SPATIAL ECONOMICS FOR GRANULAR SETTINGS

Jonathan I. Dingel
Felix Tintelnot

Spatial Economics for Granular Settings
Jonathan I. Dingel and Felix Tintelnot
NBER Working Paper No. 27287
May 2020, Revised January 2021
JEL No. C25,F16,R1,R13,R23

ABSTRACT

We examine the application of quantitative spatial models to the growing body of fine spatial data used to study economic outcomes for regions, cities, and neighborhoods. In "granular" settings where people choose from a large set of potential residence-workplace pairs, idiosyncratic choices affect equilibrium outcomes. Using both Monte Carlo simulations and event studies of neighborhood employment booms, we demonstrate that calibration procedures that equate observed shares and modeled probabilities perform very poorly in such settings. We introduce a general-equilibrium model of a granular spatial economy. Applying this model to Amazon's proposed HQ2 in New York City reveals that the project's predicted consequences for most neighborhoods are small relative to the idiosyncratic component of individual decisions in this setting. We propose a convenient approximation for researchers to quantify the "granular uncertainty" accompanying their counterfactual predictions.

Jonathan I. Dingel
Booth School of Business
University of Chicago
5807 South Woodlawn Avenue
Chicago, IL 60637
and NBER
jdingel@chicagobooth.edu

Felix Tintelnot
Kenneth C. Griffin Department of Economics
University of Chicago
5757 South University Avenue
Chicago, IL 60637
and NBER
tintelnot@uchicago.edu

1 Introduction

Economists increasingly use quantitative spatial models to evaluate urban policies such as infrastructure investments and land-use planning decisions. The growing availability of economic data observed at increasingly finer spatial scales offers tremendous potential for new insights. However, if policymakers are to rely on these models to inform their decisions, researchers must establish that they reliably capture relevant features of the data and perform well in making counterfactual predictions.

Quantitative spatial models characterize spatial linkages by the gravity equation, which says that the volume of interactions between any two locations increases with their sizes and decreases with the bilateral frictions between them (Redding and Rossi-Hansberg, 2017). The gravity equation, in turn, is derived under the assumption that the volume of interactions between any two locations is large enough to wash out any idiosyncratic choices of individuals. But when locations are defined at fine spatial scales, the interactions between them are "granular." That is, individual decision makers are large relative to the economic outcomes being analyzed. In such settings, the idiosyncratic choices of individual decision makers do not wash out. By assuming a continuum of individuals, the extant literature has neglected granularity's influence on its counterfactual predictions.

In this paper, we examine the application of quantitative spatial models to granular settings. In doing so, we make three main contributions. First, we show that granularity is common in the spatial settings that researchers study and that the most popular procedure used to compute counterfactual outcomes suffers from severe overfitting problems in granular settings. We demonstrate this procedure's inability to predict counterfactual outcomes in both Monte Carlo simulations and event studies of neighborhood employment booms in New York City. Second, we develop a quantitative spatial model that features a finite number of individuals and is suitable for counterfactual analysis. While incumbent models assume a continuum of individuals as a modeling convenience, our granular framework is just as tractable. Third, we show that counterfactual predictions are accompanied by "granular uncertainty," owing to the idiosyncratic components of equilibrium outcomes, and propose a simple approximation of this dispersion. Applying this approach to an evaluation of Amazon's proposed HQ2 in New York City suggests that the predicted counterfactual changes for most census tracts are small compared to the granular uncertainty of these outcomes.

Section 2 reviews the conventional continuum model and describes two methods for computing counterfactual outcomes in quantitative spatial models. The covariates-based approach makes the cost of commuting between a residence and a workplace a function of observed characteristics. In this approach, the model does not fit the observed data perfectly. By contrast, the popular calibrated-shares procedure for computing counterfactual outcomes selects commuting costs and other parameters to exactly fit the initial data. As a result, any idiosyncratic noise in observed outcomes is incorporated into the calibrated parameter values.

Section 3 examines the performance of quantitative spatial models with a continuum of individuals in granular empirical settings. We first establish that empirical settings of interest are

often granular. Tract-to-tract commuting matrices used in prior research are sufficiently granular that they are sparse: half or more of tract pairs have zero commuters. Small flows constitute a substantial share of total commuting: more than 40% of New York City commuters have a residence-workplace pairing populated by five or fewer individuals. We then show that the calibrated-shares procedure, which equates observed shares and underlying probabilities, predicts counterfactual outcomes poorly in these settings. A Monte Carlo simulation shows that granularity alone is sufficient to seriously impair the calibrated-shares procedure's predictive power when the model is otherwise correctly specified. While the procedure works in Monte Carlo simulations as the number of individuals becomes arbitrarily large, it produces uninformative predictions in simulations using the actual number of individuals in New York City.

Next, we examine how well these approaches predict changes in commuting flows using event studies. Commuting flows are the key spatial linkage between neighborhoods in many models of cities. We study events in which there is a large, discrete increase in employment in a single census tract, often arising from the arrival or expansion of a large employer. We increase the productivity parameter for that tract to match the increase in total employment and compare the predicted changes in bilateral commuting flows to those observed in the data. When examining 83 tract-level employment booms in New York City, we find that the covariates-based approach——using transit times as the only covariate——predicts the observed changes in commuting flows better than the conventional calibrated-shares procedure in 80 of these events. While predictive performance varies across events, regressing the observed change in commuters on the covariates-based approach's prediction typically yields an estimated slope of one, whereas the median estimated slope for the calibrated-shares procedure is negative. In other words, the counterfactual predictions from the calibrated-shares procedure are *negatively* correlated with the observed changes in commuting in more than half of the 83 events.

At first glance, it may be surprising that the covariates-based approach predicts changes in commuter counts better than the procedure that calibrates parameters to perfectly fit the pre-event flows. For intuition, note that each observation in the pre-event cross section has three parts: a time-invariant component correlated with included covariates, a time-invariant component uncorrelated with those covariates, and a time-varying component. On the one hand, the advantage of the calibrated-shares procedure is its potential to capture real spatial linkages not predicted by covariates like transit times, such as the fact that a very large share of Columbia University employees reside in nearby university-owned residences. On the other hand, the calibrated-shares procedure matches every bit of variation in the data, including transitory individual idiosyncrasies. This risks overfitting the third component of the pre-event observations and failing to predict post-event outcomes. In our event studies, the overfitting problem dominates. Pooling three years of data or aggregating predictions to much larger spatial units does not alter the finding that the covariates-based approach predicts changes in commuting flows better than the popular calibrated-shares procedure.

Fundamentally, while a continuum model can produce predictions for granular settings, it

2

cannot produce granular outcomes. In a granular economy, idiosyncratic individual decisions affect both equilibrium quantities and prices. To capture the role of these idiosyncrasies, one needs a model that admits a finite number of individuals. In Section 4, we propose a granular model in which equilibrium outcomes depend in part on the idiosyncratic component of individuals' choices. Where the continuum model would yield a single equilibrium allocation, our framework produces a *distribution* of economic outcomes. Notably, a single realization of the stochastic process often produces zeros for equilibrium outcomes that have strictly positive probabilities.

Our granular model retains the general-equilibrium approach employed in prior quantitative spatial models (e.g., Redding and Rossi-Hansberg, 2017), but the number of individuals in the economy is finite. The key modeling challenge is that individual decisions affect wages and rents. We assume that individuals act as price takers and optimize given their beliefs about wages and rents. Employing standard discrete-choice methods (e.g., Train, 2009), we model individuals choosing residence-workplace pairs on the basis of wage beliefs, rent beliefs, commuting costs, and an individual-specific idiosyncratic preference shock. It is computationally infeasible to enumerate all the possible labor allocations and resulting prices. To overcome this combinatorial challenge, we assume that individuals have point-mass wage and rent beliefs that are equal to the equilibrium prices in a model with the same parameters but a continuum of individuals. This simplification of agents' beliefs to solve an otherwise intractable decision problem is similar to the assumptions used in industrial organization (Weintraub, Benkard, and Van Roy, 2008) or macroeconomics (Krusell and Smith, 1998). In our application to New York City, decisions based on these beliefs are optimal for the overwhelming majority of individuals. At realized wages and rents, which differ from the point-mass beliefs about these prices, 96% of individuals would not change their residence-workplace choice if given the opportunity. For the 4% who would want to switch, the foregone benefits of switching are typically equivalent to an income increase of less than 1%.

Our granular model can be estimated using the same data used to estimate the covariates-based continuum model. When individuals' price beliefs are the continuum-case equilibrium prices, the estimated parameters of this model are identical to those of the covariates-based continuum model. Given the same parameter values, the mean equilibrium labor allocation of the granular model is equal to the equilibrium labor allocation of the covariates-based continuum model. By contrast, the granular model's equilibrium prices solve a system of non-linear equations, so their mean values are not necessarily equal to the continuum-case prices. Crucially, unlike the continuum model, the granular model characterizes the dispersion in quantities and prices that arises from idiosyncrasies.

In Section 5, we illustrate how counterfactual predictions depend on both the estimation procedure and granular uncertainty in the context of evaluating the local economic effects of Amazon's proposed—but later abandoned—HQ2 in New York City. The (mean) changes in the labor allocation predicted by the granular model and covariates-based continuum model differ considerably from the predictions of the calibrated-shares procedure. The granular model and covariates-based continuum model predict similar (mean) changes in land rents, and, for some neighborhoods, these predictions differ considerably from those of the calibrated-shares procedure.

We also find that most predicted tract-level changes are small relative to the accompanying granular uncertainty. We propose that researchers use a simple approximation based on the number of decision makers and the (change in) modeled probabilities to quantify the magnitude of granular uncertainty accompanying their predictions about counterfactual changes in residents and workers. Our finding that granular uncertainty is large relative to the predicted neighborhood-level outcomes induced by a headline-grabbing headquarters decision suggests that granular uncertainty is likely important for counterfactual outcomes in many granular settings.

We contribute to the rapidly expanding literature on quantitative spatial economics.[1] A growing body of research aims "to provide an empirically relevant quantitative model to perform general equilibrium counterfactual policy exercises" concerning topics such as transportation-infrastructure investment, neighborhood revitalization programs, and local business tax incentives (Redding and Rossi-Hansberg, 2017, p.23).[2] As anticipated by Holmes and Sieg (2015, p.106), "[a]s research in urban and regional applications takes advantage of new data sets at high levels of geographic resolution, it permits the study of interactions at narrow levels, where there may be relatively few decision makers," but quantitative spatial models have largely abstracted from discreteness in the underlying economic environment.

Our paper introduces a quantitative spatial model suitable for applications using small spatial units, such as census tracts. Social scientists study neighborhoods to understand housing markets, intergenerational mobility, racial segregation, and many other phenomena. Spatially precise satellite imagery (Donaldson and Storeygard, 2016) and phone movement data (Couture et al., 2020) will allow even finer investigations. Granular data are valuable in part because a growing body of evidence shows highly localized agglomeration economies (Rosenthal and Strange, 2020). Arzaghi and Henderson (2008) estimate that productivity gains from interactions among advertising firms occur primarily within 500 meters. Ahlfeldt et al. (2015) estimate that production and residential externalities are highly localized, decaying by more than half at distances of two minutes and one minute of travel time, respectively. Rossi-Hansberg, Sarte, and Owens (2010) estimate that externalities from non-market interactions decline by half approximately every 1,000 feet. We provide tools for applying quantitative spatial models to these granular empirical settings.

Our event studies of neighborhood employment booms contribute to a much smaller literature assessing the predictive power of quantitative spatial models using observed economic changes. Ahlfeldt et al. (2015) demonstrate their quantitative model's ability to capture how changes in floor prices associated with Berlin's division and reunification correlated with distance to the pre-war central business district. Monte, Redding, and Rossi-Hansberg (2018) verify a qualitative

---

[1]In gravity-based quantitative models, a location's access to all other locations' goods and factors, which is impeded by bilateral frictions, influences the magnitude and character of its economic activity. See Proost and Thisse (2019) for a recent survey of these models, Donaldson (2015) on market access in goods trade, Redding and Turner (2015) on transport costs for both goods and commuters, and Dingel (2017) on the income composition of goods market access.

[2]Heblich, Redding, and Sturm (2020); Severen (2019); Tsivanidis (2019); and Zárate (2019) study transportation-infrastructure investments. Allen, Arkolakis, and Li (2015) compute optimal zoning for each block of Chicago. Owens, Rossi-Hansberg, and Sarte (2020) compute the value of revitalizing each of 52 vacant tracts in Detroit. Berkes and Gaetani (2020) compare the effects of siting Amazon's HQ2 in four different New York City neighborhoods.

implication of their quantitative spatial model—larger employment responses to a local productivity increase in counties more open to commuting—using the million-dollar-plant openings examined by Greenstone, Hornbeck, and Moretti (2010). Adão, Arkolakis, and Esposito (2020) compare US commuting zones' actual and predicted changes in employment and wages in response to the "China shock" of Autor, Dorn, and Hanson (2013) for different calibrations of spatial links. Kreindler and Miyauchi (2020) use a gravity model and cell-phone–derived commuting flows to predict workplace wage levels in developing-economy cities that lack conventional data coverage. We explore the predictive power of quantitative spatial models in granular settings. Examining the "out-of-sample" performance of quantitative spatial models is an important step in establishing their value as an input into urban planning and other policymaking decisions.

Finally, our framework is related to prior research about granular economies. A growing literature examines the importance of granularity for aggregate fluctuations in macroeconomics (Gabaix, 2011; Carvalho and Grassi, 2019) and international economics (di Giovanni and Levchenko, 2012; di Giovanni, Levchenko, and Mejean, 2014; Gaubert and Itskhoki, 2018; Daruich, Easterly, and Reshef, 2019). In addition to its relevance for aggregate fluctuations, firm-level granularity has been studied as one explanation for zeros in international trade (Eaton, Kortum, and Sotelo, 2013; Armenter and Koren, 2014). Ellison and Glaeser (1997) and Schoefer and Ziv (2020) address the role of granular manufacturing plants when computing the geographic concentration of industries and geographic variation in productivity, respectively. Mogstad et al. (2020) raise concerns about inferring ranks in granular settings. Our paper makes two contributions to these various strands of the literature. First, we demonstrate the importance of addressing granularity when making predictions about counterfactual outcomes. Second, we introduce a tractable general-equilibrium modeling approach that quantifies the granular uncertainty accompanying predictions.

2 Computing counterfactual outcomes in continuum models

This section describes two methods for computing counterfactual outcomes in quantitative spatial models. We use a baseline model featuring a continuum of individuals making residential and workplace location choices, as in Ahlfeldt et al. (2015) and Redding and Rossi-Hansberg (2017).

2.1 Setup

We consider a closed economy populated by a measure of L individuals who each supply one unit of labor. There is a set of locations indexed by k or n. Each location has a fixed quantity of land, T_k, owned by immobile landlords who consume only goods.[3] Each location has the technology to produce a differentiated good (i.e., the Armington assumption). For simplicity, we abstract from trade costs.

Individuals have Cobb-Douglas preferences over these traded goods and land, devoting α of their expenditure to the latter. They have constant elasticity of substitution (CES) preferences

---

[3]This simplifying assumption follows Monte, Redding, and Rossi-Hansberg (2018).

over the set of differentiated goods, with elasticity of substitution $\sigma > 1$. Individuals (indexed by i) have idiosyncratic preferences for pairs of residential and workplace locations, such that i's indirect utility from residing in k and working in n is

$$U_{kn}^i = \underbrace{\epsilon \ln \left(\frac{w_n}{r_k^\alpha P^{1-\alpha} \delta_{kn}} \right)}_{\equiv U_{kn}} + \nu_{kn}^i, \tag{1}$$

where w_n denotes the wage in location n, r_k denotes the land rent in k, P denotes the common CES price index for goods, $\delta_{kn} \geq 1$ denotes the commuting cost between k and n, U_{kn} denotes the mean utility of choice kn, and ν_{kn}^i is the idiosyncratic preference shock of individual i for residing in k and working in n. The commuting elasticity ϵ governs the importance of mean utility relative to the idiosyncratic preference shock. We assume that ν_{kn}^i is drawn from a type-1 extreme value distribution.

Production of each location's differentiated good is linear in labor. The goods market is perfectly competitive.[4] Commuting is costly because time spent commuting is not spent working, so that individuals residing in k and working in n earn only w_n/δ_{kn} because they only spend $1/\delta_{kn}$ of their time working.[5] Goods output in location n is $q_n = A_n L_n$, where A_n is that location's productivity and L_n is the quantity of labor supplied by workers working in n. Given this production function and perfect competition, the price of location n's output is w_n/A_n for all consumers. Thus, the CES price index is $P = \left[\sum_n (w_n/A_n)^{1-\sigma} \right]^{1/(1-\sigma)}$.

2.2 Equilibrium

Let ℓ_{kn} denote the measure of people residing in k and working in n. Using the type-1 extreme value distribution assumption, one can integrate over the idiosyncratic preference shocks to obtain the fraction of people residing in k and working in n:

$$\frac{\ell_{kn}}{L} = \frac{w_n^\epsilon \left(r_k^\alpha \delta_{kn} \right)^{-\epsilon}}{\sum_{k',n'} w_{n'}^\epsilon \left(r_{k'}^\alpha \delta_{k'n'} \right)^{-\epsilon}}. \tag{2}$$

Goods market clearing equates each location's output to the quantity demanded. Labor supplied in location n is $L_n = \sum_k \ell_{kn}/\delta_{kn}$, and thus output there is $A_n \sum_k \ell_{kn}/\delta_{kn}$. Each individual devotes $1-\alpha$ of their expenditure to differentiated goods and α of their expenditure to land, while immobile landlords spend all of their income on differentiated goods, such that total expenditure on differentiated goods equals aggregate income. The demand for each differentiated good stemming

---

[4]Given established isomorphisms between perfect-competition and monopolistic-competition models of economic geography (e.g., Allen and Arkolakis 2014), this assumption is not crucial to our results.

[5]Alternatively, we could assume that δ_{kn} represents a utility shifter that does not affect working time. Modeling δ_{kn} as reduced working time facilitates the commuting-elasticity estimation discussed in Section 3.2.

from CES preferences means that equating quantity supplied and quantity demanded requires

$$A_n \sum_k \frac{\ell_{kn}}{\delta_{kn}} = \frac{(w_n/A_n)^{-\sigma}}{P^{1-\sigma}} \sum_{k',n'} \frac{\ell_{k'n'}}{\delta_{k'n'}} w_{n'} \qquad \forall n. \tag{3}$$

Note that goods market clearing implies labor market clearing.

Similarly, land market clearing equates the fixed land endowment T_k to the quantity demanded by individuals, who devote a constant fraction α of their expenditure to land:

$$T_k = \frac{\alpha}{r_k} \sum_n \frac{\ell_{kn}}{\delta_{kn}} w_n \qquad \forall k. \tag{4}$$

Equilibrium is a set of wages $\{w_n\}$, rents $\{r_k\}$, and labor allocation $\{\ell_{kn}\}$ such that equations (2), (3), and (4) hold. Appendix A.1 shows that if $\left(\frac{1+\epsilon}{\sigma+\epsilon}\right)\left(\frac{\alpha\epsilon}{1+\alpha\epsilon}\right) \leq \frac{1}{2}$, this equilibrium exists and is unique by application of Theorem 1 of Allen, Arkolakis, and Li (2020).

2.3 Procedures for counterfactual predictions

Researchers employing quantitative spatial models aim to provide a parsimonious general-equilibrium framework in order to predict counterfactual outcomes.[6] The value of these quantitative models for policy analysis therefore depends largely on their ability to predict the spatial distribution of economic consequences of a shock.[7]

Two approaches have been used to compute counterfactual outcomes in this class of models. They differ in how they parameterize commuting costs and whether equation (2) is satisfied. The first approach parameterizes the commuting costs δ_{kn} as functions of observed characteristics. In this covariates-based approach, equation (2) is generically not satisfied by the observed values of ℓ_{kn} and the parameterization of δ_{kn}.[8] After estimating the model, one can compute the counterfactual equilibria associated with any counterfactual values of A_n, T_k, and δ_{kn}. A prominent example of this covariates-based approach to counterfactual computations is Ahlfeldt et al. (2015), who observe only a small sample of commuters relative to the population they study.[9]

A second approach infers the values of δ_{kn} and other model parameters by assuming that

---

[6]Redding and Rossi-Hansberg (2017, p.23): "[T]his research does not aim to provide a fundamental explanation for the agglomeration of economic activity, but rather to provide an empirically relevant quantitative model to perform general equilibrium counterfactual policy exercises." Waddell and Sarte (2016, p.190): "The development of the new quantitative equilibrium models has initiated a more robust and realistic framework with which to model cities...it will also allow for more robust counterfactual policy exercises that can inform practitioners and policymakers regarding strategies for urban development."

[7]Bryan, Glaeser, and Tsivanidis (2019, p.31): "If quantitative models are to provide useful policy insights, their results have to be trusted. First, researchers must establish that their model captures relevant features of the data or (ideally) can replicate the real-world response to a policy change." More broadly, Kehoe (2005) argues that "it is the responsibility of modelers to demonstrate that their models are capable of predicting observed changes, at least ex post."

[8]One justification for the model not fitting the observed data would be that the observed set of commuters is a sample from the population.

[9]Similarly, when using a sample describing domestic trade transactions, Holmes and Stevens (2014) employ a covariates-based approach for trade costs.

equation (2) is satisfied by the observed labor allocation. As shown in Appendix A.2, one can compute the counterfactual equilibrium outcomes associated with proportionate changes in the productivity, land-endowment, and commuting-cost parameters using only data on the initial labor allocation and wages. This approach is known as "calibrated share form" or "exact hat algebra" in the international trade literature (Rutherford, 1995; Dekle, Eaton, and Kortum, 2008; Costinot and Rodríguez-Clare, 2014). This calibrated-shares procedure has been widely used in spatial economics recently, often in granular settings where the number of residence-workplace pairs is large relative to the number of decision makers.[10]

3 Counterfactual analysis in granular empirical settings

We assess the predictive performance of quantitative spatial models in granular empirical settings. In Section 3.1, we show that granularity is common in empirical settings in which researchers and policymakers are interested in the incidence of local economic shocks. After estimating the model in Section 3.2, we present a Monte Carlo exercise in Section 3.3 that reveals that the calibrated-shares procedure performs poorly in these settings because it overfits granular data. We further assess the two approaches' predictions using event studies of neighborhood employment booms in New York City in Section 3.4. While the approach using commuting costs based on transit times predicts the change in the number of people from each residential census tract who work in the booming tract quite well, the calibrated-shares procedure's predicted changes are in fact negatively correlated with the observed changes in the majority of events.

3.1 Granularity in tract-level commuting flows

Urban economists often study empirical settings in which the fine spatial resolution of the geographic units employed makes the commuting matrix necessarily granular. Owens, Rossi-Hansberg, and Sarte (2020) study 1.3 million people commuting between 1.3 million pairs of tracts in the Detroit urban area. Roughly three-quarters of the cells in the Los Angeles metropolitan commuting matrix studied by Severen (2019) are empty. Tsivanidis (2019) models the behavior of Bogotá's 8 million residents commuting between almost 7 million pairs of tracts. Zárate (2019) examines Mexico City, which has 9 million people commuting between 13 million pairs of tracts. At finer spatial resolution, Ahlfeldt et al. (2015) model the behavior of about 3 million Berliners choosing among about 254 million pairs of city blocks. In this section, we document the granularity of tract-level commuting flows in New York City. Appendix B reports comparable summary statistics for Detroit and Minneapolis-St. Paul.

