STATE OF CALIFORNIA • DEPARTMENT OF TRANSPORTATION
TECHNICAL REPORT DOCUMENTATION PAGE
TR0003 (REV 10/98)

1. REPORT NUMBER	2. GOVERNMENT ASSOCIATION NUMBER	3. RECIPIENT'S CATALOG NUMBER
CA20-3415		

4. TITLE AND SUBTITLE
Optimal Density Restrictions in the Los Angeles-Long Beach CSA

5. REPORT DATE
August 2020

6. PERFORMING ORGANIZATION CODE

7. AUTHOR
Andrii Parkhomenko;Matthew J. Delventhal;Eunjee Kwon

8. PERFORMING ORGANIZATION REPORT NO.

9. PERFORMING ORGANIZATION NAME AND ADDRESS
University of Southern California
3500 S. Figueroa Street, Suite #102
Los Angeles, CA 90089-8001

10. WORK UNIT NUMBER

11. CONTRACT OR GRANT NUMBER
65A0674 TO 15

12. SPONSORING AGENCY AND ADDRESS
California Department of Transportation
Division of Research, Innovation and System Information, MS-83
1727 30th Street, 3rd Floor
Sacramento CA94273-0001

13. TYPE OF REPORT AND PERIOD COVERED
Final Report, Aug 1, 2019 - July 31, 2020

14. SPONSORING AGENCY CODE
Caltrans DRISI

15. SUPPLEMENTARY NOTES

16. ABSTRACT
We build a quantitative general equilibrium model of residence and employment choices under municipal density limits. Developers decide where to build, businesses decide where to offer jobs, and workers decide where to live and work given exogenous location characteristics, transport infrastructure, and zoning restrictions. We use employment, real estate, and commuting data to identify effective density restrictions for 3,917 Census tracts in the Los Angeles metropolitan area. We then compute two counterfactual scenarios. In the first, zoning restrictions on density are relaxed to the level of downtown L.A. in all urban tracts. In the second, massive improvements to transport infrastructure eliminate congestion-related delays. Each change yields large welfare gains. The first scenario leads to larger increases in output and much larger decreases in real estate prices, while the second scenario brings larger reductions in average commuting time.

17. KEY WORDS	18. DISTRIBUTION STATEMENT
	The readers can freely refer to and distribute this report. If there is any questions, please contact one of the authors.

19. SECURITY CLASSIFICATION *(of this report)*	20. NUMBER OF PAGES	21. COST OF REPORT CHARGED
No security issues	50	Free for E-copy

Disclaimer Statement

Zoning and the Density of Urban Development

Aug 2020

A Research Report from the Pacific Southwest Region University Transportation Center

Matthew J. Delventhal, The Robert Day School of Economics and Finance, Claremont McKenna College

Eunjee Kwon, Department of Economics, University of Southern California

Andrii Parkhomenko, Department of Finance and Business Economics, University of Southern California

[page intentionally left blank]

TABLE OF CONTENTS

Contents

About the Pacific Southwest Region University Transportation Center

The Pacific Southwest Region University Transportation Center (UTC) is the Region 9 University Transportation Center funded under the US Department of Transportation's University Transportation Centers Program. Established in 2016, the Pacific Southwest Region UTC (PSR) is led by the University of Southern California and includes seven partners: Long Beach State University; University of California, Davis; University of California, Irvine; University of California, Los Angeles; University of Hawaii; Northern Arizona University; Pima Community College.

The Pacific Southwest Region UTC conducts an integrated, multidisciplinary program of research, education and technology transfer aimed at *improving the mobility of people and goods throughout the region.* Our program is organized around four themes: 1) technology to address transportation problems and improve mobility; 2) improving mobility for vulnerable populations; 3) Improving resilience and protecting the environment; and 4) managing mobility in high growth areas.

U.S. Department of Transportation (USDOT) Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

California Department of Transportation (CALTRANS) Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the United States Department of Transportation's University Transportation Centers program, in the interest of information exchange. The U.S. Government and the State of California assumes no liability for the contents or use thereof. Nor does the content necessarily reflect the official views or policies of the U.S. Government and the State of California. This report does not constitute a standard, specification, or regulation. This report does not constitute an endorsement by the California Department of Transportation (Caltrans) of any product described herein.

Disclosure

Principal Investigator, Co-Principal Investigators, others, conducted this research titled, “Zoning and the Density of Urban Development” at the department of Economics and Finance, Robert Day School, Claremont McKenna College; Department of Economics, Dornsife School, University of Southern California; and Department of Finance and Business Economics, Marshall School, University of Southern California. The research took place from August 1, 2019 to July 31, 2020 and was funded by a grant from the Caltrans in the amount of $35,000. The research was conducted as part of the Pacific Southwest Region University Transportation Center research program.

Acknowledgements

The authors acknowledge generous grant funding from METRANS-PSR and the Lusk Center for Real Estate.

Abstract

We build a quantitative general equilibrium model of residence and employment choices under municipal density limits. Developers decide where to build, businesses decide where to offer jobs, and workers decide where to live and work given exogenous location characteristics, transport infrastructure, and zoning restrictions. We use employment, real estate, and commuting data to identify effective density restrictions for 3,917 Census tracts in the Los Angeles metropolitan area. We then compute two counterfactual scenarios. In the first, zoning restrictions on density are relaxed to the level of downtown L.A. in all urban tracts. In the second, massive improvements to transport infrastructure eliminate congestion-related delays. Each change yields large welfare gains. The first scenario leads to larger increases in output and much larger decreases in real estate prices, while the second scenario brings larger reductions in average commuting time.

Zoning and the Density of Urban Development

Executive Summary

The goal of this project is to identify zoning policy adjustments which could improve access to jobs and quality of life for the residents of the Los Angeles-Long Beach Combined Statistical Area. An important auxiliary goal is to make neighborhood-by-neighborhood projections for how changes to policy would affect a broad range of variables, including commuting flows, urban congestion, and property values, in order to guide the design and implementation of policy adjustments.

To this end, we build a quantitative general equilibrium model of internal city structure. Locations differ in local productivity, employment and residential amenities and are linked by a transportation network that determines commuting times. Zoning policies increase the cost of building in some locations, constraining the equilibrium supply of floor space. We model density restrictions as an endogenous function of existing density.

We use our model to back out local characteristics for the nearly 4,000 census tracts of the Los Angeles-Long Beach Combined Statistical Area, given data on the price of floorspace, wages at place of employment, the density of employment and residence, and commuting times.

We then conduct two counterfactual experiments. In the first experiment, we reduce density zoning restrictions to the level of downtown L.A. in all urban tracts in the metropolitan area. The second experiment simulates a tremendous improvement in transport infrastructure: we suppose that all commuters can drive directly to their destination at 65 miles per hour without slowing down either for traffic or for curves in the road.

We find that relaxing zoning increases output per worker by 35% and welfare by 57% (Table 1, second column). At the same time, it reduces residential and commercial floor prices by 54% and 60%, respectively, on average. It also slightly reduces the average daily commute--presumably because increased concentration means there is less need to commute long distances. Figure 1 illustrates how the relaxation in density limits changes residential and employment density in the metro area.

Improving automobile infrastructure to the maximum--indeed, probably an impossible or impossibly costly project in practice--increases output per worker by 2.2%, and nearly halves the average daily one-way commute to just 19 minutes. Average welfare goes up by 30% (Table 1, third column). At the same time, it slightly raises the mean costs of residential and commercial floorspace by about 1% and 9% respectively--presumably because improved transportation increases the demand for floorspace in the most attractive tracts pushing the density there to the limit. Figure 2 illustrates how the increase in speed changes residential and employment density.

The results of these two experiments suggest that the potential productivity and housing affordability gains from re-zoning may be more substantial. They also suggest that gains from improving transport infrastructure may have an upper bound, as even the best possible automobile-based improvement had limited impact on variables other than time spent commuting.

	Benchmark	Counterfactual 1: Upzoning	Counterfactual 2: Increase speed
Output per worker	100.0	135.4	102.2
Employment	100.0	100.0	100.0
Welfare	100.0	156.6	130.1
Mean wages	100.0	135.4	102.2
Mean residential floorspace price	100.0	45.6	101.2
Mean commercial floorspace price	100.0	40.1	109.4
Mean commuting time, min	36.1	35.6	19.4
Mean commuting distance, km	26.7	24.6	33.9
Median tract density	1.34	1.39	1.34
Mean house size	100.0	264.6	101.6

Table 1. Counterfactual Result

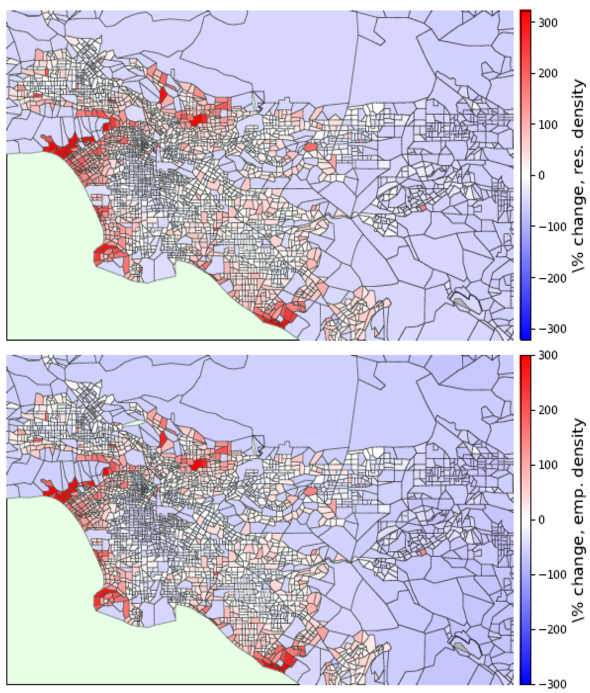


Figure 1. Impact of the increase density limits to the level of DTLA in all urban tracts on residential density (upper) and employment density (lower).

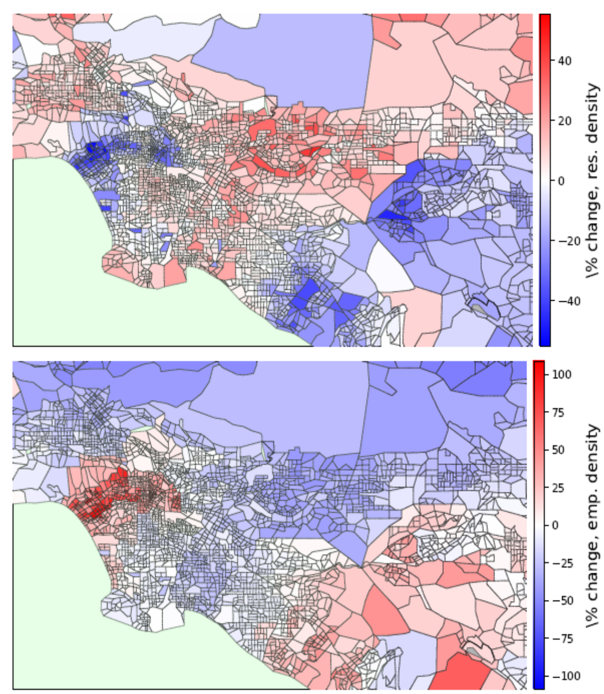


Figure 2. Impact of the increase commuting speed to 65 miles per hour on residential density (upper) and employment density (lower).

Introduction

Density-limiting zoning is a pervasive force shaping the American urban landscape. By limiting the supply of floorspace in commercial districts and the supply of housing nearby, these regulations have a powerful impact on the organization of production, commuting patterns, and real estate prices. As pundits talk of a housing affordability crisis and proposals to limit or eliminate single-family zoning circulate among city councils and state capitols, the importance of understanding the consequences of constrained urban density is as high as ever.

To this end, we build a quantitative general equilibrium model of internal city structure. Locations differ in local productivity, employment and residential amenities and are linked by a transportation network that determines commuting times. Zoning policies increase the cost of building in some locations, constraining the equilibrium supply of floor space. We model density restrictions as an endogenous function of existing density. Businesses decide where to offer jobs given local productivity, price of floor space, and access to workers. Workers choose where to live given local residential amenities, price of floor space, and access to jobs. Local productivity and residential amenities are each affected by positive spillovers, which are increasing in the density of nearby employment and residence.

We choose Los Angeles and its orbit as our laboratory for policy experiments. Increasingly restrictive zoning codes pushed the potential total population of the municipality of Los Angeles down from 10 million in 1965 to 4 million in 1992, with only slight increases since then[1]. Real estate developers struggle to supply enough housing to meet the demand[2]. As a result, America's second most-populous urban area is known not for its skyscrapers but for its bungalows and car culture. We use our model to back out local characteristics for the nearly 4,000 census tracts of the Los Angeles-Long Beach Combined Statistical Area, given data on the price of floorspace, wages at place of employment, the density of employment and residence, and commuting times. We then conduct two counterfactual experiments. In the first experiment, we reduce density zoning restrictions to the level of downtown L.A. in all urban tracts in the metropolitan area. The second experiment simulates a tremendous improvement in transport infrastructure: we suppose that all commuters can drive directly to their destination at 65 miles per hour without slowing down either for traffic or for curves in the road.

We find that relaxing zoning increases output per worker by 35% and welfare by 57%. At the same time, it reduces residential and commercial floor prices by 54% and 60%, respectively, on average. It also slightly reduces the average daily commute--presumably because increased concentration means there is less need to commute long distances. In the second scenario,

[1] See Morrow (2013).

