Two-Sided Sorting and Spatial Inequality in Cities

Matthias Hoelzlein*

University of California, Berkeley

January 10, 2020

Job Market Paper

Click here for most recent version

Abstract

This paper studies a new economic force underlying the spatial sorting of rich and poor households in cities. On the demand side, households with different incomes choose neighborhoods and differ in their expenditures across various local services. On the supply side, service establishments sort into neighborhoods while taking into account proximity to their consumers. This two-sided sorting leads to endogenous differences in the local price index that amplify the concentration of household groups. A recent literature in urban economics has rationalized spatial sorting of households that is left unexplained by local incomes or housing costs by modeling pure amenity spillovers. In this paper, I quantify the contribution of endogenous price indices to spatial sorting that is usually projected onto such reduced-form spillovers, and study the implications of two-sided sorting for urban policy. To do so, I develop a quantitative equilibrium model of the city that features two-sided sorting and nests many urban models. I estimate the key parameters of the model using detailed microdata for Los Angeles from 1990-2014. I find that spatial variation in local price indices decreases the estimates of reduced-form spillovers by about 30-50 percent. To shed light on the policy implications, I simulate policy counterfactuals, and compare the effects to the existing framework with only reduced-form amenity spillovers. By studying a number of prominent place-based policies in Los Angeles, I find substantially different effects on neighborhood composition and welfare between both models.

*Email: hoelzlein@berkeley.edu. I am indebted to my advisors Benjamin Faber, Cecile Gaubert, Andres Rodriguez-Clare and Victor Couture for their continuing guidance. I would like to thank Thibault Fally, Nick Tsivanides and the many seminar participants at the UC Berkeley Economics Department and Haas School of Business for many helpful discussions. I am grateful that I had such great colleagues, in particular, Marc Dordal i Carreras and Nicholas Sander. I would also like to thank the Fisher Center for Real Estate and Urban Economics at UC Berkeley for financial support. Researcher(s) own analyses calculated (or derived) based in part on data from The Nielsen Company (US), LLC and marketing databases provided through the Nielsen Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the Nielsen data are those of the researcher(s) and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein. All errors are mine.

1 Introduction

Spatial inequality in cities and the clustering of rich and poor households across neighborhoods have attracted widespread attention by policy makers and given rise to public debate. These sorting patterns are usually rationalized by initial differences in access to employment, housing costs or natural amenities. Over time, differences in the social composition of neighborhoods are then reinforced by endogenous changes in the urban landscape that reflect the preferences and resources of each population. For instance, richer neighborhoods provide residents with more and different consumption options, greater safety, or access to better schools.

In the recent literature, these effects are typically summarized and modeled as reduced-form amenity spillovers[1]; utility of different groups is assumed to depend directly on the composition of residents, a reduced-form way to capture many channels through which neighborhood composition feeds back into utility of its residents. Since such amenity spillovers abstract from household or firm behavior, researchers and policy makers cannot directly observe them in the data. However, many urban policies target the behavior of households and firms to affect inequality or efficiency outcomes of a city. Hence, studying the microfoundations underlying the spatial sorting of household groups is crucial for the qualitative and quantitative implications of urban policies.

In this paper, I set out to open up the black box of amenity spillovers. On the supply side, business establishments in retail and services decide where to locate while taking into account proximity to their consumers. On the demand side, households in different skill (or income) groups vary in their expenditure shares across these local services due to non-homothetic preferences. These basic forces give rise to two-sided sorting in cities whereby the spatial sorting of high and low skilled households[2] is a function of endogenous differences in access to consumption services that enter the local price index.

In this setting, the paper aims to answer two main research questions. First, to what extent does accounting for differences in the local price index across neighborhoods affect existing estimates of reduced-form amenity spillovers within and across household groups? Second, what are the implications of allowing for two-sided sorting for widely used urban policies that address spatial inequality in cities? In answering these questions, the paper makes three main contributions. First, I propose a quantitative general equilibrium model of the city that introduces two-sided sorting of skill groups and firms in various local service sectors but nests other forces present in workhorse urban models. Second, I combine my theory with rich microdata from Los Angeles in order to quantify the fraction of spatial clustering that is due to endogenous price indices arising from two-sided sorting. Third, I simulate counterfactuals in the estimated model to study the implications of two-sided sorting for a number of prominent urban policies.

[1] See, for example, Diamond (2016), Su (2018b), Tsivanidis (2018), Guerrieri *et al.* (2013), Brueckner *et al.* (1999), Fajgelbaum & Gaubert (2019)

[2] For the remainder of the paper I refer to "high skilled" as someone with at least a bachelor's degree and everyone else I classify as "low skilled". When I use the term "high skilled household" I refer to a household with a high skilled head.

The analysis proceeds in four steps. In the first step, I document a number of motivating stylized facts. First, expenditure shares across retail and service sectors vary considerably with household income or skill. For example, high-income (or high skilled) households spend a much larger fraction of income on recreation providers such as gyms, education services, and restaurants than households with lower income. Second, the spatial distribution of establishments by sector is systematically correlated with the local skill or income composition. Some local service sectors concentrate in rich neighborhoods; whereas others are more equally spread out. Third, income elasticities of demand predict which sectors collocate with skill groups. Establishments in income-elastic sectors, such as gyms, are more likely to be found in neighborhoods with many high-skilled residents.

In the second step, I develop a quantitative spatial model that captures these moments in the data by incorporating three main ingredients. First, households and firms in various sectors simultaneously choose where to locate. Second, demand by skill groups and profits by sectors are linked through non-homothetic preferences. Third, spatial frictions limit consumers' access to firms and vice versa.

In the model, high skilled and low skilled households are identical except for incomes. They choose where to live based on residential rents, the local price index of consumption amentities and non-pecuniary amenities. Households have non-homothetic CES preferences with sector-specific income elasticities of demand. Following Hanoch (1975), this non-homothetic demand system has recently been applied by Comin *et al.* (2018), Borusyak & Jaravel (2018) and Matsuyama (2019); however, my paper is the first to leverage its convenient properties in the context of spatial sorting. The non-homotheticity at the sector level implies that high skilled households spend relatively more on income-elastic goods. Hence, profits in income-elastic sectors rise disproportionally in neighborhoods with many high skilled residents. More firms in income-elastic sectors locate in rich neighborhoods. Since high skilled households value firms in such sectors more than those who are low skilled, everything else being equal, the price index of the high skilled is lower in rich neighborhoods than in poor ones.[3] In equilibrium, high skilled and low skilled households face different price indices in the same neighborhood and this difference is a function of the local skill composition. Two-sided sorting of firms and households leads to skill-location-specific pecuniary externalities that result in segregated neighborhoods. Since my model allows for reduced-form amenity spillovers, as in Diamond (2016), Su (2018b) and Tsivanidis (2018), I can trace out the qualitative and quantitative contributions of pecuniary externalities generated by two-sided sorting and reduced-form spillovers in explaining observed spatial inequality in Los Angeles.

I begin this task by characterizing how these pecuniary externalities operate between different skill groups and neighborhoods through the lens of the model. The key to understanding

[3] In more formal terms, expenditure shares are log-supermodular in income and the sector income elasticity of demand. Therefore, firm profits are log-supermodular in the share of high skilled residents and the income elasticity. Both taken together, implies that the price index of goods consumption is log-supermodular in income and the share of high-skilled residents.

how different populations affect each other through price indices lies in the covariance of their expenditure shares across goods sectors. First, the impact of the externality is stronger within than across skill groups because expenditure shares are more correlated within groups. Second, the impact is stronger for any two neighborhoods that are geographically located close to one another, since the residents in both locations buy goods from similar shopping destinations. Furthermore, I show that the strength of these pecuniary externalities across skill groups varies with the initial income inequality, precisely because expenditure shares are not constant when preferences are non-homothetic. Hence, this dependence on initial conditions suggests that the pecuniary externality generated by two-sided sorting cannot be captured with constant exogenous amenity spillovers.

In the next step of the paper, I apply the model to detailed microdata from Los Angeles. To capture the spatial distribution of households by skill, I combine tract-level data from the National Historical Geographic Information System (NHGIS) and household-level microdata from IPUMS covering the years 1990-2014. I use this dataset to estimate key elasticities on the household side of the model. On the firm side, the National Establishment Time-Series Database (NETS) provides detailed, geo-coded information on the near-universe of establishments that allows me to estimate the spatial supply elasticity of firms from exogenous shocks to firm density. To discipline the strength of non-homotheticities in the model, I estimate income elasticities of demand for 28 local service and retail sectors[4] with household-level expenditure data from the Consumer Expenditure Survey (CEX) and Nielsen Consumer Panel. Taken together, the estimated elasticities and spatial distributions of households and firms characterize the key ingredients of the model.

In the main empirical part of this paper, I assess the extent to which allowing for two-sided sorting affects previous estimates of within and cross-group amenity spillovers in accounting for spatial clustering in the city. To this end, I estimate two different models. First, I relate population changes by skill to a tract's exposure to changes in the surrounding skill-mix, but assume that price indices of goods consumption do not vary by skill group and location. A similar relationship has been used to infer the strength of within and cross-group spillovers in Diamond (2016) and Su (2018b). Using this model, I find large reduced-form spillover elasticities similar to previous estimates. Second, I estimate the same relationship but this time accounting for variation in price indices of goods. This allows me to jointly recover the key supply elasticity on the household side of the model and a set of unbiased reduced-form spillover elasticities. The comparison of the two sets of spillover elasticities shows that accounting for variation in price indices reduces the relative importance of spillovers for low and high skilled households by 30-50%. This result indicates that two-sided sorting is a quantitatively important driver of spatial sorting by skill groups.

[4] I categorize sectors as local if consumers physically go to an establishment to purchase a good or service. In my sector definitions, I try to account for quality differences as much as possible given the constraint that sectors need to match to industry codes in the establishment microdata and expenditure categories in the expenditure data. For example, I differentiate fast food restaurants and full service restaurants or department and dollar stores.

My estimation also makes a methodological contribution. Since data on household expenditures and establishment-level prices is not available for most services at the level of skill groups and tracts, I cannot directly construct price indices at this level of disaggregation. Instead, I rely on the demand structure of the model to overcome this issue. Conditional on observed changes in income and residential rents, variation in the expenditure share on goods that I observe by skill groups at the tract-level, provides a sufficient statistic for changes in the price index of goods.

For identification, I use plausibly exogenous variation in changes in access to service establishments across census tracts, since changes in population and price indices may be correlated with unobserved shocks to the attractiveness of a location. In particular, I construct a shift-share instrument for changes in the local retail environment by interacting the initial sector shares of establishments in a location with sector growth rates in the total citywide number of establishments from other large urban centers in California. Pre-existing local sector shares are able to capture that locations provide different sector-specific supply-side advantages such as access to distribution networks, worker pools, or natural characteristics. These initial differences lead to variation in the exposure of locations to overall differences in plausibly exogenous growth across sectors. With this instrument, I exploit exogenous shifts in the availability of local consumption varieties leading to changes in the relative price index of goods that inform changes in real income of a neighborhood. To estimate reduced-form spillover elasticities in the same regression, I require a second source of exogenous variation. I construct a relative shift-share instrument that uses the same sector growth rates, but I interact the initial shares of establishments with the difference in citywide expenditure shares between high and low skilled households for each sector. Relative sector-level expenditure shares inform how growth across sectors differently enters into price indices of high and skilled households surrounding a tract.

After recovering the resident supply (mobility) elasticity, I proceed by estimating the spatial supply elasticity of firms, which is identified from the sector-specific relationship of individual establishment profits and the number of establishments in a location. The estimation poses two challenges to identification. First, local profits and the number of establishments are correlated with unobserved sector-location-specific productivity. Therefore, I exploit the differential exposure of sectors to plausibly exogenous variation in local demand that is driven by households' preference for the steepness of a location, a natural amenity highly valued by households. Second, I address selection bias due to sorting of firms on idiosyncratic productivity differences across locations by comparing establishments that belong to the same multi-establishment firm.

In the final step of this paper, I assess the implications of two-sided sorting and relative price indices for our understanding of urban policies. To this end, I simulate two place-based policies in Los Angeles, a new place-based tax incentive to firms and social housing. In both counterfactual exercises, I compare outcomes from two versions of the model. In the baseline model, skill groups sort on relative price indices and my unbiased estimate of reduced-form spillovers. In

the model without price index effects, sorting is a result of only reduced-form amenity spillovers (which are biased upwards in the calibration when omitting the price index channel from the model).[5]

In my first counterfactual, I shock the firm distribution by simulating a new tax incentive to invest in economically disadvantaged areas, so-called Opportunity Zones (OZ).[6] I implement this policy by subsidizing profits of firms in the 257 OZs in Los Angeles and assume that the city's government finances the subsidy with lump-sum taxes on households. Firms respond strongly to the subsidy by moving operations into these zones. In the baseline model, the increased supply of consumption varieties induces households to locate in or close to OZs; however, high skilled households respond more to the now lower price index of consumption. As a result, the policy leads to gentrification of these initially disadvantaged areas. In the model without price index effects, the policy does not trigger any sizable mobility response of households since the location of firms has no bearing on the price index of consumption. Although the policy leads to modest average welfare losses for both skill groups of around .1-.2% of consumption, I find that welfare losses are smallest in the baseline model. With price index effects, the policy benefits local OZ residents, a population with high marginal utility and lack of access to firms in the initial equilibrium.

In the second exercise, I assess the effects of social housing on the spatial distribution of households and firms in LA. Using a newly collected dataset on address-level rent savings from social housing, I assume that the benefits of Social Housing accrue to low skilled households in the form of a rent subsidy financed by the city government. In both model versions, social housing leads to an inflow of low skilled households due to the subsidy and a corresponding decrease in skilled residents because of higher market rents. In the baseline model with price index effects, firms leave neighborhoods that have a large presence of social housing but more so in income-elastic sectors amplifying the clustering of the low skilled in treated neighborhoods due to an increase in relative price indices. In the model without price index effects, firms in all sectors do not respond thereby muting the effect of the policy on households.

The rest of the paper is organized as follows. Section 2 discusses the paper's contribution to the existing literature. Section 3 describes the data. Section 4 presents stylized evidence on the joint location of firms and skill groups. Section 5 introduces the model. Section 6 builds intuition for the model. Section 7 takes the model to the data. Section 8 presents policy counterfactuals. Section 9 concludes.

[5] I can turn off price effects without changing the preference structure of the model by removing spatial frictions. Without such frictions the location choices of households and firms are unrelated.

[6] Opportunity Zones were implemented as part of the 2017 Tax Cuts and Jobs Act. The policy offers generous tax benefits to investors if they invest capital gains from previous investments in businesses located in roughly 8,700 designated tracts. According to U.S. Treasury Secretary Steven Mnuchin the total investment in Opportunity Zones will exceed $100B in 2019.

2 Related Literature

In addition to the work discussed above, this paper relates to several strands of the literature. First, there is a growth in literature that studies how heterogeneous preferences for consumption amenities and differential access to services, such as restaurants and retail, lead to sorting of households within cities.[7] In contemporaneous work, Couture *et al.* (2019) model competitive neighborhood developers who choose the local supply of a representative service sector and non-homothetic demand for housing and services. Their framework generates endogenous differences in access to services that induce sorting of households in different income groups. My approach nests their mechanism, but I extend it by modeling how non-homothetic preferences across many types of services reinforce sorting patterns. This additional layer of heterogeneity allows me to jointly capture the spatial distributions of heterogeneous firms and households in the data, and to study how the response of specific types of firms amplifies the effects of urban policies. Hence, my paper adds a new dimension to recent work on place-based policies in cities, for example Busso *et al.* (2013), Diamond & McQuade (2019), Diamond *et al.* (2018), and Davis *et al.* (2018).

Second, extensive literature documents large spatial differences in the availability and variety of goods and services associated with the size and social composition of a local population.[8] In particular, Handbury (2013) shows that when accounting for non-homothetic preferences across food items, income-specific price indices across cities are systemically correlated with local income. Specifically, poor households face higher food price indices in rich relative to poor cities and vice versa for rich households. My paper makes the analogous argument for the relative price indices of local services across neighborhoods. To formally account for these findings, I provide a general equilibrium framework that features a market for consumption amenities where heterogeneous service firms cater to local residents with different incomes and non-homothetic preferences across services.

Third, my paper contributes to a smaller literature that studies the spatial sorting of heterogeneous firms. Motivated by the uneven distribution of productivity across space, work in this area aims to separate local agglomeration externalities in production from the sorting of firms that are ex-ante heterogeneous in productivity. For example, Behrens *et al.* (2014) and Gaubert (2018) find that firm sorting explains a sizable share of the productivity premium of large cities. Brinkman *et al.* (2015) and Ziv (n.d.) study how agglomeration forces and firm sorting interact within cities. Different from these contributions, my paper focuses on demand-side complementarities between local resident composition and the determinants of firm demand, such as income elasticities. Hence, I add to this literature by evaluating how firm sorting contributes to the uneven distribution of household groups within cities.