We use data on commuting between census tracts in New York City taken from the Longitudinal Employer-Household Dynamics, Origin-Destination Employment Statistics (LODES).[11] One impli-

---

[10]Recent studies in spatial economics employing this technique include Allen, Arkolakis, and Li (2015); Perez-Cervantes (2016); Waddell and Sarte (2016); Monte, Redding, and Rossi-Hansberg (2018); Heblich, Redding, and Sturm (2020), Krebs and Pflüger (2019); Severen (2019); and Owens, Rossi-Hansberg, and Sarte (2020).

[11] Davis et al. (2019) use the 2010 wave to compute the joint distribution of residences and workplaces for New

cation of granularity is that some data providers perturb observations to protect confidentiality. For example, the published LODES tract-level workplace employment counts are noise-infused. Moreover, the published LODES commuting matrices report synthetically generated residence-workplace pairs.[12] These confidentiality-protecting procedures are both a symptom of the granular setting and an additional cause for caution when interpreting the value of any individual observation.[13]

Small flows constitute a large share of tract-to-tract commuting in New York City. New York City has about 2.5 million resident-employees and 4.6 million tract pairs. Thus, its commuting matrix is almost necessarily sparse, as the average cell contains about one-half of a commuter. In fact, 85% of tract pairs have zero commuters between them, so the average positive cell contains about three people. As shown in Figure 1, more than half of the tract pairs with a positive number of commuters have only one commuter. Small flows constitute a substantial share of total commuting: 41.1% of New York City commuters have five or fewer commuters in their cell of the commuting matrix. Furthermore, the commuting flows are impersistent over time and exhibit asymmetric zeros, as we document in Appendix B.

Figure 1: Number of commuters between pairs of tracts in New York City
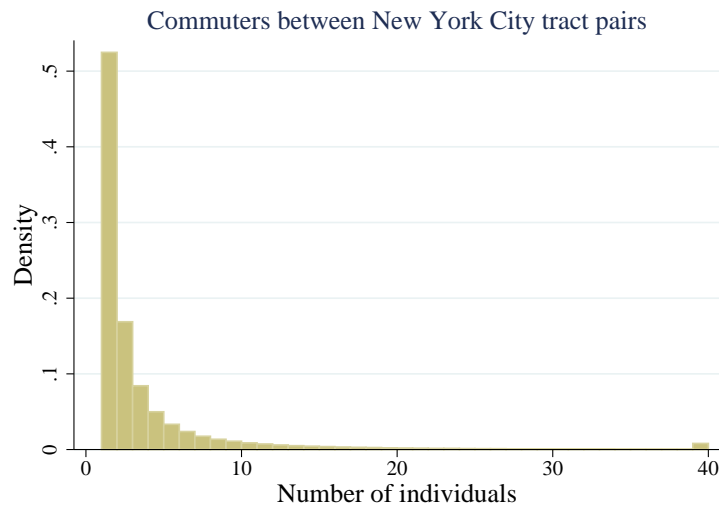
NOTES: This histogram reports the number of tract pairs in New York City by the number of individuals who reside in the origin tract and work in the destination tract in 2010 LODES data. This histogram restricts the sample to the 15% of tract pairs that have a strictly positive number of commuters. Observations with more than 40 commuters are winsorized to 40. There are 2,160 residential census tracts in New York City.

York City. Owens, Rossi-Hansberg, and Sarte (2020) use the 2014 wave of these data for the Detroit urban area to estimate a gravity model of commuting. Following Owens, Rossi-Hansberg, and Sarte (2020), we use the primary-job counts in LODES.

[12]Graham, Kutzbach, and McKenzie (2014): "For each job in a workplace cell, LODES draws from a Dirichlet multinomial posterior distribution of possible residential locations...The prior adds uncertainty, so that even commutes with few or no observed flows may appear to have a job. Conversely, even when there are commuters from an origin in the likelihood, that residence may not be drawn and thus would not appear in LODES."

[13]The econometric consequences of noise infusion and synthetic observations are specific to each use case. For example, see the discussion of these issues in Couture and Handbury (2020). Calibration approaches that perfectly match every (synthetic) observation are likely the research design that is most sensitive to these procedures.

The sparsity of urban commuting matrices is driven in part by the fact that tracts are defined by residential population, independent of employment. Tracts are defined by the US Census Bureau to be relatively uniform in population size, such that the typical tract has 4,000 residents. Since employment is spatially concentrated, many workplace destinations have fewer workers than the number of residential origins. The median tract in New York City has 248 employees working in it. Since New York City has 2,160 residential tracts, at least 89% of locations must have zero residents commuting to this workplace.

Even when studying larger geographic units, the issue of granularity may be relevant. For example, counties in the United States vary tremendously in their population sizes. While granularity is unlikely to be relevant when examining US counties with millions of residents, Appendix B documents that many county pairs have small commuting flows that represent granular outcomes.

3.2 Estimation

We estimate the continuum model using the tract-to-tract commuting data for New York City in 2010. We set the values of a few parameters prior to estimating those remaining. Following the literature, we impose $\alpha = 0.24$, based on Davis and Ortalo-Magne (2011), and $\sigma = 4$, following Monte, Redding, and Rossi-Hansberg (2018).

For commuting costs, we distinguish between observed and unobserved components in order to encompass both the covariates-based approach and the calibrated-shares procedure:

$$\delta_{kn} = \underbrace{\bar{\delta}_{kn}}_{\text{observed}} \times \underbrace{\lambda_{kn}}_{\text{unobserved}}$$

We construct the observed part of the commuting costs by assuming that each worker has H hours that are spent either working or commuting. We compute $\bar{\delta}_{kn} = \frac{H}{H - t_{kn} - t_{nk}}$, where $H = 9$ hours and t_{kn} is the transit time from k to n according to Google Maps.[14] Consistent estimation of ϵ requires that the unobserved component λ_{kn} is appropriately mean-independent of $\bar{\delta}_{kn}$, but, beyond that, a variety of assumptions are possible.[15] To emphasize the contrast with the calibrated-shares procedure, we estimate the covariates-based model assuming that there is no unobserved component, $\lambda_{kn} = 1 \ \forall k, n$.

One way to estimate the continuum model using the covariates-based approach is to interpret observed data on residence-workplace choices as a finite sample from the continuum model. In that case, given observed values of commuting costs $\delta_{kn} = \bar{\delta}_{kn}$, the remaining parameters can be

---

[14] Given H hours, $1/\bar{\delta}_{kn}$ is the share of that time spent working if the individual resides in k and works in n. Using $H = 8$ or $H = 10$ yields very little change in the model fit relative to the $H = 9$ results reported in Table 1. For New York City, we use Google Maps public-transit times collected by Davis et al. (2019). We impute missing observations by predicting transit times using physical distance, imputing transit times for fewer than 4.1% of tract pairs in New York City. For both the covariates-based approach and the calibrated-shares procedure, we assume that labor supply in location n is $L_n = \sum_k \ell_{kn}/\bar{\delta}_{kn}$. In other words, λ_{kn} affects the utility of commuting between k and n but not the hours of commuting, which are observed and captured by $\bar{\delta}_{kn}$.

[15] Among many intermediate cases, one might assume that λ_{kn} has a random-effects or an interactive-fixed-effects structure. Similarly, one could enrich $\bar{\delta}_{kn}$ by introducing more observed covariates or more flexible functions of transit times.

estimated by maximum likelihood. Interpreting the right side of equation (2) as a probability and denoting the number of individuals who chose the kn pair by ℓ_{kn}, the log likelihood function is

$$\ln \mathcal{L} \equiv \sum_{k,n} \ell_{kn} \ln \left[\Pr(U_{kn}^i > U_{k'n'}^i \ \forall (k', n') \neq (k, n)) \right] = \sum_{k,n} \ell_{kn} \ln \left[\frac{w_n^\epsilon \left(r_k^\alpha \bar{\delta}_{kn} \right)^{-\epsilon}}{\sum_{k',n'} w_{n'}^\epsilon \left(r_{k'}^\alpha \bar{\delta}_{k'n'} \right)^{-\epsilon}} \right]. \quad (5)$$

This is the canonical conditional-logit likelihood of McFadden (1974) applied to location choices, as in McFadden (1978). The k- and n-specific terms are captured by residence and workplace fixed effects, respectively. As shown by Guimarães, Figueirdo, and Woodward (2003), the maximization of this likelihood function is numerically equivalent to a Poisson pseudo-maximum-likelihood estimator (PPMLE) that is available for a variety of software packages.[16]

To estimate the commuting elasticity ϵ for use in the calibrated-shares procedure, two distinct mean-independence assumptions about the unobserved component λ_{kn} might be made. If one assumes $\mathbb{E}\left(\lambda_{kn}^{-\epsilon} | r_k, w_n, \bar{\delta}_{kn} \right) = 1$, the constant-elasticity function can be estimated using PPMLE, as shown in Silva and Tenreyro (2006). This yields the same estimate of the commuting elasticity ϵ as the maximum likelihood estimate of equation (5). If one takes the logarithm of each side of equation (2), assumes $\mathbb{E}\left(\ln \lambda_{kn} | r_k, w_n, \bar{\delta}_{kn} \right) = 0$, and restricts the estimation sample to observations for which ℓ_{kn} is strictly positive, then the commuting elasticity can be estimated by ordinary least squares (OLS). As emphasized by Silva and Tenreyro (2006), these two estimators can yield very different parameter estimates.

We apply both the PPML and OLS estimators to New York City in 2010. Our estimate of the commuting cost elasticity is presented in column 1 of Table 1. The estimate of $\epsilon \approx 8$, which is comparable to the value of 6.8 estimated by Ahlfeldt et al. (2015) for commuting within Berlin, implies that idiosyncratic preference shocks are modestly dispersed. Estimating the commuting elasticity using OLS yields a much lower elasticity, likely because of the well-understood selection bias associated with omitting four-fifths of observations from the estimation sample.[17] In what follows, we use the estimate reported in column 1 for both the covariates-based approach and the calibrated-shares procedure.

For the covariates-based approach, the remaining economic primitives $\{T_k\}$ and $\{A_n\}$ can be inferred from the estimated fixed effects and the market-clearing conditions. In particular, the residence fixed effect is proportionate to $r_k^{-\alpha\epsilon}$, and the workplace fixed effect is proportionate to w_n^ϵ. Given α, σ, $\{\delta_{kn}\}$, ϵ, $\{r_k\}$, $\{w_n\}$, and equation (2), equations (3) and (4) can be solved to yield values of $\{T_k\}$ and $\{A_n\}$.

As described in Appendix A.2, the calibrated-shares procedure does not separately identify the parameters $\{T_k\}$, $\{A_n\}$, and $\{\lambda_{kn}\}$. Rather, one can compute counterfactual outcomes based on

---

[16]See Sotelo (2019) for a discussion of the relationship between the multinomial and Poisson pseudo-maximum-likelihood estimators in the context of gravity models. In practice, we use the Stata package `ppmlhdfe` of Correia, Guimarães, and Zylkin (2020).

[17]This selection bias also has substantial implications for the estimated fixed effects, as discussed in Appendix C. Recent research in industrial organization draws attention to the problems associated with zero-valued market shares (Quan and Williams, 2018; Gandhi, Lu, and Shi, 2019; Dubé, Hortaçsu, and Joo, 2020).

Table 1: Commuting elasticity estimates

| | NYC (2010) | |
	MLE	OLS
Commuting cost	-7.986	-2.307
	(0.307)	(0.0516)
Model fit (R^2 or pseudo-R^2)	0.662	0.561
Location pairs	4,628,878	690,673
Commuters	2,488,905	2,488,905

NOTES: All specifications include residence fixed effects and workplace fixed effects. The "MLE" columns present the results from the maximum likelihood estimation described in the text. The "OLS" columns present the results of estimating the log version of equation (2) by ordinary least squares, omitting observations in which $\ell_{kn} = 0$. The model-fit statistic is the pseudo-R^2 for MLE and R^2 for OLS.

combinations of these parameters that are identified under the assumption that the continuum model exactly replicates the observed commuter counts.

The calibrated-shares procedure is more data demanding than our covariates-based parameterization of the model. Let I denote the number of individuals and N the number of locations in the economy. The covariates-based model estimates two parameters (A_n and T_n) for each of the N locations, whereas the calibrated-shares procedure introduces an unobserved component for each pair of locations and therefore implicitly estimates more than N^2 parameters to rationalize the observed commuting flows. There is a sharp contrast between the two approaches' number of parameters relative to the number of individuals whose decisions they model. For the case of New York City, I/N exceeds one thousand, while I/N^2 is about one-half. In this empirical setting, we have limited concerns about residence and workplace fixed effects overfitting the data and much greater concern about the calibrated-shares procedure overfitting the observed commuting flows.

3.3 Monte Carlo: The calibrated-shares procedure in a granular setting

We use a Monte Carlo simulation to assess how well the calibrated-shares procedure for computing counterfactual outcomes performs in a granular setting. In anticipation of examining New York City neighborhood employment booms in the next section, the data-generating process is the estimated covariates-based model of New York City in 2010, and we impose a "counterfactual" productivity increase in one workplace. As detailed in Appendix D.1, we assume commuting costs are a function of transit times, take a finite number of draws from a multinomial distribution with probabilities given by the right side of equation (2) to simulate the "2010" labor allocation, impose a counterfactual 18% increase in productivity in one tract, and then take the same finite number of draws to simulate the "counterfactual" labor allocation. We study how well the calibrated-shares procedure predicts the counterfactual allocation when applied to the finite "2010" realization drawn from the data-generating process. In short, our simulations examine the finite-sample behavior of the calibrated-shares procedure using the actual number of individuals in New York City ($I \approx$

2.5 million). Given infinite draws, equation (2) would hold true for the realized labor allocation ℓ_{kn}, and the calibrated-shares procedure would perfectly predict the changes in commuting flows associated with the productivity increase.

Figure 2: Monte Carlo: Calibrated-shares procedure performs poorly with granular data

A. Regression of observed on predicted changes   B. Ratio of models' prediction errors
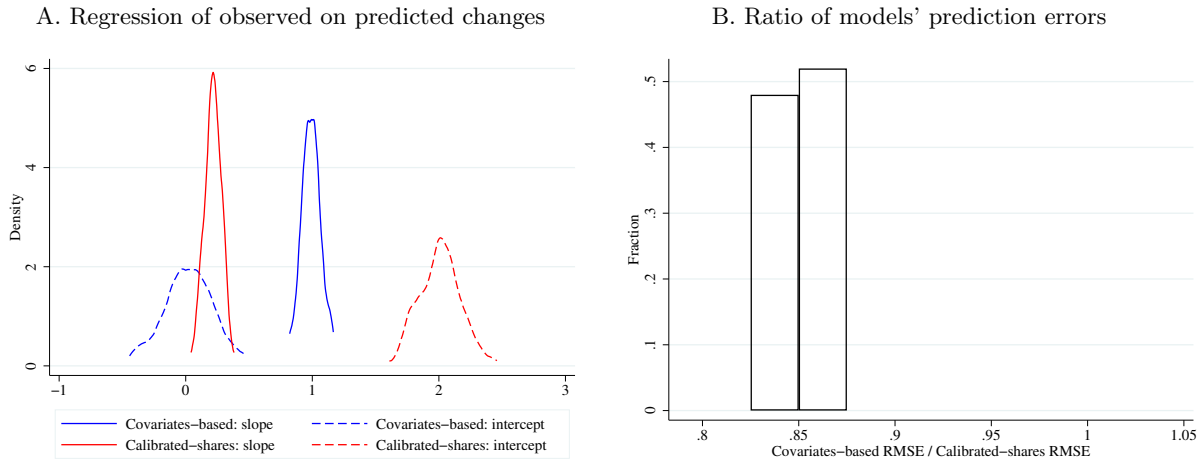
NOTES: This figure depicts the regression coefficients and RMSEs from 100 simulations in which $I = 2,488,905$, the data-generating process is the estimated covariates-based model of New York City in 2010, and the counterfactual change is an 18% increase in productivity in one workplace tract. See Appendix D.1 for a detailed description of the simulation procedure.

To assess the calibrated-shares procedure's predictive power, we compare the "observed" change in the number of residents working in the "treated" tract to the change predicted by the procedure. We do so by computing the root mean squared error (RMSE) and by regressing the observed changes on the predicted changes. With an unbiased forecast, the regression of observed on predicted changes would yield a slope coefficient of one and an intercept of zero. Of course, in a granular setting, the estimated coefficients in a single simulation will also reflect finite-sample noise. Simulating this process 100 times yields a distribution of regression coefficients and RMSEs, as shown in Figure 2. The number of individuals in the simulated economy is large enough compared to the number of locations that the covariates-based approach, which is correctly specified because it matches the underlying data-generating process, delivers accurate predictions. Its slope and intercept coefficients are centered on one and zero, respectively. The calibrated-shares procedure has very limited predictive power when applied to finite samples ($I \approx 2.5$ million) drawn from the data-generating process. Regressing simulated observed changes on the change predicted by the calibrated-shares procedure yields a slope of only 0.22 in the median simulation. The slope from the covariates-based approach is centered on one, as expected, and its RMSE is 85% of the calibrated-shares forecast error in the median simulation.

Since the only element of the data-generating process at odds with the assumptions of the calibrated-shares procedure is the finite number of individuals, these simulation results demonstrate that granularity alone can severely limit that procedure's predictive power. The procedure suffers

13

Table 2: Calibrated-shares procedure's performance and granularity

I	2.5	5	12.5	25	50	125	250	2560
slope	0.22	0.55	0.82	0.90	0.95	0.98	0.99	1.00
intercept	2.01	1.15	0.47	0.25	0.12	0.05	0.03	0.00

NOTES: This table reports the mean value of the slope and intercept coefficients from 100 simulations as we vary I, the number of individuals in the simulated economy. The column titles denote the number of individuals in millions. The $I = 2.5$ million case corresponds to $I = 2,488,905$, which is the number of individuals who reside and work in New York City in the 2010 LODES data.

from an overfitting problem when the number of individuals is small relative to the number of parameters. As shown in Table 2, the calibrated-shares procedure performs better as we increase the number of individuals.[18] But only when the number of draws is one to two orders of magnitude larger do the calibrated-shares procedure's predictions become closely correlated with outcomes.[19] In short, the calibrated-shares procedure is very data demanding because it implicitly estimates millions of parameters. Thus, we have reason to expect that the calibrated-shares procedure will perform poorly in granular empirical settings.

3.4 Event studies: Predicting commuting responses to a local shock

Given the concern that the calibrated-shares procedure may perform poorly in granular settings, we now assess the two approaches' predictions using neighborhood employment booms in New York City. As in the Monte Carlo simulations, we investigate changes in commuting patterns to workplace tracts that had large increases in employment. In particular, we conduct an "event study" for each of the 83 workplace tracts in New York City that had a two-year increase in total employment from 2010 to 2012 of at least 400 employees and at least 12.5% from a 2010 level of at least 400 employees. We focus on local employment booms because these large changes are likely driven by workplace-specific shocks, such as new office openings or an expansion by a large employer, rather than resident-workplace-specific shocks. For example, in 2011, Tiffany & Co. moved its corporate headquarters to 260,000 square feet of office space at 200 Fifth Avenue. In late 2010, Google acquired a building of nearly 3 million square feet at 111 Eighth Avenue. These locations are two of the 83 workplace tracts we examine.[20]

---

[18]Table D.1 in Appendix D.1 shows that, for a given number of individuals, the calibrated-shares procedure's performance improves relative to the covariates-based model as the share of variation in δ_{kn} orthogonal to the included covariates increases.

[19]This finding also cautions against spatial aggregation as a strategy for addressing granularity. Reducing the number of location pairs by two orders of magnitude would require reducing the number of locations by an order of magnitude. We judge that using location pairs' observed characteristics is a more attractive method of dimensionality reduction than eliminating variation by grouping locations into larger units. We demonstrate the viability of the covariates-based approach applied to real data from a granular setting in the following section. The difficulty of estimating tract-pair-specific parameters seems little reason to ignore the tract-level variation studied by vast numbers of social scientists.

[20]Figure D.1 depicts the employment changes for these two tracts.

We estimate or calibrate the model for New York City based on 2010 data, as described in Section 3.2. The calibrated-shares procedure perfectly rationalizes observed commuting flows and wages, as described in Appendix A.[21] Next, we compute the increase in productivity required to match the observed 2010–2012 change in employment for the "treated" tract. Since the two procedures fit the 2010 data differently, this productivity increase need not be the same, but in practice they are very similar. Since the productivity increases are defined so that both procedures match the observed increase in employment, we examine their predictions for the change in bilateral commuter counts.
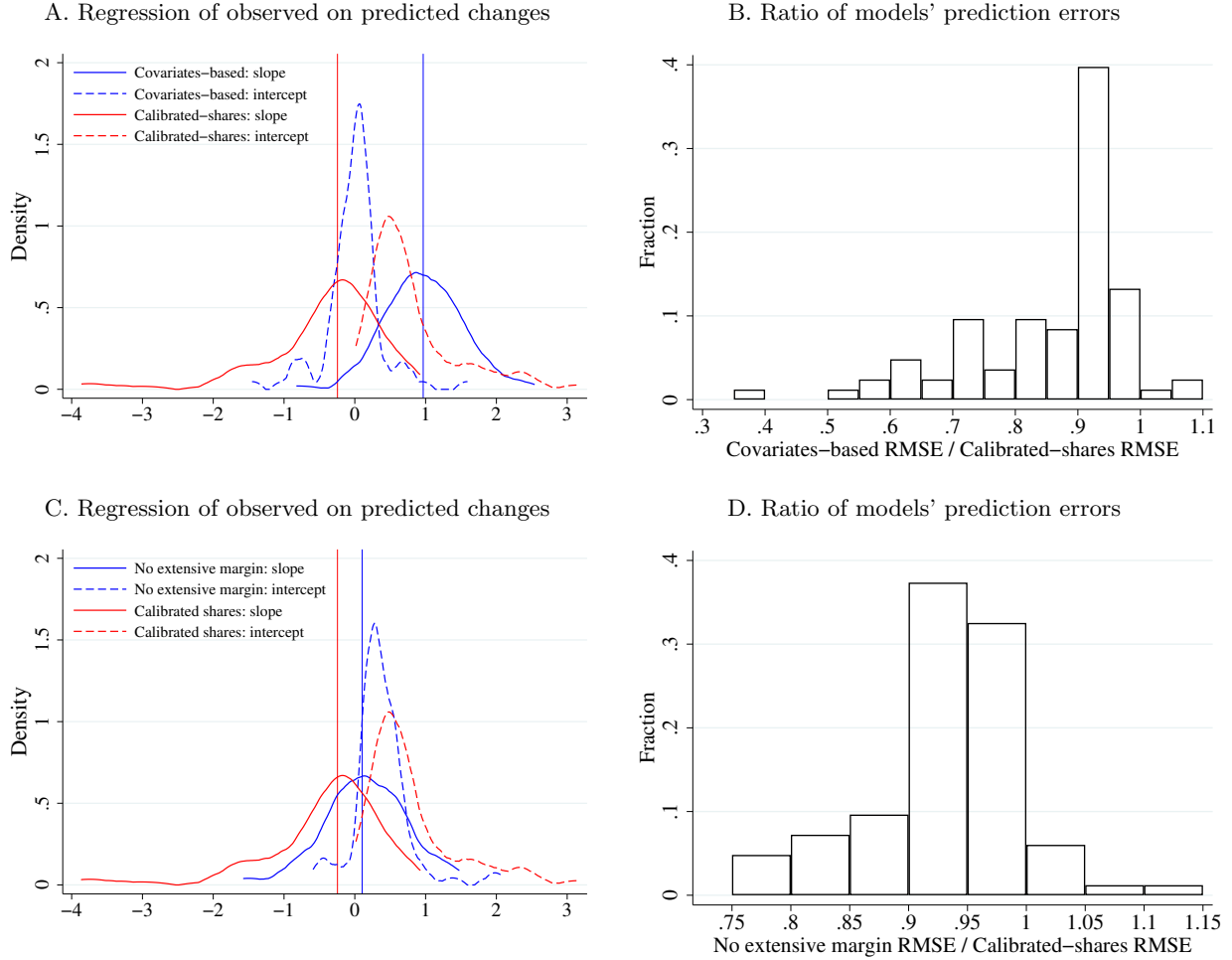
Figure 3 contrasts the two procedures' predictions for the changes in commuter flows associated with all 83 employment booms from 2010 to 2012. For each event and each procedure, we regress the observed change in the number of residents from each residential tract who work in the treated workplace tract on the predicted change. An unbiased prediction procedure should yield a slope of one and an intercept of zero. Panel A of Figure 3 depicts the distribution of these coefficients for both procedures. For the covariates-based approach, the slope coefficients are roughly centered on one (median of 0.97) and the intercept coefficients are roughly centered on zero (median of 0.01). The calibrated-shares procedure does not perform as well. Across the 83 events, the median slope coefficient is -0.25 and the median intercept coefficient is 0.60. That is, the calibrated-shares procedure's predictions are negatively correlated with observed outcomes in more than half of the events. As a result, the covariates-based approach typically has a lower forecast error. Panel B of Figure 3 contrasts the two models' RMSEs for each event.[22] The covariates-based approach has a lower RMSE than the calibrated-shares procedure in 80 of the 83 events. While the calibrated-shares procedure necessarily has a better in-sample fit, it performs poorly when predicting changes in commuting flows.

