows a typical phase diagram for a city with a unique stationary equilibrium. The right panel shows how the phase diagram shifts in response to increases in remote work productivity, from the light blue line to the black line. In both panels, the dashed gray line is the 45-degree line.

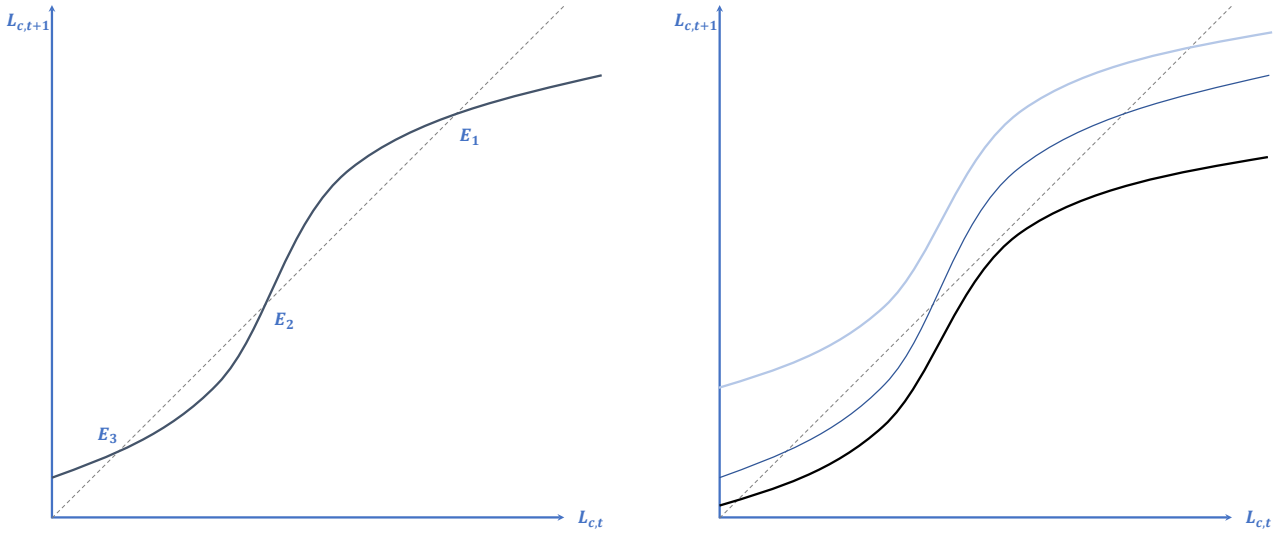
Increases in office productivity A and amenity benefits $\bar{B}$ reduce the static gains of remote work, lessening the share of commuters making the transition. These changes would move the blue line up and to the left, increasing the number of steady-state commuters. Conversely, improvements in the efficiency of remote work Z or its amenity benefit, T, or higher commuting frictions $\bar{\tau}$ increase the motivation to move to remote work, pushing the line down and to the right, and reducing the equilibrium number of commuters. Panel (b) of Figure 2 shows the consequences of a progressive improvement in remote work productivity.

The discussion above suggests that multiplicity can arise when an increase in the number of commuters strengthens the incentive to switch to commuting. These conditions occur when static agglomeration forces overcome dynamic congestion forces. The left panel of Figure 3 illustrates a typical phase diagram for this scenario with three stationary equilibria, E_1, E_2, and E_3. If a city is located between E_1 and E_2, the number of commuters tends to increase towards E_1. However, if a city is situated between E_2 and E_3, the number of commuters will decrease until it reaches the stationary equilibrium E_3. Thus, E_2 is an unstable stationary equilibrium, and any perturbation away from it will cause the city to depart from that equilibrium.

Perhaps surprisingly, strong agglomeration forces alone are not sufficient for multiplicity. The right panel of Figure 3 shows that the phase diagram shifts down and to the right, from lighter to darker colors, as remote work efficiency improves. If remote work is not efficient, having fewer commuters does not increase the incentive to work from home enough to induce more workers to work remotely. Conversely, if remote work is highly efficient, the "high commuting" equilibrium

Figure 3: Phase Diagram for Multiple Stationary Equilibria

(a) Multiple Stationary Equilibria (b) Increasing Remote Work Productivity



The left panel of this figure shows a typical phase diagram for a city with multiple stationary equilibria. The right panel shows how the phase diagram shifts in response to increases in remote work productivity, from the light blue line to the black line. In both panels, the dashed gray line is the 45-degree line.

vanishes as the incentive to switch to remote work is consistently strong. Proposition 3.3 makes this intuition formal and characterizes the complete set of circumstances in which we obtain multiplicity of stationary equilibria.

Proposition 3.3 *Suppose $F_{cc} = F_{mm} = 0$. Then, there exist a finite threshold $\eta_{min} > \theta + \gamma/\mu$, and a set $\mathcal{Z} \subset \mathbb{R}_{++}$, such that:*

i) for $\delta + \xi > \eta_{min}$, $\mathcal{Z}$ is a non-empty interval (Z_{min}, Z_{max}) and there are multiple stationary equilibria if $z/A \in \mathcal{Z}$; further, Z_{min} and Z_{max} grow with L and $\bar{B}$, and fall with T and $\bar{\tau}$;

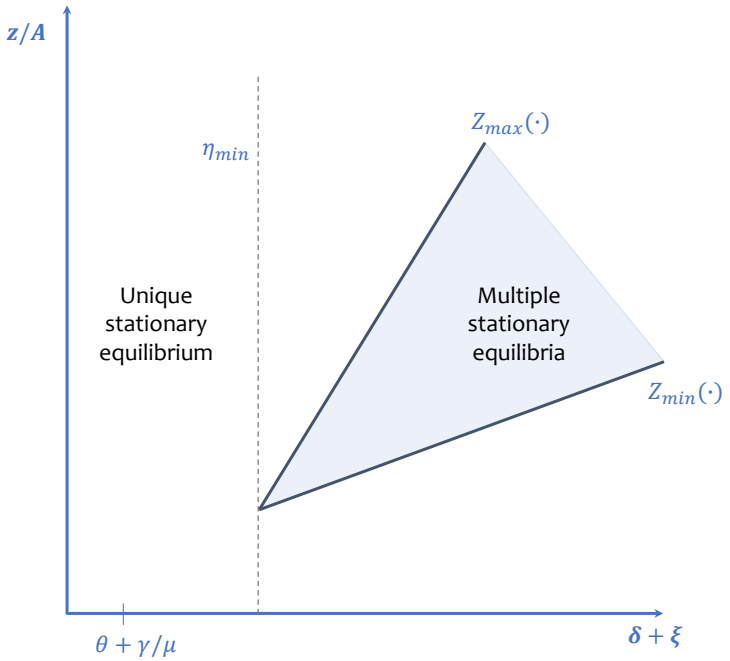
ii) there is a unique stationary equilibrium in all other cases.

Proof. *See Appendix B.4.* ■

Figure 4 provides a visual representation of this Proposition. The figure displays the combinations of $(\delta + \xi, z/A)$ for which a city exhibits multiple stationary equilibria, holding the remaining parameters constant. When $\delta + \xi < \theta + \gamma/\mu$, no values of z/A can generate multiple stationary equilibria. This is essentially a re-statement of Proposition 3.2.[8] If agglomeration forces are high

[8]Proposition 3.3 imposes $F_{cc} = F_{mm} = 0$, a slightly more restrictive assumption. Without it, $\mathcal{Z}$ need not be an interval for low values of $\delta + \xi$. Instead, there are two thresholds, η_{min} and $\eta_{max} \geq \eta_{min}$; $\mathcal{Z}$ is generally the union of disjoint intervals, and it is exactly one interval only when $\delta + \xi > \eta_{max}$. We believe it is reasonable to assume no fixed costs to remain in the same labor delivery mode. Therefore, we focus on this case in the main text.

Figure 4: Cone of Multiplicity



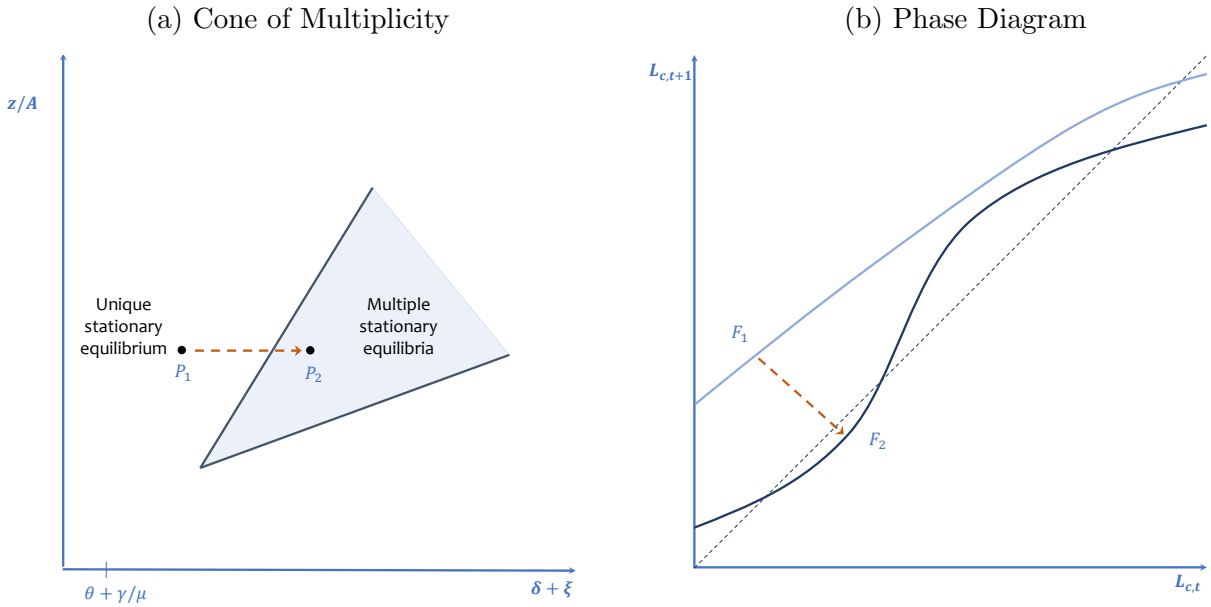
This figure represents the combinations of z/A and $\delta+\xi$ for which, given the remaining parameters, the city displays multiple stationary equilibria. Changes in the remaining parameters move the cone. The dashed gray line indicates the minimum value of agglomeration externalities that generate multiple stationary equilibria.

enough ($\delta + \xi > \eta_{min} > \theta + \gamma/\mu$), there is a range of z/A where the city exhibits multiplicity. As in the right panel of Figure 3, the efficiency of remote work can neither be too low nor too high.

An implication of Proposition 3.3 is that for some values of z/A, increasing agglomeration forces will change the structure of the city from a single to multiple stationary equilibria. This fact is easily seen in Figure 5. In the left panel, we consider a movement from point P_1, outside the cone of multiplicity, towards P_2, inside the cone. We know this movement can be generically constructed because $\mathcal{Z}$ is non-empty for at least some values of $\delta+\xi$. This comparison corresponds to the movement from the phase diagram F_1 to F_2 in the right panel. An increase in the strength of agglomeration forces exacerbates the role of coordination mechanisms and can create multiple stationary equilibria.

These results can speak to the organization of production in cities both in the long and short run. In the long run, secular trends alter the position of a city with respect to its cone. For example, the trends in remote work discussed in Section 2 suggest that a city with a high enough $\delta+\xi$ may traverse its cone from a single equilibrium to a region of multiplicity, and suddenly lose the high-commuting equilibrium in favor of the low-commuting one. From a shorter-run perspective, lockdown policies during the recent COVID-19 pandemic can be seen as a negative exogenous shock to the number of commuters. Our theory envisions heterogeneous adjustments across cities,

Figure 5: Implication of Higher Agglomeration Forces



This figure describes a comparison between low and high values of δ. The left panel shows that for some values of z/A, it is possible to move from a unique (P_1) to multiple (P_2) stationary equilibria. The right panel displays how the phase diagram changes from F_1 to F_2 in association with this move. The dashed gray line in the right panel shows the 45-degree line.

depending on whether they find themselves inside or outside the cone when the shock occurs. In the next section, we turn to a theoretical analysis of these shocks. After quantifying our model for different cities in Section 4, we relate the current evidence to these results in Section 5.

3.5 Heterogeneous Responses to Lockdown Shocks

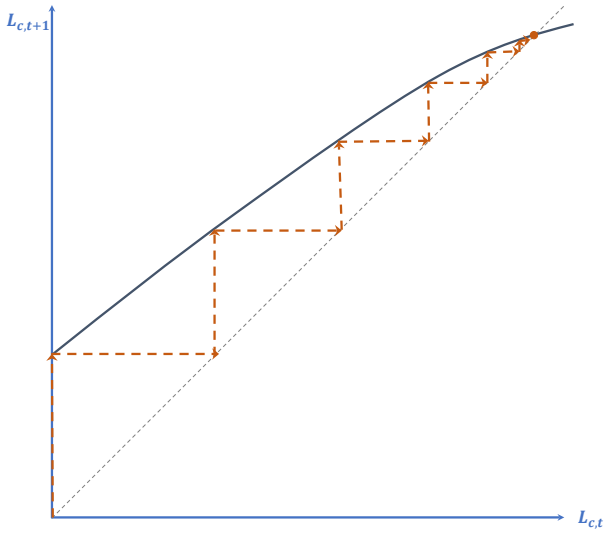
Both individual and policy-level responses to the COVID-19 pandemic have led, during March 2020, to large shifts in labor delivery mode towards remote work (Brynjolfsson et al., 2020; Aksoy et al., 2022; Barrero et al., 2021). Our model provides a natural framework to analyze the subsequent adjustment toward the resulting stationary equilibrium for different types of cities. In particular, in Figure 6, we compare a city with a single stationary equilibrium (left panel) to a city with multiple stationary equilibria, initially in the high-commuting stationary equilibrium (right panel).

We can induce a city lockdown by setting F_{cc} and F_{mc} to arbitrarily high values for a specific number of periods. From the stationary equilibrium E_1, this shock initially reduces the measure of commuters $L_{c,t}$ to almost zero. Subsequently, as the restrictions are lifted and public confidence returns (that is, F_{cc} and F_{mc} go back to their original values), individuals re-optimize their labor delivery mode. The set of incentives workers face, however, is now different between the two city structures.

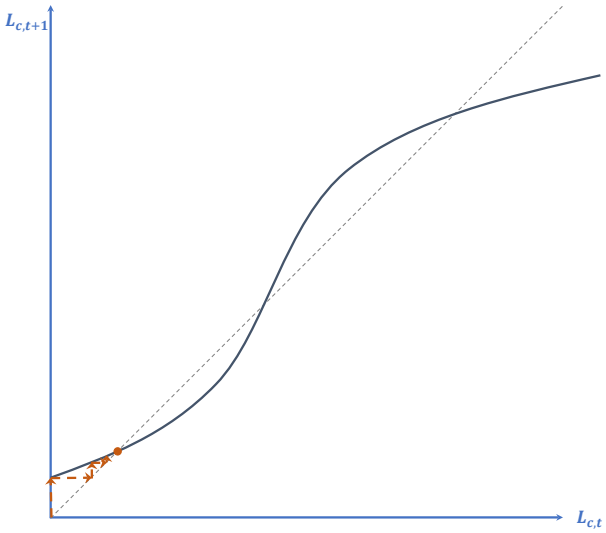
In cities with a unique equilibrium (e.g., where $\xi + \delta < \theta + \gamma/\mu$), the aggregate number of

Figure 6: Heterogeneous Responses to Lockdown Shocks

(a) Unique Stationary Equilibrium



(b) Multiple Stationary Equilibria



This figure shows the adjustment towards the stationary equilibrium in the case of unique vs. multiple stationary equilibria. In both panels, the orange line depicts the adjustment towards a stationary equilibrium following an initial shock that brings the number of commuters $L_{c,t}$ to zero. The dashed gray line in the right panel shows the 45-degree line. When the stationary equilibrium is unique (left panel), the city returns to the initial stationary equilibrium. With multiple stationary equilibria, the city converges to the low-commuting one.

commuters $L_{c,t}$ has limited influence on the net individual benefit of working in the office. In this environment, technology is the primary determinant of wage differentials, and the amenity value of being downtown is largely independent of the central business district's overall vibrancy. Because externalities have little impact on individual choices, the city's fundamentals ultimately restore the initial equilibrium, as depicted by the path of the orange line in Panel (a). Even if the willingness to pay for proximity to the CBD initially falls, the rent gradient in eq. (13), which is tied to the measure of the commuting population, returns to its initial value.

In contrast, the economy's fundamentals are not the sole determinant of the final aggregate state of cities with multiple stationary equilibria. In this case, the current number of commuters powerfully shapes the incentives to return to the office. A lockdown pushing the city into low commuting levels significantly alters the individual productivity differential in favor of remote work. While travel congestion may decrease, the allure of the downtown area also diminishes, and the amenity value of working in the central business district vanishes. As illustrated by the orange line in Panel (b) of Figure 6, the city finally settles into a low commuting stationary equilibrium. Since the number of commuters is permanently lower, the rent gradient in eq. (13) is also flatter and closer to zero.[9] Equipped with these results, we turn to quantify our model for hundreds of

[9]It is in principle possible for these models to exhibit multiplicity of equilibria, not only of stationary equilibria. For example, it could be the case that, although the city starts from low commuting levels, individual expectations

CBSAs, assemble a set of reduced-form pieces of evidence, and relate them to predictions of the quantified cities.

4 Quantification

When cities can potentially feature multiple equilibria, quantifying the model requires either taking a stand on which equilibrium the data is a realization of, or relying on relationships that hold in any equilibrium. For this reason, we rely entirely on data that was produced before 2019, a period during which every city was arguably in the high commuting equilibrium. More importantly, we only use predictions of the model that do not depend on which equilibrium—stationary or not—is being realized.

We quantify our model by estimating parameters that are specific to each Core-based Statistical Area (CBSA) in the United States when possible. We proceed in four steps. First, we estimate the dispersion of idiosyncratic preferences, which is inversely related to the elasticity of transitions into remote work with respect to earnings and the fixed cost of transitioning between delivery modes. Second, we construct estimates of the agglomeration externality in production. Third, we estimate the parameters of the remote work technology. Fourth, we estimate the transportation cost elasticity from city-specific measures of housing rent distance gradients.

4.1 Estimation of the Transition Elasticity and Transition Costs

We start with the identification of the variance of idiosyncratic preference shocks for labor delivery modes, the fixed costs, and a time trend for the amenity value of remote work. As in Proposition 3.3, we assume $F_{cc} = F_{mm} = 0$. We further assume that $F_{cm} = F_{mc} = F$, reducing the quantification of fixed transition costs to the identification of F. Following the work of Artuç et al. (2010) on the estimation of discrete choice models, and of Arcidiacono and Miller (2011) in the structural labor literature on conditional choice probabilities (CCPs), we exploit the log-linear relationship between observed transition probabilities across labor delivery modes and workers' payoffs, including transition costs, income, and continuation values. Through a linear regression, the elasticity of transitions between labor delivery modes is identified from the responsiveness of transitions to income changes, while fixed transition costs are identified from the overall level of transitions.[10] We use the following result, which provides an expression of forward-looking relative transition probabilities as a function of the within-period flows of utility, transition costs F, and transition elasticity s only.

of a return to high commuting put the system on a self-fulfilling path that converges to a high commuting stationary equilibrium. We have searched for these cases extensively in our simulations and have found no instances where this is the case. See Appendix B.6 for details.

[10] See Traiberman (2019) for an application to occupational and sectoral transitions.

Lemma 4.1 *Suppose* $F_{cc} = F_{mm} = 0$ *and* $F_{cm} = F_{mc} = F$. *Then*

$$y_{\ell\ell',t} \equiv \ln \frac{\lambda_{\ell\ell',t}\lambda^{\beta}_{\ell'\ell',t+1}}{\lambda_{\ell\ell,t}\lambda^{\beta}_{\ell\ell',t+1}} = \frac{\beta}{s}\left(u_{\ell',t+1} - u_{\ell,t+1}\right) - \frac{(1-\beta)F}{s}. \tag{20}$$

Proof. See Appendix B.5. ■

This approach considers one-period deviations from an otherwise common sequence of labor delivery modes to difference out continuation values. We can write the difference in the utility flows from (14) as

$$u_{m,t+1} - u_{c,t+1} = \ln \frac{w_{m,t+1}}{w_{c,t+1}} + \mu \ln T_{t+1} - \mu \ln B(\tilde{L}_{c,t+1}) + \mu \ln \tau(\tilde{L}_{c,t+1}) + \gamma\mu \ln L_{c,t+1}. \tag{21}$$

In order to obtain a specification that can be estimated in a linear regression, we maintain Assumption 1 about the parametric specification of the city-center amenity $B(\tilde{L}_c)$ and congestion $\tau(\tilde{L}_c)$, and assume that the remote work amenity takes the form of a log-linear function of time. That is, we let $T_t = \exp\{\alpha_0 + \alpha_1 t\}$. Our estimating equation can then be written as:

$$y_{\ell\ell',t} = \eta_0 \cdot \ln \frac{w_{\ell',t+1}}{w_{\ell,t+1}} + \eta_1 + d_{\ell\ell'} \cdot (\eta_2 + \eta_3 \cdot (t+1) + \eta_4 \cdot \ln L_{c,t+1} + \eta_5 \cdot \ln L_{t+1}) + \varepsilon_{\ell\ell',t}, \tag{22}$$

where $d_{\ell\ell'} = 1$ is a discrete variable taking the value 1 if the transition is from commuting to remote, i.e. if $\ell\ell' = cm$ and takes the value -1 if the transition is from remote to commuting, i.e. $d_{\ell\ell'} = -1$ if $\ell\ell' = mc$.

With this specification, the estimated coefficients can be recovered according to

$$\begin{aligned}
s &= \frac{\beta}{\eta_0}, && (23)\\
F &= -\frac{\beta}{1-\beta}\frac{\eta_1}{\eta_0}, && (24)\\
\alpha_0 - \ln \frac{\bar{B}}{\bar{\tau}} &= \frac{\eta_2}{\eta_0\mu}, && (25)\\
\alpha_1 &= \frac{\eta_3}{\eta_0\mu}, && (26)\\
\xi - \theta &= \frac{1}{\mu(1-\mu)}\frac{\eta_5}{\eta_0}, && (27)\\
\gamma &= \frac{1}{\eta_0}\left(\frac{\eta_4}{\mu} + \frac{\eta_5}{1-\mu}\right). && (28)
\end{aligned}$$

In principle, we should compute $y_{\ell\ell',t}$ for each city. However, the NLSY sample does not contain enough observations of these transitions to obtain them accurately for each city. This is especially true if we define remote work as spending three days per week or more at home, a choice that

was rare before the COVID-19 pandemic.[11] For this reason, we impose additional assumptions to implement the estimation of (22). In our main specification, we partition cities based on four U.S. regions, the most precise geographic units recorded in the NLSY public-use data.[12] We define transition shares for each of these four regions and for two groups of educational attainment (with or without a college degree). That is, for each year we construct the empirical equivalent of $y^r_{\ell\ell',t}$ in (22) for eight region-education groups, indexed by r. We define a remote worker as someone who spends at least 1 day per week at home to maximize the number of observed transitions.[13]

The construction of $y^r_{\ell\ell',t}$ for each group r is based on the fraction of individuals $\lambda^r_{\ell\ell',t}$ who make a transition to or from remote work between two consecutive periods of time. To compute these shares for each group r, we first control for the role of covariates other than education and region in driving the transition shares of each group that are not directly captured by our model. These include differences in the industrial and occupational composition across groups that could drive differences in transition shares. Specifically, we estimate a probit model for the share of remote workers, controlling for year-group r, t indicators, the role of age and experience, and occupation and industry fixed effects. We use as measures of $\lambda^r_{\ell\ell',t}$ the predicted transition shares from our probit model for each group r, setting all other covariates to their mean value. We then assemble the yearly transition shares into the relative shares $y^r_{\ell\ell',t}$, setting β to 0.96 on an annual basis, corresponding to a discount rate of about 4 percent.

To construct wages $w^r_{c,t}$ and $w^r_{m,t}$, we follow a similar approach by removing the contribution of industrial and occupational determinants of the wage levels. We regress individual log hourly wages on age, tenure and their square, year-group r, t indicators, and person, industry, and occupation fixed-effects. We use the residuals from these regressions to obtain a measure of wages for each year-group r, t after controlling for individual variation. The wage for a region is then the average residual wage for all the observations in a region-time. Finally, we construct measures of total employment L^r_t and of the number of commuters $L^r_{c,t}$ in each year-group r, t from the ACS.

Table 1 reports the results of the estimation of (22) and the implied structural parameters.[14] In the first column, we report the estimates from the specification in (22), with both the number of commuters $L^r_{c,t+1}$ and the total employment L^r_{t+1} as covariates. In practice, however, these two variables are almost collinear, and the coefficients on these two covariates are poorly identified. For this reason, we report in the second column the estimates by including only the number of

[11] Out of the 4,847 individuals observed between 2000 and 2018 in our sample, only 386 ever reported working remotely for 3 days per week, while 650 and 1,074 worked remotely two days or more per week, and one day or more per week, in at least one of the survey years, respectively.

[12] The four regions are Northeast, North-Central, South, and West.

[13] In Appendix A.3.2, we conduct an alternative exercise assuming that cities can be partitioned into subsets that offer similar remote work conditions and define remote work as spending at least two days per week at home. We then combine the decisions of individuals in cities in a given subset and form transition shares at that level. We obtain similar parameter estimates.