[7] For example, Couture & Handbury (2017) and Baum-Snow & Hartley (2016) document that changing tastes for services over the last couple of decades are important drivers of the observed movement of college graduates into downtown neighborhoods.

[8] Waldfogel (2008), Schiff (2014), Couture (2016), and Davis *et al.* (2019) study variety and density of restaurants. Glaeser *et al.* (2018) look at several categories of local services.

Lastly, my model builds on the quantitative spatial economics literature that studies the rich structure of cities (Ahlfeldt *et al.* (2015); Allen *et al.* (2015)). The focus of this literature is primarily on the trade-off between job location and residence. Moreover, it features homogeneous households with homothetic preferences.[9] I complement these papers by modeling spatial linkages within the city that are driven by consumption patterns of heterogeneous households with common non-homothetic preferences and the endogenous location choices of firms.

3 Data

In this section, I provide an overview of data sets I use to characterize the Los Angeles Metropolitan Area and to estimate the model.[10] In Appendix A, I provide further information on data sources, summary tables, and details on imputation steps.

Throughout my analysis, I focus on outcomes for high-skilled and low-skilled households. A high-skilled household is defined as having a household head with at least a bachelor's degree. In 2014, Los Angeles consisted of approximately 1.1 million high-skilled and 2 million low-skilled households.

I use the 2010 Census tracts as the geographic definition of neighborhoods. The urban part of Los Angeles county, which is the basis of my analysis and what I refer to as Los Angeles from this point forward, consists of 2235 tracts with a total population just under 10M in 2014. The National Historical Geographic Information System (NHGIS) provides data on tracts for the Census 1990, 2000 and American Community Survey (ACS) 2012-2016[11]. All census tract data are interpolated to constant 2010 census tract boundaries using the Longitudinal Tract Data Base (LTDB). The primary information I extract from NHGIS are income distributions and distributions of expenditure shares on housing by income at the tract-level. The US Census and the ACS specifically provide household counts within defined income bins and household counts within income-expenditure share of housing, rent, and owner-cost bins. For each year, I combine this tract-level information with sample microdata from IPUMS at the level of Public Use Microdata Areas (Puma) to impute counts of households by skill, household income by skill and expenditure shares on housing/rent/owner cost by skill for each census tract and year. Since IPUMS microdata reports only pre-tax income of households, I compute the tax liability for each household using NBER's TAXSIM software and adjust tract income and housing expenditure share by group accordingly.

To capture the location and size of firms, I use the National Establishment Time-Series Database (NETS), collected by Duns and Bradstreet (D&B). This dataset provides annual information on exact geographic location, employment, and sales, as well as NAICS six-digit industry code and

[9] A notable exception is Tsivanidis (2018) who evaluates the distributional effects of infrastructure investment in a model of commuting by skill groups with Stone-Geary preferences.

[10] I will use Los Angeles Metropolitan Area and Los Angeles interchangeably. In the data I treat all urban contiguous Census tracts in Los Angeles County as the Los Angeles Metropolitan Area.

[11] I will refer to the ACS 2012-1016 as 2014 for the remainder of the paper.

business characteristics of 2 million establishments in Los Angeles from 1990-2014.[12]

In order to map establishments into sectors, I first create 28 separate "local" sectors. I define a "local" sector based on the idea that households physically go to an establishment to purchase/consume different goods and services. In defining these sectors, I account for quality differences as far as possible. For example, households can eat out at fast food establishments versus full-service restaurants or buy groceries at supermarkets or specialty food stores. In both cases, I allow for two different sectors. However, due to data limitations mostly in household expenditure microdata, I cannot account for finer quality differences, like the difference between a regular supermarket and a specialty or upscale supermarket, such as Whole Foods Market. Next, I create a crosswalk between NAICS six-digit sectors in the NETS data and my 28 local sectors, as well as a crosswalk between my sectors and items in household expenditure microdata. In all crosswalks, I assign firms to local sectors based on where a typical household buys a good or service versus in which sectors certain goods are produced. For example, most food items are produced in agricultural sectors but predominantly purchased by consumers in grocery stores. All establishments or expenditure items that cannot be mapped to a local sector are assigned to a "frictionless" sector. I assume firms in the frictionless sector to be equally accessible to all consumers.

To estimate how demand by high skilled and low skilled households varies for local sectors, I use three datasets on household-level expenditures. First, I capture expenditure on service sectors in the Consumer Expenditure Survey (CEX) Interview data, which provides quarterly expenditures across roughly 700 unique expenditure categories. Second, I can break up some sectors that are aggregated in the quarterly data like "food away from home" using biweekly expenditures in the Diary data of the CEX. This breaks food away from home into restaurants, fast food, and bars. Lastly, I use the Nielsen Consumer Panel data to capture demand patterns across retail sectors. This dataset provides detailed expenditures across retail chains and is organized in retail channels, which correspond to 13 of my 28 local sectors for around 40-60k households between 2004-2017.

Data on the geographic distribution of housing comes from the Los Angeles County Tax Assessor. The tax assessor collects a variety of information on parcel size, building square footage, number of living units, number of stories in a building, year built, and usage for every parcel in the county. I compute the residential housing stock by aggregating the square footage of the main building on every residential parcel in a census tract. Since available tax assessor data only goes back to 2006, I impute the housing stock for 1990 and 2000 by removing all buildings built after the respective census year in a census block. I then assign the average size of the remaining units to all units that are reported in the census block data from NHGIS.[13]

Lastly, I use data from Lee & Lin (2017) to account for natural amenities like average slope,

[12] Throughout this paper I will use the terms establishment and firm interchangeably.

[13] For 2014, the number of living units in the ACS data corresponds almost perfectly to the number of living units in the tax assessor data. To keep the housing stock consistent over time I also assign the average unit size in a tract to all units reported in the 2014 ACS data.

temperatures or distance to shore for each tract.

4 Motivating Evidence

To provide context and to motivate my modeling choices, I document stylized evidence on how households with a college-educated head and firms that provide goods and services preferred by richer households collocate throughout Los Angeles. In doing so, I point to endogenous differences in access to consumption varieties for high and low skilled consumers.

First, residence of high and low skilled households are strongly segregated in LA. Figure 1 plots the ratio of high-skilled residents over low skilled residents in a census tract, henceforth skill ratio, in Los Angeles in 2014. For example, South and East Los Angeles are almost exclusively populated by low skilled residents whereas college educated residents can be found along the coast (Santa Monica or Malibu) and in the hilly parts of Los Angeles. Figure 2 shows a similar relationship for the number of firms operating in local sectors for which I estimate income elasticities above the median over the number of firms in sectors with below median income elasticity.[14] Although the spatial distribution of firms is noisier, we can observe a similar pattern: locations with more high skilled households tend to be locations with more firms in sectors that are disproportionately preferred by richer households (or they are highly income-elastic). In Table 1, I report the results from regressing the log ratio of establishment counts by income elasticity on the skill ratio in a tract. A doubling of the skill ratio is associated with 12.6% higher ratio of firms in income-elastic sectors over firms in inelastic sectors.

To give a more specific example of how firms and households collocate based on demand patterns, I compare the two sectors with the highest income elasticities, recreation and education services, with the two sectors that I find to be the least income-inelastic, liquor/tobacco stores and convenience stores. Figure 3 plots the log number of establishments for both pairs of sectors against the log skill ratio in each tract. Recreation and education services are much more prevalent in locations with more high skilled residents as compared to liquor and convenience stores. In columns 2 and 3 of Table 1, I document the same relationship in linear regressions. In richer census tracts, the number of establishments in recreation and education is four times larger than the number of liquor and convenience stores. Furthermore, in columns 4 and 5, I find that the likelihood to observe any establishment in the two highly income-elastic sectors relative to the two inelastic sectors is six times higher as a function of the local skill ratio.

Figure 4 reports coefficients and 95% confidence intervals in black for all sectors from regressing log establishment counts on the log local skill ratio in a tract. I order the point estimates by my sector-level income elasticity estimates on the vertical axis. The ranking of regression coefficients and the ranks of the income elasticities are quite correlated (Spearman Rank Cor-

[14] I estimate income elasticities by sector with household expenditure microdata in the estimation section below. However, the ordering of sectors by income elasticity is intuitive: examples of highly income-elastic sectors are recreation, education, amusement or apparel stores. Liquor stores, dollar stores, fast food restaurants, or gas stations, I find to be income-inelastic.

relation: $.495$ with p-value of $.007$). The positive relationship implies that the number of firms in sectors that offer goods and services preferred by rich consumers, e.g. high income elastic sectors, is associated with locations populated by high skilled households. A major concern in this stylized analysis is that location choices of firms are differently impacted by supply factors instead of demand factos, such as having access to high skilled workers. To alleviate this concern, albeit imperfectly, I report coefficients from the same regression in Figure 4, but I control for the log ratio of high skilled to low skilled employees, total employment, and population density in each tract. The positive association between the effect of the skill ratio on the number of firms across sectors in a tract and the ordering of how much sectors are preferred by richer households remains stable. To sum up the stylized evidence, residential location choices of skill groups and location choices of firms operating in sectors that are preferred differently by skill groups are correlated. This points to endogenous differences in access to consumption varieties for low and high skilled households, a key ingredient to my model.

5 Model

Motivated by these correlations in the data, I develop a spatial general equilibrium model. It characterizes the forces leading to the spatial sorting of skill groups in a city, and it also guides my theoretical and empirical analysis. In particular, the model features two key skill-location-specific agglomeration forces: endogenous relative local price indices due to two-sided sorting and reduced-form spillovers.

5.1 Setup

The city consists of N neighborhoods, indexed n. It is populated by K types of heterogeneous households, indexed k, with fixed mass L_k. There are J sectors whose products differ in income elasticities of demand in household preferences. Households choose the location of their residence, consume housing h_{kn} and a bundle of goods across J sectors, $C_{kn}(g)$.[15] Housing is consumed in the neighborhood of residence whereas goods can be consumed everywhere in the city at iceberg trade costs, which I refer to as shopping frictions. Within each goods-sector there is a continuum of profit-maximizing firms that produce differentiated varieties and decide in which neighborhood to locate. Households can work anywhere in the city and provide labor inelastically for which there are no commuting costs.

5.2 Household Problem

A household ι of type k with preference draw $b_{kn}(\iota)$ choosing to live in n has utility,

$$\mathcal{U}_{kn}(\iota) = U_{kn} b_{kn}(\iota) = \frac{I_{kn}}{P_{kn}} b_{kn}(\iota) \tag{1}$$

[15] When using the term "goods", I am referring to all goods and services other than housing for brevity.

where U_{kn} is real consumption, I_{kn} is type k's income and P_{kn} a type-neighborhood specific price index. Each household ι draws an idiosyncratic preference $b_{kn}(\iota)$ for every neighborhood n that is distributed Fréchet, $b_{kn}(\iota) \sim e^{-B_{kn}z^{-\kappa}}$. Conditional on their preference draws, households choose a neighborhood of residence n that provides the highest utility.

Importantly, real consumption U_{kn} follows a *non-homothetic* CES aggregator between housing and goods and is implicitly defined as

$$\left(a_h U_{kn}^{\epsilon_h}\right)^{\frac{1}{\eta}} h_{kn}^{\frac{\eta-1}{\eta}} + \left(a_g U_{kn}^{\epsilon_g}\right)^{\frac{1}{\eta}} C_{kn}(g)^{\frac{\eta-1}{\eta}} = 1 \tag{2}$$

where η is the elasticity of substitution between housing and goods.
Depending on the relative size of ϵ_h and ϵ_g consumers shift expenditure between housing and goods when real consumption changes. For example, if housing is a necessity ($\epsilon_h < \epsilon_g$) then consumers with a higher level of real consumption spend a larger fraction of income on goods at given relative prices.[16] Comin *et al.* (2018), Borusyak & Jaravel (2018), and Matsuyama (2019) provide detailed discussions of non-homothetic CES preferences. I assume that the non-homotheticity in the model operates only on real market consumption in a neighborhood n, U_{kn}, as opposed to idiosyncratic utility $\mathcal{U}_{kn}(\iota)$.[17]

Consumption across goods sectors j also follows a *non-homothetic* CES aggregator with elasticity of substitution γ,

$$C_{kn}(g) = \left(\sum_{j=1}^{J} \left(\alpha_j U_{kn}^{\nu_j}\right)^{\frac{1}{\gamma}} c_{kn}(j)^{\frac{\gamma-1}{\gamma}}\right)^{\frac{\gamma}{\gamma-1}}.$$

As for the upper nest, the implied CES weights are functions of total real consumption U_{kn}. Given prices, as a consumer's real consumption increases, she shifts expenditures to goods from sectors with higher income elasticity parameters ν_j.[18]

Within each sector j there is an endogenous set of differentiated varieties $\Omega_n(j)$ being offered in neighborhood n.[19] A variety is denoted by ω. Households aggregate varieties from all neighborhoods in each sector j with a homothetic CES aggregator and elasticity of substitution σ,

$$c_{kn}(j) = \left(\sum_{n'=1}^{N} \left(\int_{\Omega_{n'}(j)} c_{knn'}(j,\omega)^{\frac{\sigma-1}{\sigma}} d\omega\right)\right)^{\frac{\sigma}{\sigma-1}}.$$

Households in n can access sector j varieties in another neighborhood n' at shopping costs $\tau_{knn'}(j)$ that takes iceberg form. The price of variety ω offered by a firm in n' faced by a household

[16] Note that when $\epsilon_h = \epsilon_g = 1 - \eta$ the expression reduces to the regular CES consumption aggregator.

[17] It is not ex-ante clear whether systematic variation in tastes for different goods is due to real market consumption or overall well-being of a consumer. I choose the former for tractability. Alternative to making this assumption directly, I could assume that households make consumption decisions before the idiosyncratic preference shocks are realized.

[18] Similarly, the aggregator takes the homothetic CES form if $\nu_j = 0, \forall j$.

[19] The household takes set $\Omega_n(j)$ as given. It is determined in equilibrium as the interaction of household and firm problem.

of type k in neighborhood n is

$$p_{knn'}(j,\omega) = \tau_{knn'}(j)p_{n'}(j,\omega)$$

where $p_{n'}(j,\omega)$ is the price of this variety at the shopping destination n'. I can write the expenditure share of a household of type k living in n on variety ω offered by a firm in n' in sector j according to

$$\tilde{s}_{knn'}(j,\omega) = \tau_{knn'}(j)^{1-\sigma}\left(\frac{p_{n'}(j,\omega)}{p_{kn}(j)}\right)^{1-\sigma},$$

where

$$p_{kn}(j) = \left(\sum_{n'=1}^{N}\left(\int_{\Omega_{n'}(j)}\tau_{knn'}(j)^{1-\sigma}p_{n'}(j,\omega)^{1-\sigma}d\omega\right)\right)^{\frac{1}{1-\sigma}} \tag{3}$$

is the price index of sector j faced by a household k in n.[20] Utility maximization implies that the expenditure share on all varieties in sector j follows

$$\tilde{s}_{kn}(j) = \alpha_j U_{kn}^{\nu_j}\left(\frac{p_{kn}(j)}{P_{kn}(g)}\right)^{1-\gamma}, \tag{4}$$

where the price index across goods is

$$P_{kn}(g) = \left(\sum_{j=1}^{J}\alpha_j U_{kn}^{\nu_j}p_{kn}(j)^{1-\gamma}\right)^{\frac{1}{1-\gamma}}. \tag{5}$$

A convenient property of non-homothetic CES preferences is that the household's expenditure shares on housing and goods as compensated demand is,

$$s_{kn}(h) = a_h\left(\frac{r_n}{I_{kn}}\right)^{1-\eta}U_{kn}^{\epsilon_h} \quad \text{and} \quad s_{kn}(g) = a_g\left(\frac{P_{kn}(g)}{I_{kn}}\right)^{1-\eta}U_{kn}^{\epsilon_g} \tag{6}$$

where r_n stands for the residential rent in n. The overall price index for type k in n is

$$P_{kn} = \left(a_h U_{kn}^{\epsilon_h-(1-\eta)}r_n^{1-\eta} + a_g U_{kn}^{\epsilon_g-(1-\eta)}P_{kn}(g)^{1-\eta}\right)^{\frac{1}{1-\eta}}. \tag{7}$$

Applying the Fréchet distribution ($b_{kn}(\iota) \sim e^{-B_{kn}z^{-\kappa}}$) to equation 1 I can write the mass of households of type k that resides in neighborhood n as

$$L_{kn} = \frac{B_{kn}U_{kn}^{\kappa}}{\Phi_k}L_k \tag{8}$$

where $\Phi_k \equiv \sum_{n'} B_{kn'}U_{kn'}^{\kappa}$. This term is related to expected utility of a household of type k given

[20] For the remainder of the paper I denote expenditure shares within a sector or within the goods bundle with tilde. Expenditure shares out of total consumption are denoted without tilde.

by

$$\bar{\mathcal{U}}_k = \Gamma\left(\frac{\kappa-1}{\kappa}\right)\left(\sum_{n'} B_{kn'} U_{kn'}^{\kappa}\right)^{\frac{1}{\kappa}} = \gamma \Phi_k^{\frac{1}{\kappa}}.$$

Finally, skill types supply different levels of efficient units of labor: type k provides ρ_k units of labor. Since labor is freely mobile throughout the city (no commuting costs), the wage w per efficient unit of labor is equalized across all job locations. As a result, income of households with the same skill-level is constant and follows $I_k = w\rho_k + T_k$ where T_k is a lump-sum transfer that is independent of the location of a household. In this model, I abstract from spatial income differences within skill groups to focus on differences in price indices of consumption, which is the new mechanism in my model.