The covariate-based approach's superior predictive power is not due solely to the presence of zeros in the commuting matrix. To illustrate this, we produce predictions using a hybrid parameterization in which residence-workplace pairs that have zero commuters in the 2010 data are assigned infinite commuting costs rather than the covariates-based values. This model with no extensive margin produces predictions that have lower RMSEs than the predictions of the calibrated-shares procedure, as shown in panel D of Figure 3. Since the two procedures depicted in that panel handle zero-commuter flows identically, the superior predictive performance is a result of the covariate-based approach's intensive-margin predictions: it better predicts commuter-count

---

[21]The largest commuting flow in the 2010 LODES data for New York City is 827 commuters who reside between 110th and 114th Streets in Morningside Heights and work at adjacent Columbia University. The model that uses transit times predicts only 70 of the 827 observed commuters, failing to capture the effect of the university's dual role as employer and landlord. The advantage of the calibrated-shares procedure is its potential to capture such linkages.

[22]Comparing mean absolute errors (MAEs) yields the same conclusion. Characterizing forecast errors using RMSE or MAE is preferable to a measure such as R^2. As a squared correlation, R^2 would not distinguish between forecasts that are positively or negatively correlated with outcomes nor reveal a constant bias in a forecast. If expressed in terms of the sum of squared prediction errors relative to the variance of the observed changes, this "R^2" would be negative when the sum of squared prediction errors is larger, obviating the analogy with the in-sample fit of a linear regression.

Figure 3: Comparison of models' predictive performance across 83 events

A. Regression of observed on predicted changes

B. Ratio of models' prediction errors

C. Regression of observed on predicted changes

D. Ratio of models' prediction errors

NOTES: "Covariates-based approach" in panels A and B refers to the parameterization described in Section 3.2. "Calibrated shares" refers to the procedure defined in detail in Appendix A. "No extensive margin" in panels C and D is a hybrid parameterization in which $\delta_{kn} = \bar{\delta}_{kn}$ if $\ell_{kn} > 0$ in the 2010 data and $\delta_{kn} = \infty$ if $\ell_{kn} = 0$ in the 2010 data. Vertical lines depict medians of slope coefficient distributions.

changes for residence-workplace pairs that have positive counts in the baseline data.

As an alternative to using observed covariates, one might try to "smooth" granular data to be fed into the calibrated-shares procedure. For example, one might hope to eliminate idiosyncratic elements of the data by averaging across multiple years of data. In many empirical settings, particularly historical contexts, consecutive years of data are not available. In the case of the LODES data, annual data are available since 2002, and pooling may average out noise introduced by the confidentiality-protecting perturbations mentioned in Section 3.1. To explore the gain from pooling multiple years of data, we average commuter and wage observations for 2008–2010 before applying the observed-covariates approach and the calibrated-shares procedure.

As depicted in Figure 4, pooling the data yields a modest improvement for the calibrated-

Figure 4: Models' predictive performance across 83 events (pooled pre-event data)

A. Regression of observed on predicted changes

B. Ratio of models' prediction errors
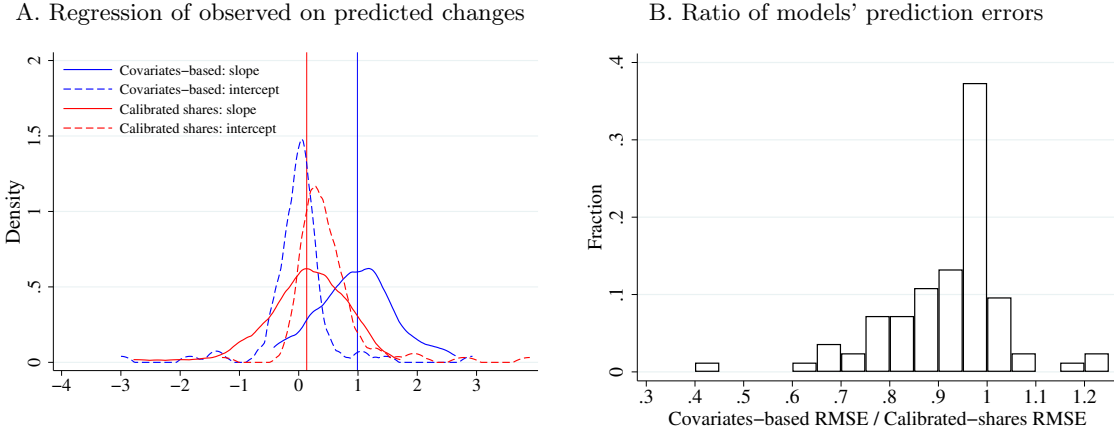
NOTES: In panel A, four outlying covariates-based observations and four outlying calibrated-shares observations are not depicted. Vertical lines depict medians of slope coefficient distributions.

shares procedure. Its predictions are now positively correlated with observed outcomes for 64% of the events. Nonetheless, the covariates-based approach typically forecasts the changes in commuter counts much better. The covariates-based approach's slope coefficients are closer to one (median of 0.99 vs. 0.14) and its intercept coefficients are closer to zero (0.01 vs. 0.39). The covariates-based approach applied to the pooled data has a lower RMSE than the pooled calibrated-shares procedure in 70 of the 83 events.

Aggregating predictions to larger spatial units has the potential to address the overfitting problem, but the covariates-based model still outperforms the calibrated-shares procedure after reducing the number of spatial units by an order of magnitude. In our view, analyzing a smaller number of locations so as to reduce the number of location *pairs* would be both costly and unnecessary. The cost is not addressing the many social science research questions that concern small spatial units like neighborhoods. That is unnecessary because the covariates-based approach reduces the dimensionality of the model parameters without incurring this cost. Regardless, when we aggregate the predicted changes in commuting flows to the destination tract using spatial units that are an order of magnitude larger, the covariates-based model still has smaller forecast errors in the majority of events. In particular, we aggregate the 2,160 residential tracts into 195 Neighborhood Tabulation Areas defined by the New York City Department of City Planning. As shown in Figure 5, we find that the covariates-based model has a lower RMSE than the calibrated-shares procedure in 55 of the 83 events.

One difference between the data requirements of the covariates-based approach and the calibrated-shares procedure is that the latter also uses data on wages. We have verified that using alternative wage values—such as equal wages across locations—does little to alter the predictions of the calibrated-shares procedure in the event studies.

The predictive power of the covariates-based approach could potentially be improved further.

17

Figure 5: Models' predictive performance across 83 events (larger spatial units)

A. Regression of observed on predicted changes

B. Ratio of models' prediction errors
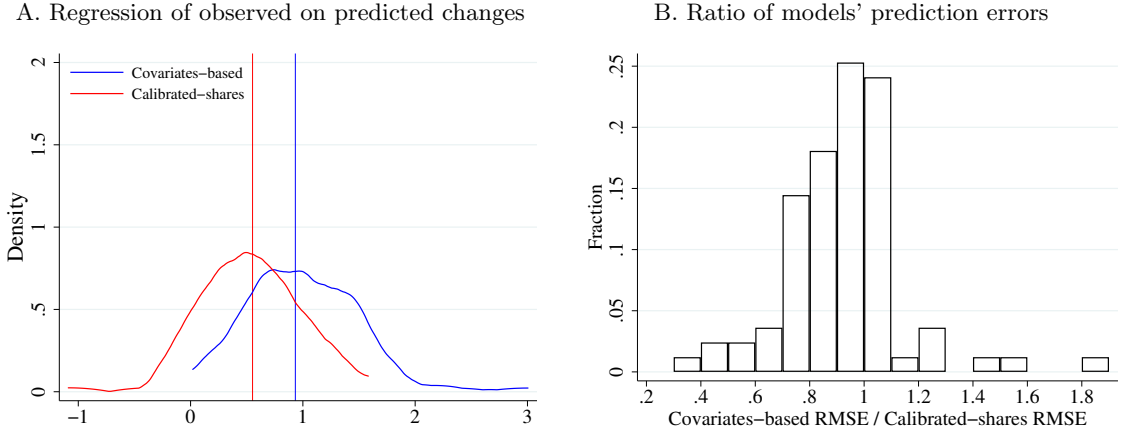
NOTES: This figure depicts predictive performance after aggregating the tract-level predictions underlying panels A and B of Figure 3 up to 195 Neighborhood Tabulation Areas defined by the New York City Department of City Planning. Panel A depicts the distributions of slope coefficients, and the vertical lines denote the median slope coefficients.

Recall that our parameterization uses only transit times to characterize commuting costs, leaving many unused degrees of freedom. Depending on the context, one could add borough-pair fixed effects, interactive fixed effects, or additional covariates such as the frequency of public-transit connections, vehicular transit times, and so forth. By contrast, the calibrated-shares procedure is the extreme case of introducing a structural error λ_{kn} for each pair of locations, thereby overfitting the observed data and failing to predict out of sample.

4 Granular model

In granular settings, individual choices affect economic outcomes. The previous section demonstrated that this granularity affects parameter estimates and therefore shapes models' counterfactual predictions. Even given parameter values, individual decisions affect both equilibrium quantities and prices in a granular setting. By contrast, the continuum model has deterministic equilibrium outcomes. To capture the role of the idiosyncratic components in individuals' decisions, one needs a model that admits a finite number of individuals.

The key challenge in modeling a finite number of people is precisely that individual decisions affect wages and rents. This raises two issues. First, do individuals internalize the effects of their own choices on local labor supplies and land demands? Second, are individuals able to enumerate the prices induced by every possible combination of others' choices? In the interest of tractability, we assume that individuals have common point-mass beliefs about wages and rents. Therefore, individuals act as price takers, as in the granular economy of Gabaix (2011). In our application, we assume that, given the model parameters, individuals have beliefs about wages and rents that are the equilibrium prices of the conventional continuum model. We label these beliefs "continuum-case

rational expectations" because they would be rational if there were a continuum of individuals.[23] In a granular economy, these point-mass beliefs are not rational because the realized prices of a granular equilibrium generically differ from them. In practice, we find that the resulting ex post regrets are small.

This section introduces a granular model with the same preferences, technology, and commuting costs as in the continuum model in Section 2. Model extensions that introduce trade costs, residential amenities, local increasing returns, and production employing land are presented in Appendix A. After defining equilibrium and discussing parameter estimation, we define ex post regrets and granular uncertainty in the context of our granular model.

4.1 Setup

The granular model differs from the continuum model in two important respects. First, we consider an economy populated by I individuals, each supplying $\frac{L}{I}$ units of labor. Second, we introduce a timing assumption: individuals first choose their residence-workplace pairs, and then all markets clear given individuals' residential and workplace locations. This timing assumption would be of no consequence if there were a continuum of individuals.

We make the following assumptions about information and expectations. All workers know the primitives of the model (L, $\{A_n\}$, $\{T_k\}$, $\{\delta_{kn}\}$, α, ϵ, σ) and have (common) expectations about the equilibrium variables r_k and w_n. We assume that the price expectations are point-mass beliefs, such that each individual assigns 100% probability to a single vector of wages and a single vector of land prices. Denote these belief vectors by $\{\tilde{w}_n\}$ and $\{\tilde{r}_k\}$. Worker i knows her idiosyncratic preference shocks $\{\nu_{kn}^i\}$.

Decisions are made and markets clear in the following order. Based on beliefs $\{\tilde{w}_n\}$ and $\{\tilde{r}_k\}$, each worker chooses the residential location and the work location that maximize expected utility

$$\tilde{U}_{kn}^i = \underbrace{\epsilon \ln \left(\frac{\tilde{w}_n}{\tilde{P}^{1-\alpha} \tilde{r}_k^\alpha \delta_{kn}} \right)}_{\equiv \tilde{U}_{kn}} + \nu_{kn}^i, \tag{6}$$

where $\tilde{P} = \left[\sum_n (\tilde{w}_n/A_n)^{1-\sigma} \right]^{1/(1-\sigma)}$. After these decisions are made, workers are immobile and cannot relocate.[24] Realized equilibrium land prices r_k and wages w_n are those that clear goods, labor, and land markets given individuals' residential and workplace locations.

The assumption that individuals have point-mass beliefs about wages, $\{\tilde{w}_n\}$, and rents, $\{\tilde{r}_k\}$,

---

[23]Similar simplifications to agents' beliefs have been made in industrial organization (e.g., oblivious equilibrium in Weintraub, Benkard, and Van Roy 2008) and macroeconomics (e.g., Krusell and Smith 1998).

[24]The assumption that individuals make irreversible decisions is common in static spatial models. For example, in the open-city model of Ahlfeldt et al. (2015), individuals choose to live in Berlin based on expected utility, which is equal to the reservation level of utility in the wider economy. Individuals who choose Berlin and realize utility below the city-wide average cannot leave. Brinkman and Lin (2019), Redding and Rossi-Hansberg (2017), Heblich, Redding, and Sturm (2020), and Owens, Rossi-Hansberg, and Sarte (2020) make the same irreversibility assumption. Similarly, Allen and Arkolakis (2014) and Tsivanidis (2019) assume that residences and workplaces are chosen sequentially.

considerably simplifies the analysis. Otherwise, individuals would need to compute the prices associated with all possible residence-workplace allocations and then solve for the fixed-point probabilities that each of these occurs. The set of feasible allocations is large. For example, a granular economy with only 10 individuals and 16 residence-workplace pairs would have more than 3 million possible allocations.[25] For empirically relevant magnitudes, computing the number of possible allocations leads to such large numbers that they cause overflow problems in standard software.

While our approach is general in the sense that it allows for any kind of point-mass beliefs about wages and rents, in our analysis below we assume that the expected prices are the equilibrium wages and rents of the continuum model with the same economic primitives (L, $\{A_n\}$, $\{T_k\}$, $\{\delta_{kn}\}$, α, ϵ, σ). The combination of these beliefs and our timing assumption is more palatable to the extent that individuals would not change their decisions if given the opportunity to re-optimize after observing realized prices. We compute the magnitude of "ex post regrets" in Section 4.4.

4.2 Equilibrium

For expositional clarity, we distinguish between a *trade equilibrium*, which clears markets taking individuals' locations as given, and a *granular commuting equilibrium*, in which a finite number of individuals choose their locations based on beliefs about the trade equilibrium that will result.

The trade equilibrium clears markets given individuals' location choices. Given a labor allocation $\{\ell_{kn}\}$, goods market clearing is identical to the continuum model and given by equation (3). Similarly, land market clearing is identical to the continuum model and given by equation (4).

Definition 4.1. Trade equilibrium. Given the labor allocation $\{\ell_{kn}\}$ and economic primitives (L, $\{A_n\}$, $\{T_k\}$, $\{\delta_{kn}\}$, α, ϵ, σ), a trade equilibrium is a set of wages $\{w_n\}$ and land prices $\{r_k\}$ such that equations (3) and (4) are satisfied.

Remark. The trade equilibrium is unique. As shown in Appendix A.3, given the labor allocation $\{\ell_{kn}\}$, there is a unique set of relative wages satisfying equation (3). Equation (4) can be rewritten as $r_k = \frac{\alpha}{T_k} \sum_n \frac{\ell_{kn}}{\delta_{kn}} w_n$, so there is a unique set of land prices associated with that wage vector.

Individuals' choices of location are made on the basis of their beliefs about the subsequent trade equilibrium. Given belief vectors $\{\tilde{w}_n\}$ and $\{\tilde{r}_k\}$ and the distribution of $\{\nu_{kn}^i\}$, equation (6) implies that the probability of any individual choosing kn as her residence-workplace pair is

$$\Pr(U_{kn}^i > U_{k'n'}^i \ \forall (k', n') \neq (k, n)) = \frac{\tilde{w}_n^\epsilon \left(\tilde{r}_k^\alpha \delta_{kn}\right)^{-\epsilon}}{\sum_{k',n'} \tilde{w}_{n'}^\epsilon \left(\tilde{r}_{k'}^\alpha \delta_{k'n'}\right)^{-\epsilon}}. \tag{7}$$

With these probabilities in hand, we define a granular commuting equilibrium as the labor allocation resulting from drawing I realizations of this process and clearing markets.

---

[25]With I individuals and N^2 residence-workplace pairs, the set of possible allocations (the support of the multinomial distribution) contains $\begin{pmatrix} I + N^2 - 1 \\ N^2 - 1 \end{pmatrix} = \frac{(I+N^2-1)!}{(N^2-1)!I!}$ elements. For $I = 10$ and $N = 4$, this is about 3.27×10^6.

Definition 4.2. Granular commuting equilibrium. Given a number of individuals I, economic primitives $(L, \{A_n\}, \{T_k\}, \{\delta_{kn}\}, \alpha, \epsilon, \sigma)$, and a set of point-mass beliefs $(\{\tilde{w}_n\}, \{\tilde{r}_k\})$, a granular commuting equilibrium is a labor allocation $\{\ell_{kn}\}$, wages $\{w_n\}$, and land prices $\{r_k\}$ such that

- $\ell_{kn} = \frac{L}{I} \sum_{i=1}^{I} \mathbf{1}\{\tilde{U}_{kn}^i > \tilde{U}_{k'n'}^i \ \forall (k', n') \neq (k, n)\}$ is the labor allocation resulting from I independent draws from the probability mass function in equation (7); and

- wages $\{w_n\}$ and land prices $\{r_k\}$ are a trade equilibrium, per Definition 4.1, given the labor allocation $\{\ell_{kn}\}$.

We now define a set of price beliefs that are the equilibrium prices of the continuum model from Section 2 with the same parameters for economic primitives.

Definition 4.3. Rational expectations for the continuum case. Given economic primitives $(L, \{A_n\}, \{T_k\}, \{\delta_{kn}\}, \alpha, \epsilon, \sigma)$, $\{\tilde{w}_n\}$ and $\{\tilde{r}_k\}$ are "continuum-case rational expectations" if $\{\tilde{w}_n\}$ and $\{\tilde{r}_k\}$ constitute a trade equilibrium for the labor allocation $\{\ell_{kn}\}$ given in equation (2).

Distinguishing between the number of individuals I and the aggregate labor supply L allows us to apply the law of large numbers to locational decisions as $I \to \infty$ without changing aggregate labor supply. As Appendix A.4 shows, since I is the number of draws from the probability mass function in equation (7), fixing L and taking $I \to \infty$ causes the labor allocation in the limiting case to be the same as in the continuum model in equation (2).

For finite I, a given set of economic primitives $(L, \{A_n\}, \{T_k\}, \{\delta_{kn}\}, \alpha, \epsilon, \sigma)$ and point-mass beliefs $(\{\tilde{w}_n, \tilde{r}_k\})$ generate a *distribution* of equilibria associated with realizations of the process generating ν_{kn}^i and hence $\{\ell_{kn}\}$.[26] Importantly, the outcomes for wages and rents are random variables in the granular model, whereas they are deterministic in the continuum model.

4.3 Estimation

The granular model can be estimated similarly to the covariates-based approach to estimating the continuum model. The log likelihood function for the model derived from equation (7) is

$$\ln \mathcal{L} \equiv \sum_{k,n} \ell_{kn} \ln \left[\Pr(U_{kn}^i > U_{k'n'}^i \ \forall (k', n') \neq (k, n)) \right] = \sum_{k,n} \ell_{kn} \ln \left[\frac{\tilde{w}_n^\epsilon \left(\tilde{r}_k^\alpha \bar{\delta}_{kn} \right)^{-\epsilon}}{\sum_{k',n'} \tilde{w}_{n'}^\epsilon \left(\tilde{r}_{k'}^\alpha \bar{\delta}_{k'n'} \right)^{-\epsilon}} \right]. \quad (8)$$

This matches equation (5), except these residence and workplace fixed effects depend on rent and wage beliefs rather than the realized equilibrium rents and wages. The maximum-likelihood estimate of the commuting elasticity ϵ is identical to that in the covariates-based continuum model (in column 1 of Table 1). The remaining economic primitives $\{T_k\}$ and $\{A_n\}$ can be inferred from the maximum-likelihood estimates using assumptions about the rent and wage beliefs. Given values

---

[26]Redding and Rossi-Hansberg (2017, p.38) anticipated this implication of granularity: "At smaller spatial scales (e.g., blocks within cities), one might expect such random idiosyncratic factors to be more important relative to the systematic deterministic components of a model (e.g., natural resource abundance) than at larger spatial scales (e.g., across regions or countries)."

of the commuting elasticity ϵ and the land expenditure share α, individuals' beliefs about land prices and wages are transformations of the estimated fixed effects. In particular, the residence fixed effect is proportionate to $\tilde{r}_k^{-\alpha\epsilon}$ and the workplace fixed effect is proportionate to $\tilde{w}_n^{\epsilon}$. If $\{\tilde{r}_k\}$ and $\{\tilde{w}_n\}$ are continuum-case rational expectations, plugging these estimated beliefs into equation (2) yields the continuum-case labor allocation. Given α, σ, $\{\tilde{r}_k\}$, $\{\tilde{w}_n\}$, $\{\delta_{kn}\}$ and that continuum-case labor allocation, equations (3) and (4) can be solved to yield values of $\{T_k\}$ and $\{A_n\}$. The estimated land and productivity parameters are identical to the estimates from the covariates-based continuum model.