[14] The total number of possible observations is the product of the number of groups r (8) with the number of distinct years (9 for every two years between 2000 and 2016), times two for each direction of transition (to remote, or to commute), i.e. 144 observations. In practice, however, there are only 48 observations with non-missing values of $y^r_{\ell\ell',t}$.

commuters $L^r_{c,t+1}$ as a covariate. The coefficients on the log hourly wage difference and on the constant term are very similar, and they provide estimates of the elasticity of transitions to wages s, and of the fixed transition cost that are statistically significant at the 1 percent level.

We obtain an estimate of s equal to 0.46. We are not aware of existing estimates of this parameter. As indicated by the coefficient in the first row of regression (1) in Table 1, this value of s corresponds to an elasticity of transitions into remote work of 1.98. This elasticity can be interpreted in two slightly different ways. First, if we return to (95) and (21), it means that an increase of 1 percent in the wage of remote workers relative to the wage of commuters can induce an "acceleration" of transitions to remote work by 1.98 percent. That is, transitions into remote work would increase in period t to benefit from these higher wages—the numerator in the definition of $y_{\ell\ell',t}$—relative to transitions that would have occurred in period $t+1$—the denominator. Alternatively, one can easily show from (18) that the elasticity of the probability of transition to remote work $\lambda_{cm,t}$ (relative to the probability of not transitioning $\lambda_{cc,t}$) with respect to the relative *value* of remote work, $V_{m,t+1} - V_{c,t+1}$, is also β/s. That is, if the wage of remote workers relative to the wage of commuters were to increase permanently by 1 percent, then transitions into remote work would increase by 1.98 percent in period t.

Our estimate for the transition cost F is 2.7. Given the specification of the utility of consumption in log terms, we can interpret the cost F as being equivalent to giving up $100(1-\exp(-F)) = 93$ percent of earnings in year t in order to transition into remote work. This large number is reminiscent of the very large transition costs estimated in structural labor models.[15] These large estimates for the cost F reflect that despite the existence of wage differentials between labor delivery modes, the mobility of workers remains quite low.[16] Of course, if transitioning between modes also implies changing homes in order to live at the appropriate distance to the CBD, large fixed costs are natural.

All other coefficients are imprecisely estimated, and we have no further guidance on the values for ξ, θ, α_1 and the combination $\alpha_0 - \ln(\bar{B}/\bar{\tau})$. Hence, we set them all to zero for the rest of our analysis, except for $\bar{\tau}$. The parameter $\bar{\tau}$ will be calibrated for each city to match the initial value of L_c/L, as described in section 4.5. In Section 4.4, we recover estimates γ_j of the elasticity of transportation costs with distance for each city j.

4.2 Agglomeration Externalities

We estimate the agglomeration externality δ_j for each Core-based statistical area (CBSA) j. The interpretation of δ_j is the percentage increase in labor value added for a one-percent increase in the number of commuters to the CBD in city j. In order to construct city-specific estimates,

[15] Artuç et al. (2010) estimate switching costs between occupations that are equal to at least four times annual wages. In their study of location decisions, Kennan and Walker (2011) estimate moving costs amounting to $312 thousand 2010 dollars on average.

[16] We plan to re-visit this estimation procedure with the restricted-use version of the NLSY data that is geocoded at the county level. With more detailed location data, we would be able to obtain more observations.

Table 1: Estimates of the Gravity Equation for Mode Transitions and Structural Parameters

Covariate	(1) $y^r_{\ell\ell',t}$	(2) $y^r_{\ell\ell',t}$	Parameter	(1) Estimates	(2) Estimates
$\ln \frac{w^r_{\ell',t+1}}{w^r_{\ell,t+1}}$	1.976***	1.984***	s	0.466***	0.464***
	(0.439)	(0.457)		(0.103)	(0.107)
$d_{\ell\ell'}$	0.103	0.442	F	2.735***	2.719***
	(1.926)	(1.512)		(0.802)	(0.811)
$d_{\ell\ell'}(t+1)$	-0.096	-.071	$\alpha_0 - \ln \frac{\bar{B}}{\bar{\tau}}$	0.209	0.445
	(0.061)	(0.050)		(3.898)	(1.526)
$d_{\ell\ell'} L^r_{c,t+1}$	-8.600	0.014	α_1	-0.195	-0.072
	(20.02)	(0.105)		(0.134)	(0.054)
$d_{\ell\ell'} L^r_{t+1}$	8.631		$\xi - \theta$	0.745	
	(20.09)			(1.522)	
$cons$	-0.459***	-0.459***	γ	0.115	
	(0.096)	(0.095)		(0.310)	
N	48	48			
R^2	0.24	0.24			

Robust standard errors in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Structural parameters estimated with the Delta method.

we rely on industry-specific estimates of agglomeration externalities obtained by Rossi-Hansberg et al. (2021). They specify a production function featuring labor productivity that is influenced by external effects stemming from the local composition of labor. They recover the endogenous levels of local productivity that rationalize observed data on wages, employment, and input-output linkages as an equilibrium. Given their measures of productivity, they estimate the degree of local externalities for four industry groups and two broad occupational categories.[17]

We construct our measure of city-specific agglomeration externalities by computing the average of the four industry-specific externality parameters weighted by each city's employment shares in the four industries. We compute these industry shares using the County Business Patterns in 2019.[18] Specifically, we compute the city-specific averages of the externality parameters τ_L in Table 1 of Rossi-Hansberg et al. (2021). The summary statistics of our estimates of agglomeration externalities are reported in Table 2. The average value of the agglomeration externality across 639

[17]These groups of industries are based on the NAICS 1997 industry classification. They are Health and education (HE), Professional and other services (PS), Manufacturing (M), Accommodation, trade, and transportation (ATT). The two occupational categories are Cognitive Non-Routine (CNR) and non Cognitive Non-Routine (non-CNR). They define CNR occupations as all occupations with SOC-2 classifications 11 to 29 and non-CNR occupations as all other occupations.

[18]We use the crosswalk provided by Rossi-Hansberg et al. (2021) to map the four groups of industries defined by NAICS 1997 into the industries used by the County Business Patterns. See Appendix Table A.16 page 80 of Rossi-Hansberg et al. (2019).

CBSAs is 0.089, which is close to estimates obtained in different contexts by other studies.[19] The estimates for all cities are depicted in Figure A4 in Appendix A.8, which shows a strong negative correlation between the cities' agglomeration externalities and the extent to which foot traffic in their CBD has recovered from pre-pandemic levels.

4.3 Remote Work and In-Office Technology

We also estimate in-office and remote productivities $A_{j,t}$ and $z_{j,t}$ for each CBSA j. We modify the exercise presented in Section 2 by estimating the remote work premium by occupation. Using data from the NLSY, we estimate the remote work premium for every incremental hour spent working from home. Specifically, we denote by $h_{remote_{i,t}}$ the number of hours worked at home, and predict $\hat{h} = 24$ hours per week, equivalent to three days per week of remote work.[20]

We estimate a national trend and an intercept for each of the 22 occupation groups, denoted by o.[21] Specifically, we estimate

$$\ln w_{i,t} = (\psi^0_{o(i)} + \psi^1 t) h_{remote_{i,t}} + X_{i,t}\beta_{o(i)} + \epsilon_{i,t}, \tag{29}$$

where $\ln w_{i,t}$ represents the log of the hourly wage for individual i at time t, $\psi^0_{o(i)}$ and ψ^1 represent the intercept for occupation group o and the national trend, respectively, and $X_{i,t}$ is a vector of individual-level control variables, including age, age squared, tenure at the job, tenure squared, industry, marital status, race, education-year, number of children, metro area-year, and individual fixed effects. We focus on full-time US-born workers in metropolitan areas, observed between 2000 and 2018.

Finally, we estimate the remote work premium in city j using

$$\psi_{j,t} = \sum_o sh^{2018}_{o,j}(\psi^0_o + \psi^1 t)\hat{h}, \tag{30}$$

where $sh^{2018}_{o,j}$ is the share of workers in city j working in occupation o in 2018. Figure 7 displays the estimates $(\hat{\psi}^0_{o(i)} + \hat{\psi}^1 t)\hat{h}$ for $t = 2018$ for each of the 22 occupation groups obtained from (29), and the estimates $\psi_{j,2018}$ of the remote work premium by CBSA obtained from (30). Both are plotted against the share of teleworkable employment estimated by Dingel and Neiman (2020).[22]

From equations (29) and (30), we can interpret $(\psi^0_{o(i)} + \psi^1 t)\hat{h}$ as an estimate of $\ln(w_{m,t}/w_{c,t})$ for each occupation, and $\psi_{j,t}$ as an estimate of $\ln(w_{m,t}/w_{c,t})$ for each CBSA. We use the specification of wages in (2), to map our estimates of the remote work premium into the remote work technology parameter, z:

$$z_{j,t} = \exp(\hat{\psi}_{j,t}/\mu) w_{cj,t}, \tag{31}$$

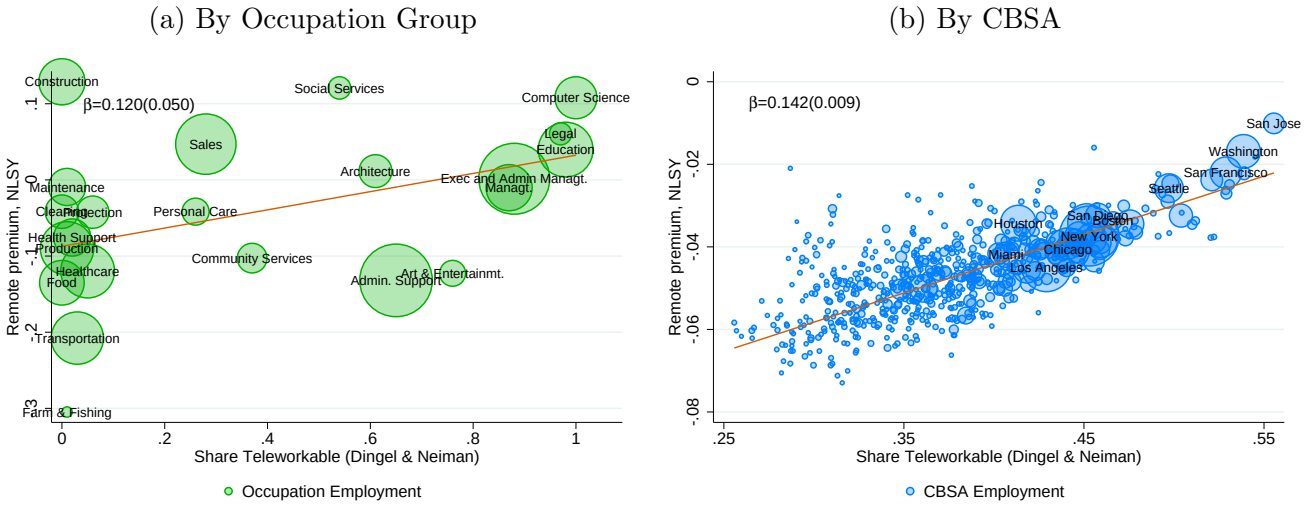
[19] Duranton and Puga (2019) estimate an overall agglomeration externality of 0.08 in the United States, once adding the static and dynamic effects of city size on individual productivity.

[20] In Appendix A.5, we estimate the remote work premium from the ACS.

[21] We describe these occupation groups in Appendix A.4.

[22] In Appendix A.5, we report the estimates for each occupation group and CBSA.

Figure 7: Estimates of the Remote Work Premium



Panel (a) reports estimates for each of the 22 occupation groups obtained from (29), against the share of teleworkable employment from Table 1 in Dingel and Neiman (2020). Panel (b) reports the estimates of the remote work premium by CBSA obtained from (30) plotted against the share of teleworkable employment at the CBSA level. We compute the CBSA-level shares of teleworkable employment as the weighted average of the occupation-specific measures from Table 1 in Dingel and Neiman (2020). The weights are given by the CBSA share of employment of each occupation from the Census sample. The coefficient β and the associated robust standard error from the regression of the remote premium on the share of teleworkable employment are reported at the top right of each figure.

where $w_{cj,t}$ is the average hourly wage of commuters in each CBSA j. We then recover the in-office productivity using

$$A_{j,t} = \tilde{L}_{cj,t}^{-\delta_j} w_{cj,t}, \tag{32}$$

given the city-level estimates of δ_j obtained in 4.2.

4.4 Transportation Costs Elasticity

We estimate the transportation costs elasticity γ_j for each CBSA j. We use the relationship that relates γ_j to the slope $\epsilon_{j,t}$ of the housing rent schedule with respect to distance to the CBD. From (13), the average elasticity of rent to distance to the CBD in period t is exactly related to the slope of the housing rent schedule according to

$$\epsilon_{j,t} = -\frac{\gamma_j}{1-\alpha}\left(1-\mu+\mu\frac{L_{cj,t}}{L_{j,t}}\right). \tag{33}$$

In order to estimate $\epsilon_{j,t}$, we use data on housing rents at the block group level from the ACS corresponding to average values between 2015 and 2019. For each CBSA, we estimate a regression at the block group level of the log of the median housing rental price p_{ij} in each block group i on the log of the distance, $dist_{ij}$, between the centroid of block group i and the centroid of the closest

CBD of CBSA j.[23] We then estimate

$$\ln p_{ij} = \epsilon_j \ln(dist_{ij}) + \beta_j X_{ij} + \alpha_j + e_{ij}, \tag{34}$$

where the controls X_{ij} contain block-level measures of the percentages of Hispanic, Black, and Asian population, dwelling characteristics such as the percentage of dwellings by type of structure, by the number of bedrooms, and by construction decade, as well as other geographic controls defined at the tract level and obtained from Lee and Lin (2017).[24]

4.5 Calibration of Remaining Parameters

We calibrate $\bar{\tau}_j$ for each city to match its observed value of L_{jc}/L_j. As shown in Appendix B.2, a stationary equilibrium value can be expressed in terms of λ_{cc}. Hence, we assume that the city is in a stationary equilibrium in 2019, infer the value of λ_{jcc} that is consistent with L_{jc}/L_j, and then solve for the value of $\bar{\tau}$ that makes such λ_{cc} a stationary equilibrium. We set the share of consumption expenditures on tradable goods to $\alpha = 0.76$, following Davis and Ortalo-Magné (2011). The share of days per week spent at home for remote workers is set to $\mu = 3/5$. As discussed in Section 4.1, our estimation does not separately identify the amenity externality ξ from the congestion externality θ. In addition, our estimates of the difference are not statistically different from zero at the 95 percent level. Therefore, without more guidance on their value, we set them each to zero.

5 Evidence

Stay-at-home policies had a dramatic, measurable, impact on the level and evolution of commuting across the country and the value of proximity to the CBD. We provide evidence of this impact using individual cell phone mobility data and estimates of the distance gradient housing price gradient to the CBD. We then relate this evidence to model-based implications of a large and abrupt shock to the measure of commuters.

5.1 Descriptive Evidence: Individual Mobility

To analyze individual mobility, we use anonymized data from SafeGraph, a data provider that tracks foot traffic to businesses relying on mobile phone GPS locations from third-party applications. We use the Neighborhood Patterns dataset, which aggregates the amount of foot traffic at the census block group level, to study the monthly evolution of trips to the CBD.[25] We first

[23] We provide details on how we identify each CBSA's CBDs in Appendix A.6.

[24] Geographic controls include the log of land area in square kilometers, number of housing units, the average age of housing units, average terrain slope, the maximum average temperature in July, the minimum average temperature in January, annual precipitations, log of distance to the closest river, to the closest lake, and to the closest shore.

[25] Chapple et al. (2023) also use SafeGraph data to document patterns of downtown foot traffic in 62 North-American Cities. They combine SafeGraph data with another private data source to obtain estimates until 2023.

Table 2: Estimates of City-Specific Parameters

	Agglomeration Externality δ_j	Remote Work Productivity z_j	Rel. Remote Work Productivity z_j/A_j	Transport Cost Elasticity γ_j
Mean	0.089	18.4	2.58	0.033
10^{th} perc.	0.081	14.9	2.18	0.003
50^{th} perc.	0.089	17.7	2.45	0.016
90^{th} perc.	0.099	22.6	3.21	0.056
Nb of CBSAs	639	639	639	274
Mean if emp. $\geq$ 1 million	0.102	25.39	4.17	0.014
Mean if emp. $\leq$ 100,000	0.089	17.44	2.48	0.048

The last two rows report the mean values of the parameters for i) the group of CBSAs with employment above 1 million (consisting of 19 CBSAs), ii) the group of CBSAs with employment below 100 thousand (consisting of 486 CBSAs). Employment is measured as the number of employed workers between 25 and 64 years old, not self-employed, listed as residents of the CBSA in the 2019 ACS. We are able to estimate the transport cost elasticity for a smaller number of CBSAs, given the data limitations on housing rents at the block group level in smaller CBSAs.

make use of the SafeGraph data to define the location of CBDs in each CBSA. The Neighborhood Patterns dataset provides information on the number of monthly visits recorded in each census block group. In particular, these visits are broken down by their most likely purpose, based on their observed characteristics.[26] We identify the center point of the CBDs as the center of a disc with a radius of 2 kilometers that receives the largest amount of work-related foot traffic. We then include all the block groups within this 2-kilometer disk into the CBD. We also allow cities to have several CBDs. We provide the details of our procedure in Appendix A.6.

The Neighborhood Patterns dataset also records the number of visits to each destination block group by block group of origin, allowing to compute bilateral flows of foot traffic at the monthly level.[27] On average, SafeGraph tracks around 30 million devices each month between January 2020 and July 2022.[28] However, the number of mobile phones recorded in a city and the set of block groups of origin in which they are located vary across months. Some of the variation in the raw counts of foot traffic is therefore due to changes in the size of the sample and in the set of origin block groups included in the dataset. We control for these variations by analyzing the change in foot traffic flows at the origin block group level. We compute the share of monthly foot

Their results confirm that, while downtown visits have recovered to pre-pandemic levels in some smaller cities, in most larger cities they have stabilized between 40 and 60 percent of their pre-pandemic levels, with little change between July 2022 and April 2023.

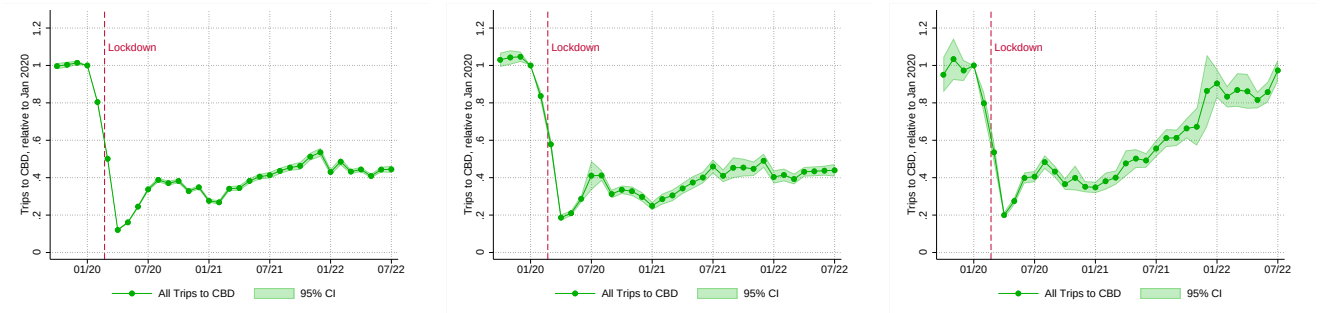
[26]Workplace visits are defined as those recorded Monday through Friday between 7:30 am and 5:30 pm and that dwelled for at least six hours.

[27]The block group of origin is imputed by assigning to each mobile phone the block group in which the phone is recorded at night as the block group of residence.

[28]In the largest cities (New York, Los Angeles, Chicago), between 400,000 and 800,000 devices are recorded each month.

Figure 8: Change in Visits to the CBD for Selected Cities

(a) New York, NY (b) San Francisco, CA (c) Madison, WI



Estimates of monthly indicators in a separate regression for each CBSA of visits to the CBD at the block group-month level. The outcome variable is the number of visits to the CBD from the specified block group, divided by the number of residing devices in that block group, as defined in (39), scaled by the value of this ratio in January 2020. The shaded areas represent the contour delimited by the consecutive 95 percent confidence intervals of each monthly indicator. Standard errors are clustered at the census tract of origin.

traffic originating from each block group that is directed to the CBD. We normalize the share of visits to the CBD from each origin block group by its value in January 2020. We then regress the normalized origin block group-level shares of visits to the CBD on month indicators.[29]

Figure 8 shows the volume of monthly trips to the CBD over the course of the pandemic, relative to its value in January 2020, for New York, NY, San Francisco, CA, and Madison, WI. For all three cities, the first months of the pandemic saw a drastic drop in trips to around 20 percent of their initial value. The recovery, however, has been quite different. While Madison, a relatively smaller city, essentially went back to its pre-pandemic levels by the end of our sample period, CBD trips for larger cities like New York and San Francisco have stabilized at around 40 percent of their pre-pandemic levels, a dramatic drop.

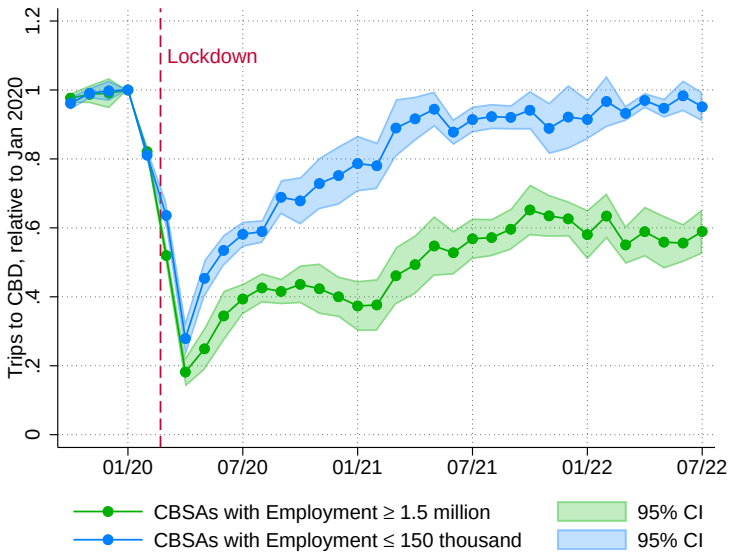
This behavior is, in fact, systematic. Figure 9 shows the monthly trips to the CBD relative to January 2020 for cities grouped by size. Our measure of size is total non-farm employment in the CBSA in 2019, published by the Bureau of Economic Analysis.[30] The green line reports the estimates of the trips to the CBD for the largest CBSAs with employment of more than 1.5 million, while the blue line collects the smallest CBSAs in our sample, with employment below 150,000.[31]

[29] Chapple et al. (2023) measure foot traffic by counting visits to downtown businesses, scaled by the sampling rate of the state. Instead, we count visits from any block group to downtown block groups, scaled by the sampling rate of each origin block group. That is, we control for changes in sampling rates across neighborhoods of each city, which we view as more robust than using the state-level sampling rate.

[30] This measure has the advantage of being available for all CBSAs, allowing us to describe patterns of recovery for the largest set of CBSAs—847 in total. In Appendix A.10, we reproduce the key graphs from Section 5 using as size the measure of employment used in the quantification, computed from Census microdata, and consistent with our measure of commuting. We confirm that the patterns described below hold for the subset of CBSAs for which we can measure employment in the Census.

[31] The smallest CBSA with employment above 1.5 million is Austin, TX, while the largest CBSA with employment below 150,000 is Rochester, MN.

Figure 9: Visits to the CBD Relative to January 2020



The figure reports the estimates of the average volume of visits to the CBD from block groups located in i) CBSAs with total employment above 1.5 million (25 largest CBSAs), ii) CBSAs with total employment below 150,000 (663 smallest CBSAs) expressed as a share of their values in January 2020. The shaded areas represent the contour delimited by the consecutive 95 percent confidence intervals of each monthly indicator. Standard errors are clustered by month and census region.

We normalize to 1 the number of trips in January 2020. Both groups experienced a sharp decline to less than 30 percent of pre-pandemic trips up to April 2020, but their paths are remarkably different afterward: while the group of small CBSAs returned to their pre-pandemic levels of visits to the CBD around early 2021, large CBSAs stabilized at around 60 percent of their initial level.

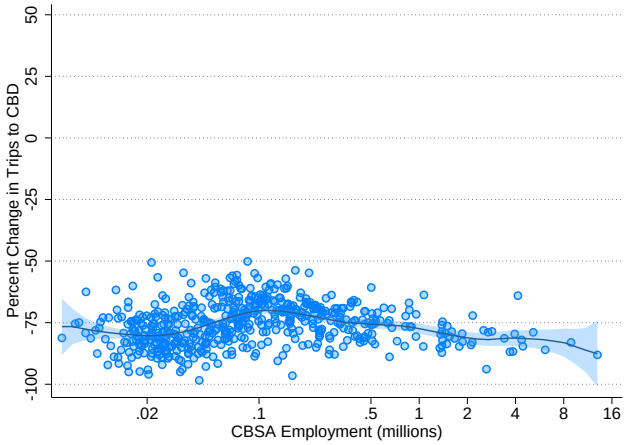
Figure 10 considers the relation of trips with city size in more detail. In the left panel, we show the extent to which these trips declined in April 2020 compared to January 2020, plotted against the city's total employment. Trips to the CBD declined by about 75 percent of the pre-pandemic level on average, with only little variation in the experiences of different cities. Importantly, this initial drop does not appear related to city size. In contrast, city size is systematically related to the return-to-office for the latest available data. While trips to the CBD rise to some extent for all cities, the right panel of Figure 10 shows that in July 2022, larger cities are substantially more likely to have trip levels way below their pre-pandemic levels.