5.3 Firm Problem

There is an infinite mass of potential entrepreneurs outside the city who have access to a single variety ω in a given sector j. To enter the city entrepreneurs have to incur fixed costs $f^e(j)$ in terms of labor. First, entrepreneurs observe expected profits from entering the city in their respective sector $E(\pi_n(j,\omega))$ and decide to do so if it exceeds the fixed entry cost $f^e(j)$. Conditional on entry, each entrepreneur ω in sector j receives an idiosyncratic productivity $z_n(j,\omega)$ to produce in neighborhood n drawn iid from a Fréchet distribution ($z_n(j,\omega) \sim e^{-A_n(j)z^{-\theta}}$ with $\theta > 1$).

An entrepreneur with variety ω in sector j and location n produces output with labor $l_n(j,\omega)$ according to

$$y_n(j,\omega) = z_n(j,\omega) l_n(j,\omega). \tag{9}$$

In order to determine the mass of firms in each sector j and neighborhood n we can write variable profits of ω choosing n as

$$\begin{aligned} \pi_n(j,\omega) &= \frac{1}{\sigma}\sum_{n'=1}^{N}\sum_{k'=1}^{K} s_{k'n'n}(j,\omega) I_{k'} L_{k'n'} \\ &\equiv z_n(j,\omega)^{\sigma-1}\tilde{\pi}_n(j). \end{aligned} \tag{10}$$

where $s_{knn'}(j,\omega)$ denotes share of total expenditure of type k living in n on variety ω located in n'. Profits can be decomposed into idiosyncratic productivity $z_n(j,\omega)$ and a neighborhood-sector-specific profit term $\tilde{\pi}_n(j)$.

Applying the Frechét distribution of $z_n(j,\omega)$ to 10 we get the mass of entrepreneurs in sector j that choose neighborhood n

$$M_n(j) = \frac{A_n(j)\tilde{\pi}_n(j)^{\frac{\theta}{\sigma-1}}}{\Pi(j)} M(j), \tag{11}$$

where $M(j)$ is the total mass of firms in sector j operating across the city and

$$\Pi(j) \equiv \sum_{n'}^{N} A_{n'}(j)\tilde{\pi}_{n'}(j)^{\frac{\theta}{\sigma-1}}.$$

I can rewrite the price index of sector j for household k in n in equation 3 as

$$p_{kn}(j) = \frac{\sigma}{\sigma-1} w\bar{\gamma}^{\frac{1}{1-\sigma}} M(j)^{-\frac{1}{\theta}} \left(\sum_{n'=1}^{N} \tau_{knn'}(j)^{1-\sigma} A_{n'}(j)^{\frac{\sigma-1}{\theta}} M_{n'}(j)^{1-\frac{\sigma-1}{\theta}} \right)^{\frac{1}{1-\sigma}} \tag{12}$$

where $\bar{\gamma} = \Gamma\left(1 - \frac{\sigma-1}{\theta}\right)$ is the Gamma function. We can directly see the effect of firm sorting on local prices: with $\theta > \sigma - 1$ neighborhoods closer to other locations (low $\tau_{knn'}(j)$) with more firms in a sector j face a lower price index due to a larger number of varieties $M_{n'}(j)$. Furthermore, as we can interpret the Frechét scale parameter $A_n(j)$ as a location-sector-specific average productivity, locations with higher $A_n(j)$ experience lower prices for j. Despite idiosyncratic productivity differences across firms and spatial mobility, all endogenous differences in sector-level prices across locations can be summarized by differences in the number of varieties.

The share of total expenditure of households in n in a given sector j on all varieties in location n' is then,

$$s_{knn'}(j) = s_{kn}(g)\tilde{s}_{kn}(j) \frac{\tau_{knn'}(j)^{1-\sigma} A_{n'}(j)^{\frac{\sigma-1}{\theta}} M_{n'}(j)^{1-\frac{\sigma-1}{\theta}}}{\sum_{n''=1}^{N} \tau_{knn''}(j)^{1-\sigma} A_{n''}(j)^{\frac{\sigma-1}{\theta}} M_{n''}(j)^{1-\frac{\sigma-1}{\theta}}}. \tag{13}$$

The shopping behaviour of households follows a gravity structure similar to workhorse trade models, but the allocation of expenditures across shopping locations is determined by the endogenous number of available varieties in a destination, shopping frictions, and average productivities.

Note that due to sorting, average profits per variety in a sector are equalized across locations in equilibrium,

$$\frac{1}{M_n(j)} \int_{\Omega_{n'}(j)} \pi_n(j,\omega) d\omega = \bar{\gamma}\Pi(j)^{\frac{\sigma-1}{\theta}}. \tag{14}$$

Lastly, we can pin down the mass of active varieties in each sector j, $M(j)$, by equalizing expected profits $E(\pi_n(j))$ with fixed costs of entry $f^e(j)$,

$$\bar{\gamma}\Pi(j)^{\frac{\sigma-1}{\theta}} = f^e(j). \tag{15}$$

5.4 Frictionless Sector

To allow for part of goods consumption that is independent of location, I include a frictionless sector (the Jth sector) within the goods bundle that operates in the same way as all other sectors. The exception being that consumers do not face shopping frictions for this sector such that $\tau_{knn'}(J) = 1, \forall n, n'$. In the absence of shopping frictions the firm location problem is exclusively

determined by productivities $A_n(J)$,

$$M_n(J) = \frac{A_n(J)}{\sum_{n'} A_{n'}(J)} M(J)$$

and the price index for sector J collapses to

$$p_{kn}(J) = p(J) = \frac{\sigma}{\sigma - 1} w \left(\bar{\gamma} M(J)\right)^{\frac{1}{1-\sigma}} .$$

Since households do not face shopping frictions the sector price index in any location is determined by the number of varieties city-wide $M(J)$, which is given by the free entry condition for the Jth sector as in 15. Since this result holds for any sector that does not face shopping frictions, I can remove relative price index differences as a force that creates sorting of households across locations by setting $\tau_{knn'}(j) = 1, \forall n, n', \forall j$.

5.5 Housing Markets

Atomistic landlords own a fixed amount of residential housing H_n in each neighborhood n. They are the claimant on the returns to housing but are fully taxed by the city government. All households in location n consume housing at rents r_n.

In equilibrium, expenditure on residential housing in n needs to equal $r_n H_n$

$$r_n H_n = \sum_k^K s_{kn}(h) I_k L_{kn} = a_h \left(r_n\right)^{1-\eta} \sum_k^K I_k^{\eta} U_{kn}^{\epsilon_h} L_{kn}.$$

5.6 City Government

The city government collects all housing expenditures in the city by fully taxing landlords. It redistributes revenues net of any expenditures D, for example place-based subsidies to firms or renters. I will describe this below in detail when I discuss policy shocks. I assume that the city government returns the leftover budget to households as lump-sum transfer proportional to household labor endowment such that

$$T_k = \rho_k \frac{\left(\sum_{n=1}^N r_n H_n\right) - D}{\sum_{k=1}^K \rho_k L_k}. \tag{16}$$

The transfer scheme ensures that relative income differences are invariant to policy shocks.[21]

[21] This formulation is isomorphic to assuming that households own a share in the city-wide housing stock proportional to their labor endowment. To finance policies, the city government taxes households lump-sum proportional to labor endowment.

5.7 Reduced-form Spillovers

Since I study the role of relative price indices and reduced-form spillovers in explaining spatial sorting of households, I allow for direct spillovers within and across skill groups and locations. Similar to the previous literature (Diamond (2016), Su (2018b), Tsivanidis (2018), Fajgelbaum & Gaubert (2019)), I model reduced-form amenity spillovers for type k as returns to the number of residents of their type k and other types k'. However, I also allow spillovers to operate across neighborhoods, for example from n' to n, to make spillovers consistent with the notion that relative price indices depend on a neighborhood's geographic location, similar to Ahlfeldt *et al.* (2015) who model reduced-form spillovers operating on population density,

$$B_{kn} = \bar{B}_{kn}\mathcal{L}_{kn} = \bar{B}_{kn}\prod_{n'}\prod_{k'} L_{k'n'}^{\delta_{k'n',kn}}, \tag{17}$$

where $\bar{B}_{kn}$ represents exogenous amenities and $\mathcal{L}_{kn}$ stands for spillovers. Elasticities $\delta_{k'n',kn}$ govern how strongly amenities on households of type k in n respond to the number of residents of type k' in neighborhood n'. This formulation nests spillovers from the local skill composition, as in Diamond (2016) or Su (2018b), by setting $\delta_{k'n',kn} = 0$ for all $n \neq n'$.

5.8 Competitive Equilibrium

The equilibrium of this economy is defined by a distribution of households by neighborhood and skill group with $\sum_{n'\in\{1,2,\ldots,N\}} L_{kn'} = L_k, \forall k$, a distribution of firms by neighborhood and sector with $\sum_{n'\in\{1,2,\ldots,N\}} M_{n'}(j) = M(j), \forall j$, mass of firms in sectors $M(j), \forall j$, prices in all sectors and neighborhoods $\{p_n(j)\}$, sector price indices $\{p_{kn}(j)\}$, neighborhood-skill goods price indices $\{P_{kn}(g)\}$, neighborhood-skill price indices $\{P_{kn}\}$, wage w, residential rents $\{r_n\}$ and transfers $\{T_k\}$ such that:

1. Each type k in a neighborhood n maximizes utility given w, r_n, $p_n(j)$, $p_{kn}(j)$, $P_{kn}(g)$, P_{kn} and T_k and chooses the neighborhood that provides the highest utility with probabilities given in equation 8.

2. Firms in sectors j in neighborhood n maximize profits in 10 taking P_{kn}, $P_{kn}(g)$, p_{kn}, w, and the distribution of households as given and choose the neighborhood that maximizes profits with probabilities given in 11.

3. In each sector j, the mass of varieties is such that fixed cost of entry equals expected profits from entering the city.

4. Markets for residential housing clears in each n.

$$\sum_{k=1}^{K} \lambda_{kn} L_k h_{kn} = H_n, \forall n. \tag{18}$$

5. The labor market clears in the city

$$\sum_k \rho_k L_k = \sum_{j=1}^{J} \sum_{n=1}^{N} \int_{\Omega_n(j)} l_n(j,\omega) d\omega + \sum_{j=1}^{J} M(j) f^e(j). \tag{19}$$

6. Transfers are given by

$$T_k = \rho_k \frac{\left(\sum_{n=1}^{N} r_n H_n\right) - D}{\sum_{k=1}^{K} \rho_k L_k}.$$

5.9 Discussion of Uniqueness of the Equilibrium

The model supports multiple equilibria if the skill-specific agglomeration externalities (two-sided sorting and reduced-form spillovers) dominate the various dispersion forces present in the model.[22] On the household side, inelastic housing supply and idiosyncratic preferences for neighborhoods ensure that all neighborhoods are populated by high and low skilled households. Similarly, firms locate in all neighborhoods due to local competition forces and idiosyncratic productivity draws. In section 7, I calibrate the parameters of the model to ensure that the household and firm distributions are unique. Although a formal proof of the necessary conditions for uniqueness is still work in progress, I perform a number of numerical simulations to test whether the my baseline model calibration supports a unique equilibrium. These tests suggest that the pecuniary externality from two-sided sorting and spillovers are weaker than the dispersion forces if preference and productivity draws are sufficiently dispersed (small κ and θ), the elasticity of substitution between housing and goods η is less than one, and real consumption is concave in expenditure.

6 Model Properties

6.1 Sorting Patterns

In the following section, I build intuition for the model's main contribution, namely, that firms in sectors with high income elasticity collocate with high-skilled households based on demand patterns arising from non-homothetic preferences in the model. This feature creates price index differences that endogenously lead to sorting patterns of households with different incomes. To keep the exposition and notation simple, I assume for now that shopping frictions outside the location of residence are infinite, $\tau_{knn'}(j) = \infty, \forall n' \neq n$, meaning households can only access varieties in their residence.[23] Furthermore, in this section and for the remainder of paper, I

[22] If externalities are stronger than dispersion forces, some neighborhoods may attract predominantly high skilled households and firms in income-elastic sectors in one equilibrium configuration; however, the same neighborhoods may be populated by low skilled households and firms in income-inelastic sectors in an alternative equilibrium. Which equilibrium is reached depends on the starting values when I compute the model.

[23] If shopping frictions do not vary by skill type $\tau_{knn'}(j) = \tau_{nn'} \forall k, j$ which I assume for the calibration of the model below, results of this section are unaffected since spatial consumption patterns are independent of skill type. How-

assume that $K = 2$, e.g. the city is populated by high and low skilled households.

The key variable summarizing underlying sorting patterns is the local expenditure share by skill group k in location n on goods from a sector j, $s_{kn}(j) = s_{kn}(g)\tilde{s}_{kn}(j)$ and is described in the following proposition:

Proposition 1. *Given prices, the expenditure share of households of skill k in location n on goods of sector j, $s_{kn}(j)$, is log-supermodular in real consumption U_{kn} and sector income elasticity parameter ν_j.*

Proposition 1 states that as households get richer[24] they value goods from sectors with higher income elasticity relatively more and that the difference is increasing with real consumption (see Matsuyama (2019) for a similar argument). As a consequence, high-skilled households' expenditure is tilted towards income-elastic sectors relative to low-skilled households. The top graph of Figure 5 shows a stylized graphical representation of this finding: I plot the log of expenditure shares for three sectors with decreasing income elasticity ($\nu_1 > \nu_2 > \nu_3$) on the vertical axis and household income on the horizontal axis. High skilled households with income I_{high} spend a larger fraction of income on the first sector and less on the other sectors in comparison to low skilled households with I_{low}.

By relating this property to firm profits in n, I can rearrange profits of all varieties in sector j and location n in equation 10 as a function of the share of high skilled residents x_n in local population L_n,

$$\frac{\Pi_n(j)}{L_n} = \frac{1}{\sigma}\left(s_{high,n}(j)I_{high}x_n + s_{low,n}(j)I_{low}(1-x_n)\right).$$

Now, I can relate proposition 1 to average profits by resident of sector j according to the following corollary:

Corollary 1. *Given prices, total profits by resident of firms in sector j in location n is log-supermodular in high-skilled share x_n and sector income elasticity parameter ν_j.*

Intuitively, since high-skilled households spend more on income-elastic sectors, locations with a larger share of high skilled residents offer larger profits to firms in income-elastic sectors relative to income-inelastic sectors. Applying equation 14 and corollary 1 it follows immediately that the number of varieties $M_n(j)$ in income-elastic relative to income-inelastic sectors in locations with more high skilled residents must be larger than in locations with a lower share thereby keeping prices and total residents equal.[25] We can conclude that $M_n(j)$ is also log-supermodular in the high-skilled share x_n and sector income elasticity parameter ν_j implying that: $\frac{M_n(j)}{M_{n'}(j)}$ is

ever, defining local demand faced by firms in a location and the price index faced by households become complex functions of geography when shopping frictions are finite outside the residence, which makes the exposition less tractable.

[24] I assume that real consumption is increasing in nominal income I_k.

[25] Profits are also increasing with total number of residents but at given expenditure shares and prices, in equal proportions for all sectors, such that only the composition of residents is relevant for relative sorting of varieties by sector.

non-decreasing in ν_j if $x_n > x_{n'}$. This result establishes that firms offering varieties in income-elastic sectors collocate with high income residents. Figure 5 summarizes how the number of varieties $M_n(j)$ increase faster in sector 1 (the highly elastic sector) than for the two less elastic sectors as the average income per resident in n on the horizontal axis increases.[26]

Next, I can combine proposition 1 and corollary 1 to characterize how residents respond to the distribution of varieties in a location. Since the price index of goods is a function of the high skilled share x_n and real consumption U_{kn}, $P_{kn}(g)$ has the following property.

Corollary 2. *Taking M_n, $M_{n'}$ and x_n as given and $\sigma, \gamma > 1$, households' price index of goods consumption, $P_{kn}(g)^{1-\gamma}$, is log-supermodular in real consumption U_{kn} and high-skilled share x_n or*

$$\frac{P_{high,n}(g)}{P_{low,n}(g)} < \frac{P_{high,n'}(g)}{P_{low,n'}(g)}$$

if $x_n > x'_n$.