4.4   Ex post regret

The granular model's equilibrium rents and wages generically differ from the point-mass beliefs about rents and wages that govern individuals' choices of residences and workplaces. Figure E.1 in Appendix E presents histograms of the standard deviation of wages and rents across 100,000 simulations of our estimated model of New York City divided by the continuum-case rational expectations of these prices. Across tracts, the median level of dispersion for wages is 0.016 and for rents is 0.032. While this dispersion may be considered sizable, these differences between expectations and realized prices do not necessarily imply that individuals would move if we relaxed the irreversibility assumption. Individuals' decisions depend on their idiosyncratic preference shocks, so many are inframarginal.

Using our estimated model of New York City, we investigate how often individuals would choose a different residence-workplace pair at the realized equilibrium prices. We define the magnitude of "ex post regret" as the increase in income an individual would require as compensation to not change their choice given the realized equilibrium rents and wages. In particular, at realized prices r and w, for individual i who chose residence-workplace pair kn, ex post regret χ_i is defined implicitly such that χ_i solves

$$\max_{k',n'}\left(\epsilon\ln\left(\frac{w_{n'}}{P^{1-\alpha}r_{k'}^{\alpha}\delta_{k'n'}}\right)+\nu_{k'n'}^{i}\right)=\left(\epsilon\ln\left(\frac{(1+\chi_i)w_n}{P^{1-\alpha}r_k^{\alpha}\delta_{kn}}\right)+\nu_{kn}^{i}\right),$$

where the left side is the individual's maximum utility at realized prices, the right side is the individual's utility from kn with their income multiplied by $(1+\chi_i)$, and by definition kn was the choice maximizing $\tilde{U}_{k,n}^{i}$. Individual i's regret is zero if kn is their optimal choice at realized prices.

In this setting, we find that ex post regret is quantitatively modest. Appendix E reports results for 10 simulations.[27] In those simulations, 96% of individuals would not want to change their residence-workplace choice. For the 4% who would want to switch, the median ex post regret χ_i is equl to 0.7%. Given the large return to making the model tractable and computationally feasible, we judge the magnitude of ex post regret when using continuum-case rational expectations to be

---

[27]As discussed in Appendix E, these particular simulations are far more computationally expensive than simulating the granular model, since they require explicitly simulating ν_{kn}^{i}. Table E.1 shows that the distribution of ex post regret is very similar across the 10 simulations.

quite modest.

4.5 Counterfactual predictions

A fundamental difference between the model with a finite number of individuals and the continuum model is that individual idiosyncrasies affect equilibrium outcomes in the granular model. Given economic primitives $(L, \{A_n\}, \{T_k\}, \{\delta_{kn}\}, \alpha, \epsilon, \sigma)$, the granular model produces a distribution of outcomes associated with the distribution of idiosyncratic preference shocks, whereas the continuum model delivers deterministic outcomes. We first compare and contrast the granular model's mean outcomes with those of the covariates-based continuum model and then turn to the consequences of the dispersion in equilibrium outcomes.

Given the same parameter values, the mean equilibrium labor allocation of the granular model is equal to the equilibrium labor allocation of the covariates-based continuum model. The multinomial distribution of residence-workplace outcomes $\{\ell_{kn}\}$ defined by the probabilities in equation (7) has expected values equal to the quantities in equation (2). By contrast, the granular model's equilibrium prices solve a system of non-linear equations, so their mean values are not necessarily equal to the continuum-case prices. However, in the estimated model for New York City, we find that the granular model's mean wages and rents are reasonably close to the equilibrium wages and rents of the continuum model with the same fundamental parameters.

Using the granular model, one can calculate the dispersion in predicted counterfactual equilibrium outcomes due to idiosyncratic preference shocks. While we consider the *ex ante* outcome to be observed (i.e., either the mean outcome across many simulations of the granular model or the covariates-based continuum model outcome), the *ex post* outcome (at counterfactual parameter values) depends on individual idiosyncrasies.[28] For example, when considering the predicted counterfactual change in residents in k (i.e., $\sum_n \ell'_{kn} - \sum_n \ell_{kn}$), the counterfactual outcome is a random variable. We can characterize this granular uncertainty by simulating the granular model at counterfactual parameter values to produce confidence intervals for outcomes of interest.[29]

While calculating these confidence intervals requires many simulations, there is a simple approximation of granular uncertainty for the labor allocation that is trivial to compute and valid when the empirical setting is not "too granular." This approximation only requires probabilities that researchers already compute when predicting counterfactual outcomes. In the granular model, the share of individuals who choose residence-workplace pair kn, ℓ_{kn}/L, is a realization of I draws from a multinomial distribution over $K \times N$ outcomes with probabilities given by equation (7). Let ℓ'_{kn} denote a counterfactual labor allocation, and let p'_k denote the counterfactual probability of choosing any pair in which the residence is k. The expected change in labor residing in k is

---

[28]The estimated residence and workplace fixed effects cause the mean numbers of residents and workers to match the initial observed values.

[29]In principle, one might hope to use the continuum model as a data-generating process from which one could draw finite-sample realizations, as we did for the labor allocation in our Monte Carlo simulations in Section 3.3. This involved interpreting the share on the right side of equation (2) as a probability and sampling from this multinomial distribution. However, equations (2), (3), and (4) in the continuum model do not define a probability distribution for prices.

$p'_k L - \sum_n \ell_{kn}$. The standard deviation of the labor residing in k, $\sum_n \ell'_{kn}$, is $\frac{L}{\sqrt{I}}\sqrt{p'_k(1-p'_k)}$, and thus the standard deviation of the change from the initial observed equilibrium, $\sum_n \ell'_{kn} - \sum_n \ell_{kn}$, when $I = L$ is $\mathfrak{s}'_k \equiv \sqrt{I}\sqrt{p'_k(1-p'_k)}$. Thus, we can approximate the 90% confidence interval for the change in the number of residents in tract k in the $I = L$ case by $p'_k L - \sum_n \ell_{kn} \pm 1.645\mathfrak{s}'_k$. Using the normal distribution to approximate the distribution of outcomes for tract k works well when $I \times p'_k$ is sufficiently large, as turns out to be the case in our application to New York City. When the expected value of the outcome is small, the approximation may suggest negative numbers of people in counterfactual equilibria. These obviously infeasible outcomes would be evidence that the empirical setting is so granular that the normal approximation is inappropriate, warranting further investigation of the granular uncertainty.

To produce confidence intervals for equilibrium rents and wages, one must repeatedly simulate the granular model. These realized prices depend on outcomes in all other locations, per equations (3) and (4). While producing these confidence intervals for prices therefore requires more computations than producing the confidence intervals for quantities using the above approximation, simulating a granular labor allocation and solving for equilibrium prices can be done quite quickly.

Note that small ex post regrets do not necessarily imply that the granular uncertainty accompanying counterfactual predictions will be small. Ex post regrets concern dispersion in realized prices in levels. Granular uncertainty accompanying counterfactual changes concerns changes in prices, which will be fractions of their initial levels in the typical counterfactual scenario. Thus, one may find meaningfully large granular uncertainty accompanying counterfactual changes despite quantitatively modest ex post regrets.

5 Application to Amazon's HQ2 in Long Island City

We now examine the predictions of quantitative spatial models for the economic consequences of Amazon's aborted second headquarters (HQ2) in Long Island City, a controversial and widely discussed proposal. We focus on understanding the different procedures' contrasting predictions and investigating the magnitude of granular uncertainty accompanying these predictions rather than conducting an exhaustive examination of the potential effects of Amazon's HQ2. We therefore employ the assumptions on preferences and technologies made in Section 2, which abstract from productivity spillovers or skill heterogeneity among workers.[30]

In September 2017, Amazon requested proposals for HQ2 locations from cities and states across North America. Amazon received 238 proposals in a year-long competition. In November 2018, the company announced two winners, saying that it would hire more than 25,000 employees each in Long Island City in New York and Arlington in Virginia (Amazon, 2018). The former was to have 4 million square feet office space, with anticipated expansion opportunities for another 4 million square feet and 15,000 more workers.[31] However, Amazon scrapped the project in February

---

[30]Berkes and Gaetani (2020) introduce a model that features these mechanisms and analyze Amazon HQ2.

[31]More details can be found in the Memorandum of Understanding between New York State, New York City and Amazon.

2019 after facing a fierce backlash from local politicians and community members concerned about corporate subsidies and gentrification.

We compare and contrast predictions about this counterfactual scenario produced by the granular model and the calibrated-shares procedure. As noted previously, the granular model's expected change for a quantity equals the change predicted by the covariates-based continuum model with the same parameter values. Section 5.1 therefore reports the predicted changes in quantities produced by the covariates-based continuum model and the mean predicted changes in prices produced by the (covariates-based) granular model. Section 5.2 presents the granular uncertainty accompanying predicted changes, which only arises in the granular model.

5.1 Predicting (mean) counterfactual outcomes

What general-equilibrium effects of Amazon HQ2 would quantitative spatial models predict? Using both the covariates-based and calibrated-shares procedures, we analyze the consequences of a productivity increase in the Long Island City tract that would cause employment to rise by 25,000 workers from the 2010 baseline levels.[32] Figure 6A depicts New York City employment by tract in 2010, which is particularly concentrated in lower Manhattan and Queens, near the center of the map. Figure 6B depicts the distribution of residences among those workers employed in the treated Long Island City tract in 2010. Notably, 49.3% of residential tracts have zero residents employed at that workplace tract. The calibrated-shares procedure rationalizes those zeros with infinite commuting costs.

Whether Amazon HQ2 is predicted to cause large changes in neighborhood residential populations depends on the method used: while the covariates-based approach predicts small changes in all tracts, the calibrated-shares procedure predicts changes that vary more and are less spatially correlated. The covariates-based model's predicted changes in the number of residents in each tract, shown in Figure 7A, are modest in size and closely related to transit times to Long Island City. The calibrated-shares procedure's predicted changes, shown in Figure 7B, are very large for some residential tracts. The calibrated-shares procedure predicts increases in residents in neighborhoods scattered across Queens and Staten Island, but very little increase in residents in Manhattan neighborhoods with short commutes to Long Island City. These predictions are closely tied to the initial numbers of residents working in the treated tract: the spatial patterns in Figure 6B and Figure 7B are remarkably similar.[33] On the whole, the calibrated-shares procedure suggests that 49% of the workers in Long Island City would reside in Queens, while the covariates-based model suggests that only 33% of workers at Amazon's HQ2 would reside in Queens.

The incidence of Amazon HQ2 on residential housing prices predicted by the two methods differs considerably. The greater variation in changes in the number of residents predicted by the calibrated-shares procedures is accompanied by greater variation in the predicted changes in rents.

---

[32] Following Berkes and Gaetani (2020), the "treated" tract is 36081000700.

[33] Figure F.1 in Appendix F depicts the close relationship between the initial number of residents working in the treated tract and the calibrated-shares procedure's predicted change in residents.

Figure 6: New York City employment in 2010

A. Number of workers

B. Number of residents working at the treatment tract
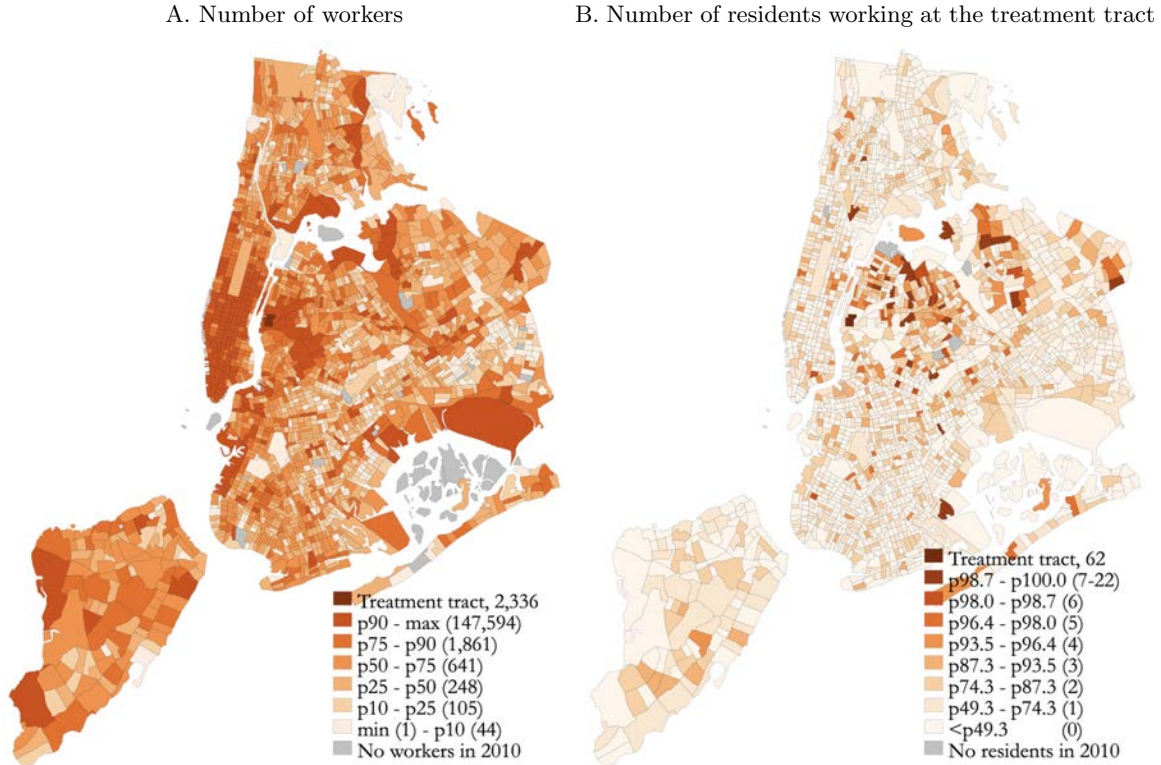
NOTES: Panel A depicts the number of workers employed in each tract. Panel B depicts the number of residents in each tract who work in the treatment tract in the LODES 2010 data. The legend of panel B reports the percentiles corresponding to the integer number of residents, excluding the treatment tract.
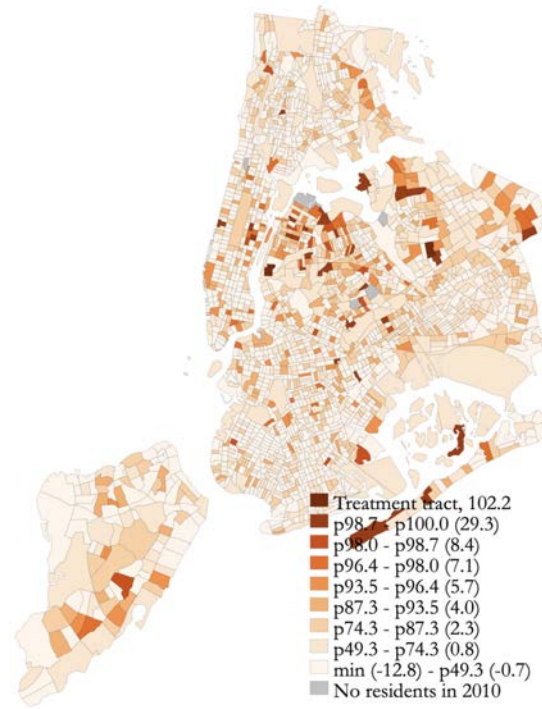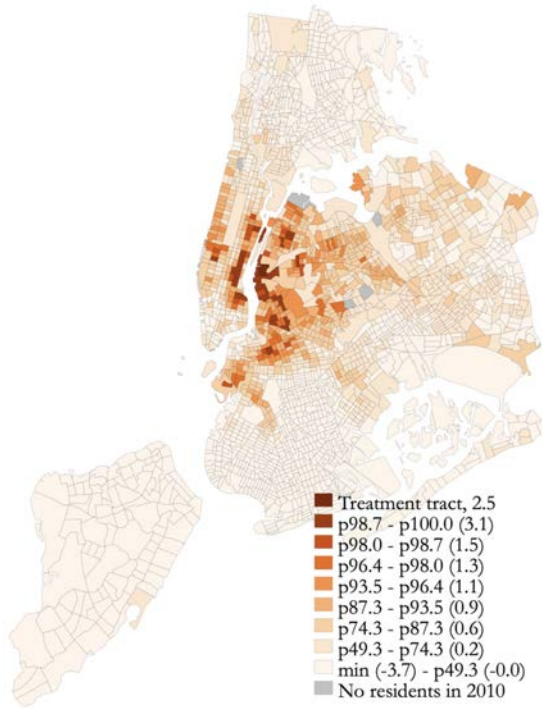
The calibrated-shares procedure predicts rent increases of 1.6% to 10.4% for the top decile of tracts (Figure 8B). By contrast, the (covariates-based) granular model predicts more modest increases of 0.8% to 1.3% for the top decile of affected tracts (Figure 8A). The geographic incidence differs notably. The calibrated shares-procedure predicts much larger rent increases in neighborhoods far from Amazon HQ2, such as Staten Island. By contrast, the political opposition to Amazon HQ2 based on concerns about gentrification largely concerned housing prices in nearby constituencies (Goodman, 2019).

Amazon HQ2 affects the predicted wages and number of workers in other workplaces through substitution and general-equilibrium effects, which are similar across the two methods in this application. The change in the number of workers by tract is depicted in Figures 7C and 7D. The two methods' predictions are very similar and are closely tied to the initial number of workers employed in each tract, depicted in Figure 6A. Relatedly, the covariates-based and calibrated-shares procedures predict somewhat similar real wage changes, all less than 1% for the non-treated tracts (Figures 8C and 8D).

Figure 7: Predicted change in the number of residents or workers

A. Change in residents, covariates-based

B. Change in residents, calibrated-shares

C. Decrease in workers, covariates-based

D. Decrease in workers, calibrated-shares
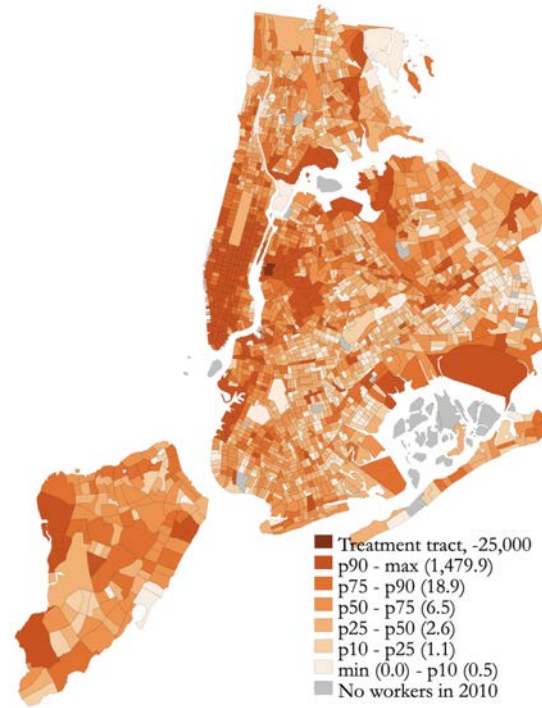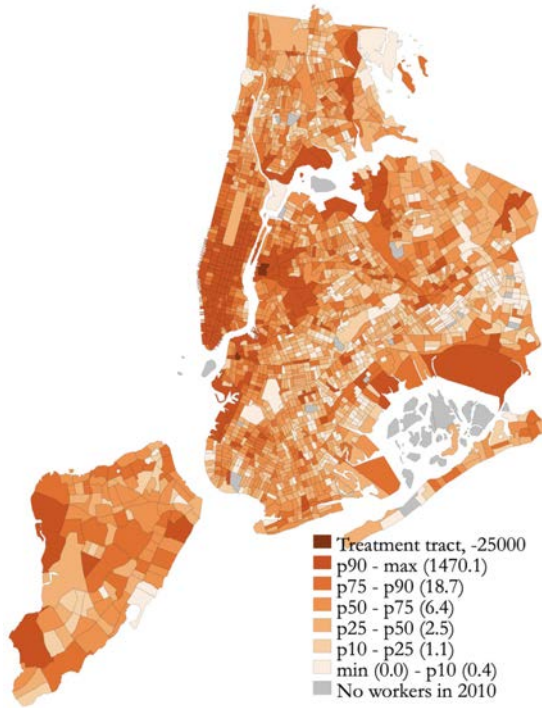
NOTES: Panels A and B depict changes in the number of residents predicted by the covariates-based model and calibrated-shares procedure, respectively. Panels C and D depict decreases in the number of workers predicted by the covariates-based model and calibrated-shares procedure, respectively. As total population is fixed, the number of workers decreases in all tracts except the Amazon HQ2 location. The legend percentile cutoffs in panels A and B correspond to those in Figure 6B, those in panels C and D to Figure 6A.

Figure 8: Predicted changes in prices

A. Rent changes, covariates-based

B. Rent changes, calibrated-shares procedure
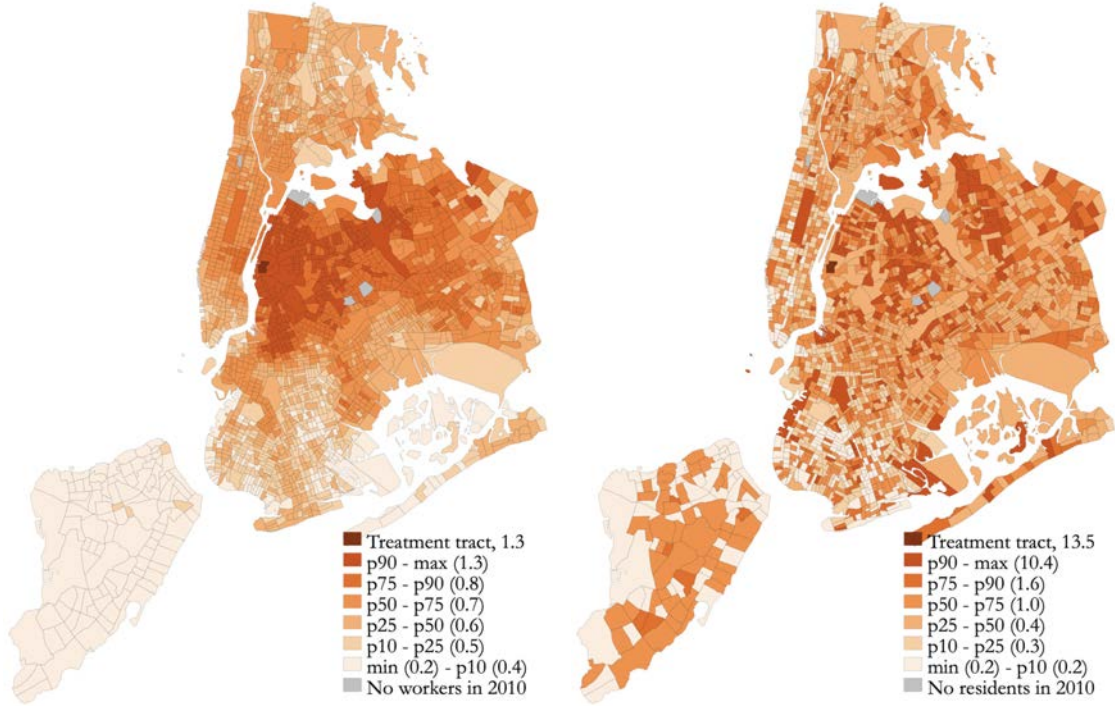
C. Wage changes, covariates-based

D. Wage changes, calibrated-shares procedure
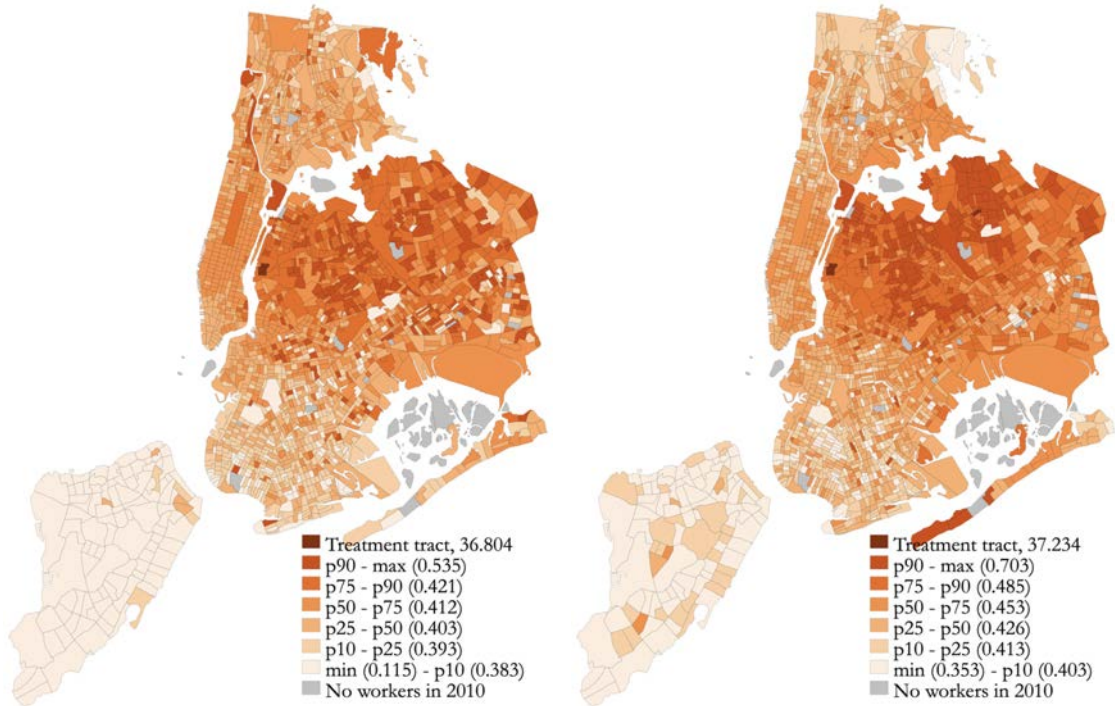
NOTES: These maps depict percentage-point changes in prices. Panels A and C depict the mean price across 100,000 simulations at the new parameter values relative to the mean price across 100,000 simulations at the old parameters. Panels B and D depict each tract's predicted price change using the calibrated-shares procedure.