5.2 Descriptive Evidence: Distance Gradients in House Prices

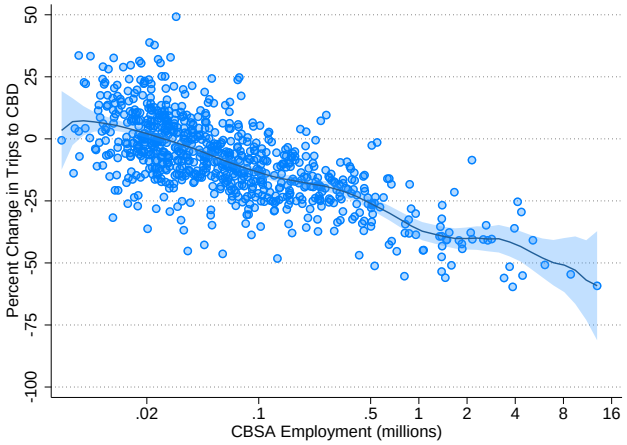
The previous section showed that the evolution of flows to the CBD has been heterogeneous across cities. Of course, changes in remote work patterns imply changes in demand for housing within cities. As workers spend less time on-site, a short commute becomes less important, leading to lower demand for housing next to centers of economic activity, and, everything else equal, a lower

Figure 10: Response of Trips to the CBD

(a) April 2020 Relative to January 2020



(b) July 2022 Relative to January 2020



Panel (a) reports the change in the average level of visits to each CBSA's CBD between January 2020 and April 2020, against employment (BEA). Panel (b) reports the change in visits to the CBD in July 2022 relative to January 2020. Each marker represents a CBSA. The solid line represents a fitted kernel-weighted local polynomial smoothing, with the shaded area corresponding to the 95 percent confidence interval around the fitted value.

premium for proximity to the CBD.

Changes in urban land gradients may be driven by a large number of factors varying at the CBSA level. Some of these include constraints on local housing development (Gyourko et al., 2008; Saiz, 2010; Lutz and Sand, 2022), the fraction of local jobs which can be done remotely (Dingel and Neiman, 2020), or the connectivity between urban and suburban parts of the CBSA.[32] For this reason, we estimate the CBSA-specific price and rent gradients in our bid-rent regression. Rather than using ACS housing rent block group level data as in Section 4.4, we turn to monthly housing price zip code level data from Zillow.[33] This higher frequency data is available for fewer CBSAs than the ACS data, but allows us to follow the monthly changes in housing prices for many cities since 2019.

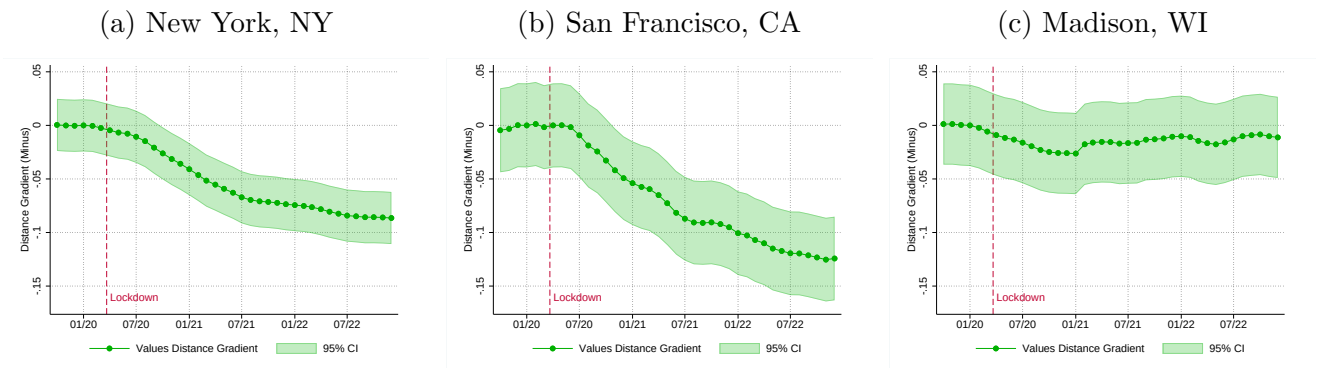
Indexing CBSAs by j, zip codes by i, and months by t, we estimate the monthly slope of the bid-rent function using an empirical specification similar to (35) but with monthly variation. Namely,

$$\ln p_{ij,t} = \delta_{j,t} \ln(dist_{ij}) + \beta X_{ij} + \alpha_{j,t} + e_{ij,t}, \tag{35}$$

[32] See Bailey et al. (2018) for measures of social connections using Facebook data.

[33] The monthly housing price dataset is the Zillow Home Value Index (ZHVI). Zillow also publishes the Zillow Observed Rent Index (ZORI), a monthly housing rent data that would be a closer analog to the ACS rent data. However, the ZORI is reported for fewer zip codes, making the estimation of the gradient imprecise for many CBSAs.

Figure 11: Change in Housing Values Gradients for Selected Cities



Estimates of monthly coefficients for log distance between zip codes and the CBD, relative to their value in January 2020. In a separate regression for each CBSA, we regress log housing values at the zip code level, on log distance by month, and geographic controls. The shaded areas represent the 95 percent confidence intervals of each monthly indicator, computed from robust standard errors. Data on housing values from Zillow's ZHVI index.

where units of observation are zip code-months. We refer to $p_{ij,t}$ as the price in zip code i of CBSA j in month t, and $dist_{ij}$ is the distance in kilometers between the centroid z_{ij} of zip code i and the CBD of CBSA j. We include CBSA-month fixed effects, $\alpha_{j,t}$, and zip code level control variables, X_{ij}. The zip code controls contain the log of median household income, the share of households in the top national income quintile, the median age, and the share of Black residents, all measured in 2019, and geographic controls.[34]

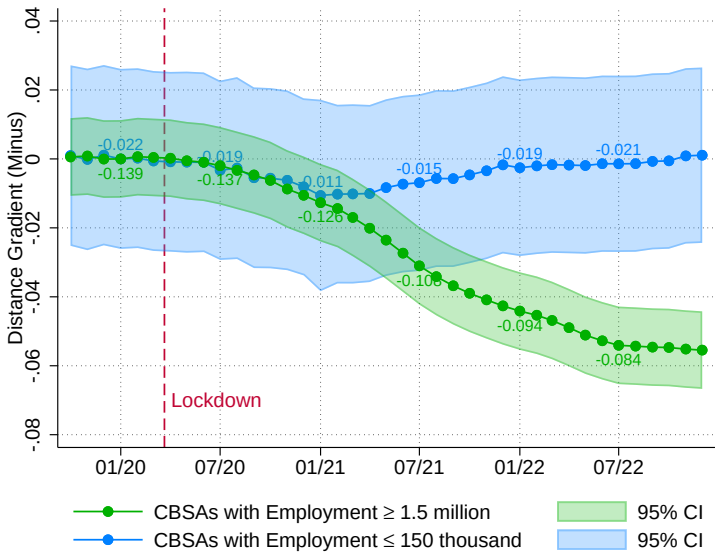
Consistent with the evidence presented in Figure 8, we find that the willingness to pay for proximity to the CBD is permanently lower in New York and San Francisco, while it returns to its initial values in Madison. Figure 11 plots the time evolution of $-(\delta_{jt} - \delta_{jt_0})$, the drop in the distance coefficient at time t relative to the baseline of January 2020: a decrease in the reported coefficients constitutes a flattening of the price gradient. For all cities, gradients fell through the first part of 2021. The drop has continued for New York and San Francisco, while the direction has reversed for Madison.

Figure 12 shows the average distance elasticity of house prices relative to January 2020 for the same group of city sizes as in Figure 9. In particular, the vertical axis reports $-(\delta_{jt} - \delta_{jt_0})$, the drop in the distance coefficient at time t relative to the baseline of January 2020: a decrease in the reported coefficients constitutes a flattening of the price gradient. Quite naturally, prices move more slowly than trips over time. All cities experience a flattening of the housing price gradients of around 0.01 by January 2021. This decline stops and reverses for small cities, which return to around their pre-pandemic level; the drop in the gradient of larger cities, instead, persists and appears to stabilize around 0.05 by the end of 2022.

Figure 13 illustrates the cross-sectional variation as a function of size in the drop of housing

[34]The geographic controls are the tract-level controls in footnote 24, averaged by zip code, using tract population as weights.

Figure 12: Distance Gradient of Housing Prices Relative to January 2020



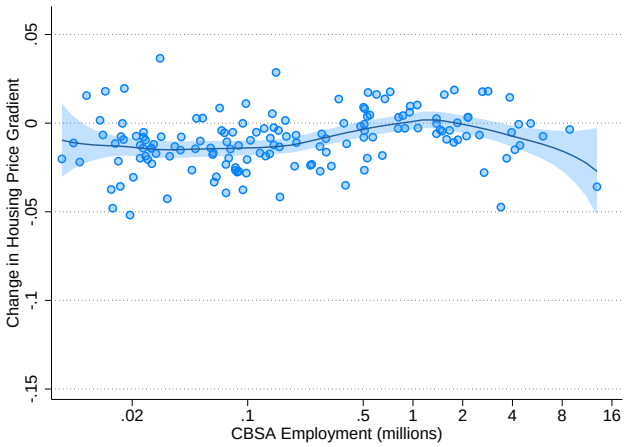
This figure reports the difference between the average housing price gradients in January 2020 and each following month, using the same employment thresholds as in Figure 9. In this figure, 25 cities have employment above 1.5 million and 102 have employment below 150,000. A negative value for the difference in housing price gradients implies a lower premium for housing located close to the CBD. The value of the average housing price gradient in each group is reported next to the marker of the associated group every six months. The shaded areas represent the contour delimited by the consecutive 95 percent confidence intervals of each monthly indicator.

price gradients at two different points during the pandemic. As above, the vertical axis represents the drop in the distance coefficient at time t relative to the baseline of January 2020. During the first year of the pandemic (January 2020 to January 2021, left panel), the distance elasticity of housing prices flattened by about 0.01 on average. This drop is sizeable relative to the average price gradient in January, equal to -0.044. The heterogeneity is moderate and again unrelated to city size. Consistently with evidence on CBD trips, however, the flattening in the price gradient becomes negatively associated with city size as the pandemic progresses and then recedes (right panel).

Taken together, the results in the last two subsections show that cities have experienced a heterogeneous response to the COVID-19 pandemic commuting shock. Appendix A.8 shows that this heterogeneity is also related to the strength of the agglomeration forces δ, the elasticity of transportation cost to distance γ, and the net strength of agglomeration forces $\delta - \gamma/\mu$. In the next section, we relate this evidence to the predicted response of our model calibrated to 274 U.S. CBSAs.

Figure 13: Response of House Price Gradients

(a) January 2021 Relative to January 2020



(b) December 2022 Relative to January 2020



This figure reports the difference between the housing price gradients in January 2020 and December 2022, against employment (BEA). Each marker represents a CBSA. The gradients are estimated using zip code level data from Zillow for 159 CBSAs. A negative value for the difference implies a lower premium for housing located close to the CBD. The solid line represents a fitted kernel-weighted local polynomial smoothing, with the shaded area corresponding to the 95 percent confidence interval around the fitted value.

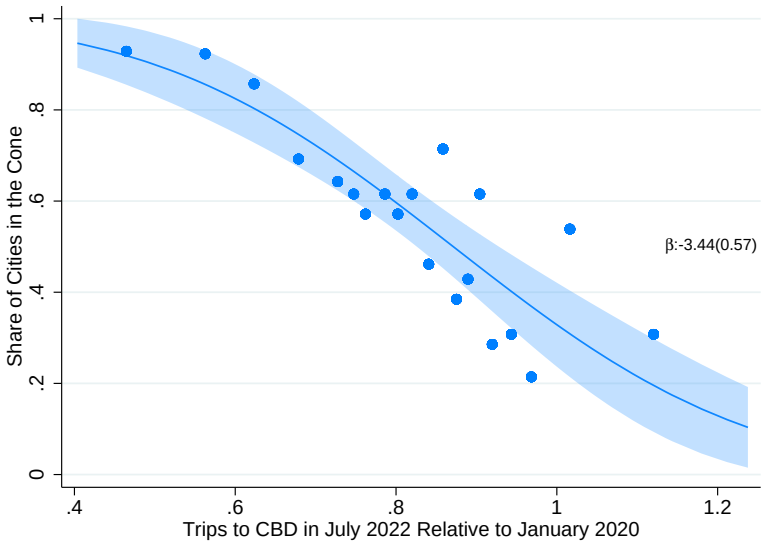
5.3 Permanent Responses to Transitory Shocks

The analysis in Subsection 3.5 indicates that a city in the cone of multiplicity could permanently switch to a path that converges towards the low-commuting stationary equilibrium when subject to a large enough shock to the number of commuters. In this new stationary equilibrium, trips to the CBD are permanently lower, and the rent and value gradients are permanently flatter.

In our data, the response heterogeneity as a function of city size of Figures 10 and 13 appears to have stabilized early on (see Figures 9 and 12). To interpret this evidence, we turn to a quantified version of our model. We use city-specific parameters retrieved in Section 4 to calculate the values of z/A, $\delta - \gamma/\mu$, and the corresponding set $\mathcal{Z}$ defined in Proposition 3.3. We construct these values for 274 CBSAs with valid information on remote work premiums and elasticities.[35] For each city, we create an indicator variable $\mathbb{1}_{cone} = 1$ if $z/A \in \mathcal{Z}$, and zero otherwise, meaning the city is inside its cone of multiplicity. We find that 155 cities were in their respective cone before the pandemic. These cities could be pushed into a low commuting equilibrium by a large enough, albeit temporary, shock to the number of commuters. Our theory appears to have explanatory power. We estimate a city-level probit of $\mathbb{1}_{cone}$ on the July 2022 measure of CBD trips, relative to January 2020. This regression returns a negative and significant coefficient: cities with a lower level of trips post-pandemic were indeed much more likely to be in their cone before the pandemic. Figure 14 reports this estimated share of cities that are inside the cone as a function of the trip

[35] The model assumes positive δ and γ, and requires a remote work wage discount for a valid calibration.

Figure 14: Share of Cities in the Cone of Multiplicity vs. CBD Trips Shortfall



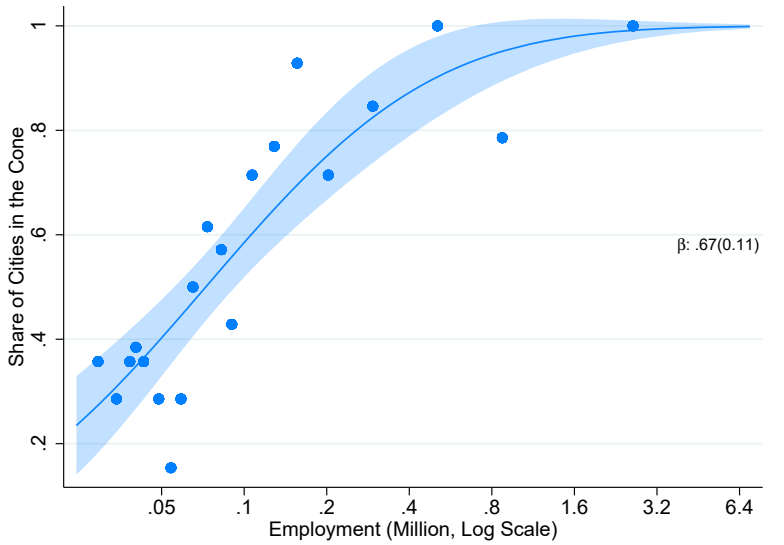
The blue line in this figure plots the predicted share of cities in their cone of multiplicity from a probit regression of the indicator variable $\mathbb{1}_{cone}$ on the volume of trips in July 2022, relative to their value in January 2020, with the shaded area representing the 95 percent confidence interval. The figure also reports the coefficient and robust standard error of the independent variable in the probit regression. One blue dot is the average value of $\mathbb{1}_{cone}$ in each of 20 quantiles of the trip shares.

shortfall, while the blue dots report the fraction of cities in the cone in each of 20 bins of the independent variable.

The propensity of being inside the cone grows with total city size. A probit regression of $\mathbb{1}_{cone}$ on the log employment of a city shows a positive and significant coefficient. Figure 15 shows this estimated share as a function of log employment, with a 95% robust confidence interval. The blue dots report the fraction of cities in the cone in each of the 20 bins of employment size. Larger cities, which in the data are more likely to display a permanent shift, are also more likely to be in the cone of multiplicity that our theory proposes. We find analogous relationships between the share of cities that are predicted to be in the cone and δ, γ, and the net strength of agglomeration externalities, $\delta - \gamma/\mu$. We report these findings in Appendix A.8.

Taken together, this evidence suggests that the observed, and by now stable, heterogeneous responses to the pandemic and related stay-at-home policies could be due to the coordination problem we have highlighted in our theory. Since this coordination happens, at least partly, outside firm boundaries, shifts in the equilibrium may be permanent in nature and firms might not have the tools to address it. Of course, city governments could implement policies that put the city back on its path to the commuting equilibrium (like taxing home work), but we do not know what would be the value of such policies and how effective they would be in actually achieving their goal. These considerations motivate an investigation of the welfare consequences of these shifts.

Figure 15: Share of Cities in the Cone of Multiplicity vs. Size



The blue line in this figure plots the predicted share of cities in their cone of multiplicity from a probit regression of the indicator variable $\mathbb{1}_{cone}$ on the log total employment of a city, with the shaded area representing the 95 percent confidence interval. The figure also reports the coefficient and robust standard error of the independent variable in the probit regression. One blue dot is the average value of $\mathbb{1}_{cone}$ in each of 20 quantiles of employment levels.

5.4 Welfare

The model gives us an expression to compute the average stationary equilibrium welfare in a city as a function of the measure of commuters. Evaluating equation (15) in a stationary equilibrium and rearranging the terms, average welfare $\mathcal{W}$ can be computed as

$$\mathcal{W} = (L_c^{ss}) = \left(\frac{L_c^{ss}}{L}\right) V_c^{ss}\left(L_c^{ss}\right) + \left(1 - \frac{L_c^{ss}}{L}\right) V_m^{ss}\left(L_c^{ss}\right), \tag{36}$$

$$\text{with } V_\ell^{ss}\left(L_c^{ss}\right) = \frac{u_\ell\left(L_c^{ss}\right) + \Omega(\bar{\varepsilon}^\ell(L_c^{ss}))}{1-\beta}. \tag{37}$$

In a city with multiple stationary equilibria, this welfare varies across stationary equilibrium. For the 155 cities inside the cone, we can then compute $\hat{\mathcal{W}} \equiv 100 \times \left(\mathcal{W}(L_{c,low}^{ss})/\mathcal{W}(L_{c,high}^{ss}) - 1\right)$ as the percentage change between the welfare in the low-commuting versus the high-commuting stationary equilibrium. All cities in the cone would lose from being permanently pushed to the low-commuting stationary equilibrium.[36] Table 3 shows some moments of the distribution of welfare

[36] Although we do not have a proof, the intuitive reason why all cities in the multiplicity cone would rather be in the commuting-intensive stationary equilibrium is that the only active elasticity in our quantification is a positive agglomeration externality. This implies that, in all equilibria, there is too little commuting compared to the efficient benchmark.

losses.[37]

Table 3: Percentage Welfare Difference Between the High and Low-Commuting Equilibrium

	Min	p10	p25	p50	p75	p90	Max	Mean	N
Welfare Loss $(-\hat{\mathcal{W}})$	1.7	2.1	2.3	2.6	2.7	2.9	3.2	2.5	155

The median potential loss from such a switch to the low-commuting stationary equilibrium is 2.6%, with a range from 1.7% to 3.2%. The loss stems from a drop in average wages, which is only partially compensated by a reduction in commuting costs and changes in the option values. A city switching equilibrium has lower wages both for commuters and remote workers since each individual interacts with a smaller average number of agents, $\tilde{L}_c$. Fewer interactions lead to diminished production externalities. Specifically, the change in average wages is given by

$$\hat{w} \equiv 100 \times \left(\frac{\left(\frac{L^{ss}_{c,low}}{L}\right) w_{c,low} + \left(1 - \frac{L^{ss}_{c,low}}{L}\right) w_{m,low}}{\left(\frac{L^{ss}_{c,high}}{L}\right) w_{c,high} + \left(1 - \frac{L^{ss}_{c,high}}{L}\right) w_{m,high}} - 1 \right). \tag{38}$$

The residual component of the welfare impact, that accounts for changes in commuting costs and option values can be constructed as $\chi \equiv \hat{\mathcal{W}} - \hat{w}$. Figure 16, right panel, plots $\hat{w}$ and χ against the net strength of agglomeration externalities, $\delta - \gamma/\mu$, for all cities estimated to be in the cone. The fall in average wages is quite severe, between 15% and 25%. This large magnitude is to be expected, however, because agglomeration externalities tend to be significant precisely for cities more likely to be in the multiplicity region. Net welfare losses are then muted by the combined effect of changes in the average commuting disutility and option values. Overall, the net strength of the agglomeration externality (Figure 16, left panel) is a strong predictor of the final welfare loss: a regression of the estimated loss on $\delta - \gamma/\mu$ delivers an R^2 of 0.85.

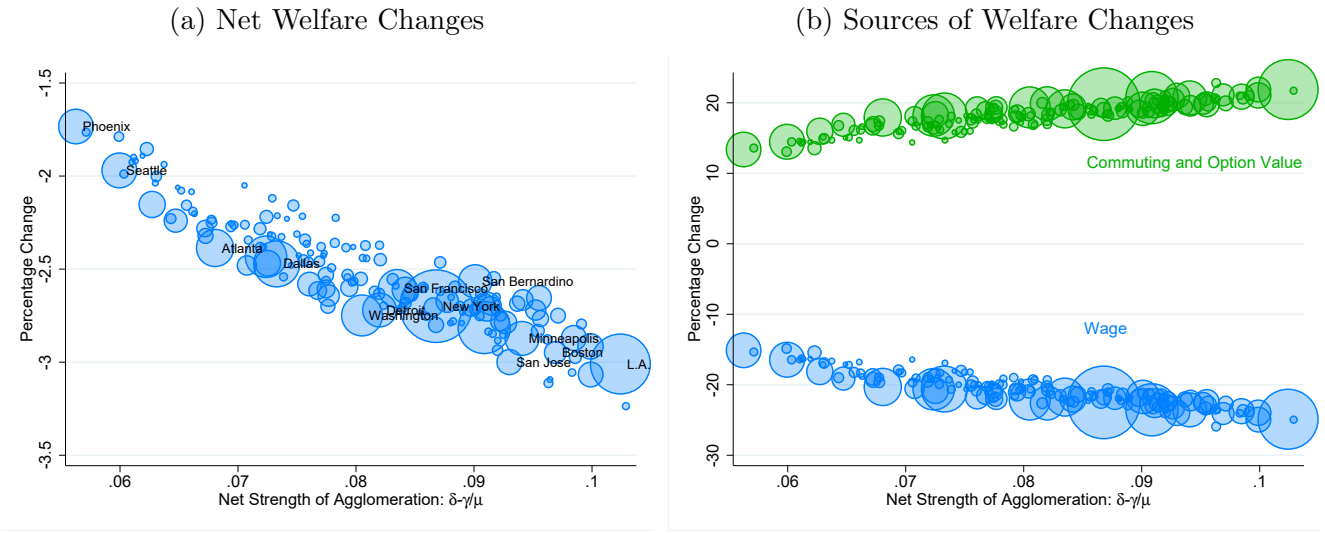
In Figure A5 in the Appendix, we also compare welfare losses to the observed drop in trips to the CBD. The largest drops in CBD trips are concentrated around welfare losses of 2.5 to 3%. In particular, we find that the welfare loss of the average resident in cities with trips to the CBD of 60% or less of the pre-pandemic level is 2.7%.

6 Conclusion

We have proposed a theory of remote work and city structure. At its core, our theory features a coordination problem: workers prefer to commute if others do as well, and prefer to work from home otherwise. This mechanism can lead to multiplicity in stationary equilibria if agglomeration forces are sufficiently larger than congestion forces and remote-work technology is decent relative

[37] Table A6 in the Appendix shows actual welfare changes for the largest U.S. cities in our calibration sample.

Figure 16: Welfare Ratio $\hat{\mathcal{W}}$ and its Decomposition



The left panel shows a scatterplot between $\hat{\mathcal{W}}$ and the net strength of agglomeration externalities, $\delta - \gamma/\mu$. The right panel shows the percentage difference in wages across equilibria, $\hat{w}$, and the residual commuting and option value gains χ, plotted against the net strength of agglomeration externalities, $\delta - \gamma/\mu$. The marker size is proportional to total city employment. These scatterplots can only be computed for the set of cities in the cone of multiplicity.

to in-office technology. In the resulting dynamic discrete choice model, large exogenous changes in the number of commuters, a state variable of the model, can put the city on an equilibrium path that makes it converge to an alternative stationary equilibrium. We have argued that the COVID-19 pandemic constituted such an exogenous shock and that, as a result, many large cities have switched from a stationary equilibrium where most workers commuted to one in which there is a large share of agents working from home. Small cities, in contrast, tend not to exhibit this multiplicity and therefore have converged back to the pre-pandemic stationary equilibrium. We find consistent evidence from our estimated model for 274 U.S. cities, from cell-phone-based individual mobility data, and from house price distance gradients. An arguably uniform shock to the labor delivery mode across cities has revealed which cities were exposed to this coordination problem – and thus, presented multiple stationary equilibria – before the pandemic, a rather unique empirical event.

The model can be used to estimate the welfare losses associated with the switch to remote work for all cities that exhibit this type of multiplicity in stationary equilibria. We find that welfare in the commuting-intensive equilibrium is between 1.7% and 3.2% larger than in the remote-work-intensive equilibrium. A significant, but not overwhelming, cost. This modest cost is, however, the result of a large drop in wages compensated by substantial utility gains from option values and commuting costs.