Corollary 2 combines the intuition of both earlier findings and is graphically depicted in the bottom picture of Figure 5. Since richer households have higher expenditure shares on income-elastic sectors and locations with a larger share of high skilled households attract disproportionately more varieties in such sectors, the relative price index of goods between high skilled and low-skilled households must be lower in such neighborhoods compared to locations with more low skilled households.

In equilibrium, the high-skilled share and real consumption are related through the location choice of households. Locations with lower relative goods prices between high and low skilled consumers attract more high skilled households. This, then, increases the high skilled share in the population and further reduces the relative price due to more local varieties in income-elastic sectors. Due to the interaction of location choice of households and firms, the model endogenously produces relative price differences that generates a pecuniary externality on residents. As we will see in the next section, this externality is separate from reduced-form spillovers as captured by $\mathcal{L}_{kn}$ in the model.

6.2 Local Decomposition of Price Index Effects and Reduced-form Spillovers

In this section, I characterize the forces in the model that link the location choice problems of different households across skill types and locations. In particular, relative price index effects and reduced-form spillovers generate externalities that amplify the mobility response of households to shocks. I consider a small shock in location n', for example an exogenous change in fixed amenities or place-based policy shock, that leads to a change in the population of skill group k' in n', which I refer to as the "shocked population." Then, I decompose the mobility response of residents of type k in location n ("target population") to this change. Taking expenditure shares

[26] Note that with fixed income by type, the average income per resident is a sufficient statistic for x_n.

as fixed,[27] I start from the expression for L_{kn} in equation 8. I take logs and differentiate with respect to $\log L_{k'n'}$ to get

$$\frac{d\log L_{kn}}{d\log L_{k'n'}} = -\kappa \underbrace{\frac{(1-\eta)}{\bar{\epsilon}_{kn}}}_{\text{Marginal Utility}} \underbrace{s_{kn}(h)\frac{d\log r_n}{d\log L_{k'n'}}}_{\text{Rent Congestion}} + \underbrace{\delta_{k'n',kn}}_{\text{Reduced-Form Spillover}}$$

$$+\kappa \underbrace{\frac{(1-\eta)}{\bar{\epsilon}_{kn}}}_{\text{Marginal Utility}} \frac{1}{\theta}\frac{I_{k'}L_{k'n'}}{Y_c}\left(\sum_j \underbrace{\frac{s_{kn}(j)s_{k'n'}(j)}{s_c(j)}}_{\text{Non-Homotheticity}}\left(1+\left(\frac{\theta}{\sigma-1}-1\right)\sum_{n''}\underbrace{\frac{\tilde{s}_{nn''}(j)\tilde{s}_{n'n''}(j)}{s_{n''}(j)}}_{\text{Spatial Frictions}}\right)\right)$$

$$+ \underbrace{c_k}_{\text{Terms independent of } n} \tag{20}$$

where $\bar{\epsilon}_{kn} = s_{kn}(h)\epsilon_h + s_{kn}(g)\left(\epsilon_g + \frac{1-\eta}{1-\gamma}\bar{\nu}_{kn}\right)$ and $\bar{\nu}_{kn} = \sum_j \tilde{s}_{kn}(j)\nu_j$ are expenditure weighted average income elasticity parameters. Y_c stands for city-wide total income, $s_c(j)$ is the city-wide average expenditure share on sector j and $s_n(j)$ is the share of expenditure on varieties in n out of total expenditure in j. The term $\tilde{s}_{nn'}(j)$ is the expenditure share of a household in n on varieties in n' out of all expenditures on j.[28]

Terms in the first line correspond to forces present in many quantitative urban models. First, if the shocked population lives in n then additional residents bid up housing rents and congest the location muting the mobility response of the target population. In a model with non-homothetic preferences, price changes are evaluated at marginal utility since changes in expenditure do not translate one-to-one into utility.[29]

Second, depending on the sign of $\delta_{k'n',kn}$, reduced-form spillovers from the shocked population make location n more or less attractive to the target population. Stated differently, if the target population likes living close to the shocked population then the attractiveness of n increases. In this sense, reduced-form spillovers are a black box as they create sorting without any specific economic force underlying them.

The second line of 20 summarizes the effect of a change in the shocked population on the goods price index of the target population, the key new sorting force in the model.[30] First, if the distribution of firms is exogenously given, then price indices are are also exogenous and the

[27] This decomposition holds only locally e.g. for a infinitesimal shock. In response to a larger shock expenditure shares adjust, making the expression intractable without necessarily conveying more information.

[28] Since I assume for simplicity that shopping frictions are the same for all skill types k, expenditures by destination do not vary with k.

[29] Under the assumption that utility in concave, but increasing in income, the importance of real consumption e.g. rents in determining the attractiveness of a location is diminishing relative to reduced-form spillovers or fixed amenities.

[30] The strength of the effect also depends, similar to rents, on the marginal utility of consumption, a set of elasticities (firm supply elasticity θ and love of variety $\frac{1}{\sigma-1}$) and the relative economic size of the shocked population, $\frac{I_{k'}L_{k'n'}}{Y_c}$.

term in the second line of 20 is zero. Thus, for prices to adjust an extensive margin in the number of varieties, e.g. sorting of firms, is necessary. When interpreting the first term in parentheses, price index effects are more pronounced for target populations that spend more on goods relative to housing (higher $s_{kn}(g) = \sum_j s_{kn}(j)$), for example for high skilled households if goods consumption is more income-elastic than housing. Price index effects are stronger if expenditure shares across sectors, $s_{kn}(j)$, of target population and shocked population are more correlated as the term in parentheses is similar to the covariance of expenditure shares. Due to non-homothetic preferences, households with similar incomes have more correlated expenditure patterns; hence, price indices respond more within skill groups than across groups.

The second sum in parentheses has a similar interpretation, but instead of variation in expenditure shares due to non-homothetic preferences, price index effects are stronger for populations that have spatially correlated expenditure shares. The "Spatial Frictions" term is large if target and shocked population both spend a lot in n'' (high $\tilde{s}_{nn''}(j)$ and $\tilde{s}_{n'n''}(j)$) relative to the overall importance of this location in j, as measured by $s_{n''}(j)$. For example, an increase in population and the associated entry of firms reduces price indices in neighboring locations more than in far away locations.

It is instructive to think about two special cases. First, if preferences are homothetic ($\epsilon_h = \epsilon_g = 1 - \eta$ and $\nu_j = 0$) and there is only one sector[31] then the expression in 20 simplifies to

$$\frac{d\log L_{kn}}{d\log L_{k'n'}} = -\kappa s_n(h)\frac{d\log r_n}{d\log L_{k'n'}} + \delta_{k'n',kn} + \kappa\frac{s_n(g)}{\theta}\frac{I_{k'}L_{k'n'}}{Y_c}\left(1 + \left(\frac{\theta}{\sigma-1} - 1\right)\sum_{n''}\frac{\tilde{s}_{nn''}\tilde{s}_{n'n''}}{s_{n''}}\right) + c_k$$

As a result of homothetic preferences, endogenous price effects on the target population are independent of skill and act solely as an agglomeration force on the local population due to love of variety and free entry, as in Krugman (1991).

Second, let us assume shopping is frictionless ($\tau_{nn'} = 1, \forall n, n'$) then

$$\frac{d\log L_{kn}}{d\log L_{k'n'}} = -\kappa s_{kn}(h)\frac{d\log r_n}{d\log L_{k'n'}} + \delta_{k'n',kn} + \kappa\frac{1}{\sigma-1}\frac{I_{k'}L_{k'n'}}{Y_c}\sum_j\frac{s_{kn}(j)s_{k'n'}(j)}{s_c(j)} + c_k$$

Without shopping frictions, price index effects only operate through entry of firms at the city border. More firms in sectors preferred by the shocked population enter the city leading to a stronger fall in prices if the target population's expenditure shares across sectors are more correlated. When I simulate counterfactuals with only reduced-form spillovers below I assume that shopping is frictionless, hence, this special case of 20 applies.

To sum up, price indices of consumption endogenously respond to changes in the spatial income or skill composition through sorting of firms. In addition, the strength of the effect depends on relative expenditure patterns of the target and shocked population leading to differential sorting of skill groups across space in response to a shock. In contrast to reduced-form

[31] There is qualitatively no difference if there is more than one sector other than exogenous differences in productivity driving the spatial distribution of firms.

spillovers, relative price index effects feature rich heterogeneity based on geography and income inequality of a city. Furthermore, the effect of relative price indices on mobility is not invariant to the initial equilibrium or context. Suppose we are able to compare the outcomes of the same policy shock for two different cities: one with little and the other with very strong income inequality. In the first city, consumption baskets of households would be similar, therefore, there would be little difference in mobility due to the price index channel. However, in the more unequal city we would observe that the price index channel leads to larger mobility responses due to more variation in expenditure patterns. Lastly, the strength and direction of the price index channel depends on the shock itself.[32] A shock to real consumption of local residents might cause variation in expenditure shares. As a result, this then changes how households' mobility responds to the shock. In the presence of non-homothetic preferences, the pecuniary externality generated by two-sided sorting cannot be captured by spillovers with constant elasticities. In other words, the price index effects, as in the second line of expression 20, are not constant such that they could be subsumed in $\delta_{k'n',kn}$.[33]

Empirically, changes in the relative price index of consumption and reduced-form spillovers are ex-ante not separable from observed mobility responses of a target population to exogenous shocks without information on changes in the price index or expenditure shares. In the next section I will separate both forces in the data.

7 Bringing the Model to the Data

In this section, I take the model to data from the Los Angeles Metropolitan Area. First, I describe a few parameters I take from the existing literature. Next, I estimate the key elasticities of the model with household-level microdata from the Nielsen Consumer Panel and the Consumer Expenditure Survey (CEX), firm level microdata from NETS, and Census tract level information. In particular, I empirically quantify how much relative price indices and reduced-form spillovers contributed to the mobility response of skill groups to shocks over time. Lastly, I discuss some additional pieces of information I need in order to invert the model to recover fundamentals and simulate counterfactuals.

7.1 Calibrated Parameters

7.1.1 Shopping frictions

In the model, shopping frictions capture how demand from households for establishments in distant locations falls relative to close locations. I assume that shopping frictions between locations n and n' are an increasing function of distance. Furthermore, I assume that this function is

[32] Generally, my model suggests that reduced-form spillovers are subject to the Lucas Critique as they are invariant to context and shocks.

[33] However, I have not yet formally proved this claim.

independent of local sector j and household type k and follows

$$\tau_{knn'}(j)^{1-\sigma} = \tau_{nn'}^{1-\sigma} = d_{nn'}^{\phi(1-\sigma)}, \forall j \in \{1, 2, ...J-1\}, \forall k$$

where $d_{nn'}$ is the straight line distance between the centroids of two tracts n and n'.[34] For the composite distance elasticity $\phi(1-\sigma)$ I choose a value of -1.5 within the range of estimates in the literature. For example, Couture *et al.* (2019) find values between -1.17 and -1.57 using smartphone movement data. Davis *et al.* (2019) use the location of consumers and restaurants from Yelp reviews to estimate a similar elasticity based on travel time and find values between -1 and -2.

7.1.2 Elasticity of Substitution within Sectors σ and across Sectors γ

The existing literature provides several estimates of the elasticity of substitution within service or retail sectors σ. Couture (2016) finds a value of 8.8 for restaurants, Atkin *et al.* (2018) 3.9 for retailers in Mexico, Dolfen *et al.* (2019) find 6.1 for offline stores, and Redding & Weinstein (2019) estimate a median σ to be 6.5 across disaggregated retail categories in Nielsen data. Su (2018a) reports values between 3.69 and 16 for disaggregated sectors, similar to my sector definition. As my sectors are quite aggregated and about half are retail sectors, which tend to have lower levels of substitution compared to services, I calibrate $\sigma = 5$ more towards the lower end of estimates. I will, however, report model results with higher σ as robustness.

To my knowledge, there exist fewer estimates for the elasticity of substitution across service or retail sectors γ.[35] For now, I rely on estimates from the trade literature and calibrate $\gamma = 2$ which is in the middle of estimates from Redding & Weinstein (2017) who estimate the elasticity of substitution across 4-digit NAICS sectors using trade data to be 1.36 and Hottman & Monarch (2018) who find 2.78 for HS4 sectors.

7.1.3 Skill Premium

To create differences in expenditure shares between high skilled and low skilled households in the model, I need to take a stance on the skill premium which, in turn, creates nominal income differences between skill groups. To this end, I regress log after-tax household income in the Los Angeles sample of the ACS 2014 on a dummy for high skilled household head controlling household size, age, sex, and survey year fixed effects. As reported in Table A.1, the coefficient is highly significant and implies that households with a high skilled head earn on average 70% higher nominal income compared to low skilled households. Hence, I set $\rho_{high} = 1.7$ and $\rho_{low} = 1$ in my model calibration.

[34] For internal distances I rely on Helliwell & Verdier (2001) who find that internal distances are well approximated by $distance_{nn} = .52\sqrt{area_n}$ for a square city. I use this approximation since census tracts are close to square.

[35] See Borusyak & Jaravel (2018) for a short discussion.

7.2 Estimation of Income Elasticity Parameters

7.2.1 Income Elasticities across Goods Sectors

The model requires two broad sets of income elasticity parameters. First, sector-specific income elasticity parameters ν_j govern how households reallocate expenditures across goods sectors as a function of real market consumption. Second, parameters ϵ_h and ϵ_g capture how households shift expenditures between housing and goods when they have higher market consumption. I begin by estimating Engel curves for each of the 28 local sectors and the frictionless sector using consumer expenditure data from the CEX and Nielsen. I can write equation 4 as the expenditure on sector j relative to the expenditure on a reference sector j^* for household i in location n at time t and taking logs as

$$\log\left(\frac{p_{n,t}(j)c_{i,n,t}(j)}{p_{n,t}(j^*)c_{i,n,t}(j^*)}\right) = \log\left(\frac{\alpha_{i,j,t}}{\alpha_{i,j^*,t}}\right) + (1-\gamma)\log\left(\frac{p_{n,t}(j)}{p_{n,t}(j^*)}\right) + (\nu_j - \nu_{j^*})\log U_{i,n,t},$$

where the demand shifters $\alpha_{i,j,t}$ can be household and time dependent. We can note that $\nu_j - \nu_{j^*}$ is the elasticity of relative expenditures with respect to $U_{i,n,t}$. Since I cannot directly observe real consumption $U_{i,n,t}$ in the data but nominal income $I_{i,n,t}$ is commonly reported, I can locally approximate $\log U_{i,n,t}$ with the product of $\log I_{i,n,t}$ and the elasticity or real consumption with respect to nominal income,

$$\log\left(\frac{p_{n,t}(j)c_{i,n,t}(j)}{p_{n,t}(j^*)c_{i,n,t}(j^*)}\right) = \log\left(\frac{\alpha_{i,j,t}}{\alpha_{i,j^*,t}}\right) + (1-\gamma)\log\left(\frac{p_{n,t}(j)}{p_{n,t}(j^*)}\right) + (\nu_j - \nu_{j^*})\frac{\partial \log U_{i,n,t}}{\partial \log I_{i,n,t}}\log I_{i,n,t}.$$

Furthermore, denoting the average elasticity of real consumption with respect to nominal expenditure by $\varepsilon = \frac{\partial \log U}{\partial \log I}$, I can write the regression specification

$$\log\left(\frac{p_{n,t}(j)c_{i,n,t}(j)}{p_{n,t}(j^*)c_{i,n,t}(j^*)}\right) = \iota_{n,j,t} + (\nu_j - \nu_{j^*})\varepsilon \log I_{i,n,t} + u_{i,n,j,t}, \tag{21}$$

where $u_{i,n,j,t} = \log\left(\frac{\alpha_{i,j,t}}{\alpha_{i,j^*,t}}\right) + (\nu_j - \nu_{j^*})\left(\frac{\partial \log U_{i,n,t}}{\partial \log I_{i,n,t}} - \varepsilon\right)\log I_{i,n,t}$ and $\iota_{n,j,t}$ is a location-sector-time fixed effect capturing relative prices between j and j^* in a given location n and time t.

Data: I estimate regression 21 using household-level expenditure data from three data sources covering 2012-2016.[36] For retail sectors, I take annual expenditures by sector in the consumer panel data in Nielsen. For restaurants, bars and fast food, I rely on biweekly data from the CEX diary survey. For all other sectors I use quarterly expenditures from the CEX Interview survey. Table 2 reports the source for each sector in parentheses after the sector description. Since total household expenditure is reported in neither the Nielsen consumer panel nor the CEX diary data, I proxy total expenditure by nominal annual household income reported in each data

[36] One concern could be that income elasticities are not stable over time. Aguiar & Bils (2015) discuss this issue and find that income elasticities are quite stable over time.

source. To make the estimates across the three samples comparable I choose grocery stores as the reference sector, since expenditure on groceries is consistently reported across all samples. I restrict the sample to households living in an MSA and with a household head aged between 25 and 64.