[2] Saiz (2010) and Baum-Snow and Han (2019) document that Los Angeles has one of the lowest housing supply elasticities in the U.S.

improving automobile infrastructure to the maximum--indeed, probably an impossible or impossibly costly project in practice--increases output per worker by 2.2%, and nearly halves the average daily one-way commute to just 19 minutes. Average welfare goes up by 30%. At the same time, it slightly raises the mean costs of residential and commercial floorspace by about 1% and 9% respectively--presumably because improved transportation increases the demand for floorspace in the most attractive tracts pushing the density there to the limit. The results of these two experiments suggest that the potential productivity and housing affordability gains from re-zoning may be substantial. They also suggest that gains from improving transport infrastructure may have an upper bound, as even the best possible automobile-based improvement had limited impact on variables other than time spent commuting.

Prior research has had little to say about the quantitative consequences of limiting local density within an urban area. Two papers by Kirti Joshi and Tatsuhito Kono (2009 and 2018) explore the question of optimal density restriction in a rich theoretical setting with both positive and negative externalities between different types of land use. Brueckner and Singh (2017) use land prices and zoning maps to measure the effective stringency of building-height restrictions in five U.S. metro areas.

There are a number of studies which explore the implications of zoning policies other than density restriction. Rossi-Hansberg (2004) analyzes the optimal division of a city into commercial and residential zones in a monocentric city model. Kantor, Rietveld, and van Ommeren (2014) extend this setting to consider the effects of congestion. Zhang and Kockelman (2016) develop a linear city model which they use to explore policies such as congestion pricing and urban growth boundaries. Allen, Arkolakis, and Li (2016) explore the division between commercial and residential land use quantitatively in a model with exogenously-given density, and develop recommendations for welfare-improving regulatory changes in the city of Chicago.

There is also a rich literature exploring the political economy of land use regulation in stylized city models and systems of cities. Helsley and Strange (1995), Calabrese, Epple, and Romano (2007), Sole-Olle and Viladecans-Marsal (2012), Hilber and Robert-Nicoud (2013), Ortalo-Magne and Prat (2014) and Parkhomenko (2018) each explore a different aspect of the political and economic forces that may bring about restrictions on land use in the first place.

Also related Ahlfeldt, Redding, Sturm, and Wolf's 2015 paper in which they use a quantitative spatial model to explore the consequences of the cold war division of Berlin and subsequent reunification--what may be thought of as a particularly extreme form of “zoning”. There have also been a number of studies, such as Albouy and Ehrlich (2018), Herkenhoff, Ohanian, and Prescott (2018) and Hsieh and Moretti (2019) which have measured the extent of variation in land-use restriction across U.S. metro areas and estimated the impact on the national economy as a whole.

Model

In this section we construct a version of Ahlfeldt, Redding, Sturm, and Wolf (2015)'s quantitative urban model to incorporate an explicit notion of density-restricting zoning. Consider an urban area embedded in a larger economy. The urban area consists of a set of I discrete locations indexed from 1 to I. Each location $i \in \{1,2, \ldots I\}$ has an exogenous supply of buildable land Λ_i. Developers use land and the consumption good to produce a supply H_i of floorspace which can be put to residential or commercial use. Firms use labor and commercial floorspace to produce the consumption good, which is traded costlessly inside the urban area. Workers supply their labor to firms and consume residential floorspace and the consumption good.

Workers are perfectly mobile between the urban area and the larger economy. Those living outside the urban area enjoy utility $\overline{U}$, which pins down the endogenous measure N who choose to live inside it. These insiders must choose a location i in which to consume residential floorspace and a location j in which to earn a wage. Workers suffer disutility from time spent commuting from their home to their work. This time is given by t_{ij} which is determined by an exogenously given transportation network. Workers' utility is affected by the local employment amenity E_j of their work location and the local residence amenity X_i of their residence location, which combines an exogenous component x_i with positive spillovers from other locations. Each worker's utility is also determined by an idiosyncratic utility shock drawn for each possible pair of residence and work locations.

Firms' total factor productivity in each location is determined by A_i, which combines an exogenous component a_i with positive spillovers from other locations. Developers' choice of how much residential and commercial floorspace to supply is influenced by $\xi_i > 0$, a zoning parameter which determines the cost of supplying commercial floorspace relative to residential. Developers' choice of how much floorspace to build in each location i is also influenced by density restriction policies in a given tract.

Workers

Worker n who resides in i and lives in j enjoys utility U_{ijn} determined by

$$U_{ijn} = \frac{z_{ijn}}{e^{\kappa t_{ij}}} \left(\frac{c_{ijn}}{1-\gamma}\right)^{1-\gamma} \left(\frac{h_{ijn}}{\gamma}\right)^{\gamma}, \quad (1)$$

where $\gamma \in (0,1)$ is the share of housing in expenditures, z_{ijn} represents an idiosyncratic shock, and the parameter $\kappa > 0$ determines the relationship between travel time and workers' disutility from commuting. Consumption of the final good, c_{ijn}, and consumption of residential floorspace h_{ijn} are subject to the budget constraint

$$w_j = c_{ijn} + q_{Ri} h_{ijn}, \tag{2}$$

where w_j represents the wage earned by working in location j and q_{Ri} is the price of residential floor space in location i. Idiosyncratic shocks z_{ijn} are drawn from a Frèchet distribution with elasticity $\epsilon > 1$ which has the following CDF:

$$F(z_{ij}) = e^{-X_i E_j z_{ij}^{-\epsilon}}. \quad (3$$

In the above formulation, X_i represents the average utility derived from living in location i, and E_j represents the average utility derived from working in location j.

Optimizing consumption choices yields the following indirect utility:

$$u_{ijn} = \frac{z_{ijn} w_j q_{Ri}^{-\gamma}}{e^{\kappa t_{ij}}} \tag{4}$$

Integrating over workers, the probability that a worker chooses to reside in block i and work in block j can be given by

$$\pi_{ij} = \frac{X_i E_j \left(e^{-\kappa t_{ij}} w_j q_{Ri}^{-\gamma}\right)^{\epsilon}}{\sum_{r=1}^{I} \sum_{s=1}^{I} X_r E_s \left(e^{-\kappa t_{rs}} w_s q_{Rr}^{-\gamma}\right)^{\epsilon}} \tag{5}$$

The probability that a worker works in j conditional on working in i is given by

$$\pi_{ij|i} = \frac{E_j \left(w_j e^{-\kappa t_{ij}}\right)^{\epsilon}}{\sum_{s=1}^{I} E_s \left(w_s e^{-\kappa t_{is}}\right)^{\epsilon}} \tag{6}$$

Let N_{Ri} represent the measure of workers residing in location i and N_{Wi} represent the measure of workers working there. These two vectors are related by the following equation:

$$N_{Wj} = \sum_{i=1}^{I} \pi_{ij|i} N_{Ri} \tag{7}$$

Define $\widetilde{w}_i$ as the average wage earned by residents of location j . This is given by

$$\widetilde{w}_i \equiv \sum_{j=1}^{I} \pi_{ij|i} w_j \tag{8}$$

Define $\widetilde{U}$ as the expected utility enjoyed by a resident of the city. This is given by

$$\widetilde{U} \equiv \Gamma\left(\frac{\epsilon-1}{\epsilon}\right) \left[\sum_{r=1}^{I} \sum_{s=1}^{I} X_r E_s \left(e^{-\kappa t_{rs}} w_s q_{Rr}^{-\gamma}\right)^{\epsilon}\right]^{\frac{1}{\epsilon}} \tag{9}$$

In equilibrium $\widetilde{U}$ must be equal to the reservation utility $\overline{U}$. The reservation utility determines the equilibrium total employment of the urban area.

Firms

Output of the final good that is produced in y_j is determined by the Cobb-Douglas production function

$$A_j N_{Wj}^{\alpha} H_{Wj}^{1-\alpha}, \tag{10}$$

where $A_j > 0$ represents total factor productivity, N_{Wj} is the total measure of workers employed and H_{Wj} is the total amount of commercial floorspace employed. Profit maximization implies that

$$N_{Wj} = \left(\frac{\alpha A_j}{w_j}\right)^{\frac{1}{1-\alpha}} H_{Wj} \tag{11}$$

$$q_{Wj} = (1-\alpha)\left(\frac{\alpha}{w_j}\right)^{\frac{\alpha}{1-\alpha}} A_j^{\frac{1}{1-\alpha}} \tag{12}$$

Developers

There is a large number of perfectly competitive floorspace developers operating in each location. Floorspace is produced using the following technology:

$$H_i = K_i^{1-\eta}(\phi_i(H_i)L_i)^{\eta}, \tag{13}$$

where $L_i \leq \Lambda_i$ and K_i are the amounts of land and the final good used to produce floorspace, and η is the share of land in production. Function $\phi_i(H_i)$ is defined as

$$\phi_i(H_i) = \bar{\phi}\left(1 - \frac{H_i}{\bar{H}_i}\right), \tag{14}$$

and determines the land-augmenting productivity of floorspace developers in location i.[3] Parameter $\bar{H}_i$ determines the density limit in tract i. When H_i approaches $\bar{H}_i$, $\phi_i(H_i)$ approaches zero. As a result, it becomes very costly to build due to regulatory or political barriers, such as zoning, floor-to-area ratios, or local opposition to development.

Developers sell floorspace at price $\bar{q}_i \equiv \min\{q_{Ri}, q_{Wi}\}$ to either residential or commercial users. However, the price that residents or firms pay for floorspace may differ from $\bar{q}_i$ due to zoning restrictions. The wedge between prices for residential and commercial floorspace is denoted by parameter $\xi_i > 0$. If $\xi_i < 1$, regulations increase the cost of supplying residential floorspace. If $\xi_i > 1$, regulations increase cost of supply commercial floorspace. Thus,

$$q_{Wi} > \xi_i q_{Ri} \Leftarrow \frac{H_{Wi}}{H_{Wi}+H_{Ri}} = 1$$

$$q_{Wi} = \xi_i q_{Ri} \Leftarrow \frac{H_{Wi}}{H_{Wi}+H_{Ri}} \in (0,1)$$

$$q_{Wi} < \xi_i q_{Ri} \Leftarrow \frac{H_{Wi}}{H_{Wi}+H_{Ri}} = 0$$

[3] This function was also used in Favilukis, Mabille, and Van Nieuwerburgh (2019) to model density limits.

Land-market clearing and profit maximization imply that the equilibrium supply of floorspace is

$$H_i = \phi_i(H_i)((1-\eta)\bar{q}_i)^{\frac{1-\eta}{\eta}} L_i. \tag{15}$$

Solving this expression for H_i, using the definition of construction efficiency in (14), yields

$$H_i = \frac{\bar{\phi}((1-\eta)\bar{q}_i)^{\frac{1-\eta}{\eta}} L_i}{1+\bar{\phi}((1-\eta)\bar{q}_i)^{\frac{1-\eta}{\eta}} L_i/\bar{H}_i}. \tag{16}$$

Externalities

Total factor productivity in location j is determined by an exogenous component, a_j, and an endogenous component that depends on the density of production in every other location s, weighted inversely by the travel time from that j to s:

$$A_j = a_j \left[\sum_{s=1}^{I} e^{-\delta t_{js}} \left(\frac{N_{Ws}}{L_s}\right)\right]^{\lambda} \tag{17}$$

The residential amenity in location i is determined by an exogenous component, x_j, and an endogenous component that depends on the density of residence in every other location, weighted inversely by the travel time from that location from i:

$$X_i = x_i \left[\sum_{s=1}^{I} e^{-\rho t_{is}} \left(\frac{N_{Rs}}{L_s}\right)\right]^{\chi\epsilon} \tag{18}$$

Urban Structure of the L.A.-Long Beach C.S.A.

Los Angeles is the second-most-populous city in the United States, yet it has a lower population density than any East Coast city of comparable size. According to Morrow (2013), the city's relatively low density is the result of progressively more stringent zoning regulations pushed into law by the "slow growth" movement between 1965 and 1992. In 1965, while the city's population stood at only 2.5 million, its relatively permissive zoning laws would have allowed the building of housing to accomodate up to 10 million residents. By 1992, tighter regulations had reduced that maximum capacity to 4 million, only barely higher than the city's population at the time. After 1992, both the maximum capacity and the population have grown slowly.

The Los Angeles-Long Beach Combined Statistical Area had a total population of 18.7 million in 2018, distributed across a total land area of 88,000 square kilometers.[4] It comprises five counties (Los Angeles, Orange, Riverside, San Bernardino, and Ventura) and 3,917 census tracts. We focus on the 5-year period between 2012 and 2016. In what follows, we describe the sources and methods we use to construct tract-level data data on residence, employment, wages, prices of residential and commercial floorspace, and commuting times.

[4] U.S. Census Bureau (2020a)

Residence, employment and wages

We take data on the number workers residing and working in each census block from the Census Bureau's LEHD Origin-Destination Employment Statistics,[5] aggregating the data to the census tract level and calculating averages over the period from 2012 to 2016. Tables 1, 2 and 3 summarize basic facts about the tracts in the LA-Long Beach CSA broken down by county. The total land area of the metro area is 87,876, approximately half of which is sparsely-populated desert in San Bernardino County. Both Los Angeles and Orange counties have positive net inflows of commuters (approximately +200,000 each), while the relatively more peripheral Riverside, San Bernardino and Ventura counties have more residents than jobs.