Although our analysis can fit the available data quite well, it inevitably abstracts from a number of important forces and issues. First, our model studies cities in isolation but not as a system of

cities with migration, trade, and investment links. This is potentially relevant since the spatial equilibrium condition across cities will determine local welfare effects and population flows. The welfare implications we estimate for some of the large U.S. cities might then be partially translated into smaller populations. Similarly, some of the small cities with unique stationary equilibria might grow and gain industries that transform them in ways that could put them in the multiplicity set. Of course, remote workers could also commute across cities, another issue we abstract from but has received some recent attention in the media. Second, although we model remote work as "partial remote work," where agents work from home only a fraction μ of days, we have kept this fraction constant rather than the agent's choice. Of course, agents in different cities, or in different parts of a city, or with different occupations or skills, might choose different shares. Understanding the importance of these choices and the resulting heterogeneity across individuals could be important but complicates the analysis significantly. Third, some of the spillovers between workers could be internalized by firms and cities, which might then proactively affect workers' delivery mode choice. Understanding the extent to which active policy can determine the dynamic path and stationary equilibrium that a city converges to, is an interesting topic for future research.

The repercussions of remote work are intriguing and potentially transforming for the organization of work in cities, firms, and labor markets. Our paper hopes to contribute to what will surely be an active area of research in the future, as the organization of work changes to accommodate the booming phenomenon of remote work.

References

Aksoy, C. G., J. M. Barrero, N. Bloom, S. J. Davis, M. Dolls, and P. Zarate (2022): "Working from Home Around the World," Working Paper 30446, National Bureau of Economic Research.

Alonso, W. (1964): *Location and Land Use*, Cambridge, MA and London, England: Harvard University Press.

Althoff, L., F. Eckert, S. Ganapati, and C. Walsh (2021): "The Geography of Remote Work," Working Paper 29181, National Bureau of Economic Research.

Arcidiacono, P. and R. A. Miller (2011): "Conditional Choice Probability Estimation of Dynamic Discrete Choice Models With Unobserved Heterogeneity," *Econometrica*, 79, 1823–1867.

Artuç, E., S. Chaudhuri, and J. McLaren (2010): "Trade Shocks and Labor Adjustment: A Structural Empirical Approach," *American Economic Review*, 100, 1008–45.

Atkin, D., A. Schoar, and S. Shinde (2022): "Worker Sorting, Worker Discipline and Development," *Working Paper*.

Bailey, M., R. Cao, T. Kuchler, J. Stroebel, and A. Wong (2018): "Social Connectedness: Measurement, Determinants, and Effects," *Journal of Economic Perspectives*, 32, 259–80.

Barrero, J. M., N. Bloom, and S. J. Davis (2021): "Why Working from Home Will Stick," Working Paper 28731, National Bureau of Economic Research.

Baum-Snow, N. and R. Pavan (2011): "Understanding the City Size Wage Gap," *The Review of Economic Studies*, 79, 88–127.

Bloom, N., R. Han, and J. Liang (2022): "How Hybrid Working From Home Works Out," Working Paper 30292, National Bureau of Economic Research.

Bloom, N., J. Liang, J. Roberts, and Z. J. Ying (2015): "Does Working from Home Work? Evidence From a Chinese Experiment," *The Quarterly Journal of Economics*, 130, 165–218.

Brueckner, J., M. E. Kahn, and G. C. Lin (2021): "A New Spatial Hedonic Equilibrium in the Emerging Work-from-Home Economy?" Working Paper 28526, National Bureau of Economic Research.

Brynjolfsson, E., J. J. Horton, C. Makridis, A. Mas, A. Ozimek, D. Rock, and H.-Y. TuYe (2023): "How Many Americans Work Remotely? A Survey of Surveys and Their Measurement Issues," Working Paper 31193, National Bureau of Economic Research.

Brynjolfsson, E., J. J. Horton, A. Ozimek, D. Rock, G. Sharma, and H.-Y. TuYe (2020): "COVID-19 and Remote Work: An Early Look at US Data," Working Paper 27344, National Bureau of Economic Research.

Cappelli, P. H. and J. Keller (2013): "A Study of the Extent and Potential Causes of Alternative Employment Arrangements," *ILR Review*, 66, 874–901.

Chapple, K., H. Moore, M. Leong, D. Huang, A. Forouhar, L. Schmahmann, J. Wang, and J. Allen (2023): "The Death of Downtown? Pandemic Recovery Trajectories across 62 North American Cities," *Research Brief, School of Cities, University of Toronto and Institute of Governmental Studies, University of California, Berkeley.*

Choudhury, P., T. Khanna, C. Makridis, and K. Schirmann (2022): "Is Hybrid Work the Best of Both Worlds? Evidence from a Field Experiment," *Harvard Business School Technology & Operations Mgt. Unit Working Paper*, No. 22-063.

Combes, P.-P., G. Duranton, and L. Gobillon (2008): "Spatial wage disparities: Sorting matters!" *Journal of Urban Economics*, 63, 723–742.

Combes, P.-P., G. Duranton, L. Gobillon, D. Puga, and S. Roux (2012): "The Productivity Advantages of Large Cities: Distinguishing Agglomeration From Firm Selection," *Econometrica*, 80, 2543–2594.

Dalton, M., M. Dey, and M. Loewenstein (2022): "The Impact of Remote Work on Local Employment, Business Relocation, and Local Home Costs," BLS Working Paper 553, U.S. Bureau of Labor Statistics.

Davis, M. A., A. C. Ghent, and J. M. Gregory (2021): "The Work-from-Home Technology Boon and its Consequences," Working Paper 28461, National Bureau of Economic Research.

Davis, M. A. and F. Ortalo-Magné (2011): "Household expenditures, wages, rents," *Review of Economic Dynamics*, 14, 248–261.

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

De La Roca, J. and D. Puga (2016): "Learning by Working in Big Cities," *The Review of Economic Studies*, 84, 106–142.

Dingel, J. I. and B. Neiman (2020): "How Many Jobs Can be Done at Home?" Working Paper 26948, National Bureau of Economic Research.

Duranton, G. and J. Handbury (2023): "Covid and Cities, Thus Far," Working Paper 31158, National Bureau of Economic Research.

Duranton, G. and D. Puga (2019): "Urban Growth and its Aggregate Implications," Working Paper 26591, National Bureau of Economic Research.

Emanuel, N., E. Harrington, and A. Pallais (2022): "The Power of Proximity: Office Interactions Affect Online Feedback and Quits, Especially for Women and Young Workers," *Working Paper*.

Etheridge, B., Y. Wang, and L. Tang (2020): "Worker productivity during lockdown and working from home: evidence from self-reports," ISER Working Paper Series 2020-12, Institute for Social and Economic Research.

Gaubert, C. (2018): "Firm Sorting and Agglomeration," *American Economic Review*, 108, 3117–53.

Gennaioli, N., R. La Porta, F. Lopez-de Silanes, and A. Shleifer (2012): " Human Capital and Regional Development *," *The Quarterly Journal of Economics*, 128, 105–164.

Gibbs, M., F. Mengel, and C. Siemroth (2023): "Work from Home and Productivity: Evidence from Personnel and Analytics Data on Information Technology Professionals," *Journal of Political Economy Microeconomics*, 1, 7–41.

Glaeser, E. and D. Maré (2001): "Cities and Skills," *Journal of Labor Economics*, 19, 316–342.

Gupta, A., V. Mittal, and S. Van Nieuwerburgh (2022): "Work From Home and the Office Real Estate Apocalypse," Working Paper 30526, National Bureau of Economic Research.

Gyourko, J., A. Saiz, and A. Summers (2008): "A New Measure of the Local Regulatory Environment for Housing Markets: The Wharton Residential Land Use Regulatory Index," *Urban Studies*, 45, 693–729.

Hansen, S., P. J. Lambert, N. Bloom, S. J. Davis, R. Sadun, and B. Taska (2023): "Remote Work across Jobs, Companies, and Space," Working Paper 31007, National Bureau of Economic Research.

Harrington, E. and N. Emanuel (2022): "'Working' Remotely? Selection, Treatment, and Market Provision of Remote Work," *Working Paper*.

Henderson, J. (2003): "Marshall's scale economies," *Journal of Urban Economics*, 53, 1–28.

Kennan, J. and J. R. Walker (2011): "The Effect of Expected Income on Individual Migration Decisions," *Econometrica*, 79, 211–251.

Künn, S., C. Seel, and D. Zegners (2020): "Cognitive Performance in the Home Office - Evidence from Professional Chess," .

Lee, S. and J. Lin (2017): "Natural Amenities, Neighbourhood Dynamics, and Persistence in the Spatial Distribution of Income," *The Review of Economic Studies*, 85, 663–694.

Lipowska, K., M. Smoter, and P. Lewandowski (2022): "Mismatch in preferences for working from home – evidence from discrete choice experiments: evidence from Poland," .

Liu, S. and Y. Su (2021): "The impact of the COVID-19 pandemic on the demand for density: Evidence from the U.S. housing market," *Economics Letters*, 207, 110010.

——— (2023): "The Effect of Working from Home on the Agglomeration Economies of Cities: Evidence from Advertised Wages," *SSRN*.

Lutz, C. and B. Sand (2022): "Highly disaggregated land unavailability," *Available at SSRN 3478900*.

Manson, S., J. Schroeder, D. V. Riper, T. Kugler, and S. Ruggles (2022): "IPUMS National Historical Geographic Information System: Version 17.0 [dataset]," http://doi.org/10.18128/D050.V17.0.

Mas, A. and A. Pallais (2017): "Valuing Alternative Work Arrangements," *American Economic Review*, 107, 3722–3759, nBER Working Paper 22708.

Matthies, B. and A. Kwan (2022): "Measuring the Impact of Remote Work Using Big Data," *Working Paper*.

Mondragon, J. A. and J. Wieland (2022): "Housing Demand and Remote Work," Working Paper 30041, National Bureau of Economic Research.

Moretti, E. (2004): "Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data," *Journal of Econometrics*, 121, 175–212, higher education (Annals issue).

Muth, R. F. (1969): *Cities and Housing*, Chicago: University of Chicago Press.

Oettinger, G. S. (2011): "The Incidence and Wage Consequences of Home-Based Work in the United States, 1980–2000," *Journal of Human Resources*, 46, 237–260.

Owens, Raymond, I., E. Rossi-Hansberg, and P.-D. Sarte (2020): "Rethinking Detroit," *American Economic Journal: Economic Policy*, 12, 258–305.

Parkhomenko, A. and M. Delventhal (2023): "Spatial Implications of Telecommuting," *Working Paper*.

Ramani, A. and N. Bloom (2021): "The Donut Effect of Covid-19 on Cities," Working Paper 28876, National Bureau of Economic Research.

Rossi-Hansberg, E., P.-D. Sarte, and F. Schwartzman (2019): "Cognitive Hubs and Spatial Redistribution," Working Paper 26267, National Bureau of Economic Research.

——— (2021): "Local Industrial Policy and Sectoral Hubs," *AEA Papers and Proceedings*, 111, 526–31.

Saiz, A. (2010): "The Geographic Determinants of Housing Supply," *The Quarterly Journal of Economics*, 125, 1253–1296.

Traiberman, S. (2019): "Occupations and Import Competition: Evidence from Denmark," *American Economic Review*, 109, 4260–4301.

Yang, L., D. Holtz, S. Jaffe, S. Suri, S. Sinha, J. Weston, C. Joyce, N. Shah, K. Sherman, B. Hecht, and J. Teevan (2022): "The effects of remote work on collaboration among information workers," 6, 43–54.

A Appendix

A.1 Data Sources

Cross-sectional Data on Remote Work. Our first source of data to measure the prevalence of remote work in the U.S. economy and the earnings of remote workers are the decennial census between 1980 and 2000 and the American Community Survey thereafter. These surveys include a question asked of those ages 16 and over who were employed and at work in the previous week on the method of transportation usually used to get to work. We define as remote workers those who reported "work from home" on this question.[1] We measure the hourly earnings rate of workers by dividing the total pre-tax wage and salary income by the number of hours worked during the year.[2] This definition of remote work allows us to obtain a consistent measure of remote work and its compensation from 1980 to 2021 through repeated cross-sectional samples. On average, every year, about 5 million workers are listed as remote workers. However, the decennial census and the ACS do not contain information on the number of hours worked at home.

Longitudinal Data on Remote Work. Our second source of data is the National Longitudinal Survey of Youth - 1979 Cohort (NLSY79), a nationally representative sample of 12,686 young men and women born between the years 1957 through 1964 in the United States. We focus on data reported from 1998 onward, a period through which the reporting of hours worked is consistent. Every two years, the survey reports the number of hours worked at home in a week. This allows us to observe variation in remote work on the intensive margin. We obtain workers' hourly rates directly from the survey. Thanks to the longitudinal design of the survey, we can observe workers' transitions into and out of remote work over time.

Foot Traffic Data. We use data from a private provider, SafeGraph, which produces foot traffic information at the business place level by collecting anonymized phone GPS tracking from third-party applications. Specifically, we use the Neighborhood Patterns dataset, which contains information on foot traffic to census block groups every month. This dataset has several key advantages. First, it captures foot traffic at a fine geographical level.[3] This allows us to specifically focus on mobility to cities' CBDs. Second, it provides a count of individual visits by category of trip, including visits that are likely to be for work. We use this feature to define CBDs. Third, the visits to any block groups are disaggregated by block group of origin. This allows us to control

[1] In the decennial census, the question on means of transportation to work was asked in the "long form" questionnaire, which was administered to approximately 1 out of every 6 housing units in the United States..

[2] The income variable is "incwage". It includes wages, salaries, commissions, cash bonuses, tips, and other money income received from an employer. Payments-in-kind or reimbursements for business expenses are not included. The number of hours worked in the year is computed by multiplying the reported number of hours worked in a typical week by the number of weeks worked in the year.

[3] There are 220,684 census block groups in the United States.

for unobserved permanent characteristics of origin block groups by studying mobility to the CBD from each block group separately.

Housing Prices and Rents. We rely on house prices and rents aggregated at the zip code level by Zillow.[4] Our measure of house prices is the Zillow House Value Index (ZHVI), which controls for common house characteristics and adjusts for seasonality. We measure rents with the Zillow Observed Rental Index (ZORI), an index of house rents obtained by averaging asking rents after controlling for house quality indicators. Both series include prices and rents of single- and multi-family units. These datasets have the advantage of being reported monthly before and during the COVID-19 pandemic, allowing us to track the changes in housing prices and rent gradients in about 200 CBSAs. While the monthly Zillow data is well suited for within-city comparison across time, the relatively low number of observed zip codes in smaller CBSAs makes it is less amenable to comparisons across cities. To obtain more precise estimates of housing rent gradients for a larger number of CBSAs, we turn to data from the ACS at the census block group level.[5] We use the median rent paid in a block group during the 2015-2019 period. This, along with other block group-level controls defined below, allows us to estimate housing rent gradients before the COVID-19 pandemic more precisely and for almost 300 CBSAs.

Local Housing Market Characteristics. We collect zip code level variables to control for variation in housing market characteristics. From the Census Bureau, we obtain the median household income, the share of households in the top national income quintile, the median age, and the proportion of Black residents. There is significant variation in housing market characteristics at the CBSA level. We complement these measures with local geographic features collected by Lee and Lin (2018). We aggregate these census tract-level indicators at the zip code level.

A.2 Definitions of Data Samples

A.2.1 Census/ACS Sample

We combine the decennial census waves in 1980, 1990, and 2000 with the yearly waves of the American Community Survey between 2005 and 2021. We start from observations from all individuals recorded in each wave, weighted by their personal sampling weight. We only include workers who are employed, of age between 25 and 64, who do not work in agriculture or in the military, and who are not self-employed.[6] We restrict our sample to individuals who have not moved to a different

[4]The datasets are publicly available from https://www.zillow.com/research/data/.

[5]The data can be obtained from IPUMS' National Historical Geographic Information System (NHGIS) portal Manson et al. (2022).

[6]Specifically, we exclude workers in 1-digit industries 1 and 14 from the 1990 Census industry classification ("Agriculture, Forestry, and Fisheries" and "Active Duty Military/NA/NC", respectively), workers in 1-digit occupations 4 and 7 from the 1990 Census occupation classification ("Agriculture, Forestry, and Fishing Occupations" and "Military Occupations/NA", respectively), and workers whose reported "class" or worker is self-employed (variable `classwkr` equal to 1).

metropolitan area since the previous year in order to allocate the labor delivery, and earnings of the individual to the appropriate labor market. To limit measurement error in earnings, we further restrict our sample to only include full-time workers, employed for the full year. We only include individuals born in the United States. Finally, we only include individuals who reside in one of the Core-based Statistical areas in the United States.

We define the main variables of interest as follows. An individual is categorized as a remote worker if they answer "work from home" to the question "What is your means of transportation to work?".[7] Our main measure of earnings is log hourly wage. We compute it by dividing yearly wage income by the estimated number of hours worked in the year. To obtain the total number of hours worked per year, we multiply the usual number of hours worked per week by the total number of weeks worked in the year.[8] For individuals located in small CBSAs (Micropolitan Statistical Areas), the CBSA is sometimes missing in IPUMS. We use a crosswalk from the PUMA (smallest geographical areas recorded in IPUMS) to CBSAs, by the intermediary of a crosswalk from PUMA to counties, and from counties to CBSAs.

Our final sample contains 17,025,788 observations, each corresponding to an individual in a given year. There are 649 distinct CBSAs covered by the sample. We report the number of observations, the share of remote workers, and the median hourly pay of commuters and remote workers per year in Table A1.

A.2.2 NLSY Sample

Our NLSY sample is constructed from the public use NLSY79 dataset.[9] We only include individuals from the main sample.[10] Besides this restriction, we include all individuals.

We define the main variables of interest as follows. When estimating the remote work premium, we use a continuous measure of remote work based on the number of hours worked at home per week. When a discrete definition is necessary, for example, when determining transitions into and from remote work, we use 24 hours of work at home per week as the threshold. Our main measure of earnings is hourly rates of pay, reported directly in the data. The geographical location is coded by the region of the United States. The four regions are Northeast, North-Central, South, and West.

Our final sample comprises, on average, 30,548 observations per year, representing observations from 5,656 individuals. We report the number of observations, average number of hours worked at home, and median hourly pay of commuters and remote workers per year in Table A2.

[7] This is coded as `tranwork` = 80 in IPUMS.

[8] The yearly income variable is `incwage`. The number of weeks is coded into bins by the variables `wkswork*`, and the usual number of hours is recorded in `uhrswork`.

[9] Note that there also exists the NLSY97, which follows individuals born between 1980 and 1984. Unfortunately, for the purposes of our analysis, the number of hours worked from home is not reported in this panel.

[10] That is, we exclude individuals from the additional samples that disproportionately sample military individuals and economically disadvantaged non-black, non-Hispanic individuals.

Table A1: Summary Statistics for the Census/ACS Sample

Year	Number of Observations	Share of Remote Workers (pct)	Median Hourly Pay of Commuting Workers (2021 $/hr)	Median Hourly Pay of Remote Workers (2021 $/hr)
1980	1,217,059	0.42	28.9	18.3
1990	2,197,452	0.64	25.7	21.4
2000	2,420,168	1.24	25.5	30.3
2005	554,602	1.38	25.3	33.8
2006	580,803	1.54	24.7	34.0
2007	585,986	1.68	25.0	33.5
2008	631,734	1.85	25.3	34.0
2009	604,571	2.02	24.7	32.9
2010	596,246	2.17	24.8	32.2
2011	585,420	2.28	24.4	32.9
2012	682,920	2.33	23.6	33.1
2013	699,979	2.47	23.6	33.6
2014	703,349	2.62	23.6	34.2
2015	717,042	2.79	23.6	33.8
2016	724,641	3.10	24.3	34.7
2017	743,963	3.32	24.9	35.8
2018	754,886	3.45	24.4	35.5
2019	764,868	3.95	25.0	35.3
2020	586,182	16.30	24.9	37.4
2021	719,188	19.22	24.9	38.8

Data from the Census/ACS. Hourly pay is adjusted for the Consumer Price Index and defined as the total annual earnings divided by the total number of hours worked in the year.

A.2.3 SafeGraph Sample

We measure visits to CBSAs' CBDs with the Neighborhood Patterns dataset, which contains monthly data between January 2019 and July 2022. Every month, the data is reported in a table that lists the number of phone devices that stopped for a visit inside a census block group by block group of origin. A visit is defined as the recording of the same device inside a polygon identified as a "place of interest" (i.e., a place of business) for more than a minute. SafeGraph also reports for each block group and each month the number of devices with a primary nighttime location inside the specified block group. We refer to this number as the number of residing devices in a block group.

Visits are categorized based on the behavior associated with the visit. Our main category of interest is the one associated with visits featuring a "work" behavior. "Work" visits are those recorded Monday through Friday between 7:30 am and 5:30 pm and that dwelled for at least six hours. Ideally, we would use "work" visits for all our analysis of visits to the CBD. However,

Table A2: Summary Statistics for the NLSY Sample

Year	Number of Observations	Average Number of Hours Worked at Home	Median Hourly Pay of Commuting Workers (2021 $/hr)	Median Hourly Pay of Remote Workers (2021 $/hr)
1998	30,555	2.09	19.9	28.2
2000	30,555	2.13	20.6	30.3
2002	30,541	1.34	21.2	28.3
2004	30,548	1.64	21.5	27.0
2006	30,545	1.62	21.4	28.7
2008	30,548	2.06	21.2	29.1
2010	30,549	2.14	21.2	29.4
2012	30,540	2.60	21.2	30.3
2014	30,548	2.97	21.2	31.2
2016	30,549	3.47	21.7	32.1
2018	30,546	3.51	21.6	31.9

Data from the NLSY79. Hourly pay is adjusted for the Consumer Price Index and is reported for the main job during the specified two-year period.

this category only accounts for a small share of the total number of visits, leading the number of recorded "work" visits to be more volatile over time and less precisely measured.

For this reason, we use two different measures of visits in our analysis. First, we use the number of "work" visits from all block groups. We use the average number of monthly "work" visits in 2019 to define the location of the CBD. Second, we use the number of visits of any category by block group of origin. This is our preferred measure to study the variation over time of foot traffic patterns to the CBD.

Once the CBDs of each CBSA are defined according to the procedure defined in Section A.6, we define mobility to the CBD at the origin block group of level. For each block group i, month t, in a CBSA j, we compute the number of visits $N_{ii',t}$ originating from block group i that have as destination a block group that is inside a CBD of CBSA j, divided by the number $N_{i,t}^{res}$ of sampled devices residing in block group i at t:

$$sh_{i,t}^{CBD} = \frac{\sum_{i' \in CBD_j} N_{ii',t}}{N_{i,t}^{res}}, \tag{39}$$

noting that we allow CBSAs to have multiple CBDs.

We exclude block groups with fewer than 300 residing devices in a month. Our sample contains data on 896 CBSAs, with 203,075 distinct census block groups. The median CBSA in the sample has 59 distinct block groups with available foot traffic data to the CBD.

A.2.4 ACS Housing Rents Sample

We assemble a dataset of housing rents and local market characteristics at the block group level from the 2015-2019 ACS. The main outcome variable in our sample of block group-level rents is the median rent paid in the block group.

Our controls include the percentage of dwellings by type of structure, number of bedrooms, and construction decade, all based on the 2015-2019 ACS data. Additionally, we include controls for block-group characteristics such as the percentage of Hispanic, Black, and Asian populations. We also include controls from Lee and Lin (2017) that are defined at the census tract level.[11] These include the average slope, the maximum temperature in July, the minimum temperature in January, annual precipitations, the distance to the closest port, the distance to the closest river, the distance to the closest lake, the distance to the closest shore, and the surface area of the block group.

Finally, we compute the distance of each block group to the closest CBD of the CBSA, using the latitude and longitude of the centroid of the block group.

A.3 Estimation of the Transition Elasticity and Transition Costs

A.3.1 Predicting Transition Shares

We estimate a probit model to control for individual characteristics in transition probabilities $\ell\ell'$:

$$Y_{i\ell\ell',t} = \Phi\left(\delta_t^{r(i)} + X_{i,t}\beta\right),$$

where $Y_{i\ell\ell',t}$ is an indicator of individual i transitioning from ℓ to ℓ', $r(i)$ is individual i's education-region group (among the eight possible groups), $X_{i,t}$ contains individual i's age, age squared, tenure at the job, tenure squared, occupation, industry. We define $\hat{\lambda}^r_{\ell\ell',t} = \Phi\left(\hat{\delta}_t^r + \bar{X}_{i,t}\hat{\beta}\right)$ the predicted transition shares evaluated at mean values of X The dependent variable in the estimation regression is then defined as

$$y^r_{\ell\ell',t} \equiv \ln \frac{\hat{\lambda}^r_{\ell\ell',t}(\hat{\lambda}^r_{\ell'\ell',t+1})^\beta}{\hat{\lambda}^r_{\ell\ell,t}(\hat{\lambda}^r_{\ell\ell',t+1})^\beta}.$$

We estimate the log hourly rates for each education group-region-year after controlling for individual characteristics

$$\ln w_{i,t} = X_{i,t}\beta + \zeta_i + \omega_{i,t},$$

where $X_{i,t}$ contains individual i's age, age squared, tenure at job, tenure squared, occupation, industry, and region. We define $\ln w^r_{\ell',t+1}$ as the average value of $\hat{\omega}_{i,t+1}$ for individuals in education-region r. With 4 regions, 2 education groups, 9 periods (every two years from 2002 to 2018), 2 transition types, we obtain a maximum of $4 \times 2 \times 9 \times 2 = 144$ observations.

[11] We assign the same value of these geographic controls to all block groups inside the same tract.

A.3.2 Transitions by City Size

In this section, we conduct an alternative estimation of (22), with transition shares defined by subsets of CBSAs with similar population rather than by groups of region-educational attainment.

While the region-education groups used in our main specification allow for differences in transition patterns across broad geographical areas, they pool observations from many different cities in the same region. These cities may offer different conditions to remote workers. Pooling observations from CBSAs with similar population recognizes that larger cities may offer different remote work conditions than smaller cities—at the expense of pooling observations from different regions.