5.2 Granular uncertainty accompanying counterfactual predictions

Counterfactual predictions come with various forms of uncertainty. For example, to capture uncertainty about the parameters used to produce counterfactual predictions, researchers can bootstrap the counterfactual predictions of a continuum model by sampling from the distribution of estimated parameters. In addition to parameter uncertainty, one important implication of granularity is that there is a distribution of possible equilibrium outcomes associated with the distribution of individuals' idiosyncratic preference shocks. That is, counterfactual predictions are subject to granular uncertainty even if the model parameters (L, $\{A_n\}$, $\{T_k\}$, $\{\delta_{kn}\}$, α, ϵ, σ) are known. The magnitude of this granular uncertainty relative to the predicted counterfactual changes depends on both the empirical setting and the counterfactual shock examined. We consider the creation of 25,000 jobs in one tract to be a large shock, as Amazon's HQ2 choice was a high-profile corporate decision that elicited numerous proposals from local policymakers. Are the counterfactual changes in outcomes across different parts of New York City large relative to the granular uncertainty driven by the idiosyncratic components of individuals' location decisions?

The changes in residents predicted by the granular model are small relative to the granular uncertainty in this setting. As shown in Figure 7A, the changes predicted by the covariates-based approach overwhelmingly involve fewer than three residents. For the number of residents in the typical census tract, the normal approximation yields a 90% confidence interval with a width of about 112 residents.[34] In this application, we find that the normal approximation is very similar to the simulated distribution of outcomes, as shown in Figure F.2 in Appendix F. As Figure 9A shows, the magnitude of granular uncertainty varies across residential tracts due to variation in p'_k, but every tract's confidence interval has a positive upper bound and a negative lower bound. In this sense, none of the changes in number of residents predicted by the granular model are distinguishable from zero. Figure 9B plots the simulated 90% confidence interval for the predicted change in prices from 100,000 simulations against the mean predicted change. For all residential tracts, the confidence interval for the predicted change in land rent includes zero.

Since the predicted changes in the number of workers for some tracts are larger, some of them are meaningfully large relative to the magnitude of granular uncertainty. As the number of workers varies across tracts much more than the number of residents, the width of the approximated confidence intervals also varies much more. Figure 9C depicts these approximated confidence intervals for the changes in the number of workers. For tracts predicted to lose more than 250 workers, the predicted change is sufficiently large relative to granular uncertainty that even the 95th percentile of changes is negative. For most workplaces, however, the approximated 90% confidence interval includes both negative and positive values. While the mean predicted wage change is less than 1% for all workplaces other than the Amazon HQ2 location, some of these predicted changes are distinguishable from zero (Figure 9D). There are 43 tracts with sufficiently small granular uncertainty accompanying the predicted wage change that the 5th percentile change is greater than

---

[34]Given 2,160 census tracts with similar residential populations, $p'_k \approx \frac{1}{2160}$ and $2 \times 1.645 \times \sqrt{I}\sqrt{\frac{1}{2160} \times \frac{2159}{2160}} \approx 112$.

Figure 9: Dispersion of predicted changes in the granular model

A. Change in residents

B. Predicted rent changes

C. Change in workers

D. Predicted wage changes
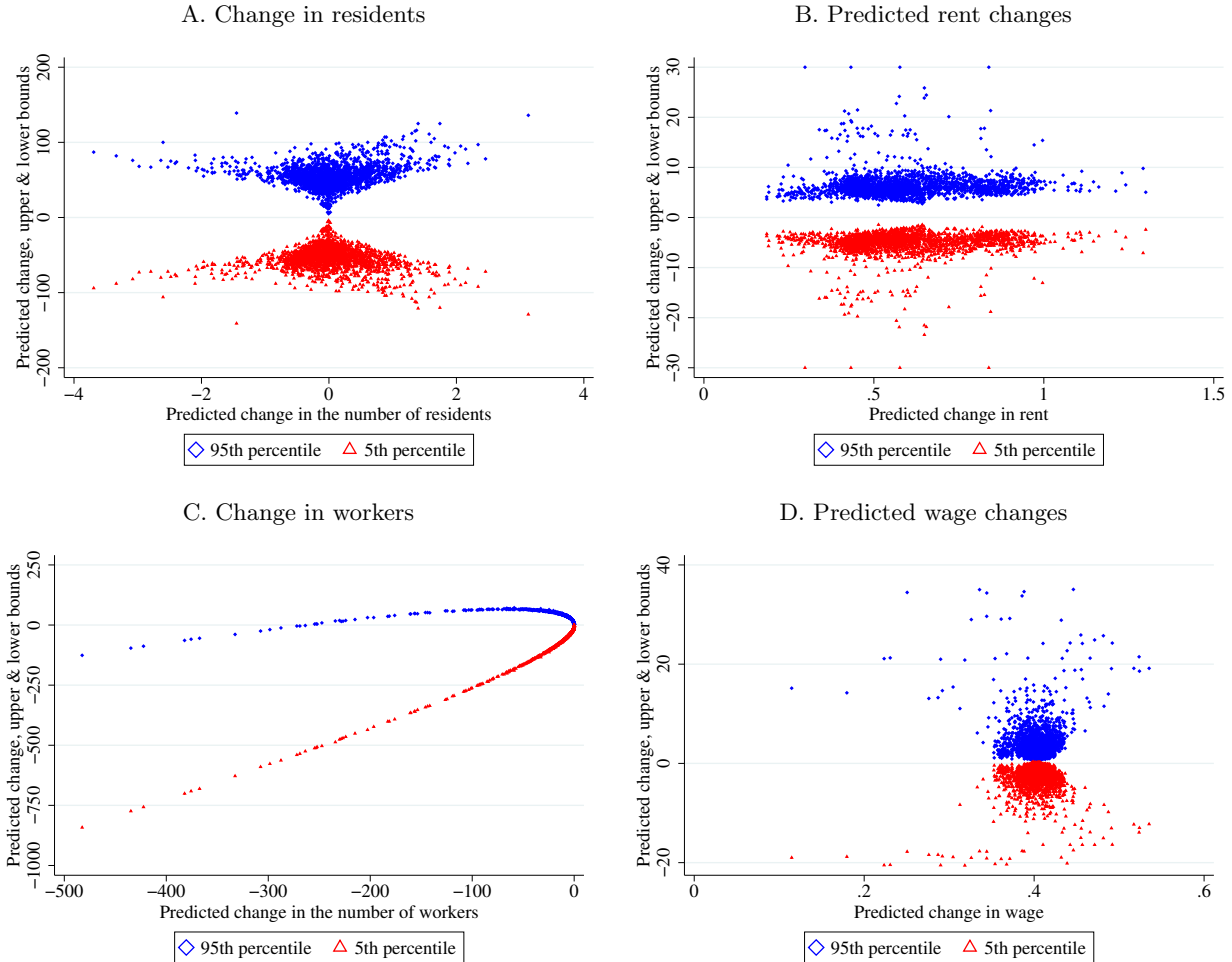
NOTES: The plots depict the 5th and 95th percentiles of predicted changes in quantities and prices across 100,000 simulations of the granular model. In panel B, upper bounds are winsorized at 30, and lower bounds are winsorized at -30. In panel C, the treated tract and one outlying tract are excluded. For the outlier tract, the granular model predicts a change of -1470 ± 612. In panel D, there are 43 tracts in which the 5th percentile wage change is greater than zero. In that panel, the treated tract is excluded.

zero. Overall, the granular uncertainty involved in studying neighborhood-level outcomes is large relative to the predicted consequences of the most hotly followed corporate headquarters decision in recent memory. This notable uncertainty would be missed by the conventional continuum approach. Given the magnitude of granular uncertainty that we document in the case of Amazon's HQ2, we believe granular uncertainty is likely relevant in many empirical applications.

6 Conclusion

Economists increasingly have access to spatially fine data and use quantitative spatial models to compute counterfactual general-equilibrium outcomes. The smaller the number of individuals

behind each economic outcome reported in these data, the less compelling the conventional modeling assumption that there is a continuum of individuals. We need to evaluate the performance of these applied general-equilibrium models in predicting observed changes if policymakers are going to rely on them to inform their decisions (Kehoe, 2005; Bryan, Glaeser, and Tsivanidis, 2019; Adão, Arkolakis, and Esposito, 2020). We document that empirical settings of interest exhibit granular commuting matrices and discuss the consequences of such granularity for predicting counterfactual outcomes. In short, applying the conventional calibrated-shares procedure to a granular setting overfits the model, causing it to fail to predict out-of-sample outcomes in both Monte Carlo simulations and event studies.

More broadly, we recommend that researchers use Monte Carlo simulations to assess the finite-sample behavior of their procedures for producing counterfactual predictions. Since all data sets are finite, in practice granularity is a matter of degree, not kind. A simulation specific to the empirical setting of interest can quantify how much the procedure may overfit the granular data it employs.

We propose a granular model in which equilibrium outcomes depend in part on the idiosyncratic component of individuals' choices. This estimated model yields a distribution of equilibrium outcomes when computing counterfactual scenarios. Evaluating the counterfactual consequences of Amazon's proposed HQ2 in Long Island City, we find that most of the predicted changes are an order of magnitude smaller than the accompanying granular uncertainty.

Since all empirical settings feature a finite number of individuals, the continuum assumption has been made in the interest of modeling convenience, not realism. Our granular model is equally tractable and leverages the same data as the continuum approach, so it can be applied in the same settings in which economists have thus far assumed a continuum.

References

Adão, Rodrigo, Costas Arkolakis, and Federico Esposito. 2020. "General Equilibrium Effects in Space: Theory and Measurement."

Ahlfeldt, Gabriel M., Stephen J. Redding, Daniel M. Sturm, and Nikolaus Wolf. 2015. "The Economics of Density: Evidence From the Berlin Wall." *Econometrica*, 83(6): 2127–2189.

Allen, Treb, and Costas Arkolakis. 2014. "Trade and the Topography of the Spatial Economy." *The Quarterly Journal of Economics*, 129(3): 1085–1140.

Allen, Treb, Costas Arkolakis, and Xiangliang Li. 2015. "Optimal City Structure."

Allen, Treb, Costas Arkolakis, and Xiangliang Li. 2020. "On the Equilibrium Properties of Network Models with Heterogeneous Agents." National Bureau of Economic Research Working Paper 27837.

Amazon. 2018. "Amazon selects New York City and Northern Virginia for new headquarters."

Armenter, Roc, and Miklós Koren. 2014. "A Balls-and-Bins Model of Trade." *American Economic Review*, 104(7): 2127–51.

Arzaghi, Mohammad, and J. Vernon Henderson. 2008. "Networking off Madison Avenue." *Review of Economic Studies*, 75(4): 1011–1038.

Autor, David H., David Dorn, and Gordon H. Hanson. 2013. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *American Economic Review*, 103(6): 2121–68.

Berkes, Enrico, and Ruben Gaetani. 2020. "Income Segregation and Rise of the Knowledge Economy."

Brinkman, Jeffrey, and Jeffrey Lin. 2019. "Freeway Revolts!" Federal Reserve Bank of Philadelphia Working Papers 19-29.

Bryan, Gharad, Edward Glaeser, and Nick Tsivanidis. 2019. "Cities in the Developing World." National Bureau of Economic Research Working Paper 26390.

Carvalho, Vasco M., and Basile Grassi. 2019. "Large Firm Dynamics and the Business Cycle." *American Economic Review*, 109(4): 1375–1425.

Correia, Sergio, Paulo Guimarães, and Tom Zylkin. 2020. "Fast Poisson estimation with high-dimensional fixed effects." *The Stata Journal*, 20(1): 95–115.

Costinot, Arnaud, and Andrés Rodríguez-Clare. 2014. "Trade Theory with Numbers: Quantifying the Consequences of Globalization." Vol. 4 of *Handbook of International Economics*, Chapter 4, 197–261. Elsevier.

Couture, Victor, and Jessie Handbury. 2020. "Urban revival in America." *Journal of Urban Economics*, 119: 103267.

Couture, Victor, Jonathan I. Dingel, Allison E. Green, Jessie Handbury, and Kevin R. Williams. 2020. "Measuring Movement and Social Contact with Smartphone Data: A Real-Time Application to COVID-19." National Bureau of Economic Research Working Paper 27560.

Daruich, Diego, William Easterly, and Ariell Reshef. 2019. "The surprising instability of export specializations." *Journal of Development Economics*, 137: 36 – 65.

Davis, Donald R., Jonathan I. Dingel, Joan Monras, and Eduardo Morales. 2019. "How Segregated Is Urban Consumption?" *Journal of Political Economy*, 127(4): 1684–1738.

Davis, Morris A., and Francois Ortalo-Magne. 2011. "Household Expenditures, Wages, Rents." *Review of Economic Dynamics*, 14(2): 248–261.

Dekle, Robert, Jonathan Eaton, and Samuel Kortum. 2008. "Global Rebalancing with Gravity: Measuring the Burden of Adjustment." *IMF Staff Papers*, 55(3): 511–540.

di Giovanni, Julian, and Andrei A. Levchenko. 2012. "Country Size, International Trade, and Aggregate Fluctuations in Granular Economies." *Journal of Political Economy*, 120(6): 1083–1132.

di Giovanni, Julian, Andrei A. Levchenko, and Isabelle Mejean. 2014. "Firms, Destinations, and Aggregate Fluctuations." *Econometrica*, 82(4): 1303–1340.

Dingel, Jonathan I. 2017. "The Determinants of Quality Specialization." *Review of Economic Studies*, 84(4): 1551–1582.

Dingel, Jonathan I, Antonio Miscio, and Donald R Davis. 2019. "Cities, Lights, and Skills in Developing Economies." National Bureau of Economic Research Working Paper 25678.

Donaldson, Dave. 2015. "The Gains from Market Integration." *Annual Review of Economics*, 7(1): 619–647.

Donaldson, Dave, and Adam Storeygard. 2016. "The View from Above: Applications of Satellite Data in Economics." *Journal of Economic Perspectives*, 30(4): 171–98.

Dubé, Jean-Pierre H, Ali Hortaçsu, and Joonhwi Joo. 2020. "Random-Coefficients Logit Demand Estimation with Zero-Valued Market Shares." National Bureau of Economic Research Working Paper 26795.

Eaton, Jonathan, Samuel Kortum, and Sebastian Sotelo. 2013. "International Trade: Linking Micro and Macro." In *Advances in Economics and Econometrics: Tenth World Congress*. Vol. III, , ed. Eddie Dekel Daron Acemoglu, Manuel Arellano. Cambridge University Press.

Ellison, Glenn, and Edward L Glaeser. 1997. "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach." *Journal of Political Economy*, 105(5): 889–927.

Gabaix, Xavier. 2011. "The Granular Origins of Aggregate Fluctuations." *Econometrica*, 79(3): 733–772.

Gandhi, Amit, Zhentong Lu, and Xiaoxia Shi. 2019. "Estimating Demand for Differentiated Products with Zeroes in Market Share Data."

Gaubert, Cecile, and Oleg Itskhoki. 2018. "Granular Comparative Advantage." National Bureau of Economic Research Working Paper 24807.

Goodman, J. David. 2019. "Why Amazon Is Caught in an Unexpected Brawl in New York." *New York Times*, 12 February: https://www.nytimes.com/2019/02/12/nyregion/amazon–nyc–hq2.html.

Graham, Matthew R., Mark J. Kutzbach, and Brian McKenzie. 2014. "Design Comparison of LODES and ACS Commuting Data Products." Center for Economic Studies, U.S. Census Bureau Working Papers 14-38.

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti. 2010. "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings." *Journal of Political Economy*, 118(3): 536–598.

Guimarães, Paulo, Octávio Figueirdo, and Douglas Woodward. 2003. "A Tractable Approach to the Firm Location Decision Problem." *The Review of Economics and Statistics*, 85(1): 201–204.

Heblich, Stephan, Stephen J Redding, and Daniel M Sturm. 2020. "The Making of the Modern Metropolis: Evidence from London." *The Quarterly Journal of Economics*. qjaa014.

Holmes, Thomas J., and Holger Sieg. 2015. "Structural Estimation in Urban Economics." In *Handbook of Regional and Urban Economics*. Vol. 5 of *Handbook of Regional and Urban Economics*, , ed. Gilles Duranton, J. Vernon Henderson and William C. Strange, 69 – 114. Elsevier.

Holmes, Thomas J., and John J. Stevens. 2014. "An Alternative Theory of the Plant Size Distribution, with Geography and Intra- and International Trade." *Journal of Political Economy*, 122(2): 369–421.

Kehoe, Timothy J. 2005. "An Evaluation of the Performance of Applied General Equilibrium Models on the Impact of NAFTA." In *Frontiers in Applied General Equilibrium Modeling: In Honor of Herbert Scarf.* , ed. Timothy J. Kehoe, T. N. Srinivasan and John Whalley, 341–377. Cambridge University Press.

Krebs, Oliver, and Michael P. Pflüger. 2019. "On the Road (Again): Commuting and Local Employment Elasticities in Germany." Institute of Labor Economics (IZA) IZA Discussion Papers 12257.

Kreindler, Gabriel, and Yuhei Miyauchi. 2020. "Measuring Commuting and Economic Activity inside Cities with Cell Phone Records." Boston University - Department of Economics Boston University - Department of Economics - Working Papers Series WP2020-006.

Krusell, Per, and Anthony A. Smith. 1998. "Income and Wealth Heterogeneity in the Macroeconomy." *Journal of Political Economy*, 106(5): 867–896.

McFadden, Daniel L. 1974. "Conditional Logit Analysis of Qualitative Choice Behavior." In *Frontiers in Econometrics.* , ed. P. Zarembka, 105–142. New York:Academic Press.

McFadden, Daniel L. 1978. "Modelling the Choice of Residential Location." In *Spatial Interaction Theory and Planning Models.* , ed. A. Karlqvist, L. Lundqvist, F. Snickars and J. Weibull. North Holland.

Mogstad, Magne, Joseph P Romano, Azeem Shaikh, and Daniel Wilhelm. 2020. "Inference for Ranks with Applications to Mobility across Neighborhoods and Academic Achievement across Countries." National Bureau of Economic Research Working Paper 26883.

Monte, Ferdinando, Stephen J. Redding, and Esteban Rossi-Hansberg. 2018. "Commuting, Migration, and Local Employment Elasticities." *American Economic Review*, 108(12): 3855–90.

Owens, Raymond, III, Esteban Rossi-Hansberg, and Pierre-Daniel Sarte. 2020. "Rethinking Detroit." *American Economic Journal: Economic Policy*, 12(2): 258–305.

Perez-Cervantes, Fernando. 2016. "Insurance Against Local Productivity Shocks: Evidence from Commuters in Mexico." Banco de México Working Papers 2016-19.

Proost, Stef, and Jacques-François Thisse. 2019. "What Can Be Learned from Spatial Economics?" *Journal of Economic Literature*, 57(3): 575–643.

Quan, Thomas W., and Kevin R. Williams. 2018. "Product variety, across-market demand heterogeneity, and the value of online retail." *The RAND Journal of Economics*, 49(4): 877–913.

Redding, Stephen, and Matthew Turner. 2015. "Transportation Costs and the Spatial Organization of Economic Activity." In *Handbook of Regional and Urban Economics.* Vol. 5, , ed. Gilles Duranton, J. V. Henderson and William C. Strange, Chapter 20, 1339–1398. Elsevier.

Redding, Stephen J., and Esteban Rossi-Hansberg. 2017. "Quantitative Spatial Economics." *Annual Review of Economics*, 9(1): 21–58.

Rosenthal, Stuart S., and William C. Strange. 2020. "How Close Is Close? The Spatial Reach of Agglomeration Economies." *Journal of Economic Perspectives*, 34(3): 27–49.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Raymond Owens. 2010. "Housing Externalities." *Journal of Political Economy*, 118(3): 485–535.

Rutherford, Thomas F. 1995. "Constant Elasticity of Substitution Functions: Some Hints and Useful Formulae." Notes prepared for GAMS General Equilibrium Workshop held December, 1995 in Boulder Colorado.

Schoefer, Benjamin, and Oren Ziv. 2020. "Productivity, Place, and Plants: Revisiting the Measurement."

Severen, Christopher. 2019. "Commuting, Labor, and Housing Market Effects of Mass Transportation: Welfare and Identification."

Silva, J. M. C. Santos, and Silvana Tenreyro. 2006. "The Log of Gravity." *The Review of Economics and Statistics*, 88(4): 641–658.

Sotelo, Sebastian. 2019. "Practical Aspects of Implementing the Multinomial PML Estimator."

Train, Kenneth. 2009. *Discrete Choice Methods with Simulation.* Cambridge University Press.

Tsivanidis, Nick. 2019. "Evaluating the Impact of Urban Transit Infrastructure: Evidence from Bogota's TransMilenio."