Identifcation: Like Aguiar & Bils (2015) and Comin *et al.* (2018), I include dummies for household size (≤ 2, 3-4, ≥ 5), age of household head (25-37, 38-50, 51-64), and number of earners (1, $\geq$2) that are interacted with sector dummies to account for heterogeneity in preferences across cells defined by household characteristics. I also control for sector-MSA-time fixed effects to capture differences in relative prices and aggregate preference shocks across regions, sectors, and time. Lastly, to deal with measurement error in nominal income and endogeneity concerns, I instrument nominal income with a dummy for high skill or whether the household head has at least a four-year college degree. To make progress, I assume that, conditional on controls and the instrument, the elasticity of real consumption with respect to nominal income is orthogonal to the average elasticity, $E\left[\frac{\partial \log U_{i,n,t}}{\partial \log I_{i,n,t}} - \varepsilon | X, Z\right] = 0$. Albouy *et al.* (2016), Aguiar & Bils (2015), and Hubmer (2018) have to make similar assumptions to estimate income elasticities in non-homothetic demand systems.

Results: Figure 6 shows the estimated income elasticities by sector relative to grocery expenditure with 95% confidence intervals. In addition, Table 2 reports these results in columns two and three. The ordering of the estimates is quite intuitive. Liquor stores, dollar stores, convenience stores, or fast food have the lowest income elasticities; whereas, education, recreation (gyms, sports activities), and clothing exhibit the highest elasticities. Aguiar & Bils (2015) and Hubmer (2018) reassuringly find similar orderings using slightly different sector definitions. Confidence intervals are fairly tight except for bars and legal services for which I rely on few observations in the data. To interpret the magnitude of the estimates, consider the expenditure on fast food relative to full-service restaurants; a doubling of nominal income reduces relative expenditure by around 45%. In Table A.3 and Figures A.1 and A.2 I report baseline results and two robustness checks. First, one implication of the model is that the location of residence matters for relative prices between sectors since local demand is likely correlated with the incomes of residents, potentially leading to biased estimates. Since I can observe the zip code of households in the Nielsen dataset, I can replace the sector-MSA-time fixed effect by a sector-zip code-time fixed effect to better account for local relative prices. Figure A.1 shows estimated income elasticities for retail sectors in Nielsen with zip code level fixed effects. The point estimates and the ordering are broadly similar. Another concern may be that the three samples are fundamentally different and, hence, would give different results if they all covered the same set of sectors. Some of the sectors I use to estimate the baseline results can be approximately found in another sample. For example, the CEX reports expenditure on apparel, which I can assign to apparel stores and estimate the elasticity using CEX instead of Nielsen. Columns 5 and 6 of Table A.3 and Fig-

ure A.2 show results for some sectors where the alternative source is either Nielsen or CEX Interview and are alternative to the letter in parentheses after the sector name. Again, the results are broadly similar with a few outliers. For example, the estimate for appliances/electronics is considerably smaller in the CEX, which can be due to the fact that those goods can be bought in a variety of stores such as discount or hardware stores.

7.2.2 Income Elasticities between Housing and Goods

Two theoretical insights are useful to better understand the calibration of the income elasticity parameters in the upper nest of the preference specification in expression 2. First, in Appendix B I show that all income elasticities ($\epsilon_g, \epsilon_h, \nu_j \forall j$) and the migration elasticity κ are defined up to a constant factor. Economic choices are unaffected if all elasticities are multiplied by a constant. At given prices, consumption and migration choices of households in response to higher nominal income are determined by their respective income elasticity parameter ($\epsilon_g, \epsilon_h, \nu_j \forall j$ and κ for location choice) relative to the elasticity of real consumption with respect to nominal income, which itself is a function of all income elasticity parameters. Second, with only expenditure data, goods-sector elasticity ϵ_g cannot be separately identified from the expenditure weighted sum of all sectoral elasticities ν_j. The reason for the latter is that when consumers get richer, they shift expenditure between housing and goods as a result of non-homothetic preferences in the upper nest. However, they also reallocate expenditures within the goods bundle due non-homothetic preferences that affect the price of goods relative to housing, leading to changing relative expenditures on goods and housing. To be able to pin down the values of $\epsilon_g, \epsilon_h, \nu_j \forall j$, I assume that preferences in the upper nest are homothetic, which implies

$$\epsilon_h = 1 - \eta \quad \text{and} \quad \epsilon_g = 1 - \eta.$$

An immediate implication of this assumption is that the non-homotheticity in housing and goods demand operates exclusively through the price index for goods relative to the price of housing. With homothetic preferences in the upper nest, the elasticity of real consumption with respect to nominal income collapses to

$$\frac{\partial \log U_{kn}}{\partial \log I_{kn}} = \frac{1-\eta}{\bar{\epsilon}_{kn}} = \frac{1}{1 + s_{kn}(g)\frac{\bar{\nu}_{kn}}{1-\gamma}}$$

where $\bar{\epsilon}_{kn} = s_{kn}(h)\epsilon_h + s_{kn}(g)\left(\epsilon_g + \frac{1-\eta}{1-\gamma}\bar{\nu}_{kn}\right)$ and $\bar{\nu}_{kn} = \sum_j \tilde{s}_{kn}(j)\nu_j$. My estimated sectoral income elasticities from equation 21 are relative to a reference sector (groceries) and relative to the average income elasticity ε. To recover specific values of ν_j for all sectors, I need values for the elasticity of substitution η and the difference between ϵ_h and the composite income elasticity parameter for goods, $\epsilon_g + \frac{1-\eta}{1-\gamma}\bar{\nu}$, evaluated at average expenditure shares by sector. For these last two pieces I rely on values from Albouy *et al.* (2016) who estimate non-homothetic CES preferences between housing and goods using variation in housing expenditure shares and returns

to skill across MSAs. I take their estimates with renters and owners, the average expenditure share on goods from IPUMS microdata for Los Angeles ($s_c(g) = .6663$), and citywide sales shares by sector from NETS for 2014 reported in Table 2. I calibrate $\eta = .493$ and $s_c(g)\frac{\bar{\nu}_c}{1-\gamma} = .839$.[37] The calibration implies that housing and goods are complements ($\eta < 1$) and that housing is a necessity relative to goods ($\epsilon_h < \epsilon_g + \frac{1-\eta}{1-\gamma}\bar{\nu}$).

In Table A.2, I report some reduced-form evidence, namely that the expenditure share on housing indeed falls with income (or skill). I regress the expenditure share on housing in the ACS microdata (colums 1 and 2), measured as housing expenditure out of after-tax HH-income, and CEX microdata (columns 3 and 4), measured as housing expenditure out of total expenditure, on a dummy for skilled household, a set of time-location fixed effects, and dummies for age, size, sex, and home ownership. Consistent with housing being a necessity, I find that high skilled households spend around 5ppt less on housing than low skilled households in the ACS data and 1-2ppt in the CEX data. Table 2 reports implied values of ν_j, based on $\gamma = 2$, and sector sales shares used in the calibration. Note that implied ν_j are negative; however, they follow the same ordering as the estimated relative income elasticities. Holding prices constant, the latter implies that expenditures on high-ν_j sectors increase with income relative to low-ν_j sectors. One implication of the former is that the goods price index increases with real consumption. The expenditure share on goods also increases with higher income, since goods and housing are complements. Moreover, an increase in the price index of goods consumption leads to an increase in the expenditure share on goods, everything else being equal.

7.3 Estimation of Resident Supply Elasticity κ and Reduced-Form Spillovers Elasticities

After characterizing the endogenous sorting channels of the model in section 6.2, I now empirically decompose the mobility response of households to exogenous shocks in real income into the price index channel and reduced-form spillovers. I estimate the resident supply elasticity κ, which governs how strongly households' location choices responds to spatial differences in real consumption and reduced-form spillover elasticities $\delta_{k'n',kn}$. The motivation behind the estimation is twofold. First, I can assess what portion of observed changes in spatial inequality, usually explained with reduced-form spillovers, can be attributed to endogenous local price index differences. Second, to perform policy counterfactuals in the model, I require values for κ and two sets of reduced-form spillover elasticities. "True" reduced-form elasticities net out the effect of relative price indices and "biased" elasticities that encompass endogenous price index differences.

Starting with the location choice of households in equation 8, reduced-form spillover defini-

[37] Albouy *et al.* (2016) estimate $\frac{1-\eta}{1-\gamma}\bar{\nu} = .6358$. Dividing by $1-\eta$ and multiplying by $s_c(g)$ gives this value.

tion in equation 17, and taking log changes over time t, denoted by the hats, I get

$$\log \hat{L}_{kn,t} = \kappa \log \hat{U}_{kn,t} + \log \frac{\hat{L}_{k,t}}{\hat{\Phi}_{k,t}} + \hat{\mathcal{L}}_{kn,t} + \log \hat{\bar{B}}_{kn}, \tag{22}$$

where $\hat{\mathcal{L}}_{kn,t}$ is a measure of the change in reduced-form spillovers. For example, Diamond (2016) assumes that spillovers $\hat{\mathcal{L}}_{kn,t}$ are a function of the local skill ratio $\frac{L_{high,n}}{L_{low,n}}$. Next, I replace the change in real consumption from expression 1 by its components and locally linearize the change in the price index in 7 around values of the last period,[38]

$$\log \hat{U}_{kn,t} = \log \hat{I}_{kn,t} - s_{kn,t-1}(h) \log \hat{r}_{n,t} - s_{kn,t-1}(g) \log \hat{P}_{kn,t}(g). \tag{23}$$

Equation 22 summarizes the drivers of a changing neighborhood population through the lens of the model. Locations become attractive if they offer higher real consumption $\hat{U}_{kn,t}$, stronger spillovers $\mathcal{L}_{n,t}$ or improving exogenous amenities $\hat{\bar{B}}_{kn}$. Real consumption itself can change due to nominal income, housing rents, or the local skill-specific price of goods. Previous work in the literature has made two broad assumptions on goods prices in estimating this type of regression. First, goods prices are independent of household type, and, second, either price indices are identical across locations or perfectly correlated with the local price of housing. With constant goods prices across locations 23 reduces to

$$\log \hat{U}_{kn,t} = \log \hat{I}_{kn,t} - s_{kn,t-1}(h) \log \hat{r}_{n,t} - s_{kn,t-1}(g) \log \hat{P}_{t}(g) \tag{24}$$

and under perfectly correlated local prices

$$\log \hat{U}_{kn,t} = \log \hat{I}_{kn,t} - \log \hat{r}_{n,t}. \tag{25}$$

As a result, in both specifications, the goods price index does not affect the differential location choice of skill groups. One contribution of this paper is to assess whether changes in goods price indices are by themselves driven by a changing neighborhood population or composition through the location choice of firms. Suppose, firms in income elastic sectors tend to collocate with high skilled households. Then, if there is an influx of high skilled households, relative goods prices falls more for high skilled than for low skilled households in the same location due to more varieties in sectors preferred by the high skilled. This leads to a negative correlation between changes in the price index and spillovers for high skilled, as well as a positive correlation for low skilled households.[39] Hence, omitting skill-specific goods price changes can lead to upward bias in the estimated spillover elasticities for high skilled and a downward bias for

[38] Note that I used $\epsilon_h = \epsilon_g = 1 - \eta$ in section 7.2.2

[39] In Figure A.3, I show evidence that changes in the skill-mix surrounding a tract are negatively correlated with changes in relative price indices for high and low skilled households. Since I cannot directly observe price indices, I use relative expenditure shares on goods in the top graph of the upper panel of A.3 as a sufficient statistic for relative price indices (see below). In the bottom graph, I construct relative CPIs broadly consistent with the model.

low skilled households, effectively overestimating the role of spillovers in explaining observed sorting patterns.[40]

Since I do not have access to data on expenditures across service sectors at skill-tract level and corresponding sector price indices, I cannot construct skill-location-specific goods price indices in the data.[41] Hence, I rely on the model to find a sufficient statistic for changes in price indices. I can rearrange equation 6 to solve for price index changes in the goods sector:

$$\log \hat{P}_{kn}(g) = \log \hat{I}_{kn} - \frac{\epsilon_g}{1-\eta} \log \hat{U}_{kn} + \frac{1}{1-\eta} \log \hat{s}_{kn}(g) - \frac{1}{1-\eta} \log \hat{a}_{kn}(g). \tag{26}$$

Intuitively, conditional on income and rent changes as well as constant relative tastes for goods and housing ($\hat{a}_{kn}(g) = \hat{a}_{kn}(h)$), all variation in the goods price index is captured by the expenditure share on goods. Hence, I can use variation in the expenditure share on goods as a sufficient statistic for variation in the price index of goods. Under the earlier assumption on the income elasticity parameter for housing and goods consumption ($\epsilon_g = \epsilon_h = 1 - \eta$) we can plug 26 into equation 23 and combine with equation 22 to arrive at the main regression specification,

$$\log \hat{L}_{kn,t} = \kappa \log \hat{I}_{kn,t} - \kappa \log \hat{r}_{n,t} - \kappa \frac{1}{1-\eta} \frac{s_{kn,t-1}(g)}{s_{kn,t-1}(h)} \log \hat{s}_{kn,t}(g) + \hat{\mathcal{L}}_{kn,t} + \iota_{k,t} + u_{kn,t}, \tag{27}$$

where I collect skill-specific terms in a skill-time fixed effect and the error terms capture changes in exogenous amenities and, potentially, changing tastes for goods and housing.[42] If housing and goods are complements ($\eta < 1$), an increase in local goods price index leads to an increase in the expenditure share on goods. Hence, exogenous variation in $\frac{s_{kn,t-1}(g)}{s_{kn,t-1}(h)} \log \hat{s}_{kn,t}(g)$ and a value of η identifies the resident supply elasticity κ.

Lastly, I assume a parametric form for reduced-form spillovers $\hat{\mathcal{L}}_{kn,t}$ in equation 17 in the model,

$$\hat{\mathcal{L}}_{kn,t} = \delta_k \sum_{n'} \frac{d^{\psi}_{nn'}}{\sum_{n''} d^{\psi}_{nn''}} \log \frac{\hat{L}_{high,n',t}}{\hat{L}_{low,n',t}}. \tag{28}$$

I assume that reduced-form spillovers operate on the distance-weighted skill ratio.[43] Despite being highly parametric this formulation has two advantages. First, it seems sensible to think that

[40] Only relative spillovers matter for differences in sorting between low and high skilled households. Consider the case when spillovers are positive, but identical for high and low skilled. Then, taking the difference of equation 22 between high and low skilled cancels spillovers. However, the level of spillovers matters for spatial sorting within group.

[41] Some datasets, for example Nielsen Homescanner data, provide barcode-level expenditures and prices for retail. However, such data is not available for most services.

[42] Note that I can write the term $\frac{s_{kn,t-1}(g)}{s_{kn,t-1}(h)} \log \hat{s}_{kn,t}(g) = -\log \hat{s}_{kn,t}(h)$ as the negative change in the housing share.

[43] I can write 17 as a Cobb- Douglas function of the number of residents in all locations

$$B_{kn} = \bar{B}_{kn} \mathcal{L}_{kn} = \bar{B}_{kn} \left(\prod_{n'} L^{\omega_{nn'}}_{high'n'} L^{-\omega_{nn'}}_{low'n'} \right)^{\delta_k}$$

and $\omega_{nn'} = d^{\psi}$. Assuming CRS with respect to the skill composition ($\sum_{n'} \omega_{nn'} = 1$) inside the parentheses and taking δ_k as the overall degree of spillovers returns equation 28.

spillovers operate beyond the skill composition of a tract. Second, the location of a neighborhood relative to the changing skill distribution across the city gives variation independent from the local skill composition and avoids a purely mechanical relationship between changes on the left and right-hand side of 27. With an estimate of δ_k and a value for ψ I can recover spillover elasticities $\delta_{k'n',kn}$ in the model with

$$\delta_{k'n',kn} = -\delta_{kn',kn} = \delta_k \frac{d^{\psi}_{nn'}}{\sum_{n''} d^{\psi}_{nn''}}. \tag{29}$$

Data: To estimate regression 27, I pool changes in the number of households by skill group between 1990-2000 and 2000-2014 in census tracts in LA. I exclude tracts with a population of less than 1000 in 1990 to avoid capturing newly developed neighborhoods. The key independent variable is the change in the expenditure share on goods measured as one minus the expenditure share on housing, which I impute from census tract data on the distribution of expenditure shares by income group (see Appendix A). In order to construct the proxy for endogenous amenities in 28, I choose the same distance elasticity as for the calibrated shopping frictions, $\psi = -1.5$ but also consider a higher value of -3 as robustness check. In the main specification, I control for changes in household income and residential rents as well as natural amenities such as log distance to the center of Los Angeles (City of Los Angeles City Hall), log average slope in a tract, and log population density in 1990.

Identification: To identify the resident supply elasticity κ in equation 27, I need a source of variation that shifts the price index on goods but is uncorrelated with unobserved exogenous amenity shocks and taste shocks in a neighborhood. For example, a neighborhood that receives a subway station might attract more residents due to better labor market access. However, firms also locate in this neighborhood since they too benefit from better access to workers and consumers, in turn, affecting access to services. Similarly, reductions in local crime rates might attract both households and firms. For the identification of the resident supply elasticity, I use plausibly exogenous variation in a tract's access to service establishments.