In order to partition cities into subsets with similar population, we first obtain information on the CBSA and county of each individual from the restricted-use NLSY Geocode dataset.[12] When the CBSA information is missing, we infer it from the county FIPS code, if available. For each CBSA j, $j = 1, \ldots, J$, we compute the average population $\overline{pop}_j$ over the 2000-2018 period. Given a number of subsets K, we order CBSAs by increasing average population, such that the CBSAs indexed by j and $j+1$ satisfy $\overline{pop}_j \leq \overline{pop}_{j+1}$ for any $j = 1, \ldots, J-1$. We then define the threshold indices $(\tau_k)_{k=0,\ldots,K}$, such that $\tau_0 = 0$, $\tau_K = J$, and for $k = 1, \ldots, K-1$, τ_k is defined recursively as the smallest index such that all CBSAs indexed between τ_{k-1} and τ_k account for at least $1/K$ of the total average population across all CBSAs,

$$\tau_k = \min_{\tau=\tau_{k-1},\ldots,J} \sum_{j=\tau_{k-1}}^{\tau} \overline{pop}_j \geq \frac{1}{K} \sum_{j}^{J} \overline{pop}_j, \qquad k = 1, \ldots, K-1.$$

For each $k = 1, \ldots, K$, we refer to the subset $\{j | \tau_{k-1} < j \leq \tau_k\}$ of CBSAs as subset k. We then define transition shares by counting the number of transitions into and out of remote work of individuals located in CBSAs that are part of the same subset k. By construction, subset $k = 1$ contains many small CBSAs, while subset K contains few large ones.

In Table A3, we describe the results from estimating (22) with $K = 5$ subsets of cities, defining remote work as spending either $d = 1$ or $d = 2$ days per week at home. We compare the OLS estimator with the Poisson Pseudo Maximum Likelihood (PPML) estimator. Given the limited number of remote workers observed per subset k and year t, the transition shares $\lambda^k_{\ell\ell',t}$ are sometimes computed based on a small number of observations. We define a minimum threshold $\bar{n}_{remote}$ of remote workers, such that observations $y^k_{\ell\ell',t}$ are included in the estimation only if more than $\bar{n}_{remote}$ remote workers are observed in subset k, year t. In the case where we define remote work as $d = 1$ day per week at home, we set $\bar{n}_{remote} = 20$, leaving 62 observations; with $d = 2$, we set $\bar{n}_{remote} = 10$, leaving 66 observations.

The estimates of s and F from the OLS and PPML estimator, with $d = 1$, are reported in Columns (1) and (2). The estimates of s are remarkably similar, and not significantly different from our main estimate of 0.46. The estimates of F are also close to our main estimate of 2.7.

[12] A description of the NLSY Geocode data is available at https://www.bls.gov/nls/request-restricted-data/nlsy-geocode-data.htm.

Table A3: Estimates of the Gravity Equation Based on City Size

	(1)	(2)	(3)	(4)
	Remote = 1 day		Remote = 2 days	
	OLS	PPML	OLS	PPML
$\ln \frac{w^r_{\ell',t+1}}{w^r_{\ell,t+1}}$	1.439***	1.384***	1.071**	0.886*
	(0.380)	(0.389)	(0.482)	(0.460)
$d_{\ell\ell'}$	1.100	1.105*	0.513	0.080
	(0.686)	(0.661)	(0.933)	(0.764)
$d_{\ell\ell'}(t+1)$	-0.051	-0.060	0.133	0.027
	(0.046)	(0.046)	(0.077)	(0.046)
$d_{\ell\ell'} L^r_{c,t+1}$	-0.152	-5.603	8.880	11.60
	(14.81)	(14.71)	(17.19)	(14.73)
$d_{\ell\ell'} L^r_{t+1}$	0.103	5.546	-8.899	-11.60
	(14.78)	(14.69)	(17.16)	(14.71)
$cons$	-0.339***	-0.231***	-0.406***	-0.272***
	(0.065)	(0.060)	(0.065)	(0.060)
Implied parameters				
s	0.640***	0.666***	0.860**	1.041*
	(0.169)	(0.187)	(0.387)	(0.540)
F	2.773***	1.962**	4.458**	3.610*
	(0.910)	(0.780)	(2.020)	(2.158)
N	62	62	66	66
R^2	0.14		0.20	

Robust standard errors in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Structural parameters estimated with the Delta method.

The estimates obtained from defining remote work as $d = 2$ days or more per week are reported in Columns (3) and (4). The OLS and PPML yield slightly higher estimates of both s and F, although our main values of s and F are contained in the 95% confidence intervals of these estimates.

A.4 Occupations

We define 22 groups of occupations based on the occupation classification used in the 2000 decennial census, denoted `OCC` in IPUMS. From the 2000 decennial census until the latest ACS, the 2000 Census classification is available. For the decennial census waves in 1980 and 1990, the 1990 Census classification is available. We use a crosswalk between the 1990 and 2000 Census classifications to obtain a consistent definition of our occupation groups.

Table A4: Summary Statistics for Occupation Groups

Code	Occupation	2019 Share of Total Employment	2019 Share That is Remote
43	Executive, Administrative, and Managerial Occupation	0.116	0.056
95	Management Related Occupations	0.062	0.065
124	Mathematical and Computer Scientists	0.019	0.087
156	Architecture and Engineering	0.009	0.021
196	Life, Physical, and Social Services	0.016	0.055
206	Community and Social Services	0.034	0.027
215	Legal	0.022	0.033
255	Education, Training, and Library	0.101	0.029
296	Art, Design, Entertainment, Sports and Media	0.022	0.071
354	Healthcare Practitioners and Technical	0.099	0.019
365	Healthcare Support	0.088	0.049
395	Protective Service	0.011	0.014
416	Food Preparation and Serving Related	0.058	0.011
425	Building and Grounds Cleaning and Maintenance	0.022	0.016
465	Personal Care and Service	0.033	0.043
496	Sales and Related	0.083	0.043
593	Office and Administrative Support	0.159	0.034
613	Farming, Forestry, and Fishing	0.000	0.000
694	Construction and Extraction	0.003	0.038
762	Installation, Repair, and Maintenance	0.002	0.035
896	Production	0.032	0.015
983	Transportation and Material Moving	0.000	0.000
999	Experienced Unemployed not Classified by Occupation	0.000	0.000

Occupation groups based on the occupation classification used in the 2000 decennial census, denoted `OCC` in IPUMS.

The NLSY uses the 2000 Census occupation classification from 2002 onward. Prior to 2002, the NLSY uses the 1970 Census occupation classification. We use a crosswalk from the 1970 Census to the 1990 Census classifications, then from the 1990 Census to the 2000 Census classifications to obtain a consistent definition.

We report descriptive statistics about these occupation groups in Table A5.

A.5 Remote Work Premium

In addition to our preferred estimation of the remote work premium based on individual longitudinal data in the NLSY, we also estimate here the remote work premium based on data from the American Community Survey. Hence, rather than defining remote work based on the number of hours worked at home, we define remote work as a binary variable. Specifically, we define

$d_{remote_{i,t}}$ as an indicator that equals one if an individual reports "work at home" as their means of transportation to work, and then predict $\hat{d} = 1$ to represent remote work.

We then estimate the national trend and an intercept for each of the 22 occupation groups, denoted by o. Specifically, we estimate the following equation:

$$\ln w_{i,t} = (\psi^0_{o(i)} + \psi^1 t) d_{remote_{i,t}} + X_{i,t}\beta_{o(i)} + \epsilon_{i,t},$$

where $\ln w_{i,t}$ is the log of the hourly wage for individual i at time t, $\psi^0_{o(i)}$ and ψ^1 represent the intercept and the national trend for occupation group o, respectively, and $X_{i,t}$ is a vector of individual-level control variables, including age, age squared, tenure at the job, tenure squared, industry, marital status, race, education-year, number of children, metro area-year. We focus on full-time US-born workers in metropolitan areas observed between 2000 and 2018.

Finally, we estimate the remote work premium in city j as follows:

$$\psi_{j,t} = \sum_o sh^{2018}_{o,j}(\psi^0_o + \psi^1 t)\hat{d},$$

where $sh^{2018}_{o,j}$ is the share of workers in city j working in occupation o in 2018.

A.6 Defining Central Business Districts

For each CBSA j, chosen radius r, we define the geographic coordinates of the k^{th} CBD of j as

$$z^{(1)}_{cbd,j}(r) = \arg\max_z \sum_{i \in cbsa_j} \mathbb{1}\{dist(z_i, z) \leq r\} \frac{N_{commute,i,2019}}{land_area_i}, \tag{40}$$

where i is a block group, z_i its coordinates (latitude and longitude), $dist(\cdot,\cdot)$ the geodesic distance, $N_{commute,i,2019}$ the total number of work-related visits recorded in block group i during the year 2019 from SafeGraph.

For $k \geq 2$, we impose that the center of the k^{th} CBD is not inside previous CBDs. We retain the k^{th} CBD if it attracts at least 80% as many commuters as the main CBD and if it does not overlap with CBDs 1 to $k-1$. We define the distance of each zip code to the CBD as the distance to the nearest CBD.

A.7 Change in Visits to the CBD for Selected Cities

In Figure A3, we report the time series of visits to the CBD of six selected cities. The estimates are the coefficients for monthly indicators in a regression at the block group-month level. The outcome variable is the number of visits to the CBD of the CBSA from the specified block group, divided by the number of residing devices in that block group, as defined in (39), scaled by the value of this ratio in January 2020. That is, our outcome variable is $y_{i,t} = sh^{CBD}_{i,t}/sh^{CBD}_{i,t_0}$, where i is a block group, t is a month, and t_0 is set to January 2020. We then estimate the following

Table A5: Remote Work Premium for Occupation Groups

	(1) NLSY	(2) ACS
Remote	-0.233***	-0.174***
Remote × Year	0.011***	0.006***
Executive, Administrative, and Managerial × Remote	0.000	0.000
Management Related Occupations × Remote	0.029	0.129***
Mathematical and Computer Scientists × Remote	0.089**	0.181***
Architecture and Engineering × Remote	0.048	0.216***
Life, Physical, and Social Services × Remote	0.176*	0.138***
Community and Social Services × Remote	0.032	-0.337***
Legal × Remote	0.001	0.035***
Education, Training, and Library × Remote	0.056	0.025***
Art, Design, Entertainment, Sports and Media × Remote	-0.086*	-0.006**
Healthcare Practitioners and Technical × Remote	-0.002	0.203***
Healthcare Support × Remote	-0.208**	-0.283***
Protective Service × Remote	0.002	0.027***
Food Preparation and Serving Related × Remote	-0.123*	0.031***
Building and Grounds Cleaning and Maintenance × Remote	-0.147**	-0.092***
Personal Care and Service × Remote	-0.134**	-0.047***
Sales and Related × Remote	0.012**	0.310***
Office and Administrative Support × Remote	-0.135**	0.150***
Farming, Forestry, and Fishing × Remote	-0.336	0.205***
Construction and Extraction × Remote	0.129**	0.075***
Installation, Repair, and Maintenance × Remote	0.034	0.038***
Production × Remote	-0.052*	0.032***
Transportation and Material Moving × Remote	-0.231*	0.047***
Indiv. Controls	Yes	Yes
Year FE	Yes	Yes
Occ. & Ind. FE	Yes	Yes
Indiv. FE	Yes	No
Observations	34,784	87,160,166

Occupation groups based on the occupation classification used in the 2000 decennial census, denoted `OCC` in IPUMS.

regression for each CBSA:

$$y_{i,t} = \sum_{s=s_0}^{S} \beta_s \mathbb{1}\{t = s\} + \epsilon_{i,t},$$

where $s_0 = 1$ for October 2019, and $S = 34$ for July 2022. We restrict the sample to block groups that are located less than 50 kilometers away from the CBD, and cluster standard errors at the census tract level, which contains 3.4 block groups on average.

Figure A1: Central Business Districts of Selected Large Cities

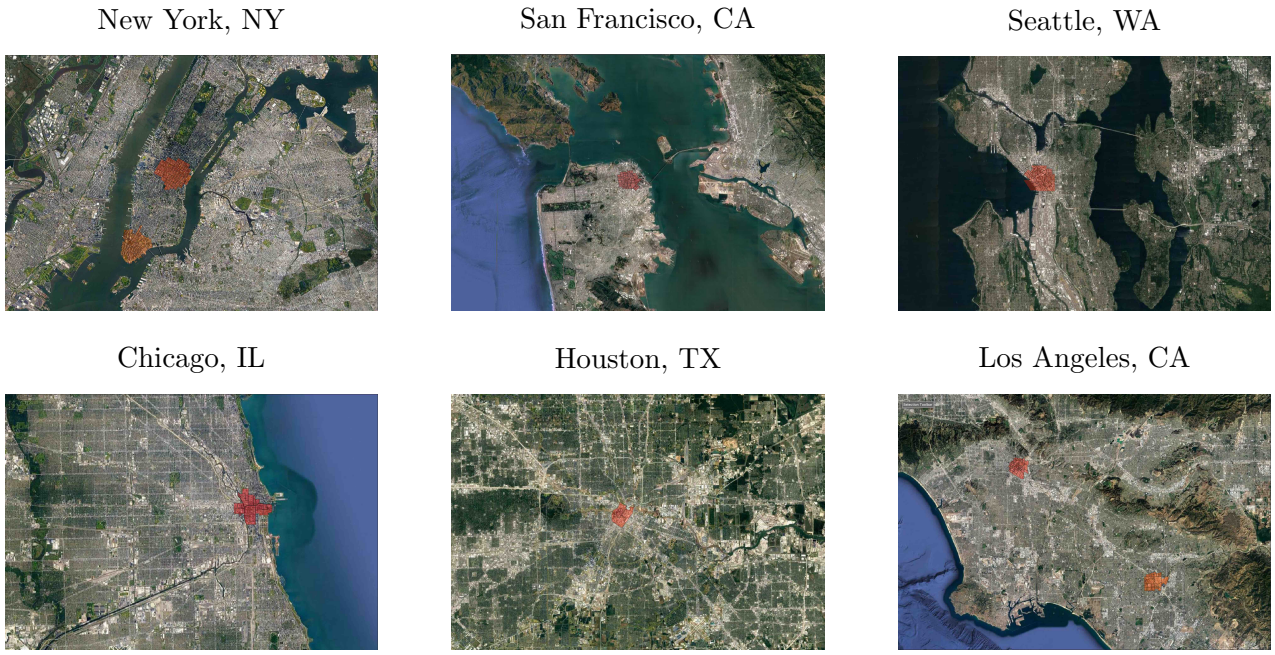
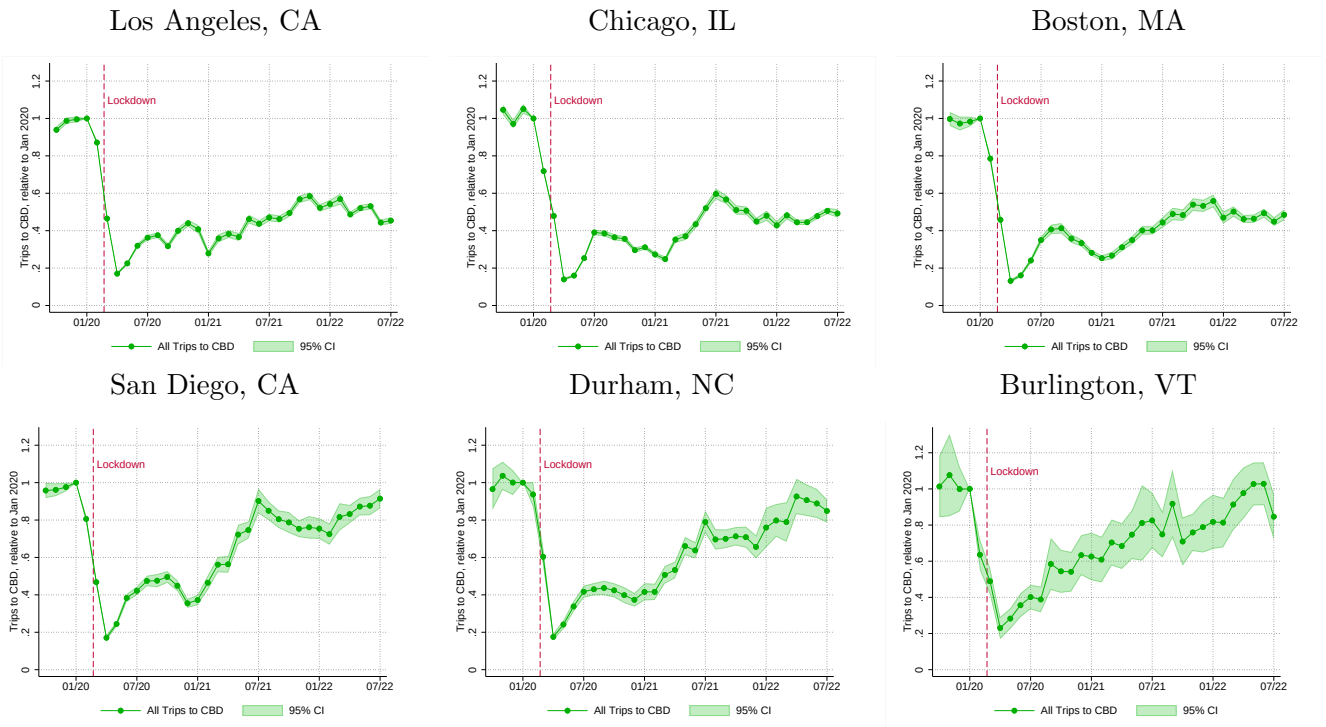


Figure A3 reports the estimates of β_s for the CBSAs of Los Angeles, CA, Chicago, IL, Boston, MA, San Diego, CA, Durham, NC, and Burlington, VT, along with the 95 percent confidence interval around each estimate. All six CBSAs experience a sudden drop in visits to the CBD in March-April 2020, corresponding to the period of nation-wide stay-at-home orders at the onset of the COVID-19 pandemic. The share of visits to the CBD drops to about 20 percent of their value just three months earlier. The patterns in the following months are similar for the three cities in the top panel. After some ebbs and flows likely explained by the various waves of COVID-19 infections, visits to the CBD in these CBSAs stabilize at, or below, 50 percent of January 2020 levels. In the three cities in the lower panel, however, the patterns of visits to the CBD is starkly different: after a small temporary upset in the winter of 2021, visits to the CBD continue to increase and return to pre-pandemic levels, or close to, by the end of the sample period.

A.8 Evidence on Remote Work, Agglomeration Externalities and Transportation Elasticities

The left side of Figure A4 reports the relationship between the change in visits to the CBD and our estimates of agglomeration externalities δ, transportation elasticities γ, and the net strength of agglomeration forces $\delta - \gamma/\mu$. Each bubble represents a city of population proportional to the bubble area. The right side of Figure A4 reports the share of cities that are predicted to be in their cone of multiplicity as a function of the same parameters. Blue dots indicate the fraction of cities in the cone in the calibrated model for 20 quantiles of the variable indicated on the horizontal axis.

Figure A3: Change in Visits to the CBD for Selected Cities



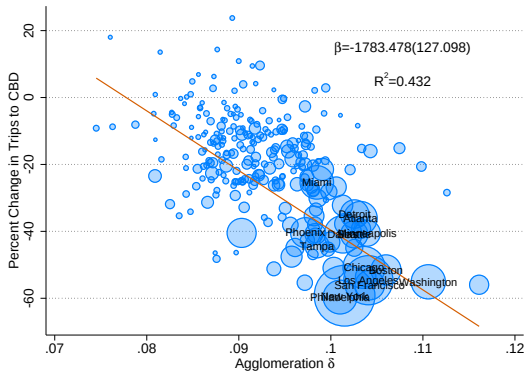
Estimates of monthly indicators in a separate regression for each CBSA of visits to the CBD at the block group-month level. The outcome variable is the number of visits to the CBD from the specified block group, divided by the number of residing devices in that block group, as defined in (39), scaled by the value of this ratio in January 2020. The shaded areas represent the contour delimited by the consecutive 95 percent confidence intervals of each monthly indicator. Standard errors are clustered at the census tract of origin.

The left side of the figure shows that cities that experienced larger declines in visits to the CBD are those that have a stronger agglomeration externality δ, and a weaker transportation cost elasticity γ, and therefore also those with a higher net strength of agglomeration $\delta - \gamma/\mu$. These relationships are quite strong, as indicated by the magnitude and significance of the population-weighted linear fit in the figures and values of R^2 between 0.11 and 0.22.

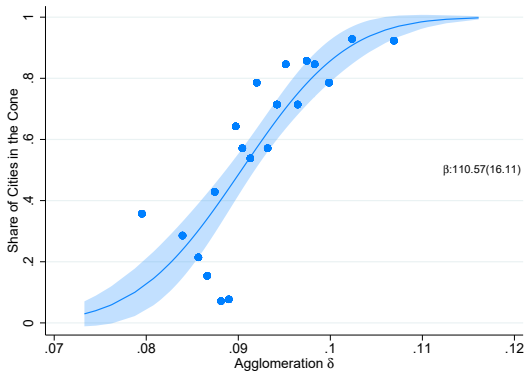
The right side of the figure shows that the relationship between the drop in visits to the CBD and the key agglomeration parameters is rooted in the workings of the model. The figures show that calibrated cities that tend to have higher agglomeration externalities δ, lower transportation elasticities γ, and higher net strength of agglomeration forces $\delta - \gamma/\mu$, are the ones that are the most likely to be inside the cone of multiplicity. Therefore, our model predicts that these are the cities that are the most likely to remain in a low-commuting equilibrium after a temporary shock such as a lockdown. Panel (d) in this figure excludes some large outliers focusing on $\gamma < 0.1$. Panel (f) is estimated using net strength values larger than zero since multiplicity cannot occur otherwise. These choices are made for the visibility of the graphs, and the estimates return unchanged messages without these restrictions.

Figure A4: Visits to CBD and Share of Cities in the Cone vs. City Parameters

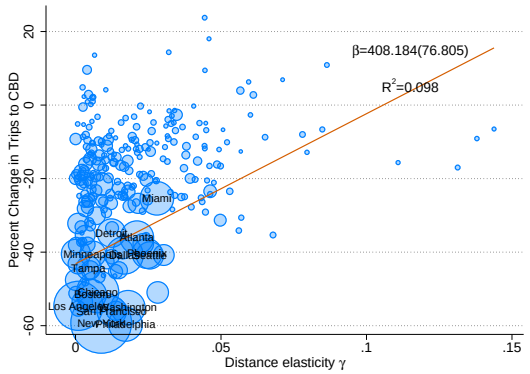
(a) Agglomeration Externality δ
Trips to CBD, mid 2022 rel. to Jan 2020



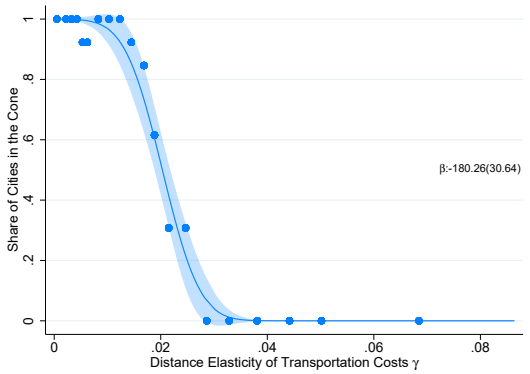
(b) Agglomeration Externality δ
Share of Cities in the Cone



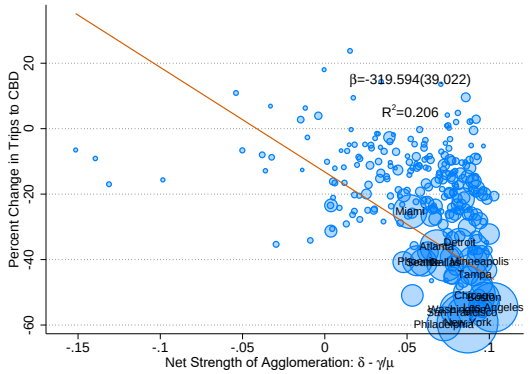
(c) Transportation Elasticity γ
Trips to CBD, mid 2022 rel. to Jan 2020



(d) Transportation Elasticity γ
Share of Cities in the Cone



(e) Net Strength of Agglomeration $\delta - \gamma/\mu$
Trips to CBD, mid 2022 rel. to Jan 2020



(f) Net Strength of Agglomeration $\delta - \gamma/\mu$
Share of Cities in the Cone

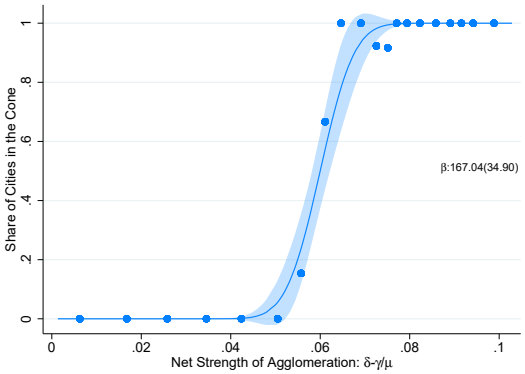
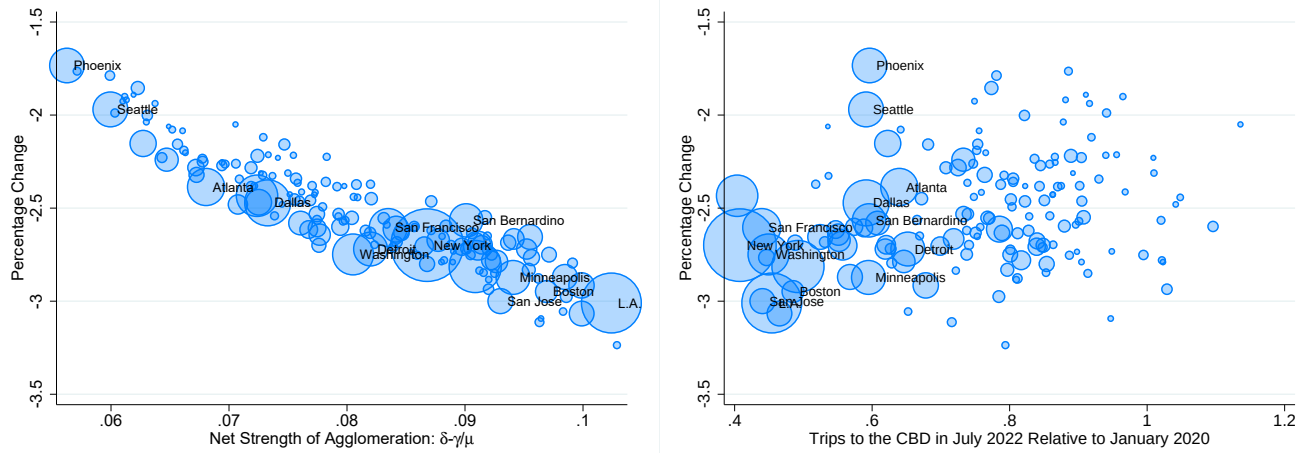


Figure A5: Correlates of Welfare Change $\hat{\mathcal{W}}$

(a) With Net Agglomeration Strength

(b) With CBD Trip Decline



A.9 Change in Trips to CBD and Welfare Change from Switching to Remote Equilibrium for Largest Cities

In Table A6, we report the values of our estimates for the change in visits to the CBD and the welfare of the low-commuting stationary equilibrium relative to the high-commuting steady-state equilibrium. Our sample is the largest 45 U.S. cities for which we can compute the relevant parameters and the commuting trips to the CBD.