Waddell, Sonya Ravindranath, and Pierre Daniel Sarte. 2016. "From Stylized to Quantitative Spatial Models of Cities." *Economic Quarterly*, 169–196.

Weintraub, Gabriel Y., C. Lanier Benkard, and Benjamin Van Roy. 2008. "Markov Perfect Industry Dynamics With Many Firms." *Econometrica*, 76(6): 1375–1411.

Zárate, Román David. 2019. "Factor Allocation, Informality and Transit Improvements: Evidence from Mexico City."

Appendix – For Online Publication

A   Theory: Proofs and extensions

Section A.1 provides a condition on parameter values sufficient for the existence and uniqueness of the continuum model's equilibrium. Section A.2 details the calibrated-shares procedure for computing counterfactual outcomes. Section A.3 shows the trade equilibrium of the granular model is unique. Section A.4 shows that the continuum model is a limiting case of the granular model. The following subsections present extensions of the bare-bones model in Section 4 that introduce trade costs, the use of land in production, residential amenities, and local increasing returns.

A.1   Existence and uniqueness of continuum model's equilibrium

We derive a set of parameter values sufficient for the existence and uniqueness of the continuum model's equilibrium, using Theorem 1 of Allen, Arkolakis, and Li (2020).

The equilibrium of the continuum model can be written as

$$\frac{r_k^{1+\alpha\epsilon} T_k}{\alpha} = \sum_n D_{kn} w_n^{1+\epsilon} \tag{A.1}$$

$$L_n w_n^{-\epsilon} = \sum_k D_{kn} r_k^{-\alpha\epsilon} \tag{A.2}$$

$$A_n^{1-\sigma} L_n w_n^{\sigma} = P^{\sigma-1} \sum_{n'} L_{n'} w_{n'}, \tag{A.3}$$

where, for notational convenience, we define

$$\Upsilon \equiv \sum_k \sum_n w_n^\epsilon (r_k^\alpha \delta_{kn})^{-\epsilon}$$

$$D_{kn} \equiv L\Upsilon^{-1} \delta_{kn}^{-(1+\epsilon)}.$$

Equations (A.1)–(A.3) are $3 \times N$ equations in $3 \times N$ unknowns $\{r_n, w_n, L_n\}_{n=1,2,\ldots,N}$ given values for Υ and P. Although Υ and P are endogenous, they do not vary across locations. Accordingly, they fall under the class of endogenous scalars discussed in Remark 2 of Allen, Arkolakis, and Li (2020), which leave the conclusion of Theorem 1 Part ii.b unchanged and can therefore be treated as if exogenous.

Following Remark 3 of Allen, Arkolakis, and Li (2020), we construct the intermediate matrices $\mathbf{\Gamma}$ and $\mathbf{B}$:

$$\mathbf{\Gamma} = \begin{bmatrix} 1+\alpha\epsilon & 0 & 0 \\ 0 & -\epsilon & 1 \\ 0 & \sigma & 1 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 0 & 1+\epsilon & 0 \\ -\alpha\epsilon & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

Then we can construct the final matrix $\mathbf{A}$ defined by $(\mathbf{A})_{hh'} \equiv \left|(\mathbf{B\Gamma}^{-1})_{hh'}\right|$:

$$\mathbf{B\Gamma}^{-1} = \begin{bmatrix} 0 & 1+\epsilon & 0 \\ -\alpha\epsilon & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{1+\alpha\epsilon} & 0 & 0 \\ 0 & -\frac{1}{\sigma+\epsilon} & \frac{1}{\sigma+\epsilon} \\ 0 & \frac{\sigma}{\sigma+\epsilon} & \frac{\epsilon}{\sigma+\epsilon} \end{bmatrix}, \qquad \mathbf{A} = \begin{bmatrix} 0 & \frac{1+\epsilon}{\sigma+\epsilon} & \frac{1+\epsilon}{\sigma+\epsilon} \\ \frac{\alpha\epsilon}{1+\alpha\epsilon} & 0 & 0 \\ 0 & \frac{\sigma-1}{\sigma+\epsilon} & \frac{1+\epsilon}{\sigma+\epsilon} \end{bmatrix}$$

where we've assumed $\sigma > 1$.

We want to show that the spectral radius, the absolute value of the maximum eigenvalue, of $\mathbf{A}$ (denoted $\rho(\mathbf{A})$) equals one. The reasoning in Appendix B.1.4 of Allen, Arkolakis, and Li (2020) demonstrates that $\rho(\mathbf{A}) \geq 1$. Thus, it remains to show that $\rho(\mathbf{A}) \leq 1$. To do so, we will use Lemma 2 from Allen, Arkolakis, and Li (2020). Consider the matrix

$$\lambda I - \mathbf{A} = \begin{bmatrix} \lambda & -\frac{1+\epsilon}{\sigma+\epsilon} & -\frac{1+\epsilon}{\sigma+\epsilon} \\ -\frac{\alpha\epsilon}{1+\alpha\epsilon} & \lambda & 0 \\ 0 & -\frac{\sigma-1}{\sigma+\epsilon} & \lambda - \frac{1+\epsilon}{\sigma+\epsilon} \end{bmatrix}.$$

Define the function $f(\lambda) \equiv \det(\lambda I - \mathbf{A})$, and denote its kth derivative by $f_k(\lambda)$:

$$f(\lambda) = \lambda^3 - \lambda^2 \left(\frac{1+\epsilon}{\sigma+\epsilon}\right) - \left(\frac{1+\epsilon}{\sigma+\epsilon}\right)\left(\frac{\alpha\epsilon}{1+\alpha\epsilon}\right)\left(\lambda - \frac{1+\epsilon}{\sigma+\epsilon}\right) + \left(\frac{1+\epsilon}{\sigma+\epsilon}\right)\left(\frac{\alpha\epsilon}{1+\alpha\epsilon}\right)\left(\frac{\sigma-1}{\sigma+\epsilon}\right)$$

$$f_1(\lambda) = 3\lambda^2 - 2\lambda\left(\frac{1+\epsilon}{\sigma+\epsilon}\right) - \left(\frac{1+\epsilon}{\sigma+\epsilon}\right)\left(\frac{\alpha\epsilon}{1+\alpha\epsilon}\right)$$

$$f_2(\lambda) = 6\lambda - \frac{2(1+\epsilon)}{\sigma+\epsilon}.$$

Lemma 2 of Allen, Arkolakis, and Li (2020) states that $\rho(\mathbf{A}) \leq 1$ if and only if $f_k(1) \geq 0$ for $k = 0, 1, 2$.

$$f(1) = 1 - \frac{1+\epsilon}{\sigma+\epsilon} - \left(\frac{1+\epsilon}{\sigma+\epsilon}\right)\left(\frac{\alpha\epsilon}{1+\alpha\epsilon}\right) + \left(\frac{1+\epsilon}{\sigma+\epsilon}\right)\left(\frac{\alpha\epsilon}{1+\alpha\epsilon}\right)\left(\frac{2+\epsilon-\sigma}{\sigma+\epsilon}\right)$$

$$f_1(1) = 3 - 2\left(\frac{1+\epsilon}{\sigma+\epsilon}\right) - \left(\frac{1+\epsilon}{\sigma+\epsilon}\right)\left(\frac{\alpha\epsilon}{1+\alpha\epsilon}\right)$$

$$f_2(1) = 6 - \frac{2(1+\epsilon)}{\sigma+\epsilon}.$$

Both $f_1(1) \geq 0$ and $f_2(1) \geq 0$ for any $\sigma > 1$ and $\alpha, \epsilon > 0$. A sufficient condition for $f(1) \geq 0$ is $1 \geq \frac{1+\epsilon}{\sigma+\epsilon}\left[1 + \frac{\alpha\epsilon}{1+\alpha\epsilon}\left(\frac{2(\sigma-1)}{\sigma+\epsilon}\right)\right]$, which can be restated as

$$\left(\frac{1+\epsilon}{\sigma+\epsilon}\right)\left(\frac{\alpha\epsilon}{1+\alpha\epsilon}\right) \leq \frac{1}{2}. \tag{A.4}$$

Note that condition (A.4) is satisfied by our baseline parameter values ($\alpha = 0.24$, $\sigma = 4$, $\epsilon = 7.986$). Under condition (A.4), $\rho(\mathbf{A}) \leq 1$, which means that $\rho(\mathbf{A}) = 1$. Thus, by Theorem 1 Part ii.b of Allen, Arkolakis, and Li (2020), if condition (A.4) is true, the equilibrium exists and is unique up

to scale.

A.2   The calibrated-shares procedure

This section describes a calibration procedure often applied in quantitative spatial economics that is known as "exact hat algebra" or "calibrated share form" in the international trade literature (Rutherford, 1995; Dekle, Eaton, and Kortum, 2008; Costinot and Rodríguez-Clare, 2014). This procedure makes a quantitative spatial model perfectly match the observed spatial distribution of economic outcomes and allows one to compute counterfactual equilibrium outcomes expressed as proportionate changes from the initial equilibrium.

The procedure relies on rewriting the equilibrium equations to describe ratios of counterfactual and initial equilibrium variables. In line with the parameterization discussed in Section 3.2, we assume that $\delta_{kn} = \bar{\delta}_{kn}\lambda_{kn}$. The term $\bar{\delta}_{kn}$ represents observed commuting times that affect the hours of labor supply (we use this variable to estimate the commuting elasticity ϵ in Section 3.2). We treat λ_{kn} as a preference shifter rather than a time cost that affects observed labor hours. Therefore, λ_{kn} does not affect hours of labor supply conditional on ℓ_{kn}. One can emulate the frequent practice of treating commuting costs as preference shifters by setting $\bar{\delta}_{kn} = 1$.[35] In this specification, the three equations that jointly define the equilibrium of the model in Section 2, equations (2), (3), and (4), can be rewritten as

$$A_n \sum_k \frac{\ell_{kn}}{\bar{\delta}_{kn}} = \frac{(w_n/A_n)^{-\sigma}}{P^{1-\sigma}} \sum_{k',n'} \frac{\ell_{k'n'}}{\bar{\delta}_{k'n'}} w_{n'}$$

$$T_k = \alpha \sum_n \frac{\ell_{kn}}{\bar{\delta}_{kn}} \frac{w_n}{r_k}$$

$$\ell_{kn} = L \times \frac{w_n^\epsilon \left(r_k^\alpha \bar{\delta}_{kn}\lambda_{kn}\right)^{-\epsilon}}{\sum_{k',n'} w_{n'}^\epsilon \left(r_{k'}^\alpha \bar{\delta}_{k'n'}\lambda_{k'n'}\right)^{-\epsilon}}.$$

The researcher knows the labor demand elasticity σ and commuting elasticity ϵ and observes the initial equilibrium labor allocation ℓ_{kn}, reduction in labor supply from commuting times $\bar{\delta}_{kn}$, and wages w_n. Denote the counterfactual equilibrium values of wages, rents, and labor allocation by w_n', r_k', and ℓ_{kn}'. Denote the counterfactual-initial ratio of a variable x by $\hat{x} \equiv \frac{x'}{x}$. Assume that $\ell_{kn}' = 0$ if $\ell_{kn} = 0$, so that, in a slight abuse of notation, we can write $\hat{\ell}_{kn} = 1$ if $\ell_{kn} = 0$.

One can solve for the counterfactual equilibrium variables associated with combinations of counterfactual-initial ratios of productivities $\hat{A}_n$, land endowments $\hat{T}_k$, and commuting costs $\hat{\lambda}_{kn}$ and $\hat{\bar{\delta}}_{kn}$. Tedious manipulation of the three equations above leads to the following three equations

---

[35]In that case, a different assumption would be needed to estimate the commuting elasticity ϵ.

written in terms of elasticities, initial equilibrium values, and counterfactual-initial ratios:

$$\hat{w}_n = \hat{A}_n^{\frac{\sigma-1}{\sigma}} \left(\sum_n \left(\frac{\hat{w}_n}{\hat{A}_n} \right)^{1-\sigma} w_n \sum_k \frac{\ell_{kn}}{\bar{\delta}_{kn}} \right)^{\frac{-1}{\sigma}} \left(\sum_k \frac{\hat{\ell}_{kn}}{\hat{\bar{\delta}}_{kn}} \frac{\frac{\ell_{kn}}{\bar{\delta}_{kn}}}{\sum_{k'} \frac{\ell_{k'n}}{\bar{\delta}_{k'n}}} \right)^{\frac{-1}{\sigma}} \left(\sum_{k',n'} \frac{\hat{\ell}_{k'n'} \hat{w}_{n'}}{\hat{\bar{\delta}}_{k'n'}} \frac{\ell_{k'n'}}{\bar{\delta}_{k'n'}} w_{n'} \right)^{\frac{1}{\sigma}}$$

(A.5)

$$\hat{r}_k = \hat{T}_k^{-1} \sum_n \frac{\hat{\ell}_{kn} \hat{w}_n \frac{\ell_{kn}}{\delta_{kn}} w_n}{\hat{\bar{\delta}}_{kn} \sum_{n'} \frac{\ell_{kn'}}{\delta_{kn'}} w_{n'}}$$

(A.6)

$$\hat{\ell}_{kn} = \begin{cases} 1, \text{ if } \ell_{kn} = 0 \\ \dfrac{\hat{w}_n^{\varepsilon} \left(\hat{r}_k^{\alpha} \hat{\bar{\delta}}_{kn} \hat{\lambda}_{kn} \right)^{-\varepsilon}}{\sum_{k',n'} \hat{w}_{n'}^{\varepsilon} \left(\hat{r}_{k'}^{\alpha} \hat{\bar{\delta}}_{k'n'} \hat{\lambda}_{k'n'} \right)^{-\varepsilon} \frac{\ell_{kn}}{L}}, \text{ if } \ell_{kn} > 0 \end{cases}$$

(A.7)

These three equations deliver $\hat{w}_n$, $\hat{r}_k$, and $\hat{\ell}_{kn}$ given the elasticities σ and ϵ, observed reductions in labor supply $\bar{\delta}_{kn}$, initial equilibrium values ℓ_{kn} and w_n, and counterfactual-initial ratios $\hat{A}_n$, $\hat{T}_k$, $\hat{\lambda}_{kn}$, and $\hat{\bar{\delta}}_{kn}$.

Note that this procedure implicitly rationalizes zero-commuters observations with infinite commuting costs, $\ell_{kn} = 0 \iff \delta_{kn} = \infty$. This procedure cannot characterize cases in which $\ell_{kn} = 0$ and $\ell'_{kn} \neq 0$ because the object $\hat{\delta}_{kn}^{-\epsilon} = \left(\frac{\delta'_{kn}}{\delta_{kn}} \right)^{-\epsilon}$ is not sensibly defined if $\delta'_{kn} \neq 0$ and $\delta_{kn} = \infty$.

Note that this procedure does not identify the initial values of the parameters A_n, T_k, and δ_{kn}. Given the elasticities σ and ϵ and the initial equilibrium values ℓ_{kn} and w_n, equations (3), (4), and (2) are insufficient to separately identify T_k and δ_{kn}. One would also need to observe land prices r_k to separate these parameters.

When implementing this procedure in Sections 3.4 and 5, we compute tract-level workplace wages using LODES and ZIP Business Patterns data following Owens, Rossi-Hansberg, and Sarte (2020).

A.3 The trade equilibrium of the granular model is unique

Recall that the trade equilibrium takes the labor allocation as given. Consequently, we just need to solve for wage and land rent vectors that satisfy

$$A_n \sum_k \frac{\ell_{kn}}{\delta_{kn}} = \frac{(w_n/A_n)^{-\sigma}}{P^{1-\sigma}} \sum_{k',n'} \frac{\ell_{k'n'}}{\delta_{k'n'}} w_{n'} \qquad \forall n$$

(A.8)

$$r_k = \frac{\alpha}{T_k} \sum_n \frac{\ell_{kn}}{\delta_{kn}} w_n \qquad \forall k$$

(A.9)

$$P = \left(\sum_n (w_n/A_n)^{1-\sigma} \right)^{1/(1-\sigma)}.$$

(A.10)

For a given labor allocation and vector of wages, equation (A.9) will determine a unique vector of land rents, so we can focus on solving for a unique vector of relative wages. Recall $L_n \equiv \sum_k \frac{\ell_{kn}}{\delta_{kn}}$ is total labor supply in location n. Substituting L_n and equation (A.10) into equation (A.8) yields

$$A_n L_n = \frac{(w_n/A_n)^{-\sigma}}{\sum_{m'}(w_{m'}/A_{m'})^{1-\sigma}} \sum_{k',n'} \frac{\ell_{k'n'}}{\delta_{k'n'}} w_{n'} \quad \forall n.$$

Note that this equation is homogeneous of degree zero in the wage vector. Select $w_1 = 1$ as the numeraire and writing the expression in terms of relative wages yields

$$w_n = \left(\frac{A_n}{A_1}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{L_n}{L_1}\right)^{-1/\sigma}.$$

Since $\sigma > 1$, this defines a unique vector of relative wages. Plugging those wages into equation (A.9) yields a unique vector of land rents.

A.4 Continuum model as limiting case of the granular model

We derive the result that the equilibrium of the granular model coincides with the equilibrium of the continuum model as the number of individuals becomes infinite, $I \to \infty$. Note that aggregate labor supply L is fixed, as each individual supplies L/I units of labor. The key step is to show that equation (2) holds as $I \to \infty$. Conditional on the labor allocation $\{\ell_{kn}\}$, granularity plays no role in the trade equilibrium, which will coincide with that of a continuum model.

Definition 4.2 says that the share of labor supplied by residents of k who work in n is

$$\frac{\ell_{kn}}{L} = \frac{1}{I} \sum_{i=1}^{I} \mathbf{1}\left\{\tilde{U}_{kn}^i > \tilde{U}_{k'n'}^i, \forall(k',n') \neq (k,n)\right\}.$$

Note that $\mathbf{1}\left\{\tilde{U}_{kn}^i > \tilde{U}_{k'n'}^i, \forall(k',n') \neq (k,n)\right\}$ is a binary random variable that is equal to one with probability $\Pr(\tilde{U}_{kn}^i > \tilde{U}_{k'n'}^i, \forall(k',n') \neq (k,n))$. Thus, $\mathbb{E}\left(\mathbf{1}\left\{\tilde{U}_{kn}^i > \tilde{U}_{k'n'}^i, \forall(k',n') \neq (k,n)\right\}\right) = \Pr\left(\tilde{U}_{kn}^i > \tilde{U}_{k'n'}^i, \forall(k',n') \neq (k,n)\right)$. Since the idiosyncratic preference shocks are independent and identically distributed random variables, the law of large numbers implies that, using equation (7),

$$\lim_{I \to \infty} \frac{1}{I} \sum_{i=1}^{I} \mathbf{1}\left\{\tilde{U}_{kn}^i > \tilde{U}_{k'n'}^i, \forall(k',n') \neq (k,n)\right\} = \mathbb{E}\left(\mathbf{1}\left\{\tilde{U}_{kn}^i > \tilde{U}_{k'n'}^i, \forall(k',n') \neq (k,n)\right\}\right)$$

$$= \frac{\tilde{w}_n^\epsilon(\tilde{r}_k^\alpha \delta_{kn})^{-\epsilon}}{\sum_{k',n'} \tilde{w}_{n'}^\epsilon(\tilde{r}_{k'}^\alpha \delta_{k'n'})^{-\epsilon}}.$$

As a result, as $I \to \infty$, $\frac{\ell_{kn}}{L} \to \frac{\tilde{w}_n^\epsilon(\tilde{r}_k^\alpha \delta_{kn})^{-\epsilon}}{\sum_{k',n'} \tilde{w}_{n'}^\epsilon(\tilde{r}_{k'}^\alpha \delta_{k'n'})^{-\epsilon}}$. Since the continuum-case rational expectations are the wages and land rents that clear markets for this labor allocation, the market-clearing prices of the granular model converge to those values.

A.5 Trade costs

Relative to the model in Section 4, we now assume that goods trade is subject to iceberg trade costs: delivering a unit of the location-n variety to location k requires producing $\tau_{nk} \geq 1$ units in n. Individuals consume differentiated goods at their residences, and the price of location n's output for consumers in location k is $\tau_{nk} w_n / A_n$. Thus, the price index in location k is $r_k^\alpha P_k^{1-\alpha}$, where the local CES price index for goods is $P_k = \left[\sum_n (\tau_{nk} w_n / A_n)^{1-\sigma} \right]^{1/(1-\sigma)}$.

In this environment, the equivalent of equation (6) is such that, based on the beliefs $\{\tilde{w}_n\}$ and $\{\tilde{r}_k\}$, each worker chooses the residential location and the work location that maximize expected utility,

$$\tilde{U}_{kn}^i = \underbrace{\epsilon \ln \left(\frac{\tilde{w}_n}{\tilde{r}_k^\alpha \tilde{P}_k^{1-\alpha} \delta_{kn}} \right)}_{\equiv \tilde{U}_{kn}} + \nu_{kn}^i,$$

where $\tilde{P}_k = \left[\sum_n (\tau_{nk} \tilde{w}_n / A_n)^{1-\sigma} \right]^{1/(1-\sigma)}$.

With trade costs, the goods-market-clearing condition, which is the equivalent of equation (3), equates quantity supplied and quantity demanded:

$$A_n \sum_k \frac{\ell_{kn}}{\delta_{kn}} = (w_n/A_n)^{-\sigma} \sum_k \left[\left(\frac{\tau_{nk}}{P_k} \right)^{1-\sigma} \left(\sum_{n'} \frac{\ell_{kn'}}{\delta_{kn'}} w_{n'} \right) \right]. \tag{A.11}$$

Equation (4), which clears the land market, is unchanged by the introduction of trade costs.

The definition of a trade equilibrium is akin to Definition 4.1, with equations (A.11) and (4) serving as the relevant market-clearing conditions.

With trade costs, the expression for the probability of choosing residential-workplace pair kn analogous to equation (7) is

$$\Pr(U_{kn}^i > U_{k'n'}^i \ \forall (k', n') \neq (k, n)) = \frac{\tilde{w}_n^\epsilon \left(\tilde{r}_k^\alpha \tilde{P}_k^{1-\alpha} \delta_{kn} \right)^{-\epsilon}}{\sum_{k',n'} \tilde{w}_{n'}^\epsilon \left(\tilde{r}_{k'}^\alpha \tilde{P}_{k'}^{1-\alpha} \delta_{k'n'} \right)^{-\epsilon}}. \tag{A.12}$$

The definition of a granular commuting equilibrium is akin to Definition 4.2 with $\{\tau_{kn}\}$ added to the list of economic primitives and expression (A.12) serving as the relevant probability mass function.

A.6 Firms produce using land

Relative to the model in Section 4, we now assume that firms produce output using labor and land inputs via a Cobb-Douglas production function. Workers produce in location n with a Cobb-Douglas production technology:

$$q_n = A_n \left(\frac{\phi_n T_n}{\beta} \right)^\beta \left(\frac{L_n}{1-\beta} \right)^{1-\beta},$$

where L_n is the labor supply of workers working in location n and ϕ_n is the share of land in location n that is used in production. As a result, the unit cost of production in location n is $c_n = \frac{1}{A_n} r_n^\beta w_n^{1-\beta}$. Thus, the CES price index is $P = \left[\sum_n c_n^{1-\sigma} \right]^{1/(1-\sigma)}$.