In particular, I construct a shift-share instrument based on cross-sectional variation in the share of establishments in service sectors in a location that I interact with sector growth rates in establishments from the other two large urban areas in California, namely San Francisco Bay Area and San Diego. The initial sector shares in the number of establishments contain information about supply-side characteristics of a location, such as access to suppliers, specific worker pools, or natural advantages. The idea is that a sector's growth in urban areas due to changing tastes or technological improvements will lead to a larger increase in the number of establishments in locations that provide such supply-side advantages to firms in this sector. For example, many establishments in the recreation sector concentrate along the beach since many recreation activities are related to water. We would expect that overall growth in the number of recreation establishments bring more businesses to locations by the water than inland tracts.

Conceptually, the instrument captures the idea that an exogenous increase in locally available consumption varieties lowers the goods price index in a location. If goods consumption and housing are complements, as discussed in Section 7.2.2, an increase in locally available varieties should lower the expenditure share on goods and attract more residents.

Formally, I construct the following average price index instrument,

$$P_{n,t}^{IV} = \sum_{j} \left(\sum_{n'} \frac{M_{n',t_0}(j) 1(distance_{nn'} < b)}{M_{t_0}(j)} \right) \log \hat{M}_t^O(j)$$

where b is a distance buffer, t_0 refers to a base period 1990 and superscript O stands for urban areas other than LA. Motivated by Agarwal *et al.* (2018), who report that consumers travel only short distances to consumption venues, I choose $b = 5km$, which includes an average of 56 tracts (median of 50) in the main specification but I also report additional results for a larger buffer of $10km$.[44]

The identifying assumption is that, conditional on controls, goods sector growth rates in San Francisco and San Diego are orthogonal to tract-level changes in amenities and tastes in Los Angeles, for example, changes in labor market access or crime rates. In other words, I argue that growth rates are as good as randomly assigned to sectors even though exposure of a location to sectors is endogenous. First, since the sector shares do not add to one, locations with initially more service establishments on average, such as downtown tracts, are more exposed to overall growth in services. They might also experience faster population growth due to changing preferences for such locations, specifically, by skilled households (Couture & Handbury (2017)). Thus, I control for the sum of establishment shares in each location interacted with time dummies to isolate the effect of the local composition of sectors. Moreover, I relate growth rates in San Francisco and San Diego with sector characteristics in Figure 7 and Table 3. To assess whether some sectors have grown faster than others due to shifting demand caused by overall economic growth leading to, for example, higher growth of sectors in initially richer tracts, I relate sector growth rates with estimated income elasticities. I find that sector growth rates are uncorrelated with estimated income elasticities. Similarly, if sectors that demand more high skilled workers grew faster, tracts with initially more high skilled residents could experience an increase in labor demand for high skilled workers. Again, I find that sector growth rates are uncorrelated with initial skill intensities.[45] Lastly, growth rates are negatively correlated with the initial citywide number of establishments across sectors, though the negative relationship is mostly due to few outliers (e.g. Dollar/Discount Stores are a relatively new sector).

At this point, it is instructive to draw parallels to previous work in this area. For example, Diamond (2016) identifies the labor supply elasticity (resident supply elasticity in my context)

[44] Agarwal *et al.* (2018) find that the median credit card transaction occurs at nine kilometers from home. However, their results do not differentiate between more rural areas and dense urban areas like LA where travel distances are likely to be shorter.

[45] I compute skill intensities by taking the national share of high skilled workers in each sector over the total number of workers in the ACS microdata for 1990 and 2000.

from local wage changes using shift-share labor demand shocks based on the initial industry composition and plausibly exogenous industry-level wage growth. In my analysis, I identify the corresponding elasticity from changes in local price indices with an instrument that is based on the initial sector composition of consumption varieties in a location and plausibly exogenous sector-level growth rates in varieties. In a nutshell, in both approaches the supply elasticity captures sensitivity to real incomes across locations. The difference is that I identify the resident supply elasticity from exogenous variation in price indices whereas Diamond (2016) identifies a similar elasticity from exogenous variation in nominal incomes.

In order to identify spillover elasticities δ_k, I require an additional source of exogenous variation that causes movements in the skill-mix surrounding a location. Intuitively, many shocks to exogenous amenities or tastes might lead to spatial correlated movements in the local population and the skill composition surrounding a tract. Suppose improvements in local school quality attract more high skilled residents into a cluster of neighborhoods. This amenity shock leads to correlation between changes in the skill-mix surrounding a tract in the cluster and changes in populations by skill in the tract itself. Hence, I interact the establishment shares in the average price shift-share instrument with the difference in sector expenditure shares by high and low skilled households derived from citywide income differences in 1990, estimated income elasticities of demand and citywide sales shares by sector $s_{city,t_0}(j)$ (see Table 2) according to

$$\Delta P^{IV}_{n,t} = \sum_j \left(\tilde{s}_{high,t_0}(j) - \tilde{s}_{low,t_0}(j)\right) \left(\sum_{n'} \frac{M_{n',t_0}(j) 1(distance_{nn'} < b)}{M_{t_0}(j)}\right) \log \hat{M}^O_t(j), \tag{30}$$

where

$$\tilde{s}_{k,t_0}(j) = \frac{s_{city,t_0}(j) \left(\frac{I_{k,t_0}}{I_{city,t_0}}\right)^{\nu_j}}{\sum_{j'} s_{city,t_0}(j') \left(\frac{I_{k,t_0}}{I_{city,t_0}}\right)^{\nu_{j'}}}.$$

The relative price instrument exploits differences in expenditure shares by skill group due to non-homothetic demand resulting in differential exposure of skill groups to growth in varieties across sectors. This leads to differential impacts on the goods price index, which affects the migration response of skill groups in and surrounding a tract. Lastly, I interact the relative price instrument with a skill dummy to recover type-specific spillover elasticities.

Results: Panel D of Table 4 reports the main regression results. First, I estimate regression 23 with both instruments and all controls under the assumption that goods prices are identical across the city as in equation 24 (column 2) or perfectly correlated with local rents in equation 25 (column 3). The main focus is on the estimate of $\delta_{high} - \delta_{low}$ since only the difference in spillover elasticities determines relative sorting patterns by skill group. My estimates of the relative spillover elasticities, $\delta_{high} - \delta_{low}$, are large in both specifications with values of around 1.7 and broadly in line with estimates from Diamond (2016) and Su (2018b).[46] The estimated spillover

[46]Despite using a different definition of amenity spillovers and working with MSAs instead of Census Tracts Dia-

elasticity on low-skilled households is small and insignificant. Columns 1 and 2 in panels B and C of Table 4 show that the relative price instrument pushes up skill ratios in surrounding tracts in the first stage for $\hat{\mathcal{L}}_{kn,t}$. Moreover, the coefficient on the average price instrument is also positive and significant since an average improvement in access to consumption is more valuable to high skilled households because they consume more services than the low skilled.

Next, I jointly estimate $\frac{\kappa}{1-\eta}$ and spillover elasticities in equation 27 with both instruments treating changes in price indices of goods as endogenous. In column 5, I report my main results controlling for changes in income and rents, as well as natural amenities and population density in 1990. My estimate of the resident supply elasticity $\hat{\kappa} = 2.4$, conditional on $\eta = .493$ from Albouy *et al.* (2016), is comparable albeit smaller than existing estimates of similar elasticities.[47] The first stages for $\frac{s_{kn,t-1}(g)}{s_{kn,t-1}(h)} \log \hat{s}_{kn,t}(g)$ in panel A of Table 4 confirm that the price index instrument lowers the expenditure share on goods consistent with an increase in local varieties under goods and housing being complements. In column 4, I report results for regression in 27 without controls and find similar coefficients though the first-stage F-Statistic is smaller.

I can now compare estimated reduced-form spillover elasticities by skill groups in both models to assess the contribution of relative price index differences to explaining spatial sorting of skill groups. When I treat changes in price indices as endogenous my estimates of the difference in spillover elasticitites between high and low skilled households, $\delta_{high} - \delta_{low}$, as reported in columns 4 and 5, fall significantly by 30-50% relative to the model with exogenous price indices of consumption. This finding implies that the importance of reduced-form externalities from the skill composition is overestimated if endogenous spatial variation in price indices is not taken into account. Moreover, it suggests that local price index differences are quantitatively important in explaining observed sorting patterns by skill groups and similar in magnitude compared to reduced-form spillovers.

The estimate of κ and the difference in spillover elasticities between both models are broadly consistent across several robustness checks reported in Appendix Table A.4. Changing the size of the distance buffer to 10km leads to almost identical results and to an even larger difference in $\delta_{high} - \delta_{low}$ in between models. Doubling the distance elasticity ψ in constructing the spillovers variable gives similar results, though the size of the spillover elasticity needs to be scaled down since spillovers are more concentrated locally. Using the skill distribution based on population over 25 instead of skill of household heads leaves the estimates unchanged. Next, I weight sector firm locations in constructing the average price instrument with sector sale shares and find similar results. Lastly, when I use expenditure shares on goods of renters instead of all households as sufficient statistic for price index changes, the difference in spillover elasticities between skill groups is comparable in both models; however, the average price instrument loses predictive power as shown by a low first stage F-Statistic.

mond (2016) finds .7 for low skilled and 1.9 for the difference between high and low skilled. Su (2018b) estimates spillovers for low skilled to be .45 and 1.4 for the difference using US Census Tracts.

[47] Tsivanidis (2018), Su (2018b), Diamond *et al.* (2018) and Couture *et al.* (2019) estimate migration elasticities of 3-4.

7.4 Estimation of Firm Supply Elasticity θ

The shape parameter θ of the sector productivity distribution becomes the elasticity subject to which firms substitute between neighborhoods in response to differences in profits and, hence, differences in the demand for a sector in a neighborhood. To derive the relationship underlying the estimation of θ I begin with the model expression 10 for profits of a firm ω in location n and sector j in time t, replace profits per unit of productivity $\tilde{\pi}_n(j)$ with equation 11, and taking logs to get

$$\log \pi_{n,t}(j,\omega) = \frac{\sigma-1}{\theta}\log M_{n,t}(j) + \frac{\sigma-1}{\theta}\log\frac{\Pi_t(j)}{M_t(j)} - \frac{\sigma-1}{\theta}\log A_{n,t}(j) + (\sigma-1)\log z_{n,t}(j,\omega).$$

Since firm profits are not directly observable in the data, I rely on the model relationship between firm profits and firm size or employment, $l_n(j,\omega) = \frac{1}{w}(\sigma-1)\pi_n(j,\omega)$. Replacing constant and sector-specific terms with a sector-time fixed effect I estimate,

$$\log l_{n,t}(j,\omega) = \frac{\sigma-1}{\theta}\log M_{n,t}(j) + \iota_t(j) + u_{n,t}(j,\omega). \tag{31}$$

where $u_{n,t}(j) = -\frac{\sigma-1}{\theta}\log A_{n,t}(j) + (\sigma-1)\log z_{n,t}(j,\omega)$ is an error term. Identification of $\frac{\sigma-1}{\theta}$ is based on the location choice of firms. If many firms are observed in a location n, then the location must offer high profits for firms conditional on average productivity $A_n(j)$ and idiosyncratic productivity $z_{n,t}(j,\omega)$ in sector j. Specifically, location n has high demand for varieties in j. Therefore, we would expect a firm in n to be larger than a firm with the same idiosyncratic productivity in a location with less firms and the same average productivity as in n.

Data: To estimate structural equation 31 in the data, I use the census tract and employment for private, for-profit establishments in Los Angeles in 28 local sectors over the period 1992-2014 from NETS. Although NETS has been shown to capture the spatial firm distribution well in the cross-section (Barnatchez *et al.* (2017), Neumark *et al.* (2005)) a large share of employment numbers are imputed and the data cannot capture employment dynamics well. To overcome these limitations, I, first, use only establishments with directly reported employment numbers and, second, I restrict my analysis to the cross-sectional relationship of establishment size and counts of establishments.

Identification: There are two main identification concerns. First, equation 11 directly states that $M_n(j)$ is positively correlated with $A_n(j)$ in the error term causing downward bias of $\frac{\sigma-1}{\theta}$. Second, the regression suffers from selection bias, namely, if few firms are observed in a location the idiosyncratic productivities $z_{n,t}(j,\omega)$ in the error term must be high conditional on average productivity $A_n(j)$ leading to downward bias in the estimate of $\frac{\sigma-1}{\theta}$. To address the first concern, I include a tract-time fixed effect that captures location supply shocks like availability of retail space, labor market access, or commercial rents. Furthermore, I instrument the log number of

establishments $\log M_n(j)$ with the log average slope in a location interacted with my estimates of the income elasticity of demand by sector, $\hat{\nu}_j$. The slope of a location is a very strong predictor of household income or skill composition in Los Angeles because households seem to prefer living in steeper locations, such as locations with a better view. Regressing log average household income or log skill ratio in a location on the log slope yields an R^2 of over 20% and highly significant positive coefficients. Conditional on tract-time and sector-time fixed effects the instrument picks up differential exposure of sectors to higher household income due to higher slope, a natural amenity highly valued by households. Since $\hat{\nu}_j$ are estimated parameters that are re-scaled as described in the previous section, I construct an alternative instrument replacing the value of $\hat{\nu}_j$ with the rank of ν_j across all sectors.

I address selection bias by comparing the size of similar establishments across locations. In particular, I assume that establishments ω^m of multi-establishment firms or chains m have a common productivity component independent of location,

$$z^m_{n,t}(j, \omega^m) = z^m_t(j, \omega^m).$$

The literature provides evidence that retail chains follow uniform pricing across stores (DellaVigna & Gentzkow (2019)), which is consistent with this assumption. By restricting the sample to chain establishments, defined as having the same headquarter in the NETS data I can include a chain-time fixed effect and identify $\frac{\sigma-1}{\theta}$ only with variation across locations serviced by the same chain.[48] Conditional on the assumption of common productivity, variation in size of establishments within chain across locations is either due to local demand, consequently due to variation in the number of establishments, or average sector productivity differences. By instrumenting the number of establishments in a location with the slope instrument, I isolate how demand differences affect relative sizes of establishments within the same chain.

Results: Panel A of Table 5 reports first stage results of regressing the log number of establishments on log average slope interacted with ν_j. Across all specifications I find a strong positive relationship between the instrument and the number of establishments in an area indicating that establishments in sectors with high income elasticity locate in tracts with high slope and, consequently, close to high income residents. Panel B shows the second stage estimates of 31 and the main estimate is reported in column 1. Doubling the number of establishments in a location implies that establishments are 25% larger on average. This estimate informs the baseline value of $\frac{\theta}{\sigma-1} = 4$ in my model calibration. With my assumed elasticity of substitution within sectors, $\sigma = 5$, the implied supply elasticity θ takes a value of 16. When I split the sample into retail and service sectors in columns 3 and 4, I find similar point estimates; although, the point estimate for retail sectors becomes insignificant. Results with the rank-based instrument are reported in column 4 and imply a slightly larger value of θ.

[48] Some multi-establishment firms operate in several sectors in the NETS data. I restrict each chain to the sector with most establishments in each year.

Table 6 summarizes all parameters of the model. I set up two different versions. In the model with endogenous relative price index differences, my baseline specification denoted with superscript 1, shopping frictions restrict access to consumption establishments. In the second calibration without endogenous price effects, akin to the previous literature, households can consume everywhere in the city at no cost; however, high skilled households are subject to higher reduced-form spillovers. Since the point estimate of the spillover elasticity for the low skilled is indistinguishable from zero in my estimation, I set $\delta_{low} = 0$ in both calibrations. In the baseline calibration, I use the estimated spillover elasticity of δ_{high}, reported in column 5 of Table 4 (panel D). For the model without relative price index effects, the corresponding estimates of δ_{high} in columns 2 or 3 fall into a parameter range for which the equilibrium is not necessarily unique. Specifically, if δ_{high} exceeds 1.3 (given all other parameters), the initial equilibrium in the data becomes unstable.[49] To ensure the model without price index effects has a unique solution, I choose a value of $\delta_{high} = 1.25$, just below this cutoff.

Equipped with the full set of model elasticities, I can invert the model to recover the fundamentals of the economy such as fixed amenities and sector productivities. Then, I can perform model-based counterfactuals of different policy shocks to assess the implications of allowing for two-sided sorting for household welfare and sorting.

7.5 Model Inversion

To be able to perform counterfactuals in the model, I require location-specific exogenous amenities by skill, $\bar{B}_{kn}$ and sector-location productivities $A_n(j)$ as well as fixed cost of entry by sector and sector demand shifters. My model falls into the set of quantitative urban economics models (e.g. Tsivanidis (2018), Monte *et al.* (2018), and Ahlfeldt *et al.* (2015)) that are fully saturated with structural residuals or "fundamentals", which cover all variation in the data unexplained by the inherent model structure. Thus, with sufficient data moments I can invert the model to recover the set of residuals as stated in the next proposition.

Proposition 2. *Given data on residents by skill and location, L_{kn}, number of firms by sector and location, $M_n(j)$, citywide revenue shares by sector, $rs_c(j)$, citywide expenditure share on goods, $s_c(g)$ and a normalized wage per unit of labor $w = 1$ there exist unique vectors of model fundamentals, namely, exogenous amenities $\bar{B}_{kn}$, composite demand and productivity shifters $\bar{A}_n(j) = A_n(j) a_g^{\theta\frac{1-\sigma}{1-\eta}} \alpha_j^{\theta\frac{1-\sigma}{1-\gamma}}$, fixed entry costs by sector $f^e(j)$ and transfers T_k that replicate the observed equilibrium in the data.*

The process follows several steps. First, using the assumption that all housing expendi-

[49] If I slightly perturb the observed household and firm distributions as starting values in the computation, the model fails to converge to the observed equilibrium. Instead, the model finds alternative configurations of the city. When I simulate counterfactuals in the model, this multiplicity makes it difficult to separate the effects of a policy from such an alternative equilibrium. I want to emphasize that the multiplicity of equilibria in an urban context is an interesting area of research; however, it is beyond the scope of this paper.

tures are redistributed to households according to the labor endowment ρ_k, normalized citywide wages transfers are characterized by

$$T_k = \rho_k \frac{1 - s_c(g)}{s_c(g)}.$$

With values for transfers and wage equalization due to free labor mobility, I can directly compute nominal income $I_k = \rho_k w + T_k$. Combining the fact that the city is closed so citywide expenditure on goods need to equal citywide income and the free entry condition of firms implies

$$f^e(j) = \frac{1}{\sigma} \frac{r s_c(j) \sum_k \rho_k w L_k}{M(j)}.$$

Next, rearranging equation 11 and combining with free entry in 15 gives a set of conditions that allows me to recover the set of demand-productivity composites $\mathcal{A} = \{\bar{A}_n(j)\}_{n,j}$ for all sector and locations, namely,

$$\begin{aligned} \bar{A}_n(j) &= \xi_n(j) \left(\sum_{n'} \bar{A}_{n'}(j) \check{\pi}_{n'}(j)^{\theta} \right) \check{\pi}_n(j)^{-\theta} && \forall j, n \\ \sum_{n'} \bar{A}_{n'}(j) \check{\pi}_{n'}(j)^{\theta} &= (f^e(j))^{\theta} && \forall j \end{aligned}$$

where

$$\check{\pi}_n(j) = c_1 \sum_{n'} \sum_{k'} \tau_{n'n}^{-1} \tilde{p}_{n'}(j)^{\frac{\sigma-\gamma}{1-\gamma}} \tilde{P}_{k'n'}(g)^{\frac{\gamma-\eta}{1-\eta}} P_{k'n'}^{-\epsilon_g - \nu_j} I_{k'}^{\eta+\epsilon_g+\nu_j} L_{k'n'}$$

and $\tilde{p}_n(j) = a_g^{\frac{1-\gamma}{1-\eta}} \alpha_j p_n(j)^{1-\gamma}$, $\tilde{P}_{kn}(g) = a_g P_{kn}^{1-\eta}$ and c_1 is a constant. Since all prices are themselves functions of $\mathcal{A}$ and moments in the data, the system maps the vector of fundamentals $\mathcal{A}$ onto itself. Intuitively, total profits by sector in the economy are given by the fixed cost of entry, hence, bounding and normalizing the set of possible parameters. A formal proof is still work in progress.

In the previous step of the inversion, I recover price indices P_{kn} and nominal income I_k from earlier, hence, real consumption by skill group and location, U_{kn}. With this information and an appropriate normalization, I can rearrange equation 8 to solve for the unique set of exogenous amenities $\bar{B}_{kn}$, similar to the previous step.

7.6 Model Fit

After estimating key parameters and recovering the fundamentals of the economy, I can exactly recover the moments used in the inversion and compare other moments produced by the simulated model to non-targeted moments in the data. The left panel of Figure 8 plots residential rents as simulated in the baseline calibration with local price effects against residential rent per square foot computed as total housing expenditure divided by housing stock in the ACS 2014. The two series are highly correlated (Correlation=.58); although, the model finds a larger spa-

tial dispersion in rents. The right panel plots model based rent per square foot against median rental prices per square foot from Zillow's zip code level data from 2014-2016. Again the series are quite correlated with a value of .43.

In Figure 9, I compare expenditure shares on goods from ACS used in the estimation earlier against model predicted expenditure shares for low skilled (left) and high skilled (right). With a correlation of .46 the model performs quite well in replicating the spatial distribution of non-targeted expenditure shares. In columns 5 and 6 of Table A.2, I report results from regressing the housing expenditure share in tract-level and model-produced data on a skill dummy and tract-FX. Although not specifically targeted, the model reproduces the non-homotheticity in housing demand, if anything understates it compared to the data. Lastly, I can compare the model's prediction for the ratio of expenditure shares between high and low skilled. The model has problems capturing the large spatial variation in the data; however, the model understates the differences between high and low skilled implying that my calibration is conservative. Reassuringly, correlations with the data are consistently smaller when the model is simulated without relative price index effects but larger external spillovers.

8 Policy Counterfactuals

With the estimated model I can now assess how urban policies interact with the two different sorting forces in the model, relative price indices and reduced-form spillovers. My counterfactual analysis tries to answer two questions. First, through the lens of my model, what are the effects of two real world urban policies, Opportunity Zones and social housing, in terms of mobility of skill groups, inequality, and aggregate welfare in Los Angeles? Second, do we miss important details of these policies if we treat the endogenous response of households (and firms) to changing neighborhoods as a reduced-form spillover as opposed to changing costs of living? To answer these questions, I perform two counterfactuals in my model.

First, I shock the spatial distribution of firms by simulating the effect of a new tax incentive to invest in a subset of locations in LA to investigate the endogenous response of skill groups to changing price indices of consumption. I do the reverse in the second counterfactual. I shock the distribution of households by adding the existing stock of Social Housing projects to the city from an initial counterfactual equilibrium without such projects in order to understand the general equilibrium reaction of firms. For each counterfactual, I recover the fundamentals of the economy (exogenous amenities and productivities) using the same moments in the census tract data for Los Angeles in 2014, but assume that sorting of skill groups is driven either by relative price indices and "true" estimated reduced-form spillovers (Calibration 1) or only "biased" reduced-form spillovers and no shopping frictions (Calibration 2).[50] Then, I simulate the same policy shocks in both versions of the model fitted to the same observed initial equilibrium.

[50] To further trace out how the predictions of my policy counterfactuals change, I also present results removing and adding other features of the model (non-homothetic preferences, spatial frictions or reduced-for spillovers).

8.1 Opportunity Zone Program

In 2017, U.S. Congress passed the Tax Cuts and Jobs Act and among several tax-related policies, created a new place-based tax credit, the so-called Opportunity Zones (OZ). The stated goal of the program is to lift living standards in economically disadvantaged urban areas by incentivizing businesses and investors to invest unrealized capital gains in designated opportunity zones.[51] U.S. Treasury Secretary Steven Mnuchin recently stated that he expects the total investment in Opportunity Zones to exceed $100B in 2019. According to Theodos *et al.* (2018) out of a total of 42,176 eligible census tracts 8,762 were designated Opportunity Zone status.[52] My Los Angeles sample includes 257 of these tracts. Figure 11 shows all 257 OZ census tracts in Los Angeles. Designated census tracts are concentrated in the center of Los Angeles with some scattered zones in the periphery. Table A.5 reports means and differences for OZ and non-OZ tracts. As expected, OZs are populated by lower-income, less educated households and host around 20% less firms in sectors with high income elasticity relative to non-OZ tracts.

The Economic Innovation Group (2019) estimates that due to the tax benefits OZ investments offer, there should be excess returns of around 30% over 10 years. Based on this estimate, I implement a 30% subsidy on variable profits $\tilde{\pi}_{OZ}(j)$ of firms in all sectors that operate in OZ tracts in the model and assume that the subsidy is financed by the city government that due to the policy redistributes a smaller transfer to all households.[53] Then I simulate the model to predict how mobile firms respond to exogenous profit subsidies in OZs and evaluate the endogenous effect on location choices of skill groups in response to the change in access to consumption varieties in general equilibrium.

Before moving on to the results, I want to discuss a few limitations of this policy counterfactual. First, since labor is freely mobile in the model, my counterfactual cannot speak to the local labor market effects of place-based tax incentives, an interesting and well-studied question. Instead, I focus on isolating the demand-side effects on household composition of attracting more consumption varieties into disadvantaged neighborhoods. Consistent with this notion, Reynolds & Rohlin (2015) argue in recent work on Federal Empowerment Zones, a broadly similar place-based policy, that firm-level incentives made targeted neighborhoods more attractive to high-income, well-educated households. They fail to explain this finding with employment effects on the initial resident population. Secondly, there have been concerns in the media that Opportunity Zones predominantly lead to more investment in high-end real estate, a channel currently absent from the model since I treat the residential housing stock as fixed. However, a recent investor survey by KPMG reports that 39% of potential investors plan on operating a business in OZs and quotes by investment fund managers indicate that developing amenities

[51] The tax cut has three parts. First, capital gains taxes from previous investments can be deferred until 2026 when reinvested in OZs. Second, the tax base of previous investments increases up to 15% depending on the duration of the OZ investment. Lastly and most importantly, after 10 years all gains from OZ investment are excluded from taxation.

[52] Eligibility is based on high poverty rate and low family income, (see Theodos *et al.* (2018) for details). State Governors propose the final selection of OZs among the eligible tracts to the U.S. Treasury for approval.

[53] Although the tax benefits are ultimately financed by all U.S. tax payers (current and future) I implicitly assume that the share of LA tax payments in the total tax bill of the reform is equal to LA's share in the population.

like retail venues is crucial for creating value in OZs and returns to business investments might be much more profitable than real estate.[54] Lastly, since the Opportunity Zone program has been implemented very recently and the tax bill requires little to no reporting on take-up or costs, I cannot validate any of my counterfactual outcomes using moments in the data.[55] With this caveat in mind I view the counterfactual as a prediction of the policy's effects.

Table 7 summarizes the main results for the calibration with endogenous price effects. In reaction to the profit subsidy, firms move into OZs such that the number of varieties increases by 80% percent, around one standard deviation of the overall variation in the number of firms across tracts.[56] However, the reallocation of economic activity is at the expense of nearby tracts as shown in Figure 13, reducing the local effect on access to varieties. Price indices of consumption fall for low and high skilled households but more so for the latter, leading to almost twice as many high skilled households relocating to OZs due to firm subsidies. The Opportunity Zones program induces gentrification in these disadvantaged areas due to demand-driven effects. There are three reasons for the stronger response of high-skilled households to the policy. First, since OZs are initially populated by relatively few high skilled households, the policy makes OZs more attractive to this group relative to other areas and this effect is larger than for the low skilled.[57] Second, high skilled households value consumption of goods more than low skilled because goods demand is more elastic than housing; hence, a fall in prices benefits high skilled more. Lastly, more firms in sectors with high income elasticity enter endogenously, as shown in Table A.8, further lowering the price index for high skilled more than for low skilled households. Finally, high skilled households are subject to positive reduced-form spillovers. Both channels amplify gentrification of OZs.

As indicated by the modest R^2 in Table 7, effects of the policy are not limited to OZ tracts. Non-OZ tracts are affected by the reallocation of firms into OZs because of two features of the model. Figure 17 plots the movements into Non-OZ tracts as a function of the share of shopping that occurs in OZs initially. At around 25% of consumption spending in OZs, the effect of the policy on non-OZ locations is the same as for OZ locations, showing a similar difference in mobility of low and high skilled households. Moreover, I set up reduced-form spillovers to be a function of the location of a tract. Hence, high skilled households in non-OZs close to targeted areas receive spillovers from the gentrification of OZs. Figures A.4 and A.5 show that changes in the price index of goods consumption and reduced-form spillovers for high skilled households are highly correlated, but they are not identical and not limited to OZs.

[54] *"..the firm plans to cluster investments in individual neighborhoods to create a critical mass of amenities, such as housing and grocery stores, that can increase property values in the area."*, Garett Bjorkman, CIM Group CEO; *"The returns on investing in a high potential company that sets up as a qualified opportunity zone business [...] could be 10 times more profitable than flipping commercial real estate."*, Brian Phillips, Founder Pearl Fund

[55] In future work I plan to validate effects based on similar past policies, like Empowerment Zones.

[56] There is also entry of firms into the city due to the policy of around 4.5% on average in all local sectors.

[57] This effect is present even without non-homothetic preferences and is due to idiosyncratic preferences for locations. As the policy targets low skilled neighborhoods expected utility falls more (or increases less) for high skilled, increasing the relative attractiveness between the targeted areas and the average larger, hence stronger movements into these areas.

In comparison, Table 8 reports the mobility effects of the same policy on households and firms for the model with no shopping frictions but larger reduced-form spillovers. This version of the model predicts almost no movement of households, since location of consumption venues has no bearing on the price index of consumption if shopping is frictionless.[58] Despite being a stark example, the counterfactual emphasizes that endogenous price index effects as opposed to reduced-form spillovers, are not equivalent drivers of spatial inequality in a counterfactual sense and can lead to first-order differences in the outcomes of a common urban policy in terms of inequality or welfare.

Table 9 summarizes spatial inequality and the welfare effect of Opportunity Zones on low and high skilled residents of LA for both calibrations and four other versions, two calibrations without spillovers and two calibrations with homothetic preferences. Spatial inequality effects are strongest in the baseline model and mostly driven by price index effects when compared to the full model without spillovers. In the models with homothetic preferences, the policy does not affect the skill ratio in OZ and Non-OZ differently such that both groups locate in OZ but in equal proportions. Lastly, I turn to welfare effects, measured as the compensating variation on expected utility by each group. The policy leads to modest welfare losses of both groups in all calibrations. The benefits of increased variety in OZs cannot compensate for the lump-sum taxes imposed by the city to finance the subsidy. In the baseline calibration, losses are slightly smaller for the high skilled because the policy leads to stronger within-group spillovers from centrally located OZs. Welfare losses are lowest in the baseline model because Opportunity Zones are targeting within-group populations with high marginal utility as the areas lack access to consumption varieties; hence, they face high consumption prices in the initial equilibrium.[59] In a model with only reduced-form spillovers, we would miss two crucial effects of Opportunity Zones. First, the reallocation of firms in the city affects relative price indices of consumption resulting in reallocation of skill groups; specifically, the influx of high skilled households leads to gentrification of OZs. Second, welfare effects of the policy are different since Opportunity Zones are targeting specific locations that are disproportionately benefiting from more consumption varieties.

8.2 Social Housing in Los Angeles

Around 100 thousand out of 3.3 million housing units in Los Angeles receive state or federal housing assistance either in the form of direct housing projects, loans or tax credits for low-income households.[60] For this counterfactual, I rely on data from the California Housing Part-

[58] The small movements into OZs in calibration 2 can be explained as follows: Overall welfare falls due to the policy making initially poor OZs more suited to the marginal resident in both groups.

[59] To a first order, the difference between homothetic and non-homothetic preferences can be explained by average marginal utility in the first four models which I calibrate around $.55$ as compared to 1 with homothetic preferences.

[60] Source: California Housing Partnership Preservation Database, June 2019. For more information, please visit chpc.net/policy-research/preservation/ ; The California Housing Partnership has provided address-level savings from social housing. The data covers housing assistance from HUD (Project-based Section 8, Project Rental Assistance Contract, Section 202 Direct Loans, Insurance Programs), Low Income Housing Tax Credit and USDA. How-

nership, an affordable housing think-tank, that provides address-level rent savings by social housing project. Since I cannot distinguish who receives the assistance in the data, but it is common in these policies that affordable housing eligibility is based on income, I aggregate rent savings to the census tract level and assign all savings to the housing expenditure of low skilled households that I observe in the ACS data. Figure 12 shows the distribution of social housing units as the share of housing costs of low skilled households covered by rent savings. Social housing in LA is fairly spread out over the city with a higher concentration in the dense center of LA. Since social housing projects tend to be large with several hundred units in small areas, intensity in terms of covered share of expenditure by low skilled households is concentrated in a few tracts.

I implement social housing in the model as subsidy on rents for low skilled households equivalent to the share of rent savings in housing expenditure (shown in Figure 12) and financed by the city government budget. First, I invert the model from the observed equilibrium assuming that social housing is present in the current equilibrium and then remove all social housing in a counterfactual. Hence, I report all outcomes going from the counterfactual to the observed equilibrium to assess the effects of the current state of social housing in LA.

In Tables 10 and 11, I report changes in population by skill and number of firms as a function of the share of rent savings for the baseline calibration and the calibration with only reduce-form spillovers, respectively.[61] In all calibrations social housing leads to higher market rents in targeted tracts. This causes high skilled households to move out because they face the market rent as opposed to the low skilled who locate in targeted areas due to subsidized rent. The skill ratio in targeted areas at mean subsidy level (3.25%) falls by around 3.5%. Figures 15 and 16 show the reallocation of firms and high skilled households away from the center of LA towards the periphery of the city. In the calibration with price effects, firms, on average, leave areas with social housing leading to an increase in price indices for both groups. However, firms in sectors with low income elasticity leave targeted areas much less than firms in income-elastic sectors as shown in Figure 18, endogenously leading to stronger price index changes for the high skilled. Responses by firms to the subsidy are quantitatively limited because of two opposing forces on local firm profits and, consequently, firm mobility. First, low-skilled households become richer in targeted areas due to the subsidy, shifting demand to goods consumption in general and towards income-elastic sectors. Second, the remaining high skilled households behave more similar to poorer households due to high market rents. The first increases profits in areas with social housing and more so in income-elastic sectors, the reverse holds for the second.

Table 12 shows that the effects on the skill ratio are broadly similar in both calibrations. It is instructive to consider the effects of the policy on neighborhood composition in the models

ever, the database does not include other state or local programs. For more detail, see https://chpc.net/affordable-housing-benefits-map/

[61] The numbers are reflecting percent changes in number of firms and HHs if a tract has 100% subsidy on rent due to social housing for the low skilled. In fact, the mean subsidy conditional on hosting any social housing is 3.25% with 19% at the 99th percentile.

without spillovers, shown in columns 3 and 4 of Table 12. The model without spillovers and price effects describes these effects, absent any endogenous amplification; whereas the model with price index effects traces out the amplification of the policy due to endogenous changes in consumption access in response to the outflow of high skilled households.[62] The firm response reinforces the shift towards less skilled neighborhoods and slightly increases the welfare loss for the high skilled. On the one hand, this welfare loss is small since high skilled households in these areas are now poorer and behave more similar to a low skilled household. On the other hand, spillovers amplify welfare losses for the high skilled since the areas with rent subsidies, which are relatively central in LA, provide less amenity spillovers to the rest of the city. Hence, the model with large spillovers overstates the welfare loss of the high skilled and understates the welfare gain of the low skilled compared to the baseline model.

9 Conclusion

Spatial inequality in cities has sparked interest by the public and policy makers. To inform urban policies, it is important to understand its sources, in particular, how the composition of local residents endogenously shapes the attractiveness of a neighborhood. This paper studies how two-sided sorting of heterogeneous households and firms generates pecuniary externalities that amplify clustering of household groups in cities. As a benchmark, I compare these forces to reduced-form spillovers that have been studied in previous literature. First, I develop a quantitative general equilibrium model of the city that features two-sided sorting of skill groups and firms in various local consumption sectors but nests previous work. Second, I combine the model with rich administrative microdata from Los Angeles to quantify the contributions of these pecuniary externalities and reduced-form spillovers to observed sorting patterns in the data. Third, I assess the implications of urban policies when allowing for two-sided sorting.

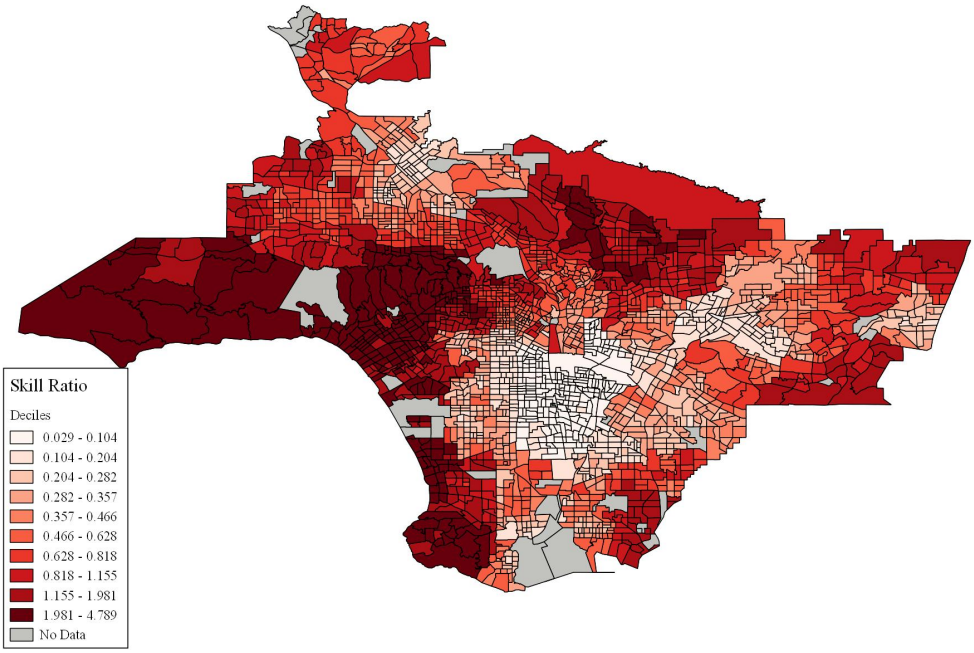
I find that two-sided sorting is an important driver of spatial inequality in Los Angeles. Spatial variation in local price indices reduces the estimates of reduced-form spillover elasticities by 30-50%. In the first policy counterfactual, I show that subsidizing firm entry in specific locations leads to heterogeneous location choices of skill groups and welfare due to differential changes in relative price indices of consumption. In response to rent subsidies that target specific locations and groups, firms also relocate thereby amplifying the sorting of households and welfare effects of the policy. In a model that relies on reduced-form amenity spillovers to generate strong sorting patterns, these effects are absent since household and firm location choices are independent. In addition, such a model overstates the welfare losses (or understates the gains) of the urban policies. My results suggest that demand linkages between different household groups as well as across neighborhoods are important determinants of spatial inequality in cities and have profound implications for our understanding of urban policies.

[62] Price index channel and reduced-form spillovers reinforce each other. Hence, a comparison of column 3 and 4 in Table 12 does not fully capture the total amplification from the price index channel.

Figures and Tables

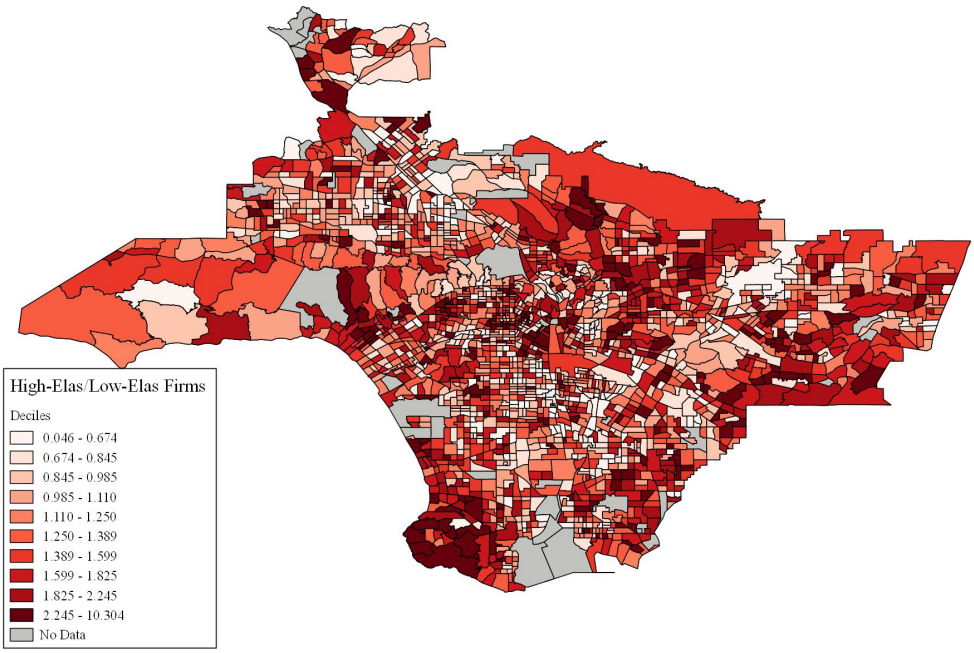
Figures

Figure 1: Spatial inequality as measured by skill ratio in Los Angeles, 2014



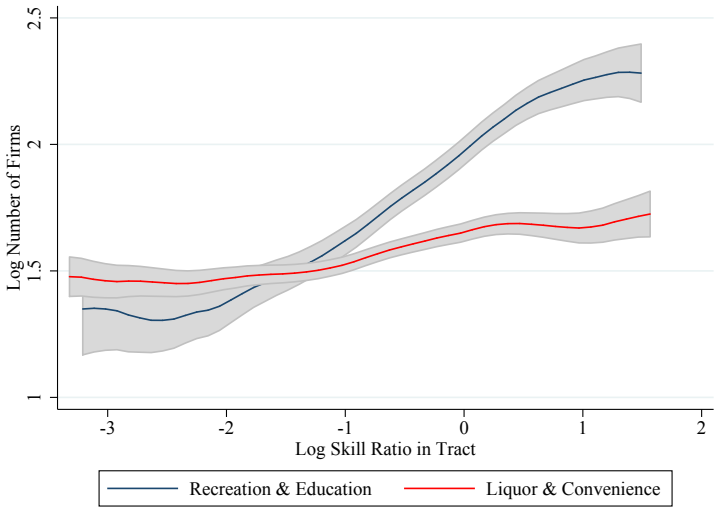
The figure plots the skill ratio in each census tract in Los Angeles, ACS 2014.

Figure 2: Distribution of establishments by income elasticity in Los Angeles, 2014



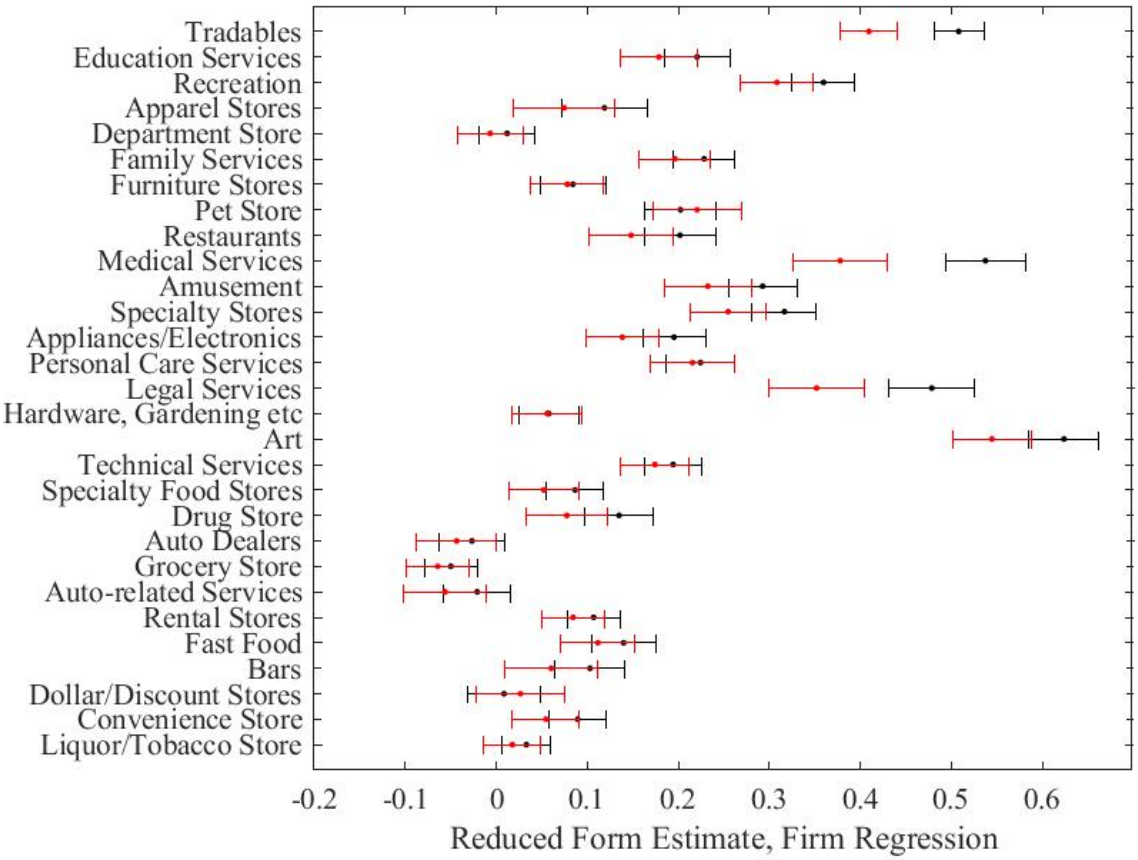
The figure plots the ratio of establishment counts in a census tract with income elasticity above median over establishment counts below median, NETS.

Figure 3: Number of establishments in Recreation & Education vs Liquor & Convenience Stores, 2014



The figure plots the log number of Recreation & Education and Liquor & Convenience Stores against the local skill ratio, NETS and ACS 2014.

Figure 4: Number of establishments in all sectors, 2014



The figure plots coefficients and 95% CI from sector-level regressions of log number of firms in tract against log local skill ratio. Regressions without controls in BLACK and with controls for log population density, ratio of skilled over unskilled employment and total employment in tract in RED. Without controls: Spearman Rank Correlation (p-value): $.495 \quad (.007)$, with controls: Spearman Rank Correlation (p-value): $.505 \quad (.006)$. Data from NETS, ACS 2014 and LODES.

Figure 5: Graphical example of sorting patterns in model

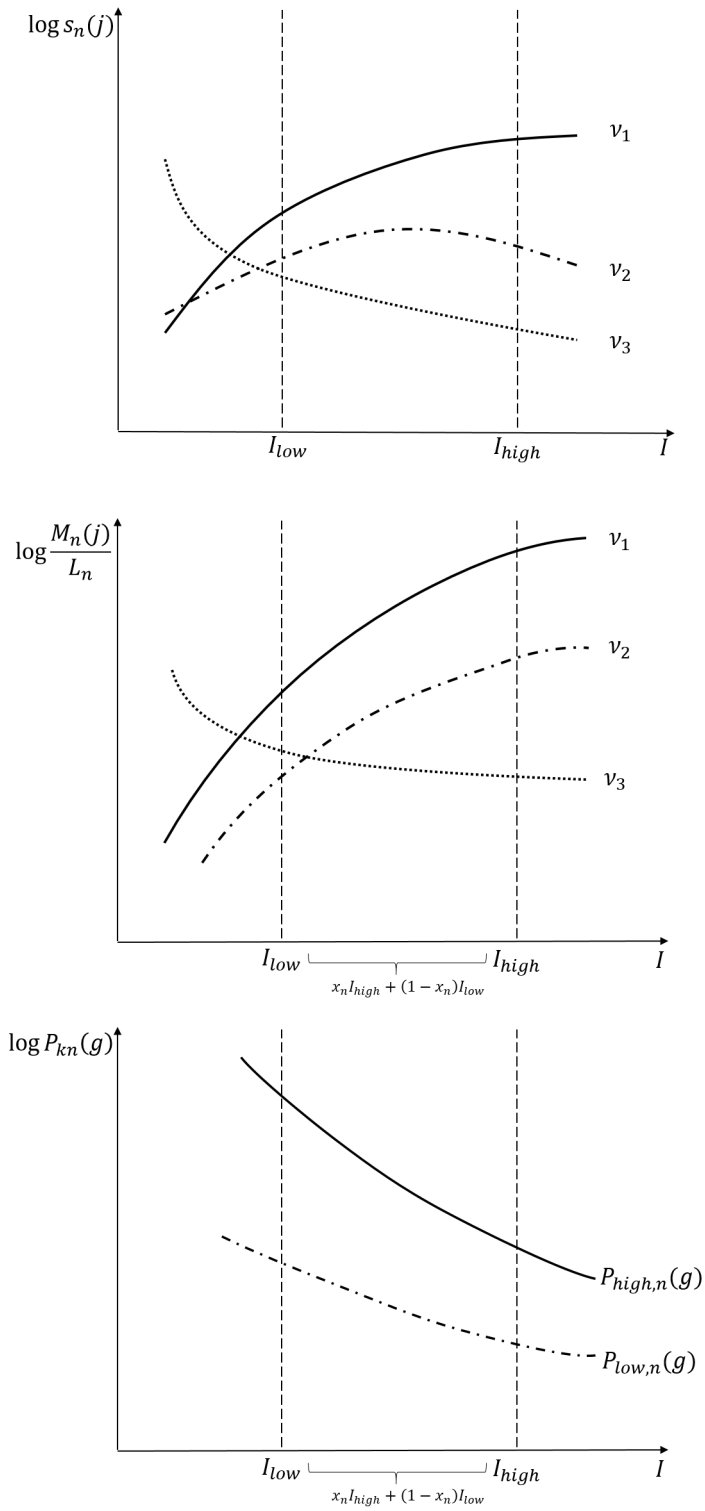