In the table, Column 2 reports the employment from the 2019 Census that we use in our calibrations, as described in Appendix A.10. Column 3 reports the change in visits to the CBD for each city from January 2020 to July 2022, as described in Appendix A.7.[13] Column 4 reports the welfare ratio $\mathcal{W}(L^{ss}_{c,low})/\mathcal{W}(L^{ss}_{c,high})$, defined in Section 5.4. These values are only computed for the CBSAs that are in their cone of multiplicity according to our estimation. We indicate by "U" the CBSAs that are outside their code, for which there exists no low-commuting stationary equilibrium.

Figure A5 shows correlates of $\hat{\mathcal{W}}$, the welfare in the low-commuting relative to the high-commuting stationary equilibrium. These scatterplots can only be computed for the set of cities in the cone of multiplicity. The left panel shows a scatterplot between $\hat{\mathcal{W}}$ and the net strength of agglomeration externalities, $\delta - \gamma/\mu$. The right panel shows a scatterplot between $\hat{\mathcal{W}}$ and the share of trips to the CBD in July 2022, relative to January 2020. The marker size is proportional to total city employment.

[13] These values are a simple transformation of those plotted on the y axis of Panel (b) of Figure 10 (the complement to 100).

Table A6: Change in Trips to CBD and Welfare Loss in Remote Equilibrium for a Group of 45 Cities

CBSA	2019 Census Employment	Trips to CBD Jul'22 / Jan'20	Welfare Change $\mathcal{W}(L^{ss}_{c,low})/\mathcal{W}(L^{ss}_{c,high})$
New York-Northern New Jersey-Long Island, NY-NJ-PA	6,983,328	0.408	0.973
Los Angeles-Long Beach-Santa Ana, CA	4,748,281	0.454	0.970
Chicago-Joliet-Naperville, IL-IN-WI	3,598,768	0.492	0.972
Dallas-Fort Worth-Arlington, TX	2,792,439	0.591	0.975
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	2,270,070	0.404	0.976
Washington-Arlington-Alexandria, DC-VA-MD-WV	2,211,377	0.449	0.973
San Francisco-Oakland-Fremont, CA	1,914,574	0.439	0.974
Atlanta-Sandy Springs-Marietta, GA	1,827,559	0.640	0.976
Seattle-Tacoma-Bellevue, WA	1,606,160	0.591	0.980
Phoenix-Mesa-Glendale, AZ	1,582,530	0.596	0.983
Minneapolis-St. Paul-Bloomington, MN-WI	1,520,268	0.595	0.971
Detroit-Warren-Livonia, MI	1,516,211	0.652	0.973
Riverside-San Bernardino-Ontario, CA	1,468,173	0.595	0.974
Miami-Fort Lauderdale-Pompano Beach, FL	1,232,340	0.746	U
Tampa-St. Petersburg-Clearwater, FL	1,095,802	0.557	0.973
Austin-Round Rock-San Marcos, TX	906,522	0.623	0.978
Portland-Vancouver-Hillsboro, OR-WA	888,928	0.490	U
Charlotte-Gastonia-Rock Hill, NC-SC	885,802	0.785	0.974
Cincinnati-Middletown, OH-KY-IN	845,880	0.678	0.971
San Jose-Sunnyvale-Santa Clara, CA	826,205	0.440	0.970
Columbus, OH	822,897	0.568	0.971
Sacramento–Arden-Arcade–Roseville, CA	811,475	0.524	0.973
Las Vegas-Paradise, NV	807,521	0.465	0.969
Pittsburgh, PA	794,500	0.592	U
Nashville-Davidson–Murfreesboro–Franklin, TN Are	735,658	0.608	0.974
Kansas City, MO-KS	701,694	0.733	0.978
Indianapolis-Carmel, IN	662,938	0.646	0.972
Virginia Beach-Norfolk-Newport News, VA-NC	662,276	0.552	0.973
Boston-Cambridge-Quincy, MA-NH	630,693	0.485	0.971
Milwaukee-Waukesha-West Allis, WI	599,389	0.551	0.974
Raleigh-Cary, NC	558,030	0.719	0.973
Richmond, VA	534,987	0.619	0.973
Oklahoma City, OK	466,449	0.817	0.972
Hartford-West Hartford-East Hartford, CT	451,318	0.621	0.973
Providence-New Bedford-Fall River, RI-MA	446,892	0.699	0.973
Buffalo-Niagara Falls, NY	412,839	0.588	0.974
Grand Rapids-Wyoming, MI	412,449	0.547	0.974
Omaha-Council Bluffs, NE-IA	389,758	0.839	0.973
Rochester, NY	385,287	0.574	0.974
Albany-Schenectady-Troy, NY	354,110	0.725	0.977
New Orleans-Metairie-Kenner, LA	341,194	0.447	0.972
Bridgeport-Stamford-Norwalk, CT	331,525	0.840	0.973
Little Rock-North Little Rock-Conway, AR	313,624	0.789	0.974
New Haven-Milford, CT	310,423	0.733	0.975
Tucson, AZ	301,333	0.488	0.973

The letter "U" in Column 4 indicates that the CBSA is outside the multiplicity cone according to our estimates.

A.10 Descriptive Evidence Using Census Employment

A.10.1 Individual Mobility

In Section 5, we characterize CBSAs by their total non-farm employment in 2019, published by the BEA. While this variable provides a measure of city size for all CBSAs, it is distinct from the measure of city size that we use in the quantification of the model in several ways. First, it is based on counting the number of jobs in the CBSAs, rather than on the number of workers, as is the case in our Census data. Since some workers hold more than one job, the BEA measure is typically larger than our Census measure. Second, because the Census data is based on individual records, we can count the number of commuters and calibrate our model to workers who are more likely to face the trade-off between remote work and commuting that our theory speaks to. For example, the Census allows us to focus on workers between 25 and 64 years of age and to exclude self-employed workers. As a result, our Census-based measure of CBSA employment is 56 percent lower than the BEA measure on average. The correlation between the log of the two employment measures is 0.89.

Using the Census individual data to measure commuting decisions, we can compute total employment for 598 out of the 847 CBSAs represented in Figure 9. In Figure A6 and A7, we reproduce the analogous figures to Figures 9 and 10 with this smaller sample of CBSAs. In Figure A6, we use lower thresholds of employment (1 million for "large" CBSAs, 100,000 for "small" CBSAs) to reflect the fact that our measure of employment from the Census is lower than the BEA measure on average. These thresholds of Census employment reproduce the partition of CBSAs in Figure 9, with the first group containing the top 3 percent of CBSAs in terms of employment, the second containing the bottom 75 percent of CBSAs.

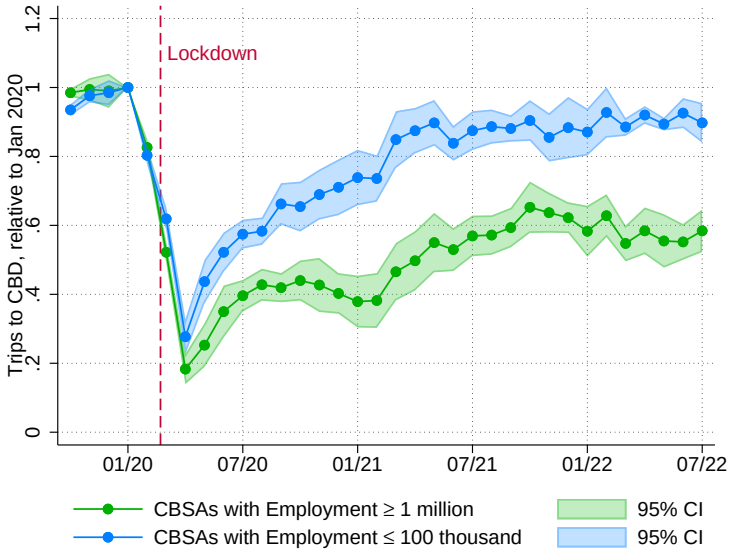
The patterns of trips to CBD over time are very similar to those described in Section 5.1. Namely, CBSAs of all sizes experience a large decline of trips to CBD of 70 to 80 percent during the stay-at-home policies in 2020, with larger CBSAs only recovering to 60 percent of their pre-pandemic values, while smaller CBSAs return at least to 90 percent on average. Figure A7 confirm that the relationship between the size of the drop in trips to CBD and CBSA employment became stronger by July 2022 than in April 2020.

A.10.2 Distance Gradients in House Prices

Figure A8 shows that the divergence in the average distance elasticity of house prices relative to January 2020 between smaller and larger cities is visible also when we measure size with Census employment.

Panel (a) of Figure A9 confirms that the flattening of the housing price gradients occurs in similar magnitude for cities of all sizes in January 2021. Panel (b) confirms that even in the sample of cities ordered by Census employment, the drop in the gradient is more strongly related to city size by December 2022.

Figure A6: Visits to the CBD Relative to January 2020, Census Employment

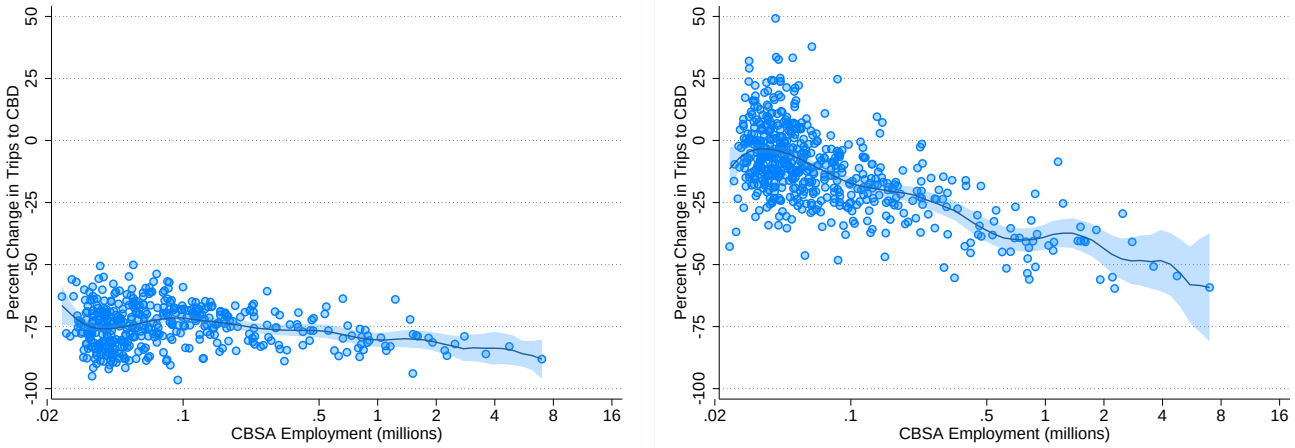


The figure reports the estimates of the average volume of visits to the CBD from block groups located in i) CBSAs with total Census employment above 1 million (19 largest CBSAs), ii) CBSAs with total employment below 100,000 (450 smallest CBSAs) expressed as a share of their values in January 2020.

Figure A7: Response of Trips to the CBD, Census Employment

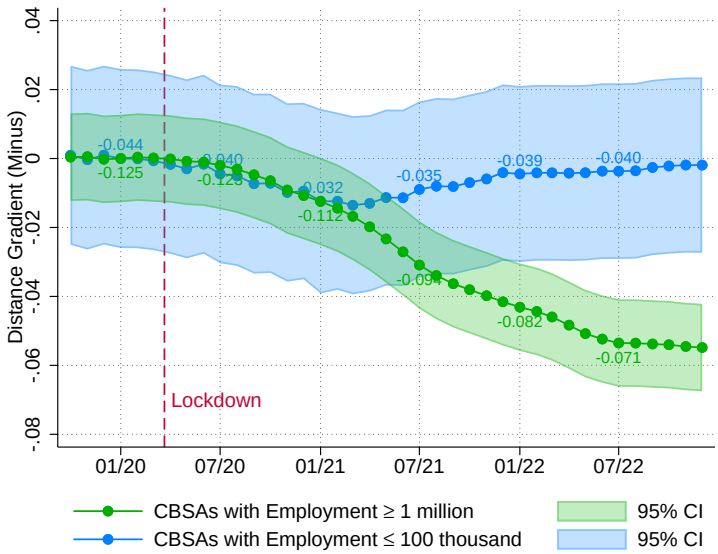
(a) April 2020 Relative to January 2020

(b) July 2022 Relative to January 2020



Panel (a) reports the change in the average level of visits to each CBSA's CBD between January 2020 and April 2020, against our measure of employment from the Census. Panel (b) reports the change in visits to the CBD in July 2022 relative to January 2020. Each marker represents a CBSA. The solid line represents a fitted kernel-weighted local polynomial smoothing, with the shaded area corresponding to the 95 percent confidence interval around the fitted value.

Figure A8: Distance Gradient of Housing Prices Relative to January 2020, Census Employment

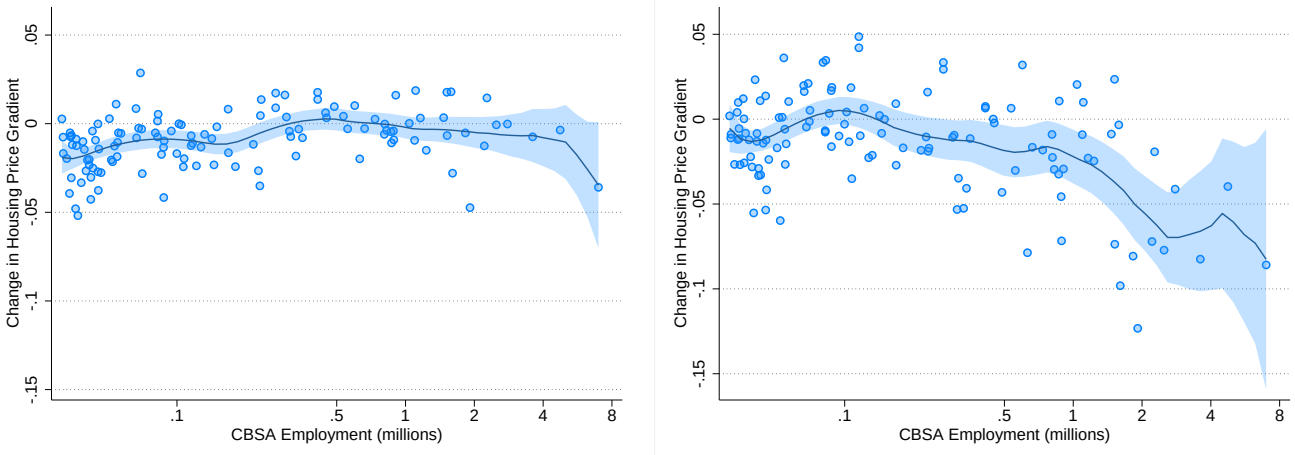


This figure reports the difference between the average housing price gradients in January 2020 and each following month, using the same Census employment thresholds as in Figure A6. A negative value for the difference in housing price gradients implies a lower premium for housing located close to the CBD. The value of the average housing price gradient in each group is reported next to the marker of the associated group every six months. In total, 19 CBSAs have employment above 1 million and 68 have employment below 100,000. The shaded areas represent the contour delimited by the consecutive 95 percent confidence intervals of each monthly indicator.

Figure A9: Response of House Price Gradients, Census Employment

(a) January 2021 Relative to January 2020

(b) December 2022 Relative to January 2020



This figure reports the difference between the housing price gradients in January 2020 and December 2022, against our measure of employment from the Census. Each marker represents a CBSA. The gradients are estimated using zip code level data from Zillow for 127 CBSAs. A negative value for the difference implies a lower premium for housing located close to the CBD. The solid line represents a fitted kernel-weighted local polynomial smoothing, with the shaded area corresponding to the 95 percent confidence interval around the fitted value.

B Theory

B.1 Rewriting the average value of a worker

By adding and subtracting the continuation value in the current labor delivery mode, the equation for U_ℓ can be rewritten as

$$\begin{aligned} U_\ell\left(\{\varepsilon_{c,t},\varepsilon_{m,t}\};L_{c,t},\omega^t\right) &= u_\ell\left(L_{c,t};A_t,T_t,z_t\right)+\max_{\ell'}\left\{\beta V_{\ell'}\left(L_{c,t+1};\omega^{t+1}\right)-F_{\ell\ell',t}+\varepsilon_{\ell',t}\right\} \\ &= u_\ell\left(L_{c,t};A_t,T_t,z_t\right)+\beta V_\ell\left(L_{c,t+1};\omega^{t+1}\right)-F_{\ell\ell,t}+ \\ &\max_{\ell'}\left\{\beta\left[V_{\ell'}\left(L_{c,t+1};\omega^{t+1}\right)-V_\ell\left(L_{c,t+1};\omega^{t+1}\right)\right]-\left(F_{\ell\ell',t}-F_{\ell\ell,t}\right)+\varepsilon_{\ell',t}\right\} \\ &= u_\ell\left(L_{c,t};A_t,T_t,z_t\right)+\beta V_\ell\left(L_{c,t+1};\omega^{t+1}\right)-F_{\ell\ell,t}+\max_{\ell'}\left\{\bar{\varepsilon}_t^{\ell\ell'}+\varepsilon_{\ell',t}\right\} \end{aligned} \tag{41}$$

where

$$\bar{\varepsilon}_t^{\ell\ell'}\left(L_{c,t+1}\right)\equiv\beta\left[V_{\ell'}\left(L_{c,t+1};\omega^{t+1}\right)-V_\ell\left(L_{c,t+1};\omega^{t+1}\right)\right]-\left(F_{\ell\ell',t}-F_{\ell\ell,t}\right) \tag{42}$$

is the average value of switching from ℓ to ℓ'. Taking the expectation of U_ℓ with respect to ε_t,

$$\begin{aligned} V_\ell\left(L_{c,t};z_t\right) &= u_\ell\left(L_{c,t};A_t,T_t,z_t\right)+\beta V_\ell\left(L_{c,t+1};\omega^{t+1}\right)-F_{\ell\ell,t}+E_t\max_{\ell'}\left\{\bar{\varepsilon}_t^{\ell\ell'}+\varepsilon_{\ell',t}\right\}\equiv \\ &\equiv u_\ell\left(L_{c,t};A_t,T_t,z_t\right)+\beta V_\ell\left(L_{c,t+1};\omega^{t+1}\right)-F_{\ell\ell,t}+\Omega\left(\bar{\varepsilon}_t^{\ell}\right) \end{aligned} \tag{43}$$

where $\bar{\varepsilon}_t^{\ell}\equiv\left\{\bar{\varepsilon}_t^{\ell c},\bar{\varepsilon}_t^{\ell m}\right\}$. $\Omega\left(\bar{\varepsilon}_t^{\ell}\right)$ is the value of the option of being able to move out of ℓ if circumstances require, or the "option value" of ℓ.

Under our distributional assumption for ε_t we can write the share of workers in ℓ state transitioning to ℓ' as

$$\lambda_t^{\ell\ell'}=\frac{\exp\left(\bar{\varepsilon}_t^{\ell\ell'}\left(L_{c,t+1}\right)/s\right)}{\exp\left(\bar{\varepsilon}_t^{\ell c}\left(L_{c,t+1}\right)/s\right)+\exp\left(\bar{\varepsilon}_t^{\ell m}\left(L_{c,t+1}\right)/s\right)} \tag{44}$$

We can use this expression to obtain the average expected switching values. In fact,

$$\frac{\lambda_t^{\ell\ell'}}{\lambda_t^{\ell\ell}}=\exp\left(\bar{\varepsilon}_t^{\ell\ell'}/s-\bar{\varepsilon}_t^{\ell\ell}/s\right)\Longrightarrow s\ln\frac{\lambda_t^{\ell\ell'}}{\lambda_t^{\ell\ell}}=\bar{\varepsilon}_t^{\ell\ell'}-\bar{\varepsilon}_t^{\ell\ell}\Longrightarrow\bar{\varepsilon}_t^{\ell\ell'}=s\ln\frac{\lambda_t^{\ell\ell'}}{\lambda_t^{\ell\ell}} \tag{45}$$

To obtain the option value of being in ℓ, recall that

$$\Omega\left(\bar{\varepsilon}_t^{\ell}\right)=s\log\left[\exp\left(\bar{\varepsilon}_t^{\ell c}/s\right)+\exp\left(\bar{\varepsilon}_t^{\ell m}/s\right)\right] \tag{46}$$

Since $\log\lambda_t^{\ell\ell}=\bar{\varepsilon}_t^{\ell\ell}/s-\log\left[\exp\left(\bar{\varepsilon}_t^{\ell c}/s\right)+\exp\left(\bar{\varepsilon}_t^{\ell m}/s\right)\right]=-\log\left[\exp\left(\bar{\varepsilon}_t^{\ell c}/s\right)+\exp\left(\bar{\varepsilon}_t^{\ell m}/s\right)\right]$, we have that

$$\Omega\left(\bar{\varepsilon}_t^{\ell}\right)=-s\ln\lambda_t^{\ell\ell} \tag{47}$$

B.2 Existence of a stationary equilibrium

To prove that a stationary equilibrium always exists, we proceed in steps. First, Lemma B.1 shows that the characterization of a stationary equilibrium can be reduced to the zeros of one equation in one unknown, the share of commuters who remain commuters next period, λ^{cc}. Lemma B.2 establishes a useful property of one component in such equation. Finally, Proposition B.1 shows that this equation always has at least one solution.

In a stationary equilibrium, the system is characterized by 11 unknowns, u_c,u_m,L_c,V_c,V_m,λ^{cc},λ^{mm}, $\bar{\varepsilon}^{cc}$,$\bar{\varepsilon}^{cm}$,$\bar{\varepsilon}^{mc}$,$\bar{\varepsilon}^{mm}$, and 11 equations: the flow utility values,

$$u_c = \log r_a^{\alpha-1}\bar{\alpha}\frac{B(\tilde{L}_c)}{\tau(\tilde{L}_c)}\frac{w_c}{\tilde{L}_c^{\gamma}} \tag{48}$$

$$u_m = \log r_a^{\alpha-1}\bar{\alpha}T^{\mu}\left(\frac{B(\tilde{L}_c)}{\tau(\tilde{L}_c)}\right)^{1-\mu}\frac{w_m}{L^{\gamma(1-\mu)}} \tag{49}$$

the law of motion for L_c,

$$L_c = L_c\lambda^{cc} + (L - L_c)(1 - \lambda^{mm}) \tag{50}$$

the average value functions by mode,

$$V_c = u_c + \beta V_c - s\ln\lambda^{cc} - F_{cc} \tag{51}$$

$$V_m = u_m + \beta V_m - s\ln\lambda^{mm} - F_{mm} \tag{52}$$

the transition probabilities,

$$\lambda^{cc} = \frac{\exp(\bar{\varepsilon}^{cc}/s)}{\exp(\bar{\varepsilon}^{cc}/s) + \exp(\bar{\varepsilon}^{cm}/s)} \tag{53}$$

$$\lambda^{mm} = \frac{\exp(\bar{\varepsilon}^{mm}/s)}{\exp(\bar{\varepsilon}^{mc}/s) + \exp(\bar{\varepsilon}^{mm}/s)} \tag{54}$$

and the expected values from switching

$$\bar{\varepsilon}^{cc} = 0 \tag{55}$$

$$\bar{\varepsilon}^{mm} = 0 \tag{56}$$

$$\bar{\varepsilon}^{cm} = \beta(V_m - V_c) - (F_{cm} - F_{cc}) \tag{57}$$

$$\bar{\varepsilon}^{mc} = \beta(V_c - V_m) - (F_{mc} - F_{mm}) \tag{58}$$

The following lemma shows that we can always reduce this system of equations to one equation in one unknown.