Since output in location n is $A_n(\phi_n T_n)^\beta \left(\sum_k \ell_{kn}/\delta_{kn} \right)^{1-\beta}$, equating quantity supplied and quantity demanded requires

$$A_n(\phi_n T_n)^\beta \left(\sum_k \ell_{kn}/\delta_{kn} \right)^{1-\beta} = \frac{(c_n)^{-\sigma}}{P^{1-\sigma}} \sum_{n'} \left(r_{n'} \phi_{n'} T_{n'} + w_{n'} \sum_{k'} \ell_{k'n'}/\delta_{k'n'} \right). \tag{A.13}$$

Equating the fixed land supply T_k to the sum of the quantities of land demanded by firms and by residents requires

$$T_k = \alpha \sum_n \frac{\ell_{kn}}{\delta_{kn}} \frac{w_n}{r_k} + \beta^{\frac{1}{1-\beta}} \left(\frac{w_k}{r_k} \right) \sum_{k'} \frac{\ell_{k'k}}{\delta_{k'k}}. \tag{A.14}$$

The definition of a trade equilibrium is akin to Definition 4.1, with equations (A.13) and (A.14) serving as the relevant market-clearing conditions. The definition of a granular commuting equilibrium is akin to Definition 4.2 with the Cobb-Douglas production parameter β added to the list of economic primitives.

A.7 Residential amenities

Relative to the model in Section 4, we now assume that each location is endowed with a residential amenity B_k. Individual i's utility from residing in k and working in n is now

$$U_{kn}^i = \epsilon \ln \left(\frac{B_k w_n}{r_k^\alpha P^{1-\alpha} \delta_{kn}} \right) + \nu_{kn}^i.$$

With residential amenities, the expression for the probability of choosing residential-workplace pair kn analogous to equation (7) is

$$\Pr(U_{kn}^i > U_{k'n'}^i \ \forall (k', n') \neq (k, n)) = \frac{(B_k \tilde{w}_n)^\epsilon (\tilde{r}_k^\alpha \delta_{kn})^{-\epsilon}}{\sum_{k',n'} (B_{k'} \tilde{w}_{n'})^\epsilon (\tilde{r}_{k'}^\alpha \delta_{k'n'})^{-\epsilon}}. \tag{A.15}$$

The definition of a trade equilibrium in Definition 4.1 remains unchanged. The definition of a granular commuting equilibrium is akin to Definition 4.2 with the residential amenities $\{B_k\}$ added to the list of economic primitives and expression (A.15) serving as the relevant probability mass function.

A.8 Local increasing returns

Relative to the model in Section 4, we now assume that production exhibits local external economies of scale. In particular, in each location n the linear production technology is $q_n = A_n L_n$, where L_n is the labor supply of workers working in location n and $A_n \equiv \bar{A}_n L_n^\eta$. Thus, the CES price index is $P = \left[\sum_n \left(w_n/(\bar{A}_n L_n^\eta) \right)^{1-\sigma} \right]^{1/(1-\sigma)}$.

Since output in location n is $\bar{A}_n \left(\sum_k \ell_{kn}/\delta_{kn} \right)^{1+\eta}$, equating quantity supplied and quantity demanded requires

$$\bar{A}_n \left(\sum_k \ell_{kn}/\delta_{kn} \right)^{1+\eta} = \frac{w_n^{-\sigma} \bar{A}_n^{\sigma} \left(\sum_k \ell_{kn}/\delta_{kn} \right)^{\eta\sigma}}{P^{1-\sigma}} \sum_{k',n'} \frac{\ell_{k'n'}}{\delta_{k'n'}} w_{n'}. \tag{A.16}$$

The remaining results are unchanged after appropriately substituting in equation (A.16). Equation (4), which clears the land market, and expression (7) are unchanged by the introduction of local increasing returns. The definition of a trade equilibrium is akin to Definition 4.1, with equations (A.16) and (4) serving as the relevant market-clearing conditions. The definition of a granular commuting equilibrium is akin to Definition 4.2 with $\{\bar{A}_n\}$ and η added to the list of economic primitives.

The model of perfect competition with local increasing returns should be isomorphic to a model of monopolistic competition with free entry, per the logic in Allen and Arkolakis (2014).

# B Characterizing granular empirical settings

B.1 Tract-level commuting flows for Detroit and Minneapolis-St. Paul

This section presents summary statistics for tract-to-tract commuting counts in the Detroit urban area and the Minneapolis-St. Paul metropolitan area in 2014 analogous to those reported for New York City in Section 3.1. We chose these examples because Owens, Rossi-Hansberg, and Sarte (2020) study the Detroit urban area and Minnesota is the only state that reports employment by establishment rather than firm in the LODES data (Graham, Kutzbach, and McKenzie, 2014).

The Detroit urban area has about 1.3 million resident-employees and 1.4 million tract pairs, so the average cell in its commuting matrix is near one. Among the tract pairs, 74% have zero commuters between them. 42.6% of Detroit commuters have five or fewer commuters in their cell of the commuting matrix.

The Minneapolis-St. Paul metropolitan area has about 1.5 million resident-employees and 0.6 million tract pairs, so the average cell in its commuting matrix has less than three commuters. Among the tract pairs, 61% have zero commuters between them.

Figure B.1: Number of commuters between pairs of tracts in Detroit and the Twin Cities
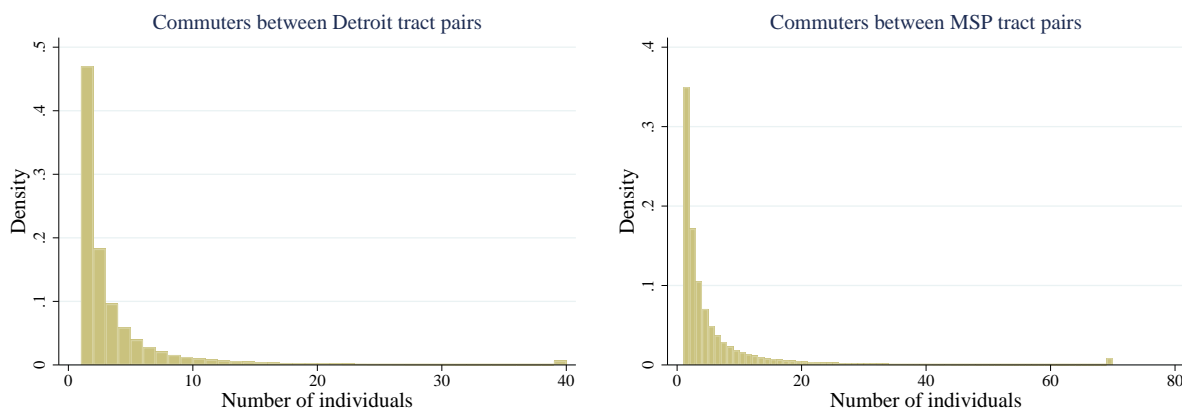
NOTES: These histograms report the number of tract pairs in the Detroit urban area (left panel) and the Minneapolis-St. Paul metropolitan area (right panel) by the number of individuals who reside in the origin tract and work in the destination tract in 2014 LODES data. These histograms restrict the samples to pairs of tracts with a strictly positive number of commuters. Observations larger than the 99th percentile are winsorized (40 commuters for Detroit urban area and 70 commuters for Minneapolis-St. Paul metropolitan area). Detroit has 1,166 residential census tracts. MSP has 789 residential census tracts.

The spatial concentration of employment contributes to the sparsity of these commuting matrices. The median tract in Detroit has 465 employees working in it. Since Detroit has 1,166 residential tracts, at least 60% of locations must have zero residents commuting to this workplace.

B.2 Granularity in county-level commuting flows

Even when studying larger geographic units, the issue of granularity may be relevant. Consider US counties, which vary greatly in population size. In the 2006–2010 American Community Survey

(ACS) data, there are 136 million commuters (with commutes less than 120 kilometers).[36] Of those 136 million, 101 million live and work in the same county, so there are 35 million cross-county commuters between 79,188 pairs of counties. Thus, the average off-diagonal element of the county-to-county commuting matrix has 445 commuters. However, the distribution of commuters is extremely uneven. The top 10 county pairs account for more than 2 million commuters alone. For the bottom 90% of off-diagonal observations, the mean value is only 40 commuters. Almost half of the county pairs report zero commuters. Among those reporting positive values, Figure B.2 shows that this distribution is skewed, so that many thousands of county pairs have small numbers of commuters. Thus, while granularity is unlikely to trouble studies examining US counties with millions of residents, many county pairs have small commuting flows that represent granular outcomes.

Figure B.2: Number of commuters between US counties
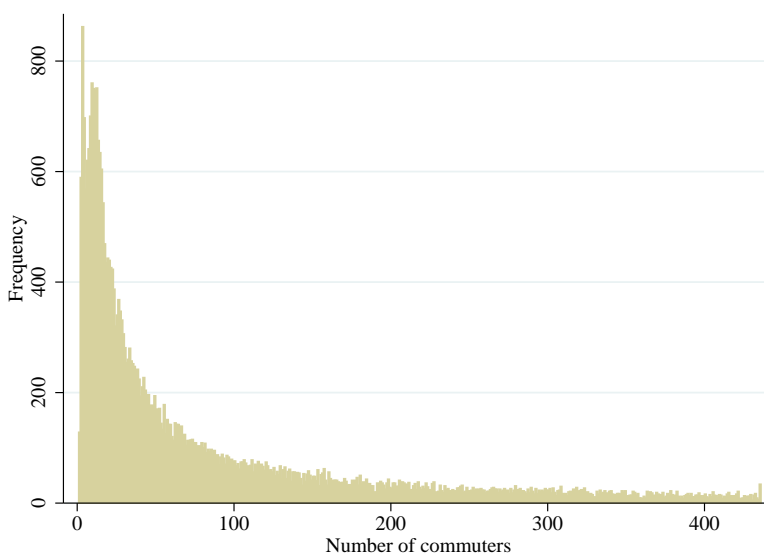
NOTES: This histogram depicts the number of county pairs by the estimated number of commuters in the 2006–2010 ACS. The sample is restricted to pairs of distinct counties within 120 kilometers, the smallest 90% of estimated commuter counts among such counties, and only those pairs reporting a strictly positive number of commuters.

In practice, the granularity of this empirical setting is severely compounded by the fact that the ACS is a 1-in-20 sample of the population. In three different five-year waves of the ACS, nearly half of the county pairs within 120 kilometers of each other are reported to have zero commuters, as shown in Table B.1. More than half of the county pairs report fewer than 100 commuters, therefore representing the behavior of five or fewer respondents. As a consequence, as shown in the third column of Table B.1, for more than one-third of the county pairs with positive commuting flows, the Census-reported margin of error exceeds the reported number of commuters.[37]

---

[36] We follow Monte, Redding, and Rossi-Hansberg (2018) by restricting attention to county pairs that are less than 120 kilometers apart.

[37] At face value, this would seem to imply that one could not reject the null hypothesis that there were no commuters

Table B.1: Zeros in US county-to-county commuting matrix

Dataset	Zero Pairs	Positive Pairs	MOE >X (percent)
ACS 2006-2010	36,403	42,785	37
ACS 2009-2013	35,547	43,641	36
ACS 2011-2015	35,002	44,096	35

NOTES: This table reports the number of county pairs with zero commuters and non-zero commuters for three editions of the ACS. The sample is restricted to pairs of counties within 120 kilometers of each other. The final column reports the share of county pairs for which the Census-reported 90% margin of error exceeds the (strictly positive) reported number of commuters.

Commuting data from other countries also exhibit granularity. For example, one-quarter of Germany's county (Kreisfreie Städte and Landkreise) pairs within 120 kilometers have fewer than 10 commuters (Krebs and Pflüger, 2019). In Brazil's 2010 Censo Demográfico, which reports estimates based on a 10% sample of the population, there are 81 million commuters (with commutes less than 60 kilometers).[38] There are 7 million cross-municipality commuters between 131,620 pairs of municipalities. Thus, the average off-diagonal element of the municipality-to-municipality commuting matrix has only 56 commuters. About three-quarters of the cells in this commuting matrix are empty.

B.3 Commuting counts are often impersistent

The integer values appearing in commuting matrices are not very persistent. A residence-workplace pair may have three commuters one year and none the next. Another pair of locations may double its number of commuters, from one to two. While the conventional continuum approach interprets these changes as substantial economic shifts, the finite-sample perspective is that these changes are not very informative if they are merely a symptom of granularity. We document substantial churn in commuting counts from year to year, suggesting that there is considerable finite-sample noise in addition to signal in these commuting counts.

Table B.2, which presents the transition matrix for pairs of tracts in the Detroit urban area and New York City between the years 2013 and 2014, demonstrates considerable impersistence. In the Detroit urban area, for pairs with one to four commuters in 2013, the percentages appearing on the diagonal of the transition matrix are quite low. A pair of tracts with one commuter in 2013 was almost three times as likely to have zero commuters in 2014 than to have one commuter. A pair with four commuters in 2013 was more likely to appear in any other column in 2014 than to report four commuters again. The 86% of pairs that had zero commuters in both years may appear to suggest persistence, but this primarily is a symptom of the fact that in both years about three-quarters of

---

for such county pairs. The unsuitability of the normal approximation for the binomial distribution in this case is another symptom of granularity.

[38]Dingel, Miscio, and Davis (2019) use these data to construct metropolitan areas based on commuting flows.

observations are zero.[39] More than 130,000 pairs of tracts that had zero commuters in 2013 had at least one commuter in 2014. At the same time, 39% of Detroit tract pairs with positive flow in 2013 were zeros in 2014 . Thus, while zeros are pervasive in this commuting matrix, they are not very persistent. The results for New York City in Table B.2 are very similar to those for Detroit. The commuter counts are so impersistent that, for many tract pairs, a gravity-based estimate predicts a tract pair's commuter count in 2014 better than its observed count in 2013 does, as reported in Appendix B.

Table B.2: The impersistence of commuting counts for tract pairs in Detroit and NYC

	(a) Detroit							(b) NYC					
			2014							2014			
2013	0	1	2	3	4	5+	2013	0	1	2	3	4	5+
0	0.86	0.10	0.02	0.01	0.00	0.00	0	0.91	0.07	0.01	0.00	0.00	0.00
1	0.60	0.22	0.10	0.04	0.02	0.02	1	0.65	0.20	0.08	0.04	0.02	0.02
2	0.37	0.25	0.16	0.09	0.06	0.08	2	0.39	0.25	0.15	0.09	0.05	0.07
3	0.23	0.22	0.18	0.13	0.08	0.16	3	0.24	0.22	0.17	0.12	0.08	0.16
4	0.15	0.17	0.17	0.14	0.11	0.26	4	0.15	0.17	0.17	0.14	0.11	0.27
5+	0.04	0.06	0.07	0.08	0.08	0.68	5+	0.03	0.05	0.06	0.07	0.07	0.71

NOTES: This table describes pairs of tracts in the Detroit urban area (left panel) and New York City (right panel) by reported number of commuters in the 2013 and 2014 LODES. It is a transition matrix, in which each cell lists the share of tract pairs in that row (number of commuters in 2013) that belong to that column (number of commuters in 2014). Each row sums to 100%, modulo rounding.

There is considerable churn even when using larger geographic units. Table B.3 shows that 22% of the county pairs reporting zero commuters in the 2006–2010 ACS reported a positive number of commuters in the following five-year interval. Conversely, about 15% of pairs reporting a positive number of commuters had zero commuters in the following five-year interval. For pairs of counties with a strictly positive number of commuters smaller than 111 in 2006–2011, the diagonal elements of the transition matrix are less than half. For example, a pair of counties reported to have 71–90 commuters in 2006–2011 has only a 14% probability of appearing in the same bin in the following five-year interval.

To the extent that the observed impersistence of commuting counts is a symptom of granularity, these findings caution against procedures that infer structural parameters from the relative magnitudes of these counts. The difference between one commuter and two commuters (or one respondent and two respondents in a finite sample) is little evidence that the latter outcome was twice as probable. Similarly, procedures that rationalize observations with zero commuters by imposing infinite commuting costs so that these are zero-probability events rule out potential margins of adjustment based on zeros that appear to be largely transitory.

---

[39]If $p \in [0, 1]$ of the pairs were randomly independently assigned zero in each period, then p^2 of those pairs would lie in the upper left cell of the transition matrix. Thus, even if zeros were randomly independently assigned to three-quarters of the tract pairs in each period, nine-sixteenths of the pairs would be zero in both periods. That would not be evidence of persistence.

Table B.3: The impersistence of commuting counts for pairs of US counties

2011-2015

2006-2010	Initial Share(%)	0	1-30	31-50	51-70	71-90	91-110	111-500	501-1,500	>1,500
0	45.97	0.78	0.18	0.02	0.01	0.00	0.00	0.00	0.00	0.00
1-30	19.66	0.35	0.45	0.10	0.05	0.02	0.01	0.02	0.00	0.00
31-50	5.47	0.16	0.36	0.19	0.12	0.07	0.04	0.07	0.00	0.00
51-70	3.43	0.08	0.26	0.18	0.15	0.10	0.08	0.15	0.00	0.00
71-90	2.50	0.05	0.16	0.15	0.16	0.14	0.12	0.23	0.00	0.00
91-110	1.86	0.02	0.11	0.12	0.15	0.12	0.13	0.35	0.00	0.00
111-500	12.03	0.00	0.02	0.03	0.04	0.05	0.05	0.74	0.07	0.00
501-1,500	5.01	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.81	0.06
>1,500	4.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.94

NOTES: This table presents a transition matrix for pairs of counties within 120 kilometers of each other by reported number of commuters in two editions of the ACS. The first column reports the percentage of county pairs that were in each row in the 2006–2010 data. The remaining columns report the share of county pairs within the row that appeared in the corresponding bin in the 2011–2015 data. By definition, the shaded cells in each row sum to one, modulo rounding. The bin boundaries are arbitrary, since Figure B.2 shows no obvious bunching. We found similar impersistence when using alternative boundaries.

Tract-pair-level commuter counts are so impersistent that a gravity model estimated using an observed bilateral characteristic like transit time or distance can sometimes predict future commuter counts better than the observed value. Table B.4 shows that, for tract pairs with fewer than ten commuters reported, a gravity-based estimate predicts the following year's value better than its current value does. The fitted values from a gravity model estimated using 2013 data have a higher R^2 for predicting observed 2014 values than the observed 2013 values in both Detroit and New York City. Using the observed 2013 values yields better predictions only for the tract pairs with the largest 2% of commuter counts.

B.4   Zeros are often asymmetric

The zeros in commuting matrices are often asymmetric. Denoting the number of commuters living in residence k and working in workplace n by ℓ_{kn}, an observed zero is asymmetric when $\ell_{nk} = 0$ and $\ell_{kn} > 0$. For US counties, $\ell_{nk} = 0$ for 22% of county pairs with $\ell_{kn} > 0$. In Detroit, $\ell_{nk} = 0$ for 66% of tract pairs with $\ell_{kn} > 0$. For Brazilian municipalities, $\ell_{nk} = 0$ for 49% of municipio pairs with $\ell_{kn} > 0$. These asymmetric flows are not explained by asymmetric numbers of total workers or residents: in Detroit, $\ell_{nk} = 0$ for 47% of tract pairs for which $\ell_{kn} > 0$ and total employment in k and n differs by 10% or less.

The fact that commuting matrices' zeros are often asymmetric poses a puzzle for calibration procedures that rationalize zero-commuter observations by infinite commuting costs. This interpretation of zeros implies severely asymmetric commuting costs, even though daily commutes are

Table B.4: Gravity-based estimates predict 2014 values better than 2013 values do

| # of commuters	Share	Gravity: time	2013 values	Gravity: distance	2013 values
Panel A: Detroit					
≤ 5	0.960	0.384	0.308	0.367	0.307
≤ 10	0.983	0.494	0.473	0.465	0.472
Panel B: NYC					
≤ 5	0.978	0.362	0.306	0.373	0.306
≤ 10	0.990	0.474	0.475	0.477	0.473

NOTES: Each row reports results for all tract pairs in Detroit (upper panel) or New York City (lower panel) with fewer than 5 or 10 commuters in the 2013 LODES data. The "share" column reports the share of tract pairs covered by the row. The "gravity" columns report the R^2 obtained by regressing the 2014 number of commuters on the number of commuters predicted by a gravity model estimated using 2013 data. The "gravity: time" column estimates the gravity equation (7) as described in Section 4.3. The "gravity: distance" column uses $\ln \delta_{kn} = \ln \text{distance}_{kn}$ rather than commuting costs defined in Section 3.2. The "2013 values" columns report the R^2 obtained by regressing the 2014 number of commuters on the 2013 number of commuters. All estimated slope coefficients are positive. The two "2013 values" columns differ because the regression sample is within observations where predictions from the gravity model are available. In the distance specification, $k = n$ observations are dropped. For NYC, there are two observations where the commute is infeasible.

round-trip journeys. If we believe that commuting from n to k is impossible because we observe $\ell_{nk} = 0$, how do the individuals who live in k and work in n commute home at the end of the day? The mechanisms generating prohibitive commuting costs would have to exhibit within-day variation. But the most plausible source of intraday variation in commuting costs — congestion caused by a large number of commuters — cannot explain residence-workplace pairs that have no commuters.

In many empirical settings, asymmetric zeros may simply be a symptom of granularity rather than evidence of very asymmetric commuting costs. When most pairs of locations have zero commuters, and the modal pair with a positive number reflects the decision of only one commuter or one survey respondent, the difference between zero and one isn't particularly informative.

C   Estimation results

Figure C.1: Destination fixed effects from tract-to-tract gravity regressions

NYC (2010)
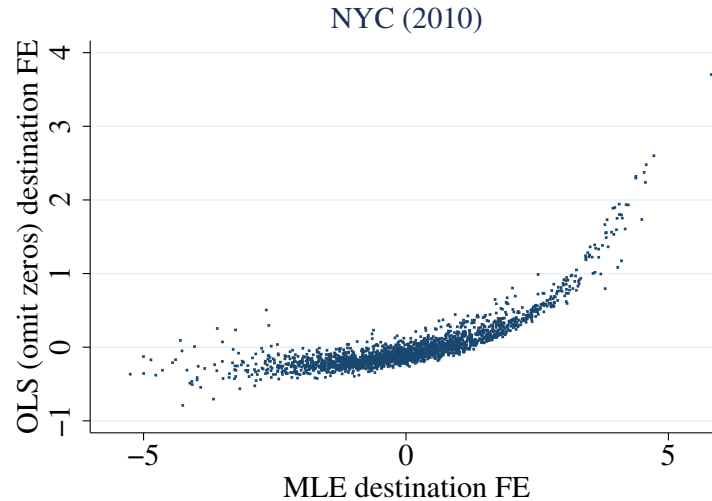
NOTES: This plot depicts the destination fixed effects estimated in Table 1's column 1 (horizontal axis) and column 2 (vertical axis). See the notes to Table 1 for details.

Figure C.1 shows that estimates of workplace fixed effects are biased by dropping observations with zero commuters. When only using positive-commuter observations, the residence and workplace fixed effects characterize the average number of commuters for a residence-workplace pair, conditional on the number of commuters being strictly greater than zero. These conditional averages are necessarily greater than the unconditional averages, and the difference is larger for workplace locations with fewer workers.[40] Figure C.1 contrasts the OLS (vertical axis) and maximum likelihood (horizontal axis) estimates of the workplace fixed effects. The difference is stark: the range of the OLS estimates is half that of the maximum-likelihood estimates because of considerable truncation from below. In essence, the popular practice of omitting zero-commuter observations attributes low-employment destinations' lower employment counts to infinite commuting costs, not lower wage beliefs (lower productivity). By contrast, our maximum likelihood estimator infers that these destinations are less attractive workplaces from the fact that many origins have zero residents working in these destinations.