Table 1. Land Area of Tracts (km^2)

County	Sum	Mean	Median	St. Dev.	Max.	N. Obs.
Los Angeles	10,508	4.5	1.0	34.0	1,028.9	2,342
Orange	2,048	3.5	1.6	10.9	169.9	582
Riverside	18,658	41.3	3.5	467.6	9,860.6	452
San Bernardino	51,947	140.8	3.4	1,078.8	18,106.9	369
Ventura	4,715	27.4	3.3	182.1	2,380.2	172
All	87,876	22.4	1.4	372.4	18,106.9	3,917

Table 2 Residence by Tract

County	Sum	Mean	Median	St. Dev.	Max.	N. Obs.
Los Angeles	4,121,308	1,759.7	1,690	659.3	4,377.8	2,342
Orange	1,379,953	2,371.1	2,281	890.5	7,652.4	582
Riverside	822,196	1,819.0	1,686	882.6	6,848.6	452
San Bernardino	766,530	2,077.3	1,902	906.7	7,710.2	369
Ventura	368,403	2,141.9	2,028	887.6	6,136.2	172
All	7,458,390	1,904.1	1,800	794.1	7,710.2	3,917

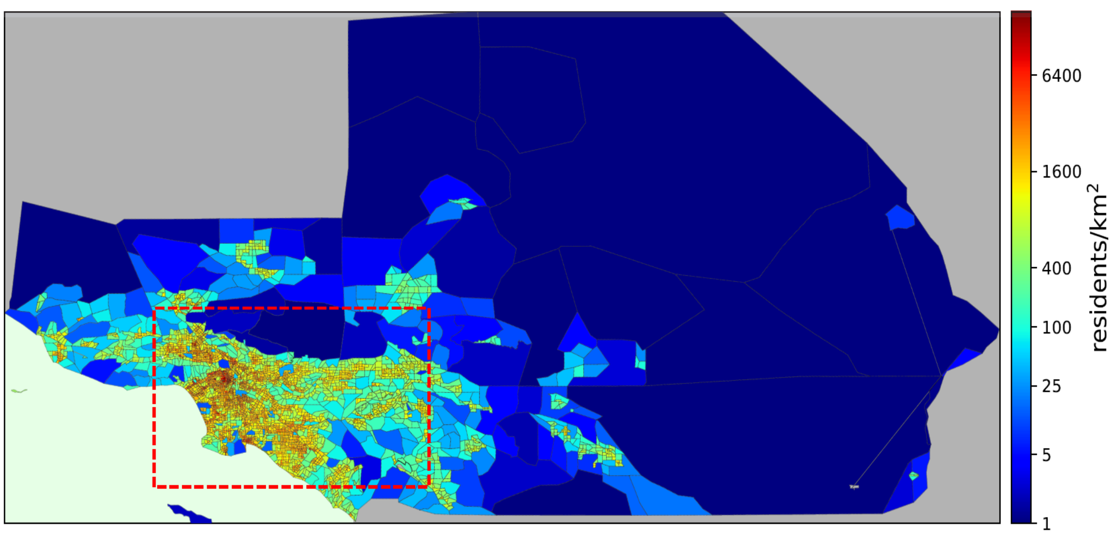
[5] U.S. Census Bureau (2020b)

Table 3 Employment by Tract

County	Sum	Mean	Median	St. Dev.	Max.	N. Obs.
Los Angeles	4,368,853	1,865.4	744	4,830.8	136,332.0	2,342
Orange	1,535,811	2,638.9	1,012	6,218.7	84,193.4	582
Riverside	623,484	1,379.4	634	2,407.1	23,929.2	452
San Bernardino	658,071	1,783.4	767	3,810.0	39,814.2	369
Ventura	301,115	1,750.7	713	2,624.0	15,751.6	172
All	7,487,333	1,911.5	768	4,706.7	136,332.0	3,917

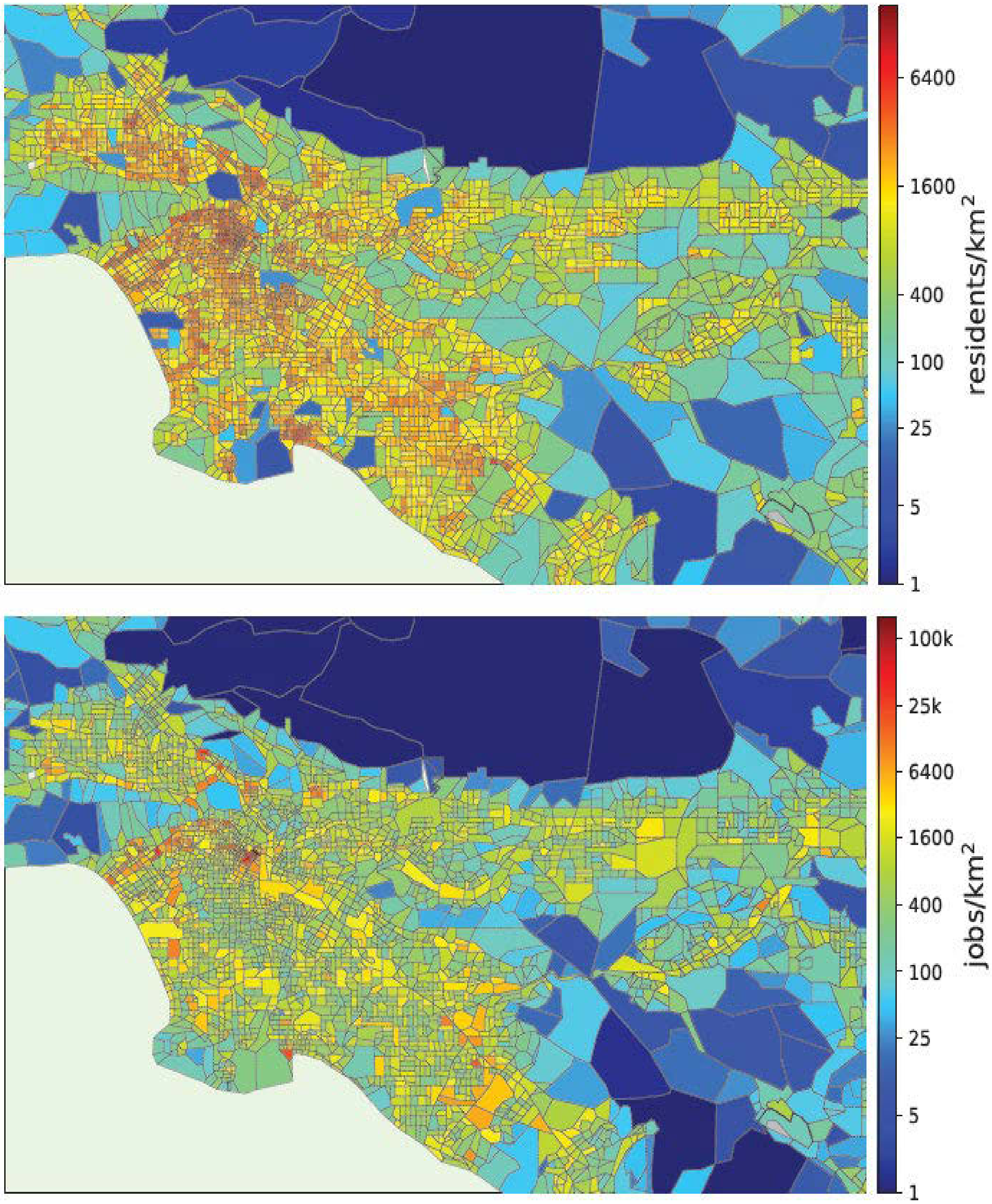
Figure 1 plots the density of residents for the entire metropolitan area including its farthest reaches in the sparsely-populated desert. The narrower core area which will be the focus of future plots can be seen outlined in red. Figure 2 plots the residential and employment density for this narrower area. Both these figures and from Tables 2 and 3 suggest that employment is much more concentrated than residence. The tract with the most jobs has more than 136,000, while the tract with the most residents has only 7,700.

Figure 1 Residential Density, Broader Metro Area



We construct tract-level average wages for 2012-2016 combining data from the CTPP database,[6] with data from the American Community Survey.[7] Further details can be found in in the appendix. Figure 3 shows data wages for each tract.

Figure 2 Residential and Employment Density



[6] U.S. Department of Transportation (2020)

[7] U.S. Census Bureau (2020a)

Figure 3 Wages

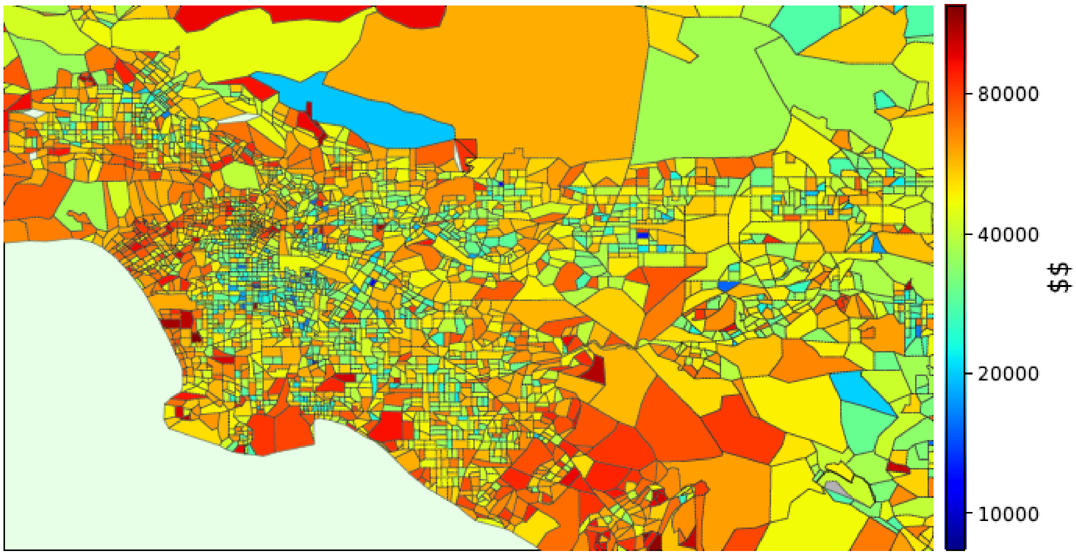


Table 4 gives summary statistics by tract for residential density, employment density, and wage weighted by employees. Here we find further evidence that employment is more concentrated than residence. The median employment density, 560 workers per square kilometer, is less than half the median residence density of 1,350 residents per square kilometer. At the same time, the maximum tract employment density is 158,000 workers per square kilometer, nearly 10 times the maximum tract residence density of 15,900 residents per square kilometer. The maximum tract average wage is $124,000 per year, firms in the median tract pay their workers only $56,000 per year.

Table 4 Data Overview

	Mean	Median	St. Dev.	Max.	N. Obs.
Residents/km ^{2}	1,586.0	1,353.7	1,375.6	15,867.8	3,917
Workers/km ^{2}	1,262.6	564.3	3,928.6	157,995.7	3,917
Wages ($$, weight by employees)	58,767	56,440	18,137	123,757	3,914

Real estate prices

Our commercial and residential property price is from DataQuick, which is transaction-level public records on property characteristics and transactions data.[8] The dataset covers 2,354,535 properties from 2007-2016 in LA-LB CSA area (i.e., Los Angeles, Orange, San Bernardino, Riverside, and Ventura counties). The data provides information such as sales price, GIS coordinates, transaction dates, property use, transaction type, number of rooms, number of baths, square-footage, lot size, year built, etc. Following Baum-Snow and Han (2019), we use

[8] This data is frequently used in the recent literature (e.g., Diamond and McQuade (2019))

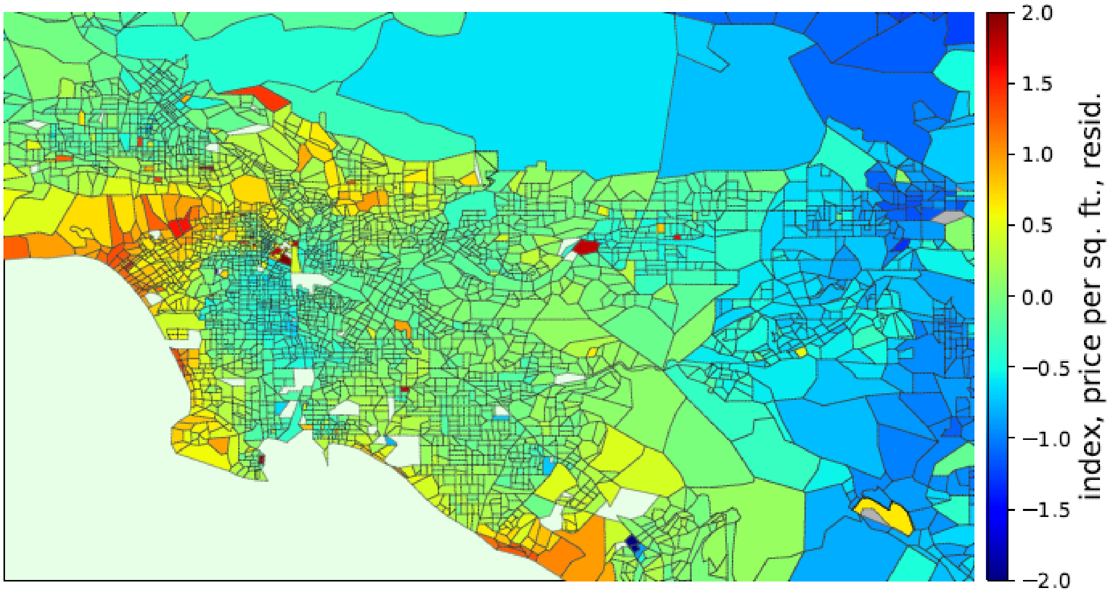
hedonic regressions to obtain floor space price indices that reflect the value of a constant-quality unit of commercial or residential floor space in a given location. Further details can be found in the appendix.

Table 5 Data Overview, Real Estate Prices

	Mean	Median	St. Dev.	Max.	N. Obs.
Res. flr. space index (weight by residents)	-0.12	-0.08	0.50	3.22	3,853
Com. flr. space index (weight by employees)	0.83	0.81	0.42	1.94	3,867

Table 5 gives summary statistics for residential floor price index weighted by residents and commercial floor price index weighted by employees.

Figure 4 Floor Space Price Indices



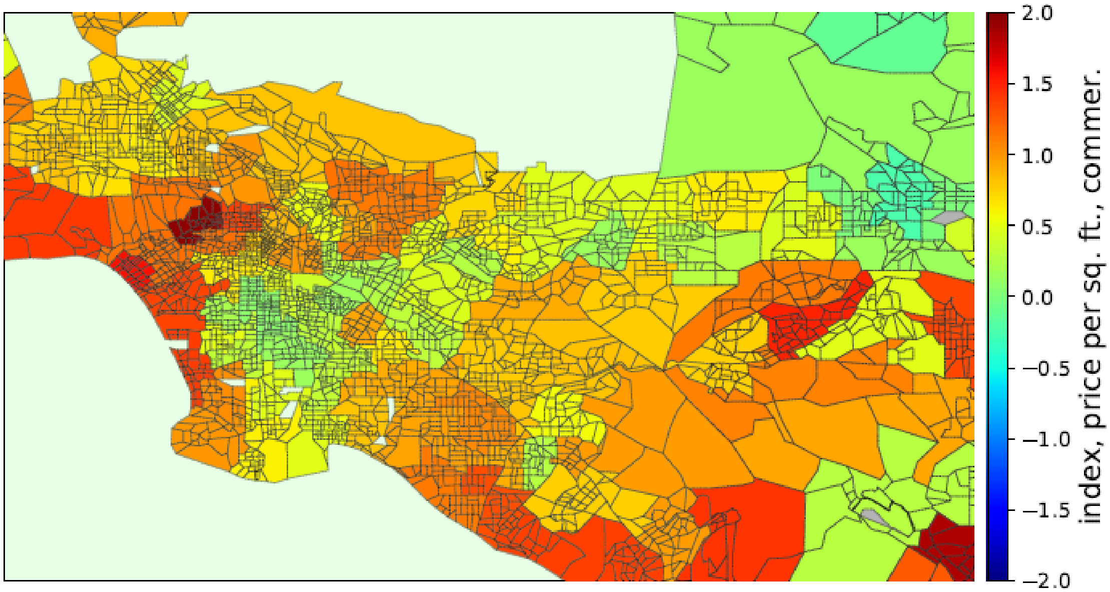


Figure 4 shows the residential and commercial floor space price indices for each tract.

Table 6 Correlations

	log Res./km ^{2}	log Emp./km ^{2}	log wage	Res.Price index	Comm.Price index.
log Res./km ^{2}	1				
log Emp./km ^{2}	0.70	1			
log wage	-0.14	0.10	1		
Res. Price index	0.34	0.39	0.34	1	
Comm. Price index	0.11	0.18	0.28	0.47	1

Table 6 shows correlations between the five primary tract-level variables. Note that the 0.70 correlation between residential density and employment density is high but much less than one. The 0.47 correlation between the prices of residential and commercial floor space is clearly positive but not very strong. And there is a weak negative correlation, -0.14, between residential density and the wages of employees in the same location, indicating that, on average, the highest-paying jobs are located in different tracts than the highest concentrations of residence.

Commuting

The CTPP database provides data on workers' reported commute times by tract-of-residence/tract-of-employment pair for the period 2012-2016, which can be used to calculate average commute times between tracts. A key advantage of this data is that is derived from responses by actual commuters, and so reflects actual commute times as influenced by all relevant facts including infrastructure, congestion, and parking availability. A challenge presented by this data is that it only directly provides information on tract-pairs sampled by the CTPP survey, which are a heavily-traveled subset of the 15.3 million possible tract-to-tract trajectories.[9]

In order to use the observed information to infer commuting times for the entire set of possible residence-employment location pairs, we take the observed links as first-order connections in a transport network. In other words, we assume that if there is not a direct link from A to C, but there are links from A to B and B to C, a commuter could arrive at C from A by using B as a waypoint. We then use Dijkstra's algorithm to calculate the quickest routes between each origin and destination. Additional details on this procedure can be found in section C in the appendix.

Table 7 gives an overview of estimated commuting times along all 15.3 million origin-destination trajectories. Distance here is the straight-line (great circle) distance between tract centroids in kilometers, and times are given in minutes. Here we see that commuting flows are lumpy across tract pairs. No commuters are observed for approximately two thirds of possible trajectories, accounting for the low mean and median commuter flow, while the most popular origin-destination pair has over 1,100 daily commuters. The mean and median commuting times are around half an hour.

Table 7 Commuting Overview

	Mean	Median	St. Dev.	Max.	N. Obs.
Commuters	0.45	0.00	2.88	1,184.60	15,342,889
Distance	26.42	16.34	29.43	472.74	15,342,889
Travel time	33.03	28.36	20.33	523.55	15,342,889

[9] These 270,436 tract-pairs with positive flows in the CTPP data for 2012-2016 contain 34.3% of the 6.9 million commuters in our sample. They represent only 4.9% of all 5.6 million positive tract-to-tract flows, but cover 82.1% of the tract-to-tract flows of 25 commuters or more, and 63.2% of the flows of 10 commuters or more.

Figure 5 Commuting Time to Downtown

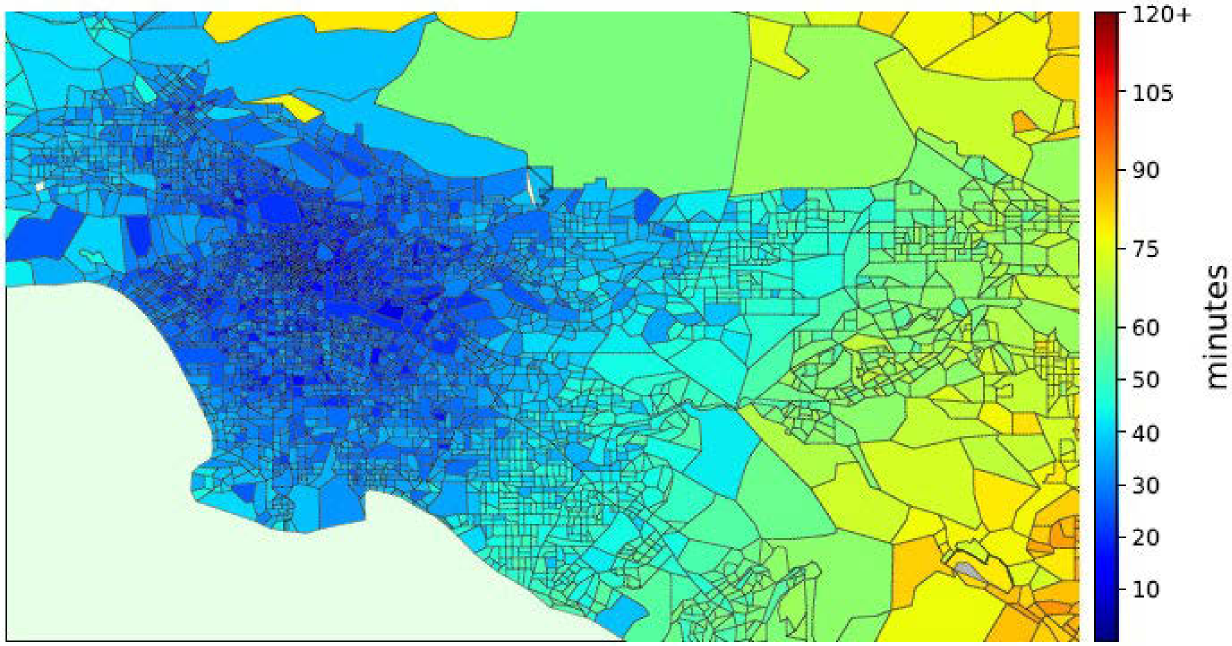


Figure 5 plots average commute times to census tract 6037207400 in the heart of downtown Los Angeles. Figure 6 plots the residences of the people who work in that same tract in the upper panel. In the lower panel, it plots the residences of workers in tract 6037701902 in Santa Monica. It is clear that the downtown tract, home to skyscraper offices that employ tens of thousands of people, draws workers from a relatively broad and evenly-spread cachement area. This can be contrasted with the Santa Monica tract, the vast majority of whose workers live relatively nearby.

Figure 6 Residence of Workers in Downtown and Santa Monica

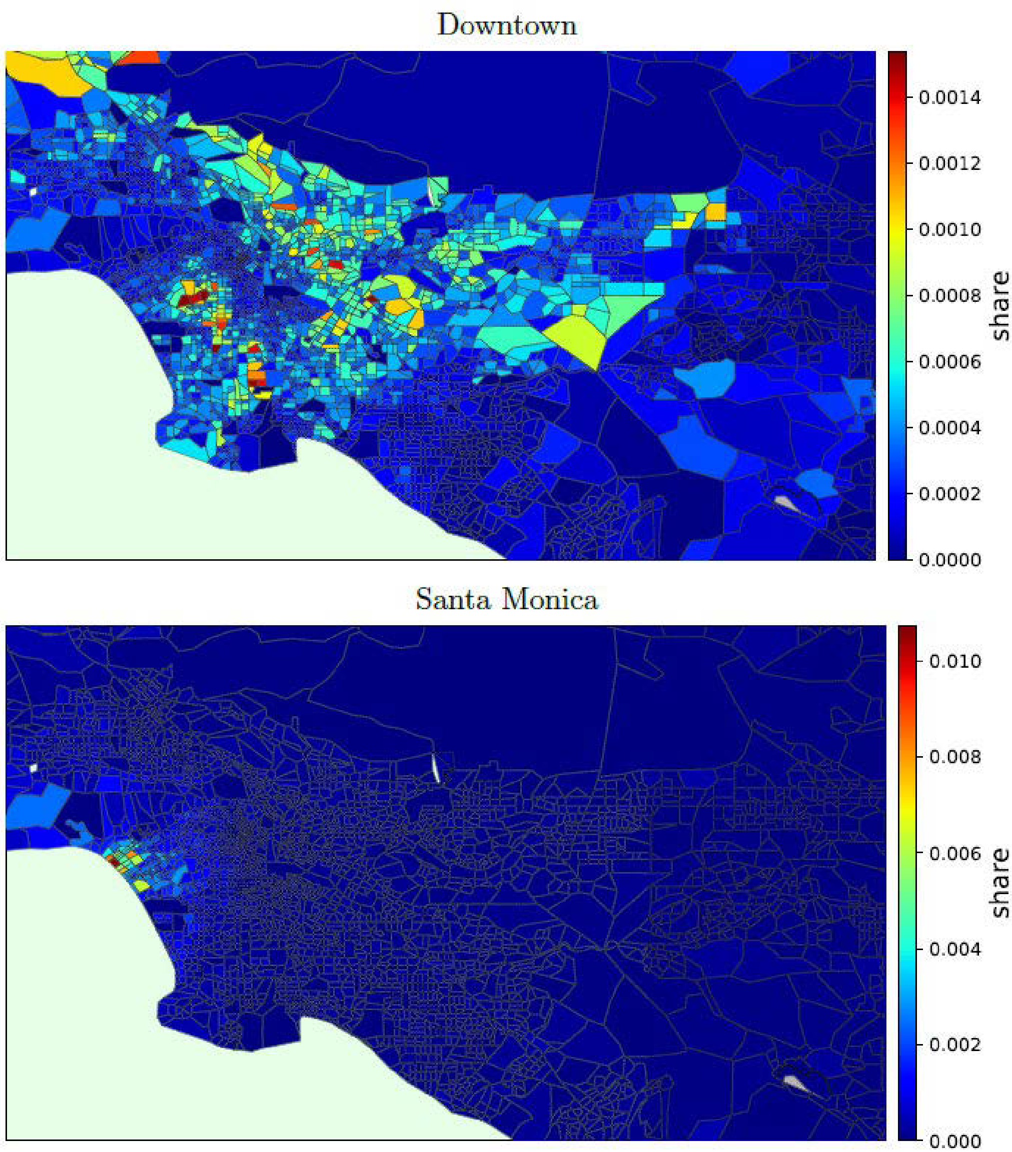


Table 8 displays correlations between commuting flows, physical distance, and commuting times at the tract pair level. Distance is strongly positively correlated with travel time, and number of commuters is weakly negatively correlated with both distance and travel time.

Table 8 Commuting Correlations

	N. of commuters	Distance	Commute time
N. of commuters	1		
Distance	-0.13	1	
Commute time	-0.13	0.96	1

Calibration

Table 9 summarizes parameters used in the quantitative model and their sources.

Table 9 Parameters

Parameter	Value	Source
Share of housing in expenditures	$\gamma = 0.25$	Davis and Ortalo-Magne (2011)
Labor share	$\alpha = 0.2$	Ahlfeldt, Redding, Sturm, and Wolf (2015)
Land share	$\eta = 0.25$	Combes, Duranton, and Gobillon (2018)
Variance of Frechet shocks	$\epsilon = 6.6491$	Ahlfeldt, Redding, Sturm, and Wolf (2015)
Disutility of commuting	$\kappa = 0.0105$	Ahlfeldt, Redding, Sturm, and Wolf (2015)
Decay of residential amenities	$\rho = 0.7595$	Ahlfeldt, Redding, Sturm, and Wolf (2015)
Decay of productive amenities	$\delta = 0.3617$	Ahlfeldt, Redding, Sturm, and Wolf (2015)
Residential amenity spillover	$\chi = 0$	
Productive amenity spillover	$\lambda = 0$	

In order to solve the model, we also need to know vectors of structural residuals: E, x, a, ξ, and $\bar{H}$. The model provides a relationship between each of these structural residuals and equilibrium prices and quantities. Using the data, we can then back out these residuals.

In order to solve for E, first, define $\hat{E}_j \equiv E_j w_j^{\epsilon}$. Then from equations (6) and (7), $\hat{E}_j$ can be defined implicitly as:

$$\hat{E}_j = N_{Wj} \left(\sum_{i=1}^{I} \frac{N_{Ri} d_{ij}^{-\epsilon}}{\sum_{s=1}^{I} \hat{E}_s d_{is}^{-\epsilon}} \right)^{-1}, \tag{19}$$

where N_{Wj} and N_{Ri} are observed tract-level employment and residential populations, and $d_{is} \equiv e^{\kappa t_{ij}}$, where t_{ij} are observed average commuting times from tract i to tract j. A vector $\hat{E}$ is solved recursively using equation (19) and then the vector of residuals E is recovered as $E_j = \hat{E}_j / w_j^{\epsilon}$, using observed tract-level wages.

A similar procedure is applied to solve for x. First, define $\hat{X}_j \equiv X_j q_{Rj}^{-\gamma\epsilon}$. $\hat{X}_j$ can be defined implicitly as:

$$\hat{X}_i = N_{Ri} \left(\sum_{j=1}^{I} \frac{N_{Wj} d_{ij}^{-\epsilon}}{\sum_{r=1}^{I} \hat{X}_r d_{rj}^{-\epsilon}} \right)^{-1}. \tag{20}$$

A vector $\hat{E}$ is solved recursively using equation (20) and then the vector of residuals X is recovered as $X_j = \hat{X}_j q_{Rj}^{\gamma\epsilon}$, using observed tract-level prices of residential floorspace. Then the exogenous part of local amenities, x_j, can be recovered using equation (18) and the data on local residential population and land area.

The vector of local productivities A can be solved for using (12) and the data on wages and commercial floorspace prices as follows:

$$A_j = \alpha^{\alpha}(1-\alpha)^{1-\alpha} w_j^{\alpha} q_{Wj}^{1-\alpha}. \tag{21}$$

Then the exogenous part a can be recovered using equation (17) and the data on local employment and land area.

Since we observe commercial and residential floorspace prices for all Census tracts, we can calculate the zoning parameter ξ_i as

$$\xi_i = \frac{q_{Wi}}{q_{Ri}} \tag{22}$$

Finally, in order to recover $\overline{H}_i$, we use market clearing conditions for land and floorspace. Combining them, we can recover $\overline{H}_i$ from the following relationship:

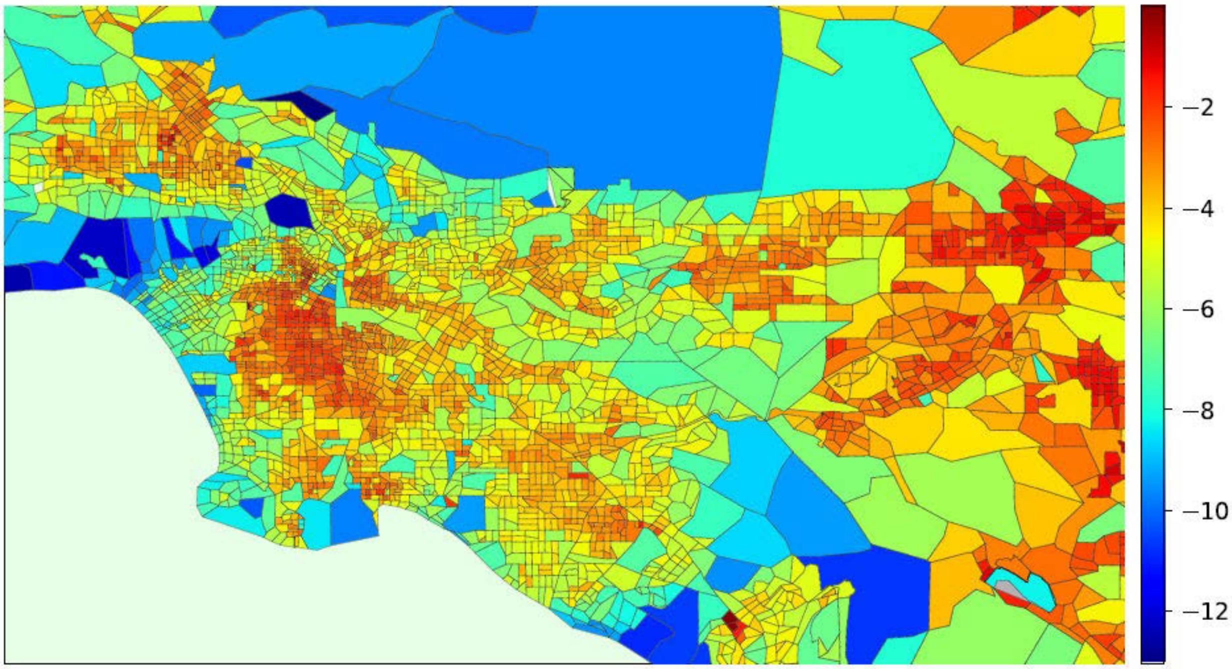
$$\overline{H}_i = \frac{\bar{\phi}((1-\eta)\bar{q}_i)^{\frac{1-\eta}{\eta}} L_i}{\bar{\phi}((1-\eta)\bar{q}_i)^{\frac{1-\eta}{\eta}} L_i / H_i^D - 1}, \tag{23}$$

where H_i^D is the total demand for floorspace in tract i given by

$$H_i^D = H_{Ri}^D + H_{Wi}^D = \frac{\gamma \tilde{w}_i}{q_{Ri}} N_{Ri} + \left(\frac{(1-\alpha)A_i}{q_{Wi}}\right)^{1/\alpha} N_{Wi}.$$

The values of the endogenous cost of construction, $\phi_i(H_i)$, implied by calibrated density limits $\overline{H}_i$ are depicted in Figure 7. They show that some of the most restricted urban areas (low $\phi_i(H_i)$) are Beverly Hills, Malibu, Santa Monica, and South Pasadena, all of which are notorious examples of stringent land use regulation and strong local opposition to development. The least restricted areas (high $\phi_i(H_i)$) are South LA, San Gabriel Valley, and most of the Inland Empire.

Figure 7 Values of Inverse Construction Costs $\phi_i(H_i)$ (logs)



Counterfactual

Some of the most acute problems of large urban areas, including Los Angeles, are high housing costs and long commutes. In 2018, median house price in Los Angeles County was about \$600,000, according to Zillow. In the same year, median household income was \$68,000, according to Census. That is, the price-to-income ratio was nearly 9, compared to the national average of 4.2.[10] Despite the shortage of affordable housing, it is extremely difficult to increase supply as most residential land in Los Angeles area is zoned for single family detached houses.[11] High costs of housing and restrictive zoning push many workers away from major employment centers towards suburbs where housing is cheaper. As a result, an average worker in Los Angeles metro area covers 32 miles and spends about 55 minutes a day commuting to and from work.

We therefore use the quantitative model to study how the economy of the Los Angeles metropolitan area would respond to changes in density restrictions and commuting speeds. To this end, we design two counterfactual scenarios. In the first one, density limits in all urban tracts are relaxed. In the second, commute speeds between all tracts are increased. Both experiments are conducted under the assumption of a closed city. That is, even though average

[10] See https://www.jchs.harvard.edu/blog/price-to-income-ratios-are-nearing-historic-highs/https://www.jchs.harvard.edu/blog/price-to-income-ratios-are-nearing-historic-highs/

[11] See https://www.nytimes.com/interactive/2019/06/18/upshot/cities-across-america-question-single-family-zoning.htmlhttps://www.nytimes.com/interactive/2019/06/18/upshot/cities-across-america-question-single-family-zoning.html

expected utility in the urban area would change relative to the reservation utility, we keep the total employment the same as in the benchmark economy.

Experiment 1: Lift Density Limits

In this experiment, we relax density restrictions by increasing $\overline{H}_i$ in all urban tracts so that the maximum allowed floorspace density, i.e. $\overline{H}_i/L_i$, is the same as in Downtown Los Angeles.[12]

City-wide results of the first counterfactual experiment are illustrated in Table 10. When density restrictions are relaxed to the level of Downtown L.A., output grows by about 35% and welfare increases by 57%. As Table 11 shows, most of the output gains result from larger supply of floorspace, though some gains are brought by better allocation of workers to jobs. The relaxation of density limits result in a massive increase in floorspace supply and a fall in floorspace prices. It also allows workers to settle closer to their jobs, thereby reducing their commute times and distances.

Figure 8 illustrates how the relaxation in density limits changes residential and employment density in the metro area. First, since density limits apply to both residential and commercial development, the areas which experience an increase in workers also experience an increase in residents. The areas most affected by this change in policy are Santa Monica, Beverly Hills, Century City, Torrance, Pasadena and Newport Beach. These areas are characterized by attractive residential and workplace amenities and high productivity. At the same time, with a few exceptions, development in these areas is not particularly dense. Opening up these areas for commercial and residential developers brings many additional workers and residents, and boost productivity and welfare in the city as a whole.

Experiment 2: Increase Commuting Speed

In the second experiment, we increase commute speeds to 65 miles an hour, the maximum speed limit on most California highways, between each pair of tracts.

The results are summarized in Table 10. As a result of increasing speed between all pairs of tracts, workers choose to locate more than 7 km further from their employers relative to the benchmark and are still able to cut their commuting time from 36 to 19 minutes. This fall in commuting time leads to large welfare gains, even though output per worker only grows by 2.2%. Most of the increase in output happens thanks to a better allocation of labor across space, as Table 11 demonstrates. When commuting is inhibited by traffic congestion, many workers choose to work in less productive tracts closer to home. When speeds increase, they are able to take jobs in more productive tracts without having to move residences.

Figure 8 shows how the increase in commuting speeds changes the allocation of workers and resident across the Los Angeles metro area. Santa Monica, Beverly Hills, Century City and, to a

[12] As a benchmark, we select tract 6037207400, the tract with highest employment density in downtown LA. An urban tract is a tract with population density of at least 1,000 inhabitants per square mile, as defined by the Census Bureau. By this definition, 3,249 out of 3,917 tracts in the Los Angeles metro area are urban.

lesser extent, Downtown Los Angeles and Irvine see a large increase in employment. At the same time, these areas lose residents. This happens since, on the one hand, more residents thoughout the metro area can access attractive jobs in these locations and, on the other hand, workers who lived there in the benchmark economy do not have to live there anymore. Instead, they can enjoy cheaper housing elsewhere and commute there for work.

Table 10 Counterfactual Experiments

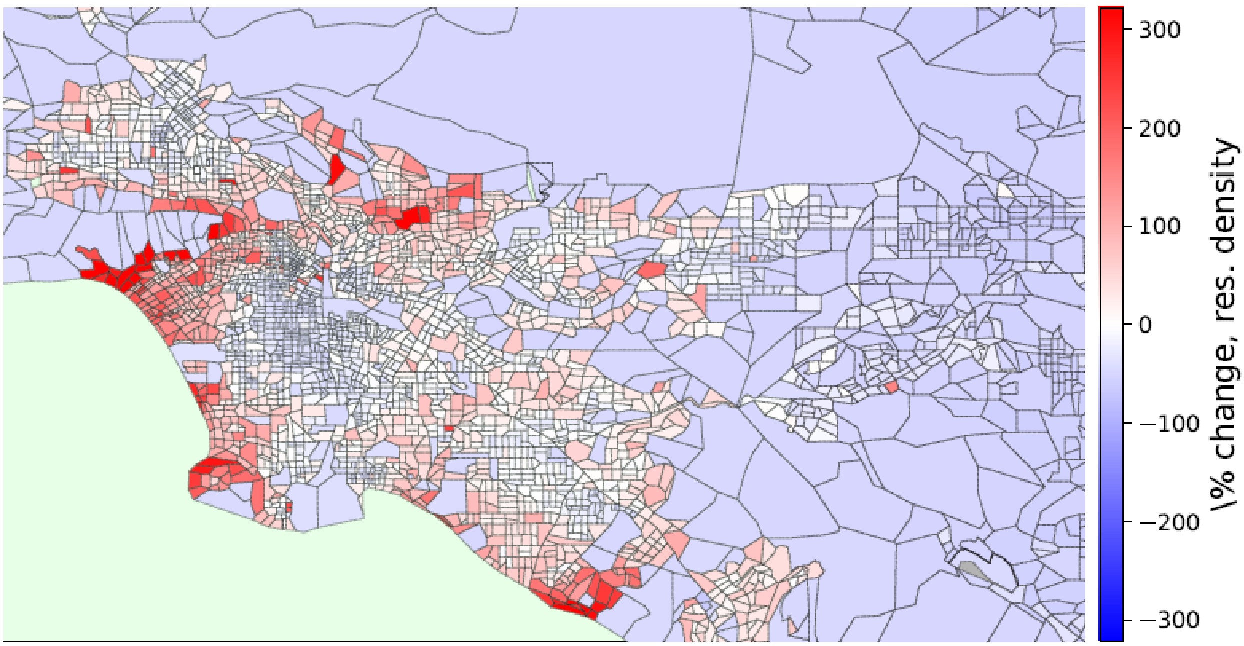
	Benchmark	**Counterfactual 1:**	**Counterfactual 2:**
		Upzoning	**Increase speed**
Output per worker	100.0	135.4	102.2
Employment	100.0	100.0	100.0
Welfare	100.0	156.6	130.1
Mean wages	100.0	135.4	102.2
Mean residential floorspace price	100.0	45.6	101.2
Mean commercial floorspace price	100.0	40.1	109.4
Mean commuting time, min	36.1	35.6	19.4
Mean commuting distance, km	26.7	24.6	33.9
Median tract density	1.34	1.39	1.34
Mean house size	100.0	264.6	101.6

Table 11 Decomposition of output gains

	Counterfactual 1:	**Counterfactual 2:**
	Upzoning	**Increase speed**
Output per worker, % chg	35.4	2.2
effect of larger floorspace supply	23.5	-0.4
effect of better allocation of labor	1.3	1.7

Note: The decomposition is performed as follows. To isolate the effect of floorspace supply, we only adjust floorspace to the counterfactual level, while keeping the distribution of employment as in the benchmark economy. To isolate the effect of labor supply, we only adjust employment to the counterfactual level, while keeping the amount of floorspace as in the benchmark economy. The total effect of a policy change on output is larger than the sum of the floorspace and labor effects due to the interaction between these two effects.

Figure 8 Residential and Employment Density. Counterfactual 1: Increase Density Limits to the Level of DTLA in All Urban Tracts.



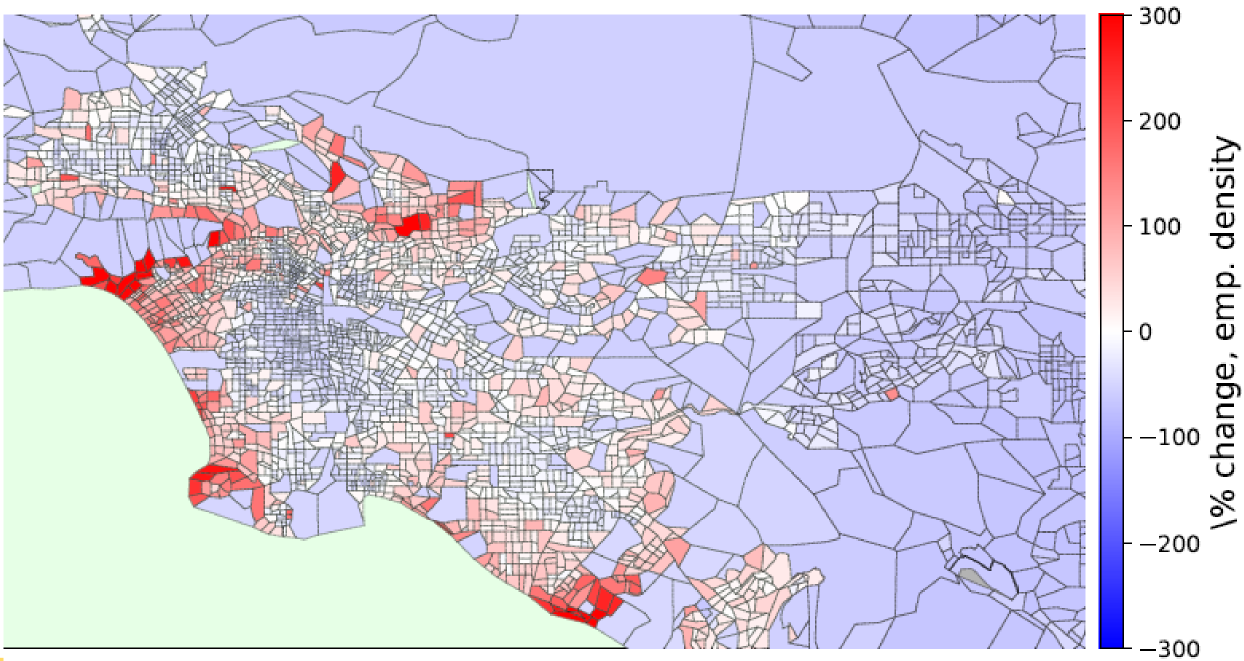
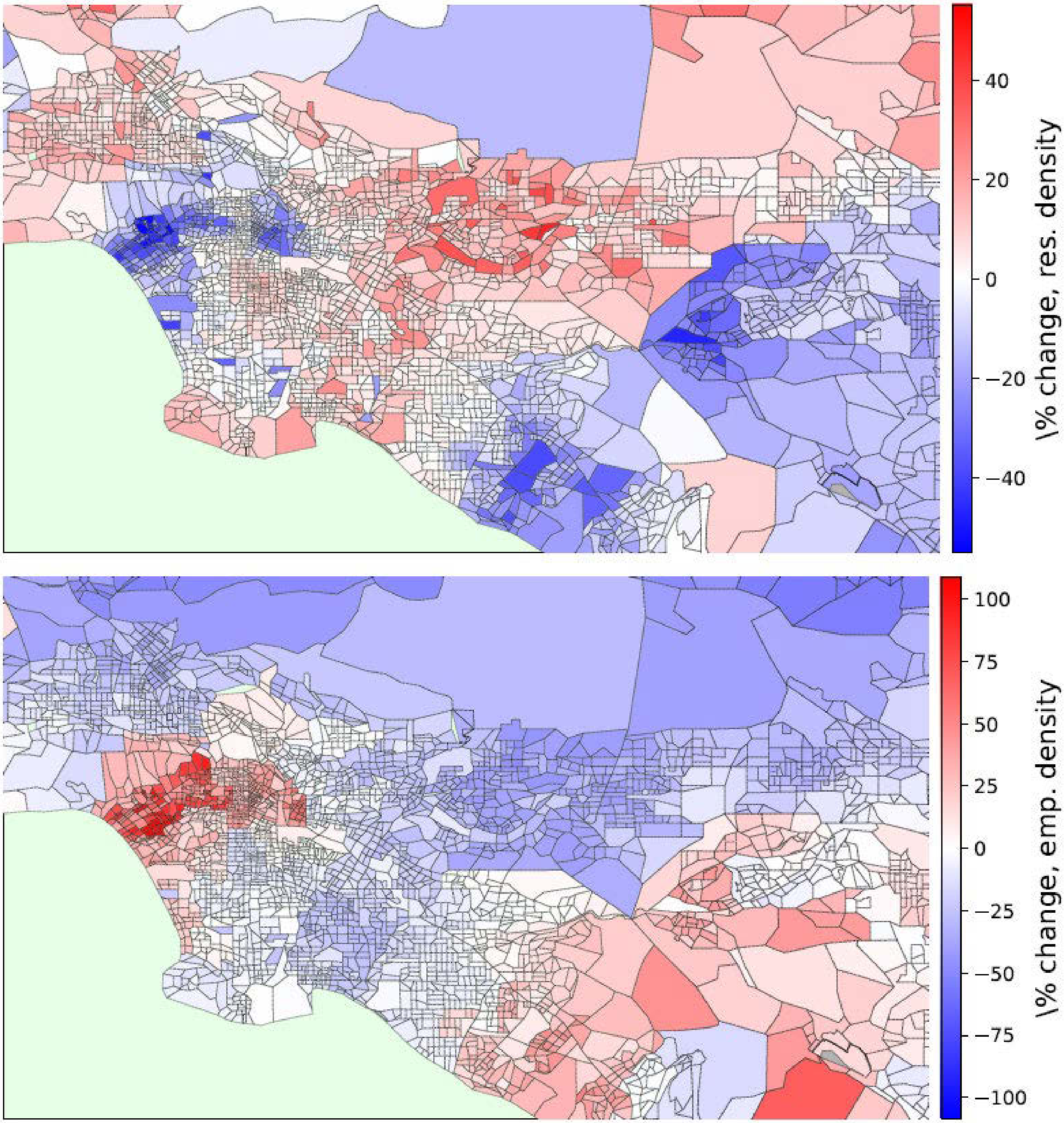


Figure 9 Residential and Employment Density. Counterfactual 2: Increase Commuting Speed to 65 miles per Hour



Conclusion

In this paper we built a quantitative general equilibrium model of residence and employment choices inside an urban area. Using the Los Angeles metropolitan area as a laboratory, we ran two simple but informative counterfactual experiments. It has long been thought that the efficiency gains from freer cross-city sorting under looser density restrictions migth be large. Our results suggest that even within a single city, the gains from rethinking density-restricting zoning may be considerable. While a maximal, almost impossible improvement in transportation infrastructure yields some gains in output and wages, a loosening of zoning restrictions yields even bigger gains, while also slashing real estate prices.

Some caveats counsel caution in interpreting the results–for example, we do not account for endogenous responses of traffic congestion. But as a first pass, these results offer strong support to the idea that the efficiency of urban economies can be significantly improved through smarter zoning.

References

Ahlfeldt, G. M., S. J. Redding, D. M. Sturm, and N. Wolf. The Economics of Density: Evidence from the Berlin Wall. *Econometrica*, 2015. 83(6): 2127-2189.

Albouy, D., and G. Ehrlich. Housing Productivity and the Social Cost of Land-use Restrictions. *Journal of Urban Economics*, 2018. 107: 101-120.

Allen, T., C. Arkolakis, and X. Li. Optimal City Structure. 2016. Working Paper.

Baum-Snow, N., and L. Han. The Microgeography of Housing Supply. 2019. Discussion paper, mimeo.

Brueckner, J. K., and R. Singh. Stringency of Land-Use Regulation: Building Heights in US Cities. 2017. Working Paper.

Calabrese, S., D. Epple, and R. Romano. On the Political Economy of Zoning. *Journal of Public Economics*, 2007. 91(1): 25-49.

Combes, P.-P., G. Duranton, and L. Gobillon. The Costs of Agglomeration: House and Land Prices in French Cities. *The Review of Economic Studies*, 2008. 86(4): 1556-1589.

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

Diamond R, T. McQuade. Who wants affordable housing in their backyard? An equilibrium analysis of low-income property development. Journal of Political Economy. 2019 Jun 1;127(3):1063-117.

Favilukis, J., P. Mabille, and S. Van Nieuwerburgh (2019) Affordable Housing and City Welfare. Working Paper 25906, 2019 National Bureau of Economic Research.

Helsley, R. W., and W. C. Strange. Strategic Growth Controls. *Regional Science and Urban Economics*, 1995. 25(4): 435-460.

Herkenhoff, K. F., L. E. Ohanian, and E. C. Prescott. Tarnishing the golden and empire states: Land-use restrictions and the U.S. economic slowdown. *Journal of Monetary Economics*, 2018. 93, 89: 109, Carnegie-Rochester-NYU Conference on Public Policy held at the Stern School of Business at New York University.

Hilber, C. A., and F. Robert-Nicoud. On the Origins of Land Use Regulations: Theory and Evidence from US Metro Areas. *Journal of Urban Economics*, 2013. 75: 29-43.

Hsieh, C.-T., and E. Moretti. Housing Constraints and Spatial Misallocation. *American Economic Journal: Macroeconomics*, 2019. 11(2): 1-39.

Joshi, K. K., and T. Kono. Optimization of Floor Area Ratio Regulation in a Growing City. *Regional Science and Urban Economics*, 2009. 39(4): 502-511.

Kantor, Y., P. Rietveld, and J. van Ommeren. Towards a General Theory of Mixed Zones: The Role of Congestion. *Journal of Urban Economics,* 2014. 83: 50-58.

Kono, T., and K. K. Joshi. Spatial Externalities and Land Use Regulation: an Integrated Set of Multiple Density Regulations. Journal of Economic Geography, 2018. 18(3): 571-598.

Morrow, G. D. The Homeowner Revolution: Democracy, Land Use and the Los Angeles Slow-Growth Movement, 1965-1992. UCLA, 2013.

Ortalo-Magne, F., and A. Prat. On the Political Economy of Urban Growth: Homeownership versus Affordability. *American Economic Journal: Microeconomics*, 2014. 6(1): 154-81.

Parkhomenko, A. Local Causes and Aggregate Implications of Land Use Regulation. 2019, Working Paper.

Rossi-Hansberg, E. Optimal Urban Land Use and Zoning. *Review of Economic Dynamics*, 2004. 7(1): 69-106.

Saiz, A. The Geographic Determinants of Housing Supply. *The Quarterly Journal of Economics*, 2010. 125(3): 1253-1296.

Sole-Olle, A., and E. Viladecans-Marsal. Lobbying, Political Competition, and Local Land Supply: Recent Evidence from Spain. *Journal of Public Economics*, 2012. 96(1): 10-19.

Spear B. Improving employment data for transportation planning. *American Association of State Highway and Transportation Officials*. 2011 Sep.

American Community Survey, 2012-2016, U.S. Census Bureau. Washington, D.C Accessed 2020.

LEHD Origin-Destination Employment Statistics Data 2002-2017, U.S. Census Bureau. Washington, D.C Accessed 2020.

Longitudinal-Employer Household Dynamics Program, U.S. Census Bureau. Washington, D.C. Accessed 2020.

Census Transportation Planning Products, U.S. Department of Transportation. Accessed 2020.

Zhang, W., and K. M. Kockelman. Optimal Policies in Cities with Congestion and Agglomeration Externalities: Congestion Tolls, Labor Subsidies, and Place-based Strategies. *Journal of Urban Economics*, 2016. 95: 64-86.

Data Management Plan

Products of Research

The existing database to which these elements will be added contains the following elements: (1) employment and residence population in each tract, (2) wages in each tract, (3) prices of residential and commercial floorspace in each tract, (4) commuting flows, times, and distance between each pair of tracts. Some detail on each of these elements follows.

- **Tract-level Employment and Residence Population** We take data on the number of workers residing and working in each census block from the Census Bureau's LEHD Origin-Destination Employment Statistics, aggregating the data to the census tract level and calculating averages over the period from 2012 to 2016.

- **Tract-level Wages** We construct tract-level average wages for 2012-2016 combining tract-level data from the Census CTPP database, with microdata at the PUMA level from the American Community Survey. Further details on how we calculate tract-level wages can be found in Appendix A of the report.

- **Tract-level Prices of Residential and Commercial Floorspace** Our commercial and residential property price is from DataQuick, which is transaction-level public records on property characteristics and transactions data. The dataset covers 2,354,535 properties from 2007-2016 in LA-LB CSA. The data provides information such as sales price, GIS coordinates, transaction dates, property use, transaction type, number of rooms, number of baths, square-footage, lot size, year built, etc. We use hedonic regressions to obtain floor space price indices that reflect the value of a constant-quality unit of commercial or residential floor space in a given location. Further details can be found in Appendix B of the report.

- **Commuting Flows, Times, and Distance between Each Pair of Tracts** The Census CTPP database provides data on workers' reported commute times by tract-of-residence/tract-of-employment pair for the period 2012-2016, which we use to calculate average commute times between tracts. A key advantage of this data is that is derived from responses by actual commuters, and so reflects actual commute times as influenced by all relevant facts including infrastructure, congestion, and parking availability.

Data Format and Content
Tract-level (data_tracts.dta)

AREAKEY	Census Tract
intplat	Census Tract Internal Point Latitude
intptlon	Census Tract Internal Point Longitude

arealand	Area(Land)
HR	Residence Total Number of Jobs
HM	Workplace Total Number of Jobs
w	Mean wage (calculated from CTPP using PUMA bin averages from ACS)
q	Commercial floorspace price, per sqft (PUMA-level)
Q	Residential floorspace price, per sqft (tract-level)

Tract-pair level (data_tractpairs.dta)

h_AREAKEY	Tract of residence
w_AREAKEY	Tract of employment
ts000	Total number of commuters, average 2012-2016
distance	Distance between tract centroids, km
commtime	Estimated commute time, min

Data Access and Sharing

- LEHD Origin-Destination Employment Statistics is publicly available at https://lehd.ces.census.gov/data/
- The Census CTPP database is publicly available at https://ctpp.transportation.org/2012-2016-5-year-ctpp/
- DataQuick is exclusively available with purchase from Corelogic (https://www.corelogic.com/solutions/configurable-real-estate-data-reports.aspx)

Reuse and Redistribution

The data can be reused and redistributed by the general public. Please cite our working paper, Delventhal, Kwon, and Parkhomenko, "Zoning and the Density of Urban Development," if the used data does not come directly from original sources but was produced from original data using our methodology.

Appendix

Appendix A. Wage Index

Our wage estimates by employment tract are based on the database of Census Transportation Planning Products (CTPP). CTPP data sets produce tabulations of American Community Survey (ACS) data. The microdata provides rich information about where people live and work, commuting patterns, socioeconomic, and travel characteristics. To compute the wage index at tract-level, we make use of the five years (2012-2016) estimates at tract-level.

Specifically, we make use of "earnings in the past 12 months (2016$), for the workers 16-year-old and over," which is based on the respondents' workplace locations. The variable provides the estimates of the number of people in each earning bin in each tract. Table 12 provides the sum of the estimates in each bin across counties.

To calculate a representative wage(earning), for an earning bin b , in a tract j , with the associated PUMA p , we calculated a weighted average of earnings ($\widehat{w}_j$) as follows:

$$\widehat{w}_j = \frac{\sum_b estimate_{b,j} * \widehat{mean}_{b,p}}{\sum_b estimate_{b,j}} \qquad (24)$$

$$\widehat{mean}_{b,p} = \frac{\sum_{i \in b,p} earning_{i,b,p}}{\sum_{i \in b,p} 1} \qquad (25)$$

where $estimate_{b,j}$ is estimates for the number of workers in a bin b in tract j , and $\widehat{mean}_{b,p}$ is a calculated mean of earning, within a bin b , of a PUMA p that the tract j is matched.

Finally, we make use of Public-Use Microdata Samples (PUMS), to have a representative value for each earning bin at PUMA-level. $\widehat{mean}_{b,p}$ is calculated by taking an average of reported earnings of individuals within earning bins at each PUMA. Table 13 is summary statistics for the estimated tract-level earnings.

Table 12 Number of Observations in Each Earnings Bins

Income Bin	Los Angeles	Orange	Riverside	San Bernardino	Ventura
$1 to $9,999 or loss	416,469	147,484	86,219	85,854	34,973
$10,000 to $14,999	279,132	90,871	51,959	52,605	21,143
$15,000 to $24,999	541,649	168,284	97,184	97,059	40,458
$25,000 to $34,999	440,298	146,337	79,994	81,911	34,829
$35,000 to $49,999	493,434	170,364	77,170	87,969	37,487
$50,000 to $64,999	387,533	138,932	57,409	62,487	27,979
$65,000 to $74,999	176,079	63,244	24,869	27,687	13,895
$75,000 to $99,999	308,994	114,436	39,159	44,409	23,871
$100,000 or more	486,179	189,108	44,925	43,158	36,346
No earnings	520	134	144	85	55

Table 13 Descriptive Statistics: Average Wage at Tract-level, by County

	Obs	Mean	Std. Dev.	Min	Max
Los Angeles	2,339	51775.9	17793.98	6762.519	134337.2
Orange	582	59388.59	18278.89	20537.82	125397.6
Riverside	453	46754.07	13552.59	12142.64	102704.7
San Bernardino	369	45558	12952.6	12170.31	85635.8
Ventura	173	56111.33	17980.33	18791.67	112409.3

Appendix B. Commercial and Residential Price Index

All the data sets are collected at tract-level. We bring together data from a variety of sources. Our commercial and residential property price is from DataQuick, which is transaction-level public records on property characteristics and transactions data[13]. The dataset covers 2,354,535 properties from 2007-2016 in LA-LB CSA area (i.e., Los Angeles, Orange, San Bernardino, Riverside, and Ventura counties). The data provides information such as sales price,

[13] This data is frequently used in the recent literature (e.g., Diamond and McQuade (2019))

GIS coordinates, transaction dates, property use, transaction type, number of rooms, number of baths, square-footage, lot size, year built, etc.

To get the commercial and residential property index at tract-level, we first define a commercial and residential property, according to the property use variable, as described in Table 14.

Table 15 provides the number of observations in each county and year, for commercial and residential properties. Note that the observations of commercial transactions are far less than residential transactions, as the sales happen less frequently for commercial properties. Also, the quality of the data varies across counties; especially commercial properties in Riverside county has a lower quality of data, as its accessor's data was less accurately collected. Table 16 provides descriptive statistics.

The tract-level residential and commercial property index is estimated with the following hedonic regressions (Baum-Snow and Han (2019)).

For a residential transaction of a property p, in tract j in year-month t,

$$ln(P_{pjt}) = \alpha + \beta X_p + \tau_t + \eta_j + \epsilon_{pjt} \qquad (26)$$

where P_{pjt} is price per square-footage, X_p is property characteristics including property use, transaction type, number of rooms, number of baths, lot size, and year built, τ_t is year-month fixed effect, and η_j is tract-fixed effect, which is used as residential property index of tract j.

To overcome the issue of the lack of commercial transactions, the commercial property index is estimated at Public Use Microdata Area (PUMA)-level[14]. Table 16 shows the number of tracts and PUMAs in each county.

For commercial property transaction of a property p, in tract j of PUMA g in year-month t, the commercial hedonic regression is:

$$ln(P_{pgjt}) = \alpha + \beta X_p + \tau_t + \zeta_g + \upsilon_{pgjt} \qquad (27)$$

where P_{pgjt} is price per square-footage, X_p is property characteristics including property use, τ_t is year-month fixed effect, and ζ_g is PUMA-fixed effect, which is used as a commercial property price index at PUMA-level.

[14] PUMA is a geographic unit used by the US Census for providing statistical and demographic information. Each PUMA contains at least 100,000 people.

Table 14 Commercial and Residential Categorization

	Residential	Commercial
Property Use	Condominium, PUD; Cooperative; Duplex; Miscellaneous Residential; Mobile /Manufactured Home; Multi-Family Dwelling; Quadraplex; Single Family Residence; Timeshare; Triplex; Residential;	Auto sales, services; Casino; Department Store; Financial Building; Food Store, Market; Hospitals, Convalescent; Hotel/ Motel; Laundry, Dry Cleaning; Medical Buildings; Miscellaneous Commercial; Nursery; Office Building; Parking Lot, Parking; Restaurant, Bar, Food; Service Station, Gas Station; Shopping Center; Store / Office Combo; Stores, Retail Outlet; Food Processing; Heavy Industrial; Light Industrial; Lumber, Building Materials; Miscellaneous Industrial; Warehouse, Storage; Winery; Bowling Alley; Clubs, Fraternal Organizations; Communications; Roadways; Theaters; Transportation, Air, Rail, Bus, Commercial, Industrial,

Table 15 Number of Transactions over Counties and Years

Panel A. Residential Properties						
Sale Year	**Los Angeles**	**Orange**	**Riverside**	**San Bernardino**	**Ventura**	**Total**
2007	85,085	27,708	46,368	29,348	9,481	197,990
2008	96,554	34,255	73,810	45,526	11,908	262,053
2009	107,969	34,134	77,973	49,734	11,350	281,160
2010	98,396	34,966	68,211	42,238	10,692	254,503
2011	100,013	32,798	57,955	40,694	10,540	242,000
2012	89,348	35,624	53,943	36,039	10,999	225,953
2013	91,292	34,180	48,080	31,960	10,071	215,583
2014	78,055	30,126	41,833	28,112	9,256	187,382
2015	84,798	33,223	44,126	29,810	10,729	202,686
2016	78,444	33,675	44,905	29,712	10,492	197,228
Panel B. Commercial Properties						
Sale Year	**Los Angeles**	**Orange**	**Riverside**	**San Bernardino**	**Ventura**	**Total**
2007	5,956	1,255	1,182	910	293	9,596
2008	4,141	790	1,282	770	311	7,294
2009	3,244	695	1,002	867	204	6,012
2010	3,784	1,151	1,414	1,021	283	7,653
2011	4,205	1,237	1,296	1,130	301	8,169
2012	4,581	1,470	1,480	1,206	350	9,087
2013	5,352	1,371	1,518	1,224	353	9,818
2014	5,417	905	1,545	1,266	404	9,537
2015	5,657	1,651	1,779	1,388	450	10,925
2016	5,071	1,559	1,547	1,317	412	9,906

Table 16 Number of Public Use Microdata Area (PUMA)s and Tracts

County	Number of PUMAs	Number of Tracts
Los Angeles	69	2334
Orange	18	582
Riverside	15	453
San Bernardino	15	368
Ventura	6	173
Total	123	3910

Table 17 Descriptive Statistics

Panel A. Residential Properties				
County	sqft (mean)	sqft (median)	sales price (mean)	sales price (median)
Los Angeles	1752.25	1499	774734.19	389000
Orange	1969.92	1578	714043.38	495000
Riverside	2046.06	1855	489885.35	246649
San Bernardino	1759.41	1584	345662.41	200000
Ventura	1860.88	1626	569042.40	410000
Panel B. Commercial Properties				
County	sqft (mean)	sqft (median)	sales price (mean)	sales price (median)
Los Angeles	20687.28	5203	5661399.99	1300000
Orange	16447.48	5329	3879699.73	1260000
Riverside	1329.38	1201	1813988.76	590000
San Bernardino	19486.08	3541	2472923.09	522000
Ventura	12087.09	4565	3513023.97	982500

Appendix C. Commuting Time Data

The CTPP database provides commuting time data for 270,436 origin-destination tract- pairs in the Los Angeles-Long Beach Combined Statistical Area for 2012-2016. There are 15,342,889 possible trajectories, and the LODES data for 2012-2016 reports positive commuting flows for 5,647,791 of them. We follow the practice recommended by Spear (2011) and use LODES data as a measure of commuting flows and CTPP data to provide information on commute times.

The CTPP data places commuting times into 10 bins: less than 5 minutes, 5 to 14 minutes, 15 to 19 minutes, 20 to 29 minutes, 30 to 44 minutes, 45 to 59 minutes, 60 to 74 minutes, 75 to 89 minutes, 90 or more minutes, and work from home. In order to get as accurate commute times as possible for the set of primitive connections of the network, we drop all home-workers, who are irrelevant for transit times. We drop workers in the top time bin, because this bin has no upper bound and so the mean may vary substantially across trajectories. We assign mean commute times to all the remaining bins as the mid-points between the bin bounds. We then drop all observations which report an average commuting speed that is either less than 8 kilometers per hour, a brisk walking pace, or more than 70 miles per hour (112.7 kilometers per hour), the standard rural freeway speed limit in the United States. Finally, we calculate tract-pair mean commuting times as the average of the mean commuting times in each bin weighted by the share of commuters on that tract-pair reporting times in each bin.

The previous cleaning steps eliminate observations for 36,279 trajectories, and we are left with commuting time data for 234,157 origin-destination pairs. We then find that there are 211,521 paths for which a commuting time estimate exists for the outbound route but not the reverse. We impute commute times for these missing return journeys, assuming that they can be completed in the same time as the outbound trajectories. This set of connections is then almost enough to connect all tracts that are still detached from the rest of the network. In order to remedy this, we create a connection at the mean travel speed of 31.3 kilometers per hour between these left-out tracts and any tracts within a radius twice as large as the hypothetical radius of tract if its land area formed a circle.15 The final directed network contains 447,277 directed paths. We use Dijkstra's algorithm to calculate the fastest path through this network for each origin-destination pair.

We assume that these calculated times represent the time require to travel from tract centroid to tract centroid. We then add time to each trajectory to represent the time need to travel from place of residence within tract to residence tract centroid, and from workplace tract centroid to workplace within the tract. Naturally, these times are proportional to tract land area-larger tracts should on average require more internal travel time. Specifically, we assume that the distance traveled on each end of the trip is equal to the hypothetical average straight-line distance from any point in the tract to the tract centroid, if the tract were a circle. We then assume that each of these distances is traveled at twice the overall average commuting speed in the cleaned data of 31.3 kilometers per hour. For the vast majority of tracts this adds a negligible amount to commuting time-two minutes or less. For a handful of very large tracts it adds considerable travel time-up to half an hour. We think that this is reasonable given the time that is required to travel within these much larger tracts. These origin-destination distribution effects are also applied to self-commute times, so that a worker that lives and works in the same tract will still have to spend some time travelling to their workplace- more time for larger tracts.

Figures 10, 11 and 12 show the results of these calculations for commuting times in and out of downtown Los Angeles, Santa Monica, and the town of Claremont in Los Angeles County.

Figure 10 Commuting Times in and out, Downtown L.A.

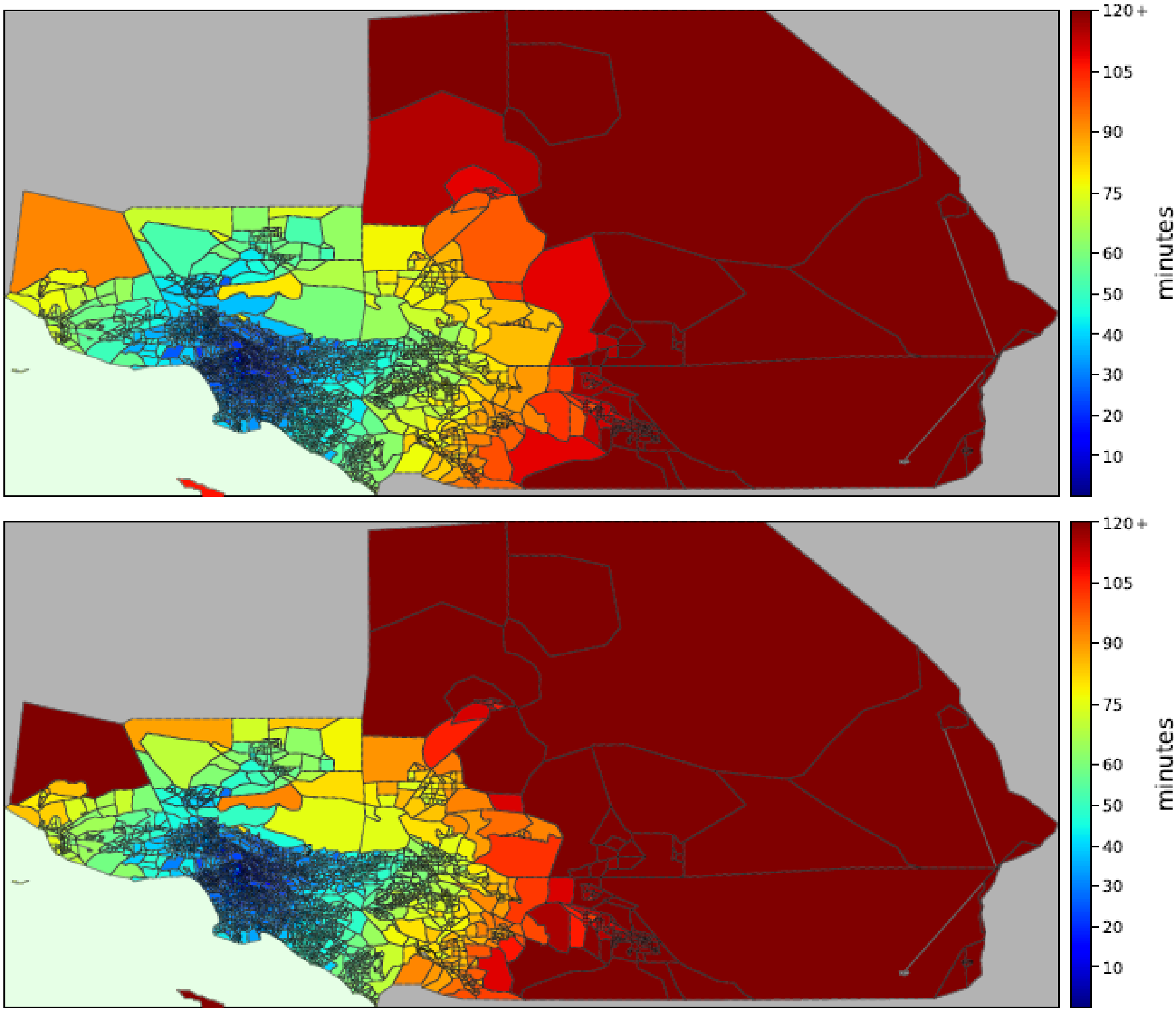


Figure 11 Commuting Times in and out, Santa Monica

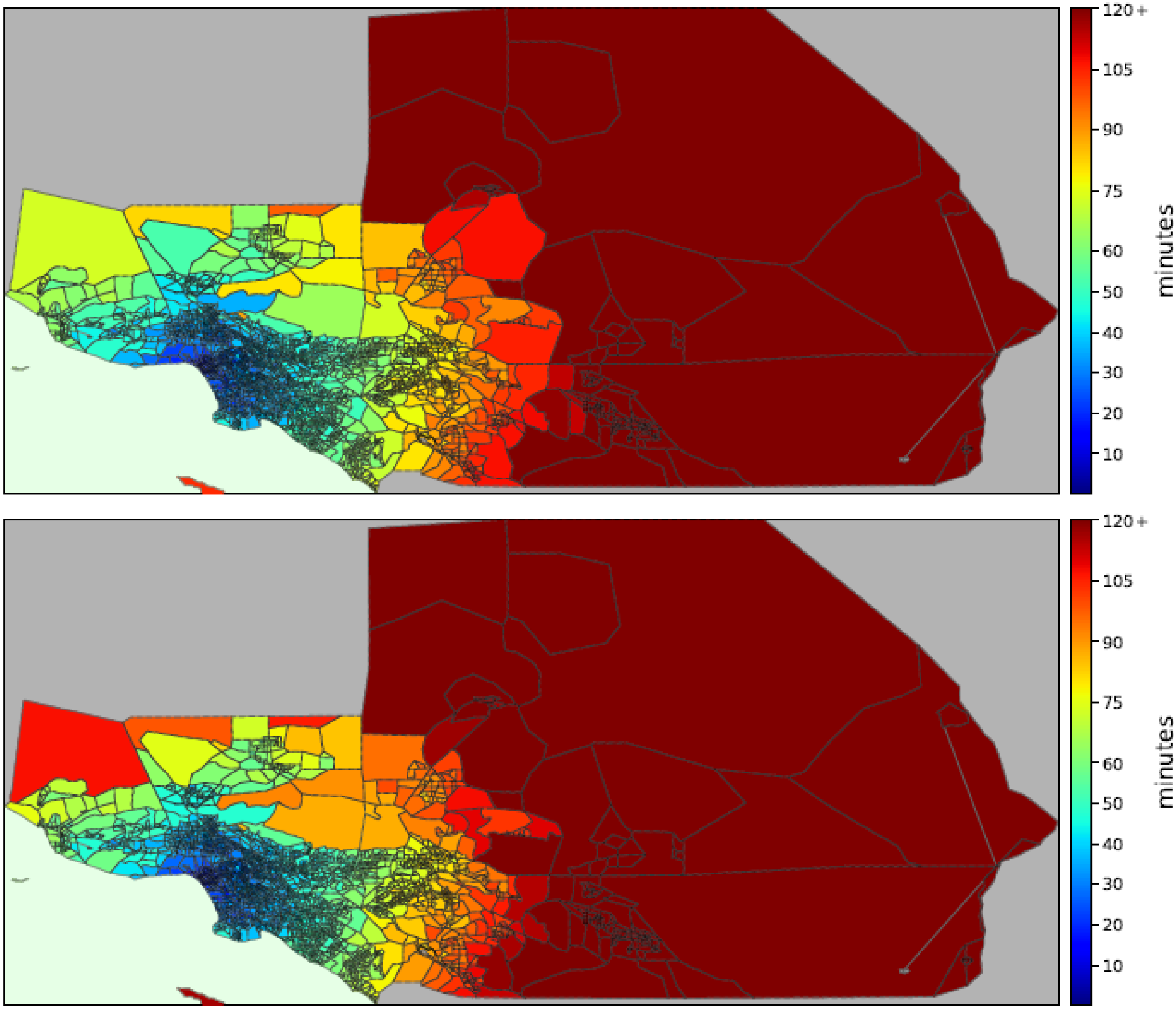
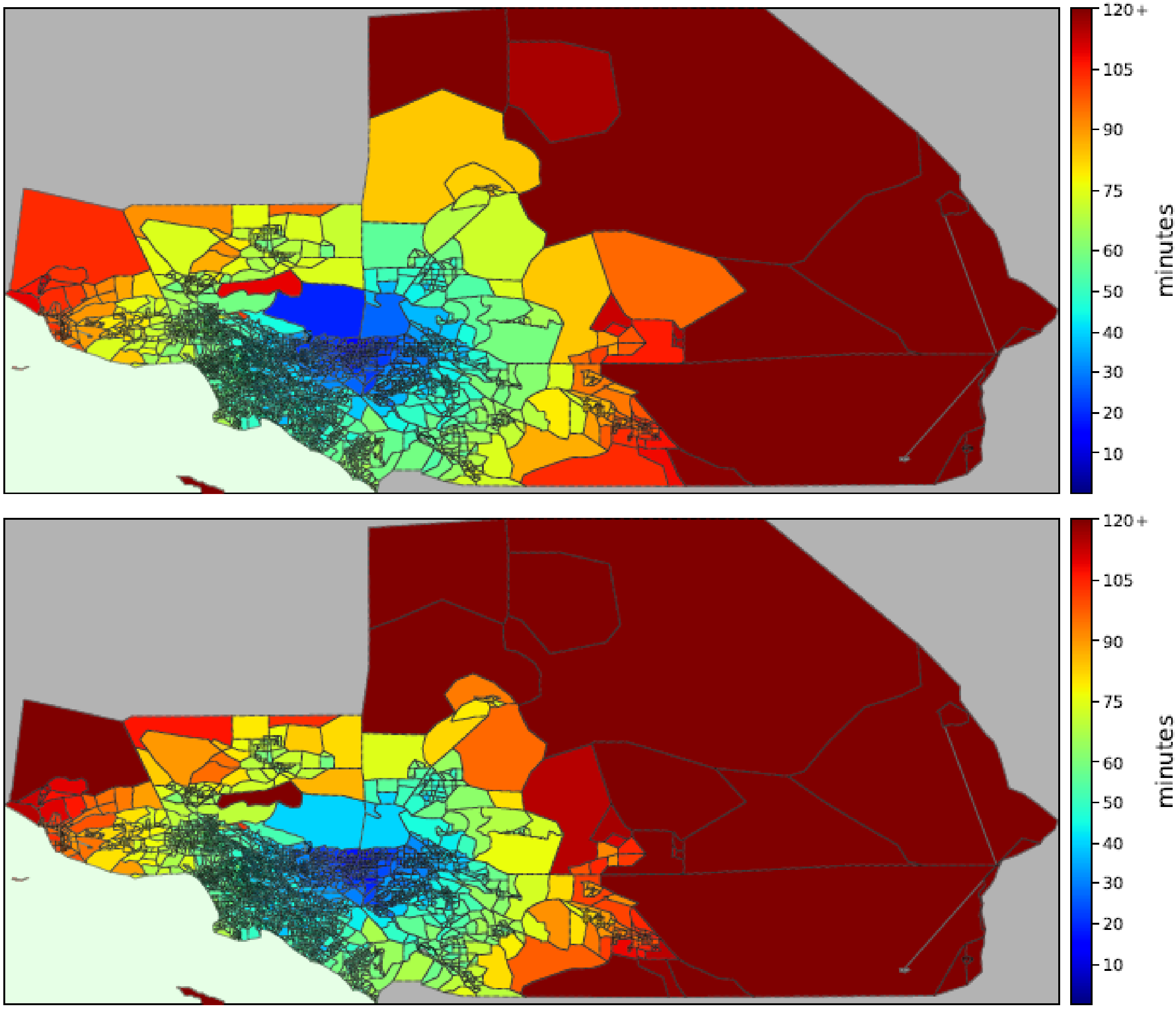


Figure 12 Commuting times in and out, Claremont



Pacific Southwest Region UTC
University Transportation Center