Lemma B.1 *The characterization of the stationary equilibrium (48)-(58) can be reduced to one equation in one unknown,*

$$f\left(\lambda^{cc}\right) \equiv \frac{1}{1+\exp\left(\bar{\varepsilon}^{cm}\left(\lambda^{cc}\right)/s\right)} - \lambda^{cc} = 0 \tag{59}$$

where

$$\bar{\varepsilon}^{cm}\left(\lambda^{cc}\right) \equiv \frac{\beta}{1-\beta}\left[e_f\left(\lambda^{cc}\right) + e_o\left(\lambda^{cc}\right) - \left(F_{mm} - F_{cc}\right)\right] - F_{cm} + F_{cc} \tag{60}$$

is the value of switching from c to m; the function

$$e_o\left(\lambda^{cc}\right) \equiv s \ln \lambda^{cc} \frac{1-\phi\lambda^{cc}}{1-\lambda^{cc}} \tag{61}$$

is the difference in the option values of staying in m versus staying in c; and the function

$$e_f\left(\lambda^{cc}\right) \equiv u_m\left(L_c\left(\lambda^{cc}\right)\right) - u_c\left(L_c\left(\lambda^{cc}\right)\right) \tag{62}$$

is the difference in the flow value of utilities, with

$$L_c\left(\lambda^{cc}\right) = \frac{\left(1-\phi\right)\lambda^{cc}}{\left(1-\lambda^{cc}\right)\left(1-\phi\lambda^{cc}\right)+\left(1-\phi\right)\lambda^{cc}} L \equiv \sigma\left(\lambda^{cc}\right) L \tag{63}$$

$$\phi \equiv \left[1 - \exp\left(\frac{F_{cc}+F_{mm}-F_{cm}-F_{mc}}{s}\right)\right] \leq 1 \tag{64}$$

Proof. *Rearrange the value of switching from c to m (eq. (57)) as* $\beta\left(V_c - V_m\right) = -\bar{\varepsilon}^{cm} - F_{cm} + F_{cc}$. *Substitute this expression into eq. (58) for* $\bar{\varepsilon}^{mc}$ *to obtain*

$$\bar{\varepsilon}^{mc} = \beta\left(V_c - V_m\right) - F_{mc} + F_{mm} \tag{65}$$

$$= -\bar{\varepsilon}^{cm} - F_{cm} + F_{cc} - F_{mc} + F_{mm} \tag{66}$$

Rearranging the average value functions (51)-(52), one can write

$$V_c = \left(u_c - s\ln\lambda^{cc} - F_{cc}\right)/\left(1-\beta\right)$$
$$V_m = \left(u_m - s\ln\lambda^{mm} - F_{mm}\right)/\left(1-\beta\right)$$

and use these expressions into (57) to obtain

$$\bar{\varepsilon}^{cm} = \beta\left(V_m - V_c\right) - F_{cm} + F_{cc} = \frac{\beta}{1-\beta}\left[u_m - s\ln\lambda^{mm} - F_{mm} - \left(u_c - s\ln\lambda^{cc} - F_{cc}\right)\right] - F_{cm} + F_{cc} =$$

$$= \frac{\beta}{1-\beta}\left[u_m - u_c + s\ln\lambda^{cc}/\lambda^{mm} - F_{mm} + F_{cc}\right] - F_{cm} + F_{cc} \tag{67}$$

We want to express the right-hand side of this expression only as a function of λ_{cc}.

To do that, note that eq. (53) and (55) together imply

$$\lambda^{cc} = \frac{1}{1+\exp\left(\bar{\varepsilon}^{cm}/s\right)} \Longrightarrow \exp\left(-\bar{\varepsilon}^{cm}/s\right) = \lambda^{cc}/\left(1-\lambda^{cc}\right) \tag{68}$$

Use (56) and (66) into eq. (54) for λ^{mm}*,*

$$\lambda^{mm} = \frac{1}{\exp\left(\left(-\bar{\varepsilon}^{cm} - F_{cm} + F_{cc} - F_{mc} + F_{mm}\right)/s\right)+1} \tag{69}$$

We can then use eq. (68) in (69) to obtain,

$$\begin{aligned}
\lambda^{mm} &= \frac{1}{\exp\left(-\bar{\varepsilon}^{cm}/s\right)\exp\left(\left(-F_{cm} + F_{cc} - F_{mc} + F_{mm}\right)/s\right)+1} = \\
&= \frac{1}{\lambda^{cc}/\left(1-\lambda^{cc}\right)\exp\left(\left(-F_{cm} + F_{cc} - F_{mc} + F_{mm}\right)/s\right)+1} = \\
&= \frac{\left(1-\lambda^{cc}\right)}{\lambda^{cc}\exp\left(\left(-F_{cm} + F_{cc} - F_{mc} + F_{mm}\right)/s\right)+\left(1-\lambda^{cc}\right)} = \\
&= \frac{\left(1-\lambda^{cc}\right)}{1-\lambda^{cc}\left[1-\exp\left(\left(-F_{cm} + F_{cc} - F_{mc} + F_{mm}\right)/s\right)\right]} \Longrightarrow \\
\lambda^{mm} &= \frac{1-\lambda^{cc}}{1-\phi\lambda^{cc}} \qquad (70) \\
\text{with } \phi &\equiv \left[1-\exp\left(\frac{F_{cc}+F_{mm}-F_{cm}-F_{mc}}{s}\right)\right] \leq 1 \qquad (71)
\end{aligned}$$

Use (70) to write the stationary equilibrium value of L_c *(from eq. (50)) as*

$$\begin{aligned}
&L_c = L_c\lambda^{cc} + \left(L - L_c\right)\left(1-\lambda^{mm}\right) \\
&L_c - L_c\lambda^{cc} + L_c\left(1-\lambda^{mm}\right) = L\left(1-\lambda^{mm}\right) \\
&L_c\left(1-\lambda^{cc} + \frac{\left(1-\phi\right)\lambda^{cc}}{1-\phi\lambda^{cc}}\right) = L\frac{\left(1-\phi\right)\lambda^{cc}}{1-\phi\lambda^{cc}} \\
&L_c\left(\left(1-\lambda^{cc}\right)\left(1-\phi\lambda^{cc}\right) + \left(1-\phi\right)\lambda^{cc}\right) = L\left(1-\phi\right)\lambda^{cc} \\
&L_c = \frac{\left(1-\phi\right)\lambda^{cc}}{\left(1-\lambda^{cc}\right)\left(1-\phi\lambda^{cc}\right) + \left(1-\phi\right)\lambda^{cc}} L \equiv \sigma\left(\lambda^{cc}\right) L
\end{aligned} \tag{72}$$

Making use of (70) and (72), we can then rewrite (67) as

$$\begin{aligned}
\bar{\varepsilon}^{cm} &= \frac{\beta}{1-\beta}\left[u_m\left(L_c\left(\lambda^{cc}\right)\right) - u_c\left(L_c\left(\lambda^{cc}\right)\right) + s\ln\lambda^{cc}/\lambda^{mm} - F_{mm} + F_{cc}\right] - F_{cm} + F_{cc} \\
&= \frac{\beta}{1-\beta}\left[u_m\left(L_c\left(\lambda^{cc}\right)\right) - u_c\left(L_c\left(\lambda^{cc}\right)\right) + s\ln\lambda^{cc}\frac{1-\phi\lambda^{cc}}{1-\lambda^{cc}} - F_{mm} + F_{cc}\right] - F_{cm} + F_{cc}
\end{aligned} \tag{73}$$

■

Lemma B.2 *The function $e_o(\lambda)$ in eq. (61) increases from -∞ to $+\infty$ as $\lambda \in (0,1)$.*
Proof. *The derivative of $e_o(\lambda) = s \ln\left(\lambda \frac{1-\phi\lambda}{1-\lambda}\right)$ with respect to λ is*

$$\begin{aligned} s\left[\frac{1}{\lambda} - \frac{\phi}{1-\phi\lambda} + \frac{1}{1-\lambda}\right] &= s\left[\frac{1}{\lambda} + \frac{-\phi(1-\lambda)+(1-\phi\lambda)}{(1-\phi\lambda)(1-\lambda)}\right] = s\left[\frac{1}{\lambda} + \frac{1-\phi}{(1-\phi\lambda)(1-\lambda)}\right] = \\ &= s\left[\frac{(1-\phi\lambda)(1-\lambda)+(1-\phi)\lambda}{(1-\phi\lambda)(1-\lambda)\lambda}\right] = s\frac{1-2\phi\lambda+\phi\lambda^2}{(1-\phi\lambda)(1-\lambda)\lambda} \end{aligned} \tag{74}$$

This function is always positive because the numerator is a convex parabola in λ, and the minimum is reached at $\lambda = 2\phi/(2\phi) = 1$, where the numerator is equal to $1 - 2\phi + \phi = 1 - \phi > 0$. Note that $\lim_{\lambda\to 0} \lambda\frac{1-\phi\lambda}{1-\lambda} = 0$, and $\lim_{\lambda\to 1} \lambda\frac{1-\phi\lambda}{1-\lambda} = +\infty$, so that the function $\lim_{\lambda\to 0} e_o(\lambda) = -\infty$ and $\lim_{\lambda\to 1} e_o(\lambda) = +\infty$. ■

Recall that $e_o(\lambda^{cc})$ is the option value of a remote worker minus the option value of a commuting worker. Lemma B.2 says that: 1) as the equilibrium number of commuters grows, the relative option value of being a remote worker grows; 2) as we approach a purely remote equilibrium ($\lambda^{cc} \to 0$), switching from c to m implies a loss in option values that approaches $+\infty$; as we approach a purely commuting equilibrium ($\lambda^{cc} \to 1$), switching from c to m implies a gain in the option values that approaches $+\infty$.

Proposition B.1 *There exists at least one stationary equilibrium $L_{c,ss}$.*
Proof. *Consider the equation $f(\lambda) = 0$ characterizing the stationary equilibria from (59). $f(\lambda)$ is continuous over $\lambda \in (0,1)$. We aim at showing that $\lim_{\lambda\to 0} f(\lambda) > 0$ and $\lim_{\lambda\to 1} f(\lambda) < 0$, so that $f(\lambda) = 0$ (and there exists at least one stationary equilibrium) for some $\lambda \in (0,1)$ by the intermediate value theorem.*

Consider the behavior of

$$\bar{\varepsilon}^{cm}(\lambda^{cc}) \equiv \frac{\beta}{1-\beta}\left[e_f(\lambda^{cc}) + e_o(\lambda^{cc}) - (F_{mm} - F_{cc})\right] - F_{cm} + F_{cc}$$

from (60). The difference in the option values $e_o(\lambda^{cc})$ behaves as $\lim_{\lambda\longrightarrow 0} e_o(\lambda^{cc}) = -\infty$ and $\lim_{\lambda\longrightarrow 1} e_o(\lambda^{cc}) = +\infty$ from Lemma B.2.

Using the functional forms in Assumption 1 in the indirect utility functions (48) and (49), we can write

$$\begin{aligned} e_f(\lambda^{cc}) &\equiv u_m(\lambda^{cc}) - u_c(\lambda^{cc}) = \mu\left(\log T - \log\frac{\bar{B}}{\bar{\tau}} - (\delta+\xi-\theta-\gamma)\log L + \log\frac{Z}{A}\right) \\ &- \mu^2(\delta+\xi-\theta-\gamma/\mu)\log\sigma(\lambda^{cc}) \end{aligned} \tag{75}$$

As $\lambda \to 1$, this difference approaches a finite number. When $\lambda \to 0$, this difference approaches $-\infty$ if $\delta+\xi < \theta+\gamma/\mu$, and $+\infty$ if $\delta+\xi > \theta+\gamma/\mu$,

These observations imply that $\lim_{\lambda^{cc}\longrightarrow 1}\bar{\varepsilon}^{cm}\left(\lambda^{cc}\right)=+\infty$. *Moreover, Assumption 2 guarantees that* $\lim_{\lambda^{cc}\to 0}\bar{\varepsilon}^{cm}\left(\lambda^{cc}\right)=-\infty$. *It follows that* $\lim_{\lambda\longrightarrow 0}f\left(\lambda\right)=1$ *and* $\lim_{\lambda\longrightarrow 1}f\left(\lambda\right)=-1$. *Hence, eq. (59) always has at least one solution.* ■

B.3 Sufficient conditions for a unique stationary equilibrium

In this subsection, we show sufficient conditions for the stationary equilibrium to be unique and how this stationary equilibrium changes with some of the model's parameters. The proof of this result in Proposition B.2 is based on establishing conditions for the monotonicity of $f\left(\lambda\right)$ in eq. (59). Lemma B.3 below preliminarily establishes the monotonicity of L_c in the share of commuters that remain commuters, λ^{cc}.

Lemma B.3 *The function* $\sigma\left(\lambda\right)$ *in eq. (63) is monotonically increasing from 0 to 1 when* $\lambda\in[0,1]$.

Proof. *Computing the derivative,*

$$\begin{aligned}\sigma'\left(\lambda\right) &= \frac{\left(1-\phi\right)\left[\left(1-\lambda\right)\left(1-\phi\lambda\right)+\left(1-\phi\right)\lambda\right]-\left(1-\phi\right)\lambda\left[\left(1-\phi\right)-\left(1-\phi\lambda\right)-\phi\left(1-\lambda\right)\right]}{\left[\left(1-\lambda\right)\left(1-\phi\lambda\right)+\left(1-\phi\right)\lambda\right]^{2}}= \\ &= \left(1-\phi\right)\frac{\left(1-\lambda\right)\left(1-\phi\lambda\right)+\left(1-\phi\right)\lambda-\lambda\left(1-\phi\right)+\lambda\left(1-\phi\lambda\right)+\phi\lambda\left(1-\lambda\right)}{\left[\left(1-\lambda\right)\left(1-\phi\lambda\right)+\left(1-\phi\right)\lambda\right]^{2}}= \\ &= \frac{\left(1-\phi\right)\left(1-\lambda^{2}\phi\right)}{\left[\left(1-\lambda\right)\left(1-\phi\lambda\right)+\left(1-\phi\right)\lambda\right]^{2}} \qquad (76)\end{aligned}$$

which is always positive. ■

Proposition B.2 *Assume* $\delta+\xi<\theta+\gamma/\mu$. *Then there is a unique stationary equilibrium value of the commuting population* $L_{c,ss}$. *This stationary equilibrium value monotonically increases with* A *and* $\bar{B}$, *and monotonically falls with* Z, T *and* $\bar{\tau}$.

Proof. *Recall eq. (75), stating that*

$$\begin{aligned}e_{f}\left(\lambda^{cc}\right) &\equiv u_{m}\left(\lambda^{cc}\right)-u_{c}\left(\lambda^{cc}\right)=\mu\left(\log T-\log\frac{\bar{B}}{\bar{\tau}}-\left(\delta+\xi-\theta-\gamma\right)\log L+\log\frac{Z}{A}\right) \\ &- \mu^{2}(\delta+\xi-\theta-\gamma/\mu)\log\sigma(\lambda^{cc}) \qquad (77)\end{aligned}$$

When $\delta+\xi<\theta+\gamma/\mu$, *this difference is increasing in* λ^{cc} *from Lemma B.3 and it follows that* $\bar{\varepsilon}^{cm}\left(\lambda^{cc}\right)$ *in eq. (60) is monotonically increasing in* λ^{cc}, *mapping the interval* $(0,1)$ *into* $(-\infty,+\infty)$. *Hence, the function* $f\left(\lambda^{cc}\right)$ *in (59) is monotonically decreasing in* λ^{cc} *and maps the interval* $(0,1)$ *into itself. Equation (59) has then a unique solution for* λ^{cc}, *which maps into a unique value for* $L_c=\sigma\left(\lambda_{cc}\right)L$.

Note also that A, Z, T, $\bar{B}$ *and* $\bar{\tau}$ *only affect the value of* $\bar{\varepsilon}^{cm}\left(\lambda^{cc}\right)$ *in (60) through* u_m-u_c; *in particular,* $\bar{\varepsilon}^{cm}\left(\lambda^{cc};A,\bar{B},Z,T,\tau\right)$ *grows with* Z, T, *and* $\bar{\tau}$ *while falling with* A *and* $\bar{B}$ *for any*

given λ^{cc}. Hence, $f(\lambda^{cc})$ in (59) shifts uniformly down with Z, T, $\bar{\tau}$ and up with A and $\bar{B}$ in its domain $\lambda^{cc} \in (0,1)$. This implies that the solution to (59) and $L_c = \sigma(\lambda_{cc}) L$ also moves in the same direction with these parameters. ■

B.4 Necessary and sufficient conditions for multiple stationary equilibria

In Proposition B.1 we have shown that there is always at least one stationary equilibrium because $f(\lambda)$ starts positive, ends negative, and is continuous. In Proposition B.3, we show that for any level of (sufficiently strong) agglomeration forces, there is a range of values z/A where the economy exhibits multiple stationary equilibria. Proposition B.3 makes use of the following four lemmas. Lemma B.4 characterizes necessary and sufficient conditions for multiplicity in terms of the function $f(\lambda)$ in eq. (59): multiple stationary equilibria exist if $f'(\lambda) > 0$ at a stationary equilibrium, which can be restated into a condition on $\partial \bar{\varepsilon}^{cm}(\lambda)/\partial\lambda$, the slope of the expected value of switching from c to m with respect to λ. Lemma B.5 considers such slope and it shows that, at a candidate $\hat{\lambda}$, one can always make $\partial \bar{\varepsilon}^{cm}(\lambda)/\partial\lambda$ arbitrarily negative by choosing an externality term δ above a threshold δ^*. This fact is useful because the slope of $\bar{\varepsilon}^{cm}$ can be used to control the slope of $f(\lambda)$. Lemma B.6 characterizes some properties of this threshold δ^*. Lemma B.7 shows that for any level of δ, we can always find a ratio of the exogenous remote work to in-office productivity, z/A, that makes an arbitrary $\hat{\lambda}$ a stationary equilibrium. Proposition B.3 uses these lemmas to show that for a high enough δ, we can always find a z/A that makes some $\hat{\lambda}$ a stationary equilibrium where $f'(\hat{\lambda}) > 0$.

Lemma B.4 *Consider an economy Ω. The economy has multiple stationary equilibria if and only if there exist one $\hat{\lambda}$ such that*

$$f(\hat{\lambda}) = 0 \tag{78}$$

$$\left.\frac{\partial f(\lambda)}{\partial \lambda}\right|_{\lambda=\hat{\lambda}} > 0 \tag{79}$$

Moreover, if such $\hat{\lambda}$ exists, then it holds true that:

$$\left.\frac{\partial \bar{\varepsilon}^{cm}(\lambda)}{\partial \lambda}\right|_{\lambda=\hat{\lambda}} < -\frac{s}{\hat{\lambda}\left(1-\hat{\lambda}\right)} \tag{80}$$

Proof. From Lemma B.1, a stationary equilibrium is characterized by $f(\hat{\lambda}) = 0$. Suppose

$$\left.\frac{\partial f(\lambda)}{\partial \lambda}\right|_{\lambda=\hat{\lambda}} > 0 \tag{81}$$

Since $f(\lambda)$ always starts at $+1$ and ends at -1, $f(\lambda)$ will cross zero at least two more times, one

to the left and one to the right of $\hat{\lambda}$. In other words, the equation $f(\lambda)$ has more than one solution, that is, the economy exhibits multiple stationary equilibria.

Note that, using the definition of f in eq. (59),

$$\begin{aligned} \frac{\partial f(\lambda)}{\partial \lambda} &= -\frac{\exp\left(\bar{\varepsilon}^{cm}(\lambda)/s\right)}{\left(1+\exp\left(\bar{\varepsilon}^{cm}(\lambda)/s\right)\right)^2}\frac{1}{s}\frac{\partial \bar{\varepsilon}^{cm}(\lambda)}{\partial \lambda} - 1 > 0 \\ \frac{\partial \bar{\varepsilon}^{cm}(\lambda)}{\partial \lambda} &< -s\frac{\left(1+\exp\left(\bar{\varepsilon}^{cm}(\lambda)/s\right)\right)^2}{\exp\left(\bar{\varepsilon}^{cm}(\lambda)/s\right)} \end{aligned} \tag{82}$$

Additionally, using the same definition, at a stationary equilibrium it holds true that

$$f\left(\hat{\lambda}\right) \equiv \frac{1}{1+\exp\left(\bar{\varepsilon}^{cm}\left(\hat{\lambda}\right)/s\right)} - \hat{\lambda} = 0 \implies \tag{83}$$

$$1+\exp\left(\bar{\varepsilon}^{cm}\left(\hat{\lambda}\right)/s\right) = 1/\hat{\lambda} \tag{84}$$

$$\exp\left(\bar{\varepsilon}^{cm}\left(\hat{\lambda}\right)/s\right) = \left(1-\hat{\lambda}\right)/\hat{\lambda} \tag{85}$$

Hence, using equations (84) and (85) in eq. (82), we have

$$\left.\frac{\partial \bar{\varepsilon}^{cm}(\lambda)}{\partial \lambda}\right|_{\lambda=\hat{\lambda}} < -\frac{s}{\hat{\lambda}\left(1-\hat{\lambda}\right)} \tag{86}$$

■

Lemma B.5 $\forall \hat{\lambda} \in (0,1), K > 0, \exists\, \delta^*\left(\hat{\lambda}, K\right) > 0$:

$$\left.\frac{\partial \bar{\varepsilon}^{cm}(\lambda,\delta)}{\partial \lambda}\right|_{\lambda=\hat{\lambda}} < -K \iff \delta + \xi > \delta^*\left(\hat{\lambda}, K\right)$$

Proof. Consider the expression for $\bar{\varepsilon}^{cm}(\lambda)$ from (60):

$$\bar{\varepsilon}^{cm}(\lambda) \equiv \frac{\beta}{1-\beta}\left[e_f(\lambda) + e_o(\lambda) - (F_{mm} - F_{cc})\right] - F_{cm} + F_{cc}$$

The first derivative of this function with respect to λ is

$$\frac{\partial \bar{\varepsilon}^{cm}(\lambda)}{\partial \lambda} = \frac{\beta}{1-\beta}\left[e_f'(\lambda) + e_o'(\lambda)\right]$$

We know that $e_o'(\lambda) > 0$. From eq. (75), we can easily compute e_f' so that our condition becomes

$$\begin{aligned} \left.\frac{\partial \bar{\varepsilon}^{cm}(\lambda)}{\partial \lambda}\right|_{\lambda=\hat{\lambda}} &< -K \iff e_f'(\lambda)\big|_{\lambda=\hat{\lambda}} < -e_o'(\lambda)|_{\lambda=\hat{\lambda}} - \frac{1-\beta}{\beta}K \iff \\ \delta + \xi &> \frac{1}{\mu^2}\left[\frac{\sigma(\lambda)}{\sigma'(\lambda)}\left(e_o'(\lambda) + \frac{1-\beta}{\beta}K\right)\right]_{\lambda=\hat{\lambda}} + \frac{\gamma}{\mu} + \theta \equiv \delta^*\left(\hat{\lambda}, K\right) \end{aligned} \tag{87}$$

■

Lemma B.6 *Define*

$$\hat{\delta}(\lambda) \equiv \frac{1}{\mu^2}\left[\frac{\sigma(\lambda)}{\sigma'(\lambda)}\left(e_o'(\lambda) + \frac{1-\beta}{\beta}\frac{s}{\lambda(1-\lambda)}\right)\right] + \frac{\gamma}{\mu} + \theta \tag{88}$$

as the threshold function $\delta^*(\lambda, K)$ for $K = \frac{s}{\lambda(1-\lambda)}$ in eq. (87). Then, $\hat{\delta}(\lambda)$ reaches a minimum value of $\eta_{min} > \gamma/\mu + \theta$. If $\phi \in (0,1)$, then $\hat{\delta}(\lambda)$ is convex in λ.

Proof. The function $\hat{\delta}(\lambda)$ is defined for $\hat{\lambda} \in (0,1)$, is continuous, and differentiable. Since the term in the square brackets is always positive, the function is also always positive. To further characterize $\hat{\delta}(\lambda)$, we study its limiting properties. To compute $\lim_{\hat{\delta}\to 0}\hat{\delta}(\hat{\lambda})$, we substitute the expressions for $\sigma(\lambda)$, $e_o'(\lambda)$ and $\sigma'(\lambda)$ from equations (72), (74), and (76) and obtain

$$\lim_{\hat{\lambda}\to 0}\hat{\delta}(\hat{\lambda}) = \frac{s}{\mu^2\beta} + \frac{\gamma}{\mu} + \theta \equiv \hat{\delta}_0 > 0$$

Further, $\lim_{\hat{\delta}\to 1}\hat{\delta}(\hat{\lambda}) = +\infty$ since, from (72) and (76), $\lim_{\hat{\lambda}\to 1}\sigma\left(\hat{\lambda}\right) = 1$ and $\lim_{\hat{\lambda}\to 1}\sigma'\left(\hat{\lambda}\right) = 1$; and from Lemma B.2, $\lim_{\hat{\lambda}\to 1} e_o'\left(\hat{\lambda}\right) = +\infty$.
The first derivative of $\hat{\delta}(\lambda)$, after some algebra, is given by:

$$\begin{aligned} \hat{\delta}'(\lambda) = &- \frac{s}{\mu^2}\frac{4\phi(1-2\phi\lambda+\phi\lambda^2)}{(1-\phi\lambda^2)(1-\phi\lambda)} \\ &+ \frac{s}{\mu^2}\frac{(1-2\phi\lambda+\phi\lambda^2)^2(1-3\phi^2\lambda^2-3\phi\lambda^2+4\phi^2\lambda^3+\phi)}{(1-\phi\lambda^2)^2(1-\phi\lambda)^2(1-\lambda)^2} \\ &+ \frac{s}{\mu^2}\frac{1-\beta}{\beta}\frac{1+2\phi(2\lambda-2\lambda^2-1)-\phi^2\lambda^2(2-4\lambda+\lambda^2)}{(1-\phi\lambda^2)^2(1-\lambda)^2} \end{aligned} \tag{89}$$

Moreover,

$$\lim_{\hat{\lambda}\to 0}\hat{\delta}'(\lambda) < 0 \iff \phi > \frac{1}{2+\beta} \tag{90}$$

Hence, the function starts sloping down if ϕ is large enough. Assume that $\hat{\delta}(\lambda)$ is convex, which

we will prove right below. In this case, the (unique) minimum is achieved for a $\eta_{min} : \gamma/\mu + \theta < \eta_{min} < \hat{\delta}_0$. If the function starts sloping up, then $\eta_{min} \equiv \hat{\delta}_0$.

It remains to be proven that $\hat{\delta}(\lambda)$ is convex. To this end, it is thus sufficient to prove the convexity of the simplified function $\hat{\delta}(\lambda)_{simple}$, which takes out the constant term $\frac{\gamma}{\mu} + \theta$ and the positive multiplicative term $\frac{s}{\mu^2}$:

$$\begin{aligned}
\hat{\delta}(\lambda)_{simple} &= \frac{\lambda\,(1 - 2\phi\lambda + \phi\lambda^2)}{1 - \phi\lambda^2}\left[\frac{1 - 2\phi\lambda + \phi\lambda^2}{(1-\phi\lambda)(1-\lambda)\lambda} + \frac{1-\beta}{\beta}\frac{1}{\lambda(1-\lambda)}\right] \\
&= \frac{(1 - 2\phi\lambda + \phi\lambda^2)}{(1 - \phi\lambda^2)(1-\lambda)}\left[\frac{1 - 2\phi\lambda + \phi\lambda^2}{(1-\phi\lambda)} + \frac{1-\beta}{\beta}\right] \\
&= \underbrace{\frac{(1 - 2\phi\lambda + \phi\lambda^2)^2}{(1 - \phi\lambda^2)(1-\lambda)(1-\phi\lambda)}}_{\text{first term } \Xi_1} + \frac{1-\beta}{\beta}\underbrace{\frac{(1 - 2\phi\lambda + \phi\lambda^2)}{(1 - \phi\lambda^2)(1-\lambda)}}_{\text{second term } \Xi_2}.
\end{aligned}$$

We argue that each of the terms Ξ_1 and Ξ_2 is convex, which will prove the convexity of their sum.

The second derivative of Ξ_1 is:

$$\frac{d^2\Xi_1}{d\lambda^2} = \frac{2}{(1-\lambda)^3} + \phi^2\left(\frac{2}{(1-\phi\lambda)^3} - \frac{8x(\phi\lambda^2+3)}{(1-\phi\lambda^2)^3}\right)$$

We prove by contradiction that $\frac{d^2\Xi_1}{d\lambda^2}$ cannot be negative. Recall that the range of values for ϕ, λ is $\phi \in (0,1), \lambda \in (0,1)$. The first term of the expression $\frac{2}{(1-\lambda)^3}$, is always positive. If $\frac{d^2\Xi_1}{d\lambda^2}$ is negative at any value of λ, then the term $\left(\frac{2}{(1-\phi\lambda)^3} - \frac{8x(\phi\lambda^2+3)}{(1-\phi\lambda^2)^3}\right)$ must be negative. If this is so, then:

$$\begin{aligned}
\frac{d^2\Xi_1}{d\lambda^2} &= \frac{2}{(1-\lambda)^3} + \phi^2\left(\frac{2}{(1-\phi\lambda)^3} - \frac{8x(\phi\lambda^2+3)}{(1-\phi\lambda^2)^3}\right) \\
&\geq \frac{2}{(1-\lambda)^3} + \left(\frac{2}{(1-\phi\lambda)^3} - \frac{8x(\phi\lambda^2+3)}{(1-\phi\lambda^2)^3}\right) \\
&\geq \frac{2}{(1-\lambda)^3} + \left(\frac{2}{(1-\lambda)^3} - \frac{8x(\phi\lambda^2+3)}{(1-\phi\lambda^2)^3}\right) \\
&\geq \frac{2}{(1-\lambda)^3} + \left(\frac{2}{(1-\lambda)^3} - \frac{8x(\lambda^2+3)}{(1-\lambda^2)^3}\right) \\
&= \frac{4}{(1+\lambda)^3} \\
&\geq 0
\end{aligned}$$

The fourth line follows because $\frac{d}{da} = -\frac{8x(\phi\lambda^2+3)}{(1-\phi\lambda^2)^3} = -\frac{16\lambda^3(5+\phi\lambda^2)}{(-1+\phi\lambda^2)^4} \leq 0$, so the term is most

negative at $\phi = 1$. We have shown that $\frac{d^2\Xi_1}{d\lambda^2}$ is weakly positive. Thus, Ξ_1 is convex.

We now move on to the second term term. The second derivative of Ξ_2 is:

$$\frac{d^2\Xi_2}{d\lambda^2} = \frac{2(1-\phi\lambda^2)^3 - 4\phi^2\lambda(\phi\lambda^2+3)(1-\lambda)^3}{(1-\lambda)^3(1-\phi\lambda^2)^3}$$

The denominator is obviously positive for $\phi \in (0,1), \lambda \in (0,1)$. We want to prove that the numerator is positive. Let us observe so the numerator, which we denote as $N_2(\lambda)$:

$$N_2(\lambda) = 2(1-\phi\lambda^2)^3 - 4\phi^2\lambda(\phi\lambda^2+3)(1-\lambda)^3$$

By inspection, $N_2(\lambda)$ is decreasing in ϕ for $\phi \in (0,1)$. As such we only need to prove that $N_2(\lambda)$ is positive for $\phi = 1$ to show that $N_2(\lambda)$ is always positive.

$$\begin{aligned} N_2(\lambda) &= 2(1-\phi\lambda^2)^3 - 4\phi^2\lambda(\phi\lambda^2+3)(1-\lambda)^3 \\ &\geq 2(1-\lambda^2)^3 - 4\lambda(\lambda^2+3)(1-\lambda)^3 \\ &= 2(1-\lambda)^6 \\ &\geq 0 \text{ for } \lambda \in (0,1) \end{aligned}$$

We have shown that both the numerator and denominator of the second derivative of Ξ_2 are positive. As such, Ξ_2 is convex.

We have shown that Ξ_1 and Ξ_2 are convex, their sum $\hat{\delta}(\lambda)_{simple}$ is also convex, which in turn implies that our original function $\hat{\delta}(\lambda)$ is convex.

■

Lemma B.7 *For each $\hat{\lambda}$ and δ, there exists a ratio z/A such that $f\left(\hat{\lambda}; \delta, z/A\right) = 0$. This ratio is described by the function*

$$\begin{aligned} \rho(\hat{\lambda}) &= \exp\left\{\frac{1}{\mu}\left[\frac{1-\beta}{\beta}(F_{cm} - F_{cc}) + F_{mm} - F_{cc}\right]\right\} \times \frac{\bar{B}}{\bar{\tau}T} \\ &\times \frac{\sigma(\hat{\lambda})^{\mu(\xi+\delta-\theta)-\gamma}}{(1-\phi\hat{\lambda})^{s/\mu}} \times \left(\frac{1-\hat{\lambda}}{\hat{\lambda}}\right)^{s/(\beta\mu)} \times L^{\xi+\delta-\theta-\gamma} \end{aligned} \tag{91}$$

Proof. Using the definition of $\bar{\varepsilon}^{cm}\left(\hat{\lambda}\right)$ from eq. (60) in the condition for a stationary equilibrium from (59),

$$\hat{\lambda} = \frac{1}{1+\exp\left(\bar{\varepsilon}^{cm}\left(\hat{\lambda}\right)/s\right)} \implies s\ln\frac{1-\hat{\lambda}}{\hat{\lambda}} = \bar{\varepsilon}^{cm}\left(\hat{\lambda}\right)$$

$$s\ln\frac{1-\hat{\lambda}}{\hat{\lambda}} = \frac{\beta}{1-\beta}\left[e_f\left(\hat{\lambda}\right) + e_o\left(\hat{\lambda}\right) - (F_{mm}-F_{cc})\right] - F_{cm} + F_{cc}$$

Using $e_f \equiv u_m - u_c$ and the result in eq. (75), after some manipulation, we obtain

$$\begin{aligned}\frac{Z}{A} &= \exp\left\{\frac{1}{\mu}\left[\frac{1-\beta}{\beta}(F_{cm}-F_{cc}) + F_{mm} - F_{cc}\right]\right\} \times \frac{\bar{B}}{\bar{\tau}T} \\ &\times \frac{\sigma(\hat{\lambda})^{\mu(\xi+\delta-\theta)-\gamma}}{(1-\phi\hat{\lambda})^{s/\mu}} \times \left(\frac{1-\hat{\lambda}}{\hat{\lambda}}\right)^{s/(\beta\mu)} \times L^{\xi+\delta-\theta-\gamma} \equiv \rho(\hat{\lambda}) \end{aligned} \tag{92}$$

L_c is a function of $\hat{\lambda}$ and is bounded between 0 and L. One can verify using eq. (61) that the sum in the curly bracket $-e_o\left(\hat{\lambda}\right) + \frac{1-\beta}{\beta}s\ln\frac{1-\hat{\lambda}}{\hat{\lambda}}$ goes to $+\infty$ for $\hat{\lambda}\to 0$ and to $-\infty$ for $\hat{\lambda}\to 1$. Hence, the image of $\hat{z}$ spans $(0,+\infty)$ as $\hat{\lambda}$ varies in $(0,1)$.

■

Proposition B.3 *Assume $F_{cc} = F_{mm} = 0$. Then, there exist a finite threshold $\eta_{min} > \theta + \gamma/\mu$, and a set $\mathcal{Z} \subset \mathbb{R}_{++}$, such that:*

i) for $\delta + \xi > \eta_{min}$, $\mathcal{Z}$ is a non-empty interval (Z_{min}, Z_{max}) and there are multiple stationary equilibria if $z/A \in \mathcal{Z}$; further, Z_{min} and Z_{max} grow with L and $\bar{B}$ and fall with T and $\bar{\tau}$.

ii) there is a unique stationary equilibrium in all other cases.

Proof. From Lemma B.4, we know that equations (78) and (80) characterize an economy with multiple stationary equilibria. We seek necessary and sufficient conditions for those equations to hold.

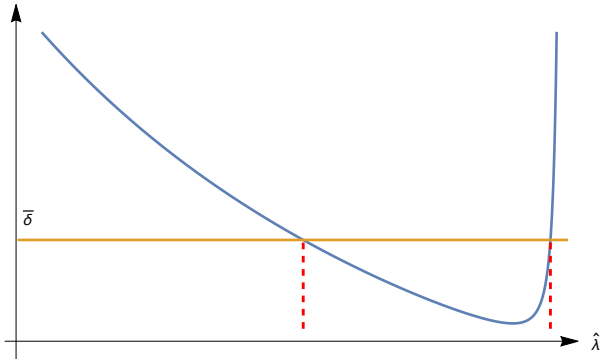
Consider the equation $\hat{\delta}(\hat{\lambda})$ in Lemma B.6. From Lemmas B.4 and B.5, $f'(\hat{\lambda}) > 0$ if and only if we pick a $\bar{\delta} : \bar{\delta} + \xi \geq \hat{\delta}(\hat{\lambda})$. Note that $\hat{\delta}(\hat{\lambda})$ is defined for $\hat{\lambda} \in (0,1)$, is continuous, and differentiable. Also, from Lemma B.6, it is convex in $\hat{\lambda}$ and it reaches a global minimum value of $\hat{\delta}_{min}$. An example of this function is plotted in Panel (a) of Figure A10.

Our proof proceeds with a taxonomic discussion of the implications of different values of $\bar{\delta}$. This discussion can be organized around features of the set

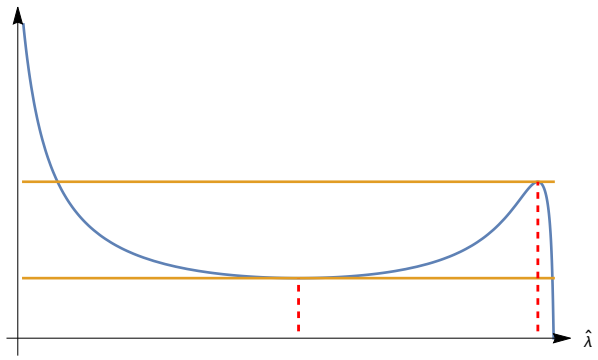
$$\Lambda(\bar{\delta}) \equiv \{\hat{\lambda} : \hat{\delta}\left(\hat{\lambda}\right) \leq \bar{\delta} + \xi\} \tag{93}$$

Note that by construction, for each $\hat{\lambda} \in \Lambda(\bar{\delta})$, eq. (80) is satisfied if $f(\hat{\lambda}, \bar{\delta}, z/A) = 0$.

Figure A10: Identifying the Range of Parameters for Multiple Stationary Equilibria



(a) The function $\hat{\delta}(\hat{\lambda})$



(b) The function $\rho(\hat{\lambda})$

Consider part *i)* of the proposition, and fix a $\bar{\delta} : \bar{\delta} + \xi > \hat{\delta}_{min} > \gamma/\mu + \theta$. Then, the set $\Lambda(\bar{\delta})$ is non-empty; moreover, since $\hat{\delta}$ is convex from Lemma B.6, the set $\Lambda(\bar{\delta})$ is an interval of the form $\Lambda(\bar{\delta}) = (\lambda_{min}(\bar{\delta}), \lambda_{max}(\bar{\delta}))$. Consider then the set

$$\mathcal{Z}(\bar{\delta}) = \{z/A : z/A = \rho\left(\hat{\lambda}\right) \text{ for some } \hat{\lambda} \in \Lambda(\bar{\delta})\} \tag{94}$$

The set $\mathcal{Z}$ is the image of the set Λ in z/A (see Panel (b) of Figure A10) By construction, for any z/A in this set, there is some $\hat{\lambda}$ for which $\hat{\lambda}, z/A$ is a stationary equilibrium, that is, eq. (78) holds. Since at the same $\hat{\lambda}$ eq. (80) also holds, the economy exhibits multiple stationary equilibria. Note also that since $\rho(\lambda)$ is continuous and Λ is an interval, $\mathcal{Z}(\bar{\delta})$ is also an interval and can be written as

$$\begin{aligned}
\mathcal{Z}(\bar{\delta}) &= \left(Z_{min}(\bar{\delta}), Z_{max}(\bar{\delta})\right) \\
Z_{min}(\bar{\delta}) &= min_{\lambda \in \Lambda(\bar{\delta})}\ \rho(\lambda) \\
Z_{max}(\bar{\delta}) &= max_{\lambda \in \Lambda(\bar{\delta})}\ \rho(\lambda)
\end{aligned}$$

To show how $Z_{min}(\bar{\delta}), Z_{max}(\bar{\delta})$ vary with L, we consider the expression for $\rho(\lambda)$ in eq. (92). If λ^- is the minimizer when the total city size is L, then λ^- is also the minimizer when the total city size is kL, for $k > 0$; additionally, $\rho(\lambda, kL) = k^{\bar{\delta}+\xi-\gamma-\theta}\rho(\lambda, L)$. Since $\bar{\delta} + \xi - \gamma - \theta > 0$ is satisfied whenever $\mathcal{Z}$ is non-empty, then Z_{min} grows with L. A similar argument shows that Z_{max} also grows with L. A similar argument can be made to show that these boundaries grow with $\bar{B}$ and fall with $\bar{\tau}$ and T.

To prove part *ii)*, note that if $\bar{\delta} + \xi > \hat{\delta}_{min}$ but $z/A \notin \mathcal{Z}(\bar{\delta})$, then by construction none of the $\hat{\lambda} \in \Lambda(\bar{\delta})$ are a stationary equilibrium where (80) holds; however, since a stationary equilibrium always exists, the equilibrium must be unique. Similarly, if $\bar{\delta} + \xi < \hat{\delta}_{min}$, then the set $\Lambda(\bar{\delta})$ in eq. (93) is empty; hence, there is no $\hat{\lambda}$ for which eq. (78) holds at a candidate stationary equilibrium.

Again, since a stationary equilibrium always exists, it must be unique.

■

Proposition 3.3 then follows from the fact that L_c is monotone increasing in λ^{cc} by Lemma B.3.

B.5 Relative Transition Shares

Lemma B.8 *Suppose $F_{cc} = F_{mm} = 0$ and that $F_{cm} = F_{mc} = F$. Then*

$$y_{\ell\ell',t} \equiv \ln \frac{\lambda_{\ell\ell',t}\lambda^{\beta}_{\ell'\ell',t+1}}{\lambda_{\ell\ell,t}\lambda^{\beta}_{\ell\ell',t+1}} = \frac{\beta}{s}\left(u_{\ell',t+1} - u_{\ell,t+1}\right) - \frac{(1-\beta)F}{s}. \tag{95}$$

Proof.

Combining (16) and (18), the share of workers in state ℓ transitioning to ℓ' is given by

$$\lambda_{\ell\ell',t} = \frac{\exp\left(\beta V_{\ell',t+1} - F_{\ell\ell',t}\right)^{1/s}}{\exp\left(\beta V_{c,t+1} - F_{\ell c,t}\right)^{1/s} + \exp\left(\beta V_{m,t+1} - F_{\ell m,t}\right)^{1/s}}, \tag{96}$$

where we denote $V_{\ell,t+1} = V_\ell\left(L_{c,t+1}; \omega^{t+1}\right)$. We first obtain an expression of $\lambda_{\ell\ell',t}$ as a function of $V_{\ell,t}$, $V_{\ell',t+1}$ and transition costs only by re-writing (15) with time indices only,

$$V_{\ell,t} = u_{\ell,t} + \beta V_{\ell,t+1} - F_{\ell\ell,t} + \Omega_{\ell,t}, \tag{97}$$

and re-writing (17) after replacing the transition share by its expression in (96):

$$\Omega_{\ell,t} = -s \ln\left(\frac{\exp\left(\beta V_{\ell,t+1} - F_{\ell\ell,t}\right)^{1/s}}{\exp\left(\beta V_{c,t+1} - F_{\ell c,t}\right)^{1/s} + \exp\left(\beta V_{m,t+1} - F_{\ell m,t}\right)^{1/s}}\right). \tag{98}$$

Substituting for $\Omega_{\ell,t}$ in (97), we get the following expression for the value of labor delivery mode ℓ:

$$V_{\ell,t} = u_{\ell,t} + s \ln\left(\exp\left(\beta V_{c,t+1} - F_{\ell c,t}\right)^{1/s} + \exp\left(\beta V_{m,t+1} - F_{\ell m,t}\right)^{1/s}\right), \tag{99}$$

or equivalently,

$$\exp\left(\beta V_{c,t+1} - F_{\ell c,t}\right)^{1/s} + \exp\left(\beta V_{m,t+1} - F_{\ell m,t}\right)^{1/s} = \exp\left(V_{\ell,t} - u_{\ell,t}\right)^{1/s}. \tag{100}$$

Recognizing the left hand side as the denominator in (96), we obtain the following expressions for the share of workers in state ℓ transitioning to ℓ' at t, in state ℓ transitioning to ℓ at t, in state ℓ' transitioning to ℓ' at $t+1$ and in state ℓ transitioning to ℓ' at $t+1$, respectively, as:

$$\lambda_{\ell\ell',t} = \exp\left(\beta V_{\ell',t+1} - F_{\ell\ell',t} + u_{\ell,t} - V_{\ell,t}\right)^{1/s}. \tag{101}$$

$$\lambda_{\ell\ell,t} = \exp\left(\beta V_{\ell,t+1} - F_{\ell\ell,t} + u_{\ell,t} - V_{\ell,t}\right)^{1/s}, \tag{102}$$

$$\lambda_{\ell'\ell',t+1} = \exp\left(\beta V_{\ell',t+2} - F_{\ell'\ell',t+1} + u_{\ell',t+1} - V_{\ell',t+1}\right)^{1/s}, \tag{103}$$

$$\lambda_{\ell\ell',t+1} = \exp\left(\beta V_{\ell',t+2} - F_{\ell\ell',t+1} + u_{\ell,t+1} - V_{\ell,t+1}\right)^{1/s}. \tag{104}$$

We get the following expression for the ratio of the discounted probabilities:

$$\frac{\lambda_{\ell\ell',t}\left(\lambda_{\ell'\ell',t+1}\right)^{\beta}}{\lambda_{\ell\ell,t}\left(\lambda_{\ell\ell',t+1}\right)^{\beta}} = \exp\left(\beta\left(u_{\ell',t+1} - u_{\ell,t+1}\right) + F_{\ell\ell,t} - F_{\ell\ell',t} + \beta\left(F_{\ell\ell',t+1} - F_{\ell'\ell',t+1}\right)\right)^{1/s}. \tag{105}$$

Using the assumption imposed in Proposition 3.3 that $F_{\ell\ell,t} = 0$ for all t and $\ell \in \{m, c\}$, and that $F_{lm} = F_{ml} = F$, the logarithm of this ratio can be written as

$$\ln \frac{\lambda_{\ell\ell',t}\left(\lambda_{\ell'\ell',t+1}\right)^{\beta}}{\lambda_{\ell\ell,t}\left(\lambda_{\ell\ell',t+1}\right)^{\beta}} = \frac{\beta}{s}\left(u_{\ell',t+1} - u_{\ell,t+1}\right) - \frac{1-\beta}{s}F. \tag{106}$$

■

B.6 Testing for multiple equilibria

It is in principle possible for models like ours to exhibit multiple equilibria, not only multiple stationary equilibria. For example, a city could start in an initial condition with low commuting, but individual expectations of a return to high commuting might put the system on a self-fulfilling path. In this subsection, we describe how we test – and find no evidence of – this possibility.

In particular, let us examine first a case where, starting from an initial condition that we expect to converge toward the high-commuting equilibrium, the system might actually converge toward the low commuting one. For cities with multiple stationary equilibria, we have always found 2 stable and one unstable stationary equilibrium. We therefore start by setting an initial condition for $L_{c,0}$ above the unstable equilibrium. The top panels of Figure A11 set $L_{c,0}/L = 0.15$ for the metro area of New York, which is above the unstable stationary equilibrium (dashed red line).

In Panel (a), we set as initial guesses for the time path of the value functions V_c and V_m their respective value in the high-commuting stationary equilibrium. We require, as a convergence criterion, both that the sup-norm relative distance between successive guesses of the value functions is less than a given threshold and that the terminal value of these value functions is within the same relative tolerance from the high-commuting stationary equilibrium value. We set this threshold to 10^{-6}. The panel shows successive time paths of L_c/L at the indicated iteration number. In this exercise, our value function iteration algorithm converges to a solution in a little more than 4,000 iterations.

Starting from the same $L_{c,0}/L = 0.15$, Panel (b) sets as initial guesses for the time path of the value functions V_c and V_m their respective value in the low-commuting stationary equilibrium. We impose the same convergence criterion, requiring this time that the terminal value of the value functions converge to the low-commuting stationary equilibrium. In this case, our value function iteration algorithm does not converge within 40,000 iterations. Panel (b) shows that along the

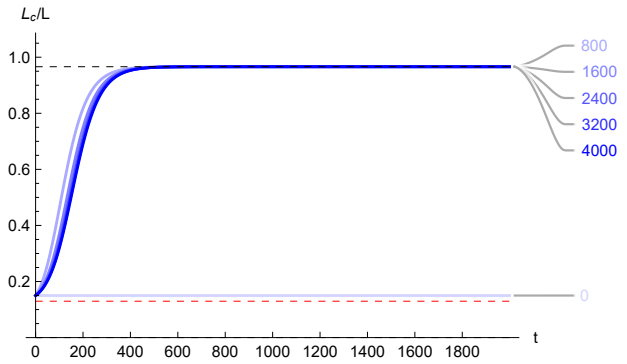
iteration process, the guesses attempt to reach the low commuting equilibrium, but they move progressively away from it.

The bottom panels of Figure A11 report analogous exercises where we instead start from an initial condition below the unstable equilibrium, namely $L_{c,0}/L = 0.10$. In Panel (c), we again attempt to converge toward the high-commuting equilibrium, with no success. In Panel (d), we successfully converge towards the low-commuting equilibrium in about 3,600 iterations.

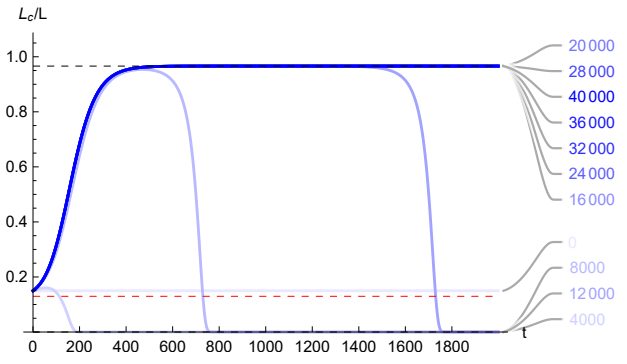
In unreported results, we have also experimented with other initial conditions farther away from either side of the unstable equilibrium, and other cities of varying sizes (like San Francisco, CA, Kansas City, MO-KS, Milwaukee, WI, Salt Lake City, UT, and Providence, RI), always reaching the same conclusion.

Figure A11: Tests for multiple equilibria

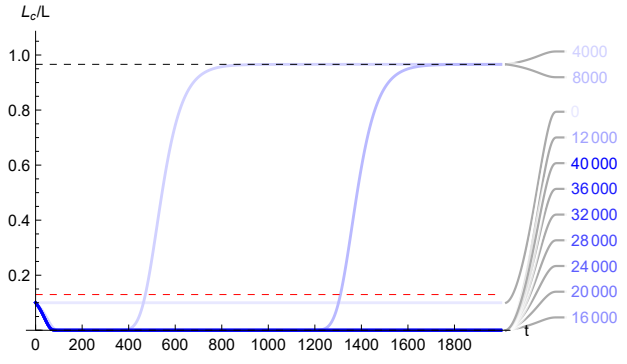
(a) $L_{c,0}/L = 0.15$, shooting toward high-commuting equilibrium



(b) $L_{c,0}/L = 0.15$, shooting toward low-commuting equilibrium



(c) $L_{c,0}/L = 0.10$, shooting toward high-commuting equilibrium



(d) $L_{c,0}/L = 0.10$, shooting toward low-commuting equilibrium