---

[40]Census tracts are defined so that the number of residents is similar across tracts, while there is tremendous heterogeneity in total employment. Thus, the selection bias is evident in the workplace fixed effects. When estimating the analogous gravity regression for county-to-county commuting flows, we find that the selection bias manifests in both the origin and destination fixed effects.

D   Details of Monte Carlo simulations and event studies

D.1   Monte Carlo exercise

A Monte Carlo exercise shows that granularity alone can substantially reduce the predictive power of the calibrated-shares procedure. In this simulation, the data-generating process is our estimated covariates-based model of New York City in 2010. We impose a counterfactual 18% increase in productivity for the tract containing 200 Fifth Avenue, which generates an increase in employment matching the observed change from 2010 to 2012. In the limiting case, as $I \to \infty$, the calibrated-shares procedure would perfectly describe the changes in commuting flows associated with this productivity increase. Thus, any predictive failure when drawing granular ($I = 2,488,905$) realizations from this data-generating process are due to problems stemming from granularity.

Given our estimated model of New York City in 2010, we simulate 100 "event studies" as in Section 3.4. For each event, we implement the following five steps:

1. To mimic "observed" data, draw one realization from the data-generating process at its 2010 parameter values and one realization at its counterfactual parameter values. Compute the difference in employment in the tract containing 200 Fifth Avenue.

2. Estimate the covariates-based model using the realization drawn from the data-generating process at its 2010 parameter values.

3. Compute the increase in productivity required to match the "observed" change in employment for the "treated" tract for both the calibrated-shares procedure and the covariates-based approach. These increases may differ from each other and the true 18% increase.

4. Compute the predicted change in commuter counts using the calibrated-shares procedure and the covariates-based approach.

5. Regress the "observed" changes in commuter counts destined for the tract containing 200 Fifth Avenue on the changes predicted by the calibrated-shares procedure. Regress the "observed" changes in commuter counts destined for the tract containing 200 Fifth Avenue on the changes predicted by the covariates-based approach.

We summarize the 100 simulations by plotting the coefficients and relative RMSEs in Figure 2.

The Monte Carlo simulation shows that our estimation procedure recovers model parameters well enough when the granular model is the data-generating process to predict counterfactual changes in commuter counts. By contrast, the calibrated-shares procedure has limited predictive power when applied to granular data, as discussed in Section 3.3.

Next, we consider a data-generating process in which commuting costs have both an observed and unobserved component in order to examine the relative predictive performance of the covariates-based model and calibrated-shares procedure in granular settings. The commuting costs are now

determined by

$$\ln \delta_{kn} = \ln \bar{\delta}_{kn} + \ln \lambda_{kn}$$

$$\ln \lambda_{kn} \stackrel{iid}{\sim} \mathcal{N}\left(0, \Lambda \times var(\ln \bar{\delta}_{kn})\right), \quad \Lambda \in \mathbb{R}^+$$

The parameter Λ governs the variance of the unobserved component of commuting costs, stated relative to the variance of the observed component. The simulations considered above are the case in which $\Lambda = 0$, so that $\lambda_{kn} = 1 \; \forall k, n$. We now consider values of $\Lambda > 0$. When $\Lambda > 0$, the covariates-based model omits an unobserved component of commuting costs that the calibrated-shares procedure may, in principle, capture. Since $\ln \bar{\delta}_{kn}$ and $\ln \lambda_{kn}$ are orthogonal by assumption, the covariates-based model's predictions will be unbiased, but its forecast errors will increase with the magnitude of unobserved variation Λ.

Table D.1 presents the results of these simulations. As established in the $\Lambda = 0$ case, the calibrated-shares procedure is very data demanding. For all values of Λ reported in Table D.1, the calibrated-shares procedure's predictions become unbiased (slope near one) only as the number of individuals becomes orders of magnitude larger than the true value. Since the calibrated-shares procedure can capture the unobserved component of commuting costs, its mean forecast error becomes smaller than that of the covariates-based model when the unobserved variation is sufficiently large. For the number of individuals in the empirical setting ($I = 2,488,905$), this occurs only when the unobserved component of commuting costs varies at least half as much as the observed component ($\Lambda > 0.5$).

D.2 Event studies

Figure D.1 depicts the time series of total employment in two New York City census tracts that contain 200 Fifth Avenue and 111 Eighth Avenue. Between 2010 and 2012, Tiffany & Co. and Google, respectively, substantially increased the employment in these two tracts. These tracts are two of the 83 examined in the event studies in Section 3.4.

Table D.1: Contrasting procedures' predictive performance in Monte Carlo simulations

Λ	I	Slope (mean) Covariates-based	Calibrated-shares	RMSE (mean) Covariates-based	Calibrated-shares
0	2.5	0.9908	0.2161	3.9716	4.6661
0	5	0.9854	0.5493	2.7999	3.2893
0	12.5	1.0047	0.8162	1.7722	2.0839
0	25	0.9988	0.9022	1.2570	1.4777
0	50	1.0006	0.9520	0.8803	1.0352
0	125	1.0011	0.9813	0.5602	0.6586
0	250	0.9998	0.9896	0.3952	0.4650
0	2560	0.9998	0.9986	0.1241	0.1458
0.1	2.5	0.9960	0.2469	3.9889	4.6698
0.1	5	0.9983	0.5771	2.8364	3.3005
0.1	12.5	1.0035	0.8220	1.8098	2.0760
0.1	25	1.0029	0.9080	1.3113	1.4718
0.1	50	0.9991	0.9495	0.9698	1.0421
0.1	125	1.0030	0.9813	0.6828	0.6583
0.1	250	1.0016	0.9892	0.5566	0.4649
0.1	2560	1.0022	0.9988	0.4124	0.1455
0.25	2.5	0.9962	0.3570	4.0949	4.6690
0.25	5	1.0163	0.6595	2.9845	3.3043
0.25	12.5	1.0052	0.8506	2.0333	2.0857
0.25	25	1.0030	0.9218	1.5989	1.4720
0.25	50	1.0058	0.9629	1.3382	1.0425
0.25	125	1.0045	0.9831	1.1425	0.6586
0.25	250	1.0057	0.9927	1.0753	0.4665
0.25	2560	1.0056	0.9991	1.0062	0.1458
0.5	2.5	1.0075	0.6097	4.5128	4.6813
0.5	5	1.0104	0.7956	3.5131	3.2914
0.5	12.5	1.0069	0.9087	2.7525	2.0895
0.5	25	1.0080	0.9575	2.4606	1.4725
0.5	50	1.0061	0.9760	2.2954	1.0399
0.5	125	1.0080	0.9920	2.1997	0.6602
0.5	250	1.0086	0.9949	2.1549	0.4664
0.5	2560	1.0074	0.9978	2.1219	0.1459
1	2.5	1.0007	0.8657	6.3818	4.6525
1	5	0.9924	0.9232	5.7525	3.3021
1	12.5	0.9940	0.9575	5.3364	2.0837
1	25	0.9985	0.9729	5.1853	1.4704
1	50	1.0002	0.9778	5.1112	1.0488
1	125	1.0003	0.9809	5.0604	0.6825
1	250	1.0007	0.9827	5.0489	0.4923
1	2560	1.0001	0.9833	5.0328	0.2210

NOTES: This table reports the mean outcomes across 100 simulated event studies for both the covariates-based model and the calibrated-shares procedure. The data-generating process is the covariates-based model estimated on data for New York City in 2010 augmented by an unobserved component of commuting costs, whose magnitude is governed by Λ, as described in the text. The value of I is stated in millions of individuals. The value labeled as 2.5 million is in fact $I = 2,488,905$.

Figure D.1: Employment in two of the event-study tracts
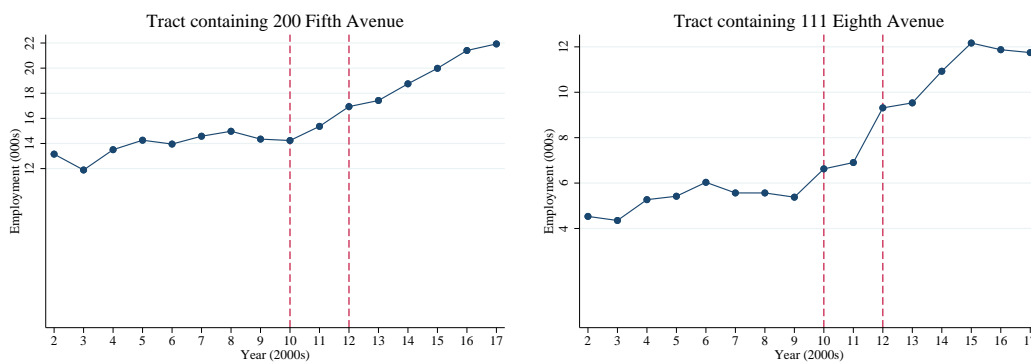
NOTES: This figure depicts the number of primary jobs held by New York City residents in tracts 36061005800 and 36061008300 in the LODES data.

E   Granular model: Price dispersion and "ex post regret"

E.1   Dispersion of real wages and rents relative to beliefs

Figure E.1 plots the dispersion of prices for each tract in New York City for the estimated granular model. We define price dispersion as the standard deviation of the granular commuting equilibrium price across 100,000 simulations divided by the continuum-case rational expectation of that price.

Figure E.1: Price dispersion across granular equilibria

A. Rent   B. Wage
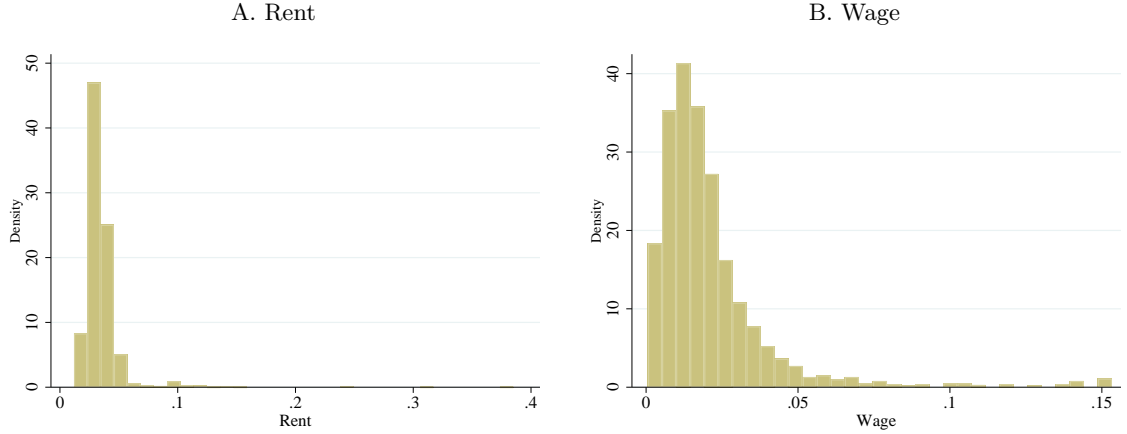
NOTES: The plots depict the dispersion of prices (r_k/P or w_n/P) for each tract in New York City using the granular model estimated on 2010 data. We define price dispersion as the standard deviation of the granular commuting equilibrium price across 100,000 simulations divided by the continuum-case rational expectation of that price. Panel A depicts dispersion in tracts' rents, which have a median value of 0.032 (p5 = 0.021, p95 = 0.051). Panel B depicts dispersion in tracts' wages, which have a median value of 0.016 (p5 = 0.004, p95 = 0.054).

The average realizations of wages and rents across 100,000 simulations of the granular model are very close to the continuum-case rational expectation of these prices. Across tracts, the median tract's absolute percentage-point deviation of mean realized price from the continuum-case rational expectations is 0.06% for wages and 0.02% for rents. The 95$^{\text{th}}$ percentiles of absolute percentage-point deviations for wages and rents are 0.68% and 0.07%, respectively.

E.2   Ex post regret

We compute the share of individuals with ex post regret and their magnitudes as follows. As discussed in the main text, we define ex post regret as the income gain an individual would accept to not switch to their best choice under realized wages and rents. For individual i who chose residence-workplace pair kn, at realized prices r and w ex post regret, χ_i is defined implicitly as

$$\max_{k',n'}\left(\epsilon \ln\left(\frac{w_{n'}}{P^{1-\alpha}r_{k'}^{\alpha}\delta_{k'n'}}\right) + \nu_{k'n'}^i\right) = \left(\epsilon \ln\left(\frac{(1+\chi_i)w_n}{P^{1-\alpha}r_k^{\alpha}\delta_{kn}}\right) + \nu_{kn}^i\right),$$

where by definition kn was the choice maximizing $\tilde{U}_{k,n}^i$. Denote $\underline{k}$ and $\underline{n}$ as the optimal choices of k' and n' for individual i. Given kn, we can solve for the ex post regret χ_i based on the following

equation:

$$\epsilon \ln \left(\frac{w_{\underline{n}}}{P^{1-\alpha} r_{\underline{k}}^{\alpha} \delta_{\underline{kn}}} \right) + \nu_{\underline{kn}}^{i} = \epsilon \ln \left(\frac{(1+\chi_i) w_n}{P^{1-\alpha} r_k^{\alpha} \delta_{kn}} \right) + \nu_{kn}^{i},$$

which implies that

$$\chi_i = \left(\frac{e^{\left(\frac{1}{\epsilon} \nu_{\underline{kn}}^{i} \right)} w_{\underline{n}} r_{\underline{k}}^{-\alpha} \delta_{\underline{kn}}^{-1}}{e^{\left(\frac{1}{\epsilon} \nu_{kn}^{i} \right)} w_n r_k^{-\alpha} \delta_{kn}^{-1}} \right). \tag{E.1}$$

We simulate 2.5 million individuals to compute this statistic. Each of these individuals uses the point-mass beliefs from the continuum model. The realized wages and rents differ because of idiosyncratic preference shocks that do not wash out in the aggregate. A practical challenge in computing this value is that for each of the 2.5 million individuals, one needs to store the 4.6-million-element vector of idiosyncratic ν_{kn}^{i} shocks. To get around this storage problem, we fix the seed for each individual and re-draw the ν_{kn}^{i} term under the same seed in the counterfactual—which is equivalent to storing the vector.

Table E.1 reports the distribution of ex post regrets.

Table E.1: Distribution of ex post regrets

s	Share with regret	Unconditional distribution					Conditional distribution	
		p95	p96	p97	p98	p99	Mean	Median
1	0.0442	0.0000	0.0011	0.0042	0.0082	0.0150	0.0106	0.0073
2	0.0433	0.0000	0.0009	0.0039	0.0078	0.0143	0.0102	0.0071
3	0.0446	0.0000	0.0012	0.0043	0.0083	0.0150	0.0106	0.0072
4	0.0446	0.0000	0.0012	0.0043	0.0084	0.0152	0.0106	0.0073
5	0.0437	0.0000	0.0010	0.0040	0.0079	0.0144	0.0103	0.0071
6	0.0444	0.0000	0.0012	0.0042	0.0083	0.0150	0.0107	0.0073
7	0.0447	0.0000	0.0013	0.0043	0.0083	0.0150	0.0105	0.0072
8	0.0445	0.0000	0.0012	0.0043	0.0084	0.0150	0.0106	0.0073
9	0.0452	0.0000	0.0014	0.0045	0.0086	0.0154	0.0109	0.0074
10	0.0444	0.0000	0.0011	0.0042	0.0082	0.0148	0.0106	0.0072
mean	0.0444	0.0000	0.0012	0.0042	0.0083	0.0149	0.0106	0.0072

NOTES: The table reports the share of individuals with ex post regret and the distribution of ex post regrets of their desired switches in simulations of our estimated granular model. The first column identifies the simulation s. The second column reports the fraction of individuals who have ex post regret and therefore would prefer a different choice given realized prices. Columns under "Unconditional distribution" report the distribution of ex post regret based on the full sample ($I = 2,488,905$). Columns under "Conditional distribution" report the distribution of ex post regret among those who would want to switch. The "mean" row reports the mean value across the 10 simulations.

Table E.2 compares the price dispersion in 100,000 simulations of granular equilibria depicted in Figure E.1 and 10 simulations of granular equilibria depicted in Table E.1. The simulations in

Figure E.1 each require one draw from a multinomial distribution with 4.6 million outcomes whose probabilities are given by equation (7). Generating the labor allocation for each simulation by this method requires less than one second of computing. The simulations in Table E.1 involve more than 11 trillion draws from the type-1 extreme value distribution. Generating the labor allocation for each simulation by this method requires about 100 hours of computing. Hence, the latter simulations are much more computationally expensive. It is heartening that the distributions of price dispersion summarized in Table E.2 are very similar.

Table E.2: Price dispersion across simulation methods

Simulation count	mean	p5	p10	p25	p50	75	p90	p95
Wage								
100,000	0.021	0.004	0.006	0.010	0.016	0.025	0.039	0.054
10	0.023	0.004	0.005	0.009	0.015	0.025	0.039	0.057
Rent								
100,000	0.035	0.021	0.024	0.027	0.032	0.038	0.045	0.051
10	0.034	0.016	0.019	0.024	0.031	0.038	0.049	0.058

NOTES: This table compares the price (r_k/P or w_n/P) dispersion generated by the simulations used in Section E.1 (100,000 simulations) and Section E.2 (10 simulations). As described in the text, each tract's price dispersion is the standard deviation of the granular commuting equilibrium price across simulations divided by the continuum-case rational expectation of that price. The first column reports the average price dispersion across all tracts. The remaining columns report the level of price dispersion for tracts at selected percentiles.

F   Amazon HQ2 counterfactual

Comparing Figures 6B and 7B reveals that the spatial pattern of changes in residents is closely related to the initial number of residents who work in the tract in which Amazon HQ2 would be located. Figure F.1 depicts the relationship between these two variables more explicitly in a scatterplot.

Figure F.1: Contrasting predicted change in residents and observed data
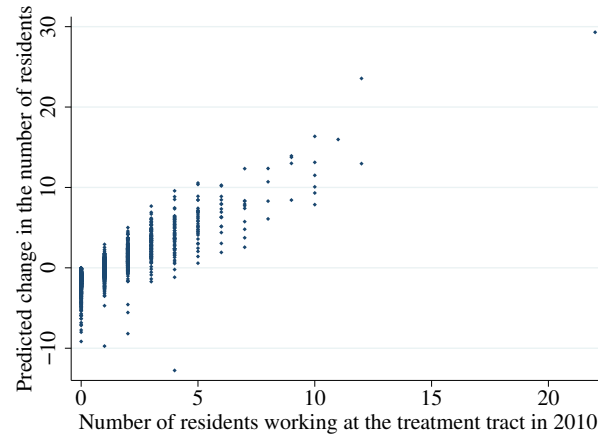
NOTES: This plot depicts the predicted change in the number of residents by calibrated-shares procedure against the number of residents working at the treatment tract in 2010. The treated tract is excluded from the plot.

Figure F.2 depicts both the normal approximation and simulated distribution of outcomes for the changes in the number of residents and workers for the application of the granular model in the Amazon HQ2 counterfactual. The values are very similar.

Figure F.2: Distribution of predicted change in the number of residents or workers
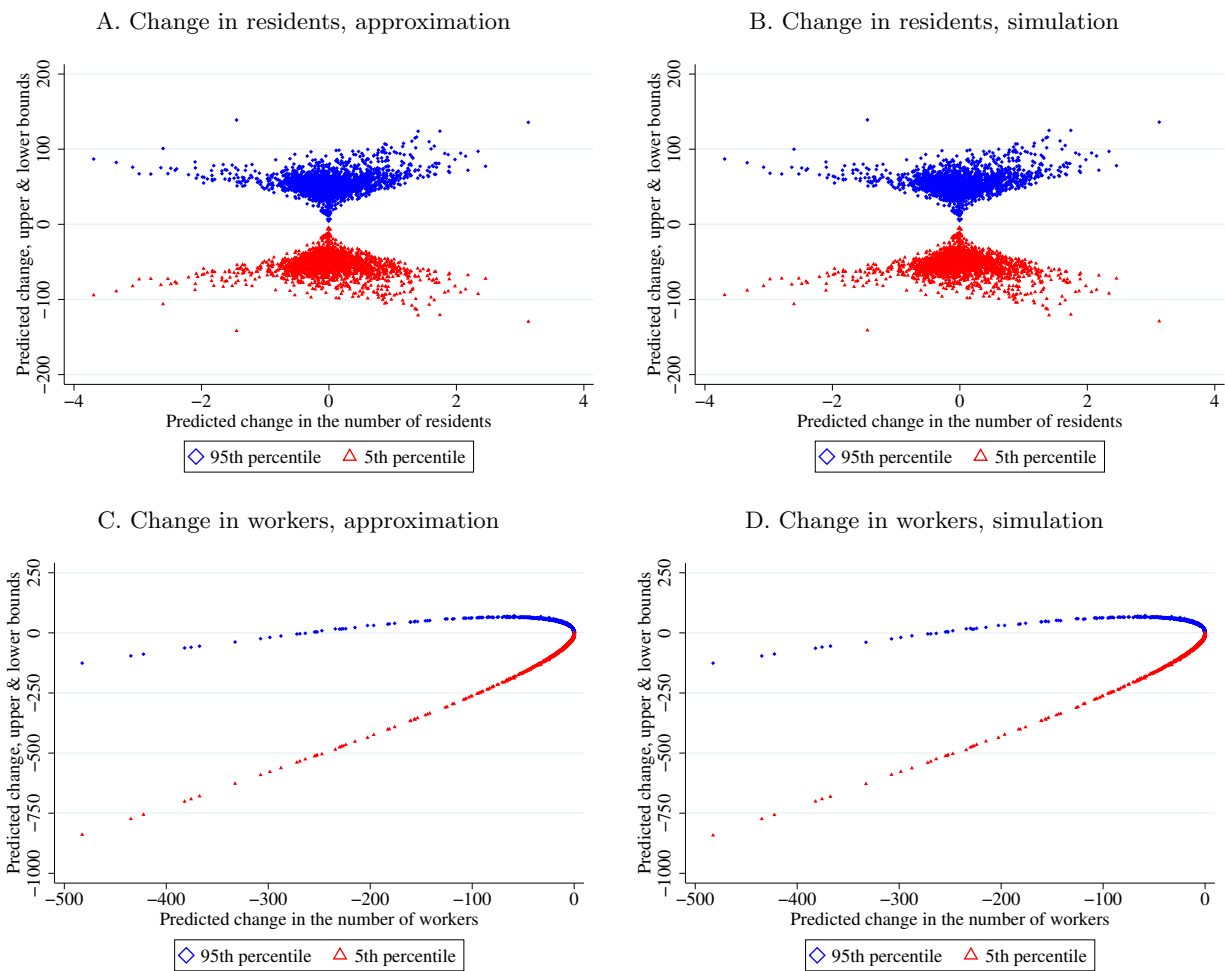
NOTES: The plots depict the 5th and 95th percentiles of predicted change in the number of residents or workers under the granular model, by the approximation method and simulation method mentioned in the main text. For predicted change in residents, panel A plots under the approximation method and panel B plots under the simulation method. For predicted change in workers, panel C plots under the approximation method and panel D plots under the simulation method. In both panels, the treated tract and one outlying tract are excluded from the plot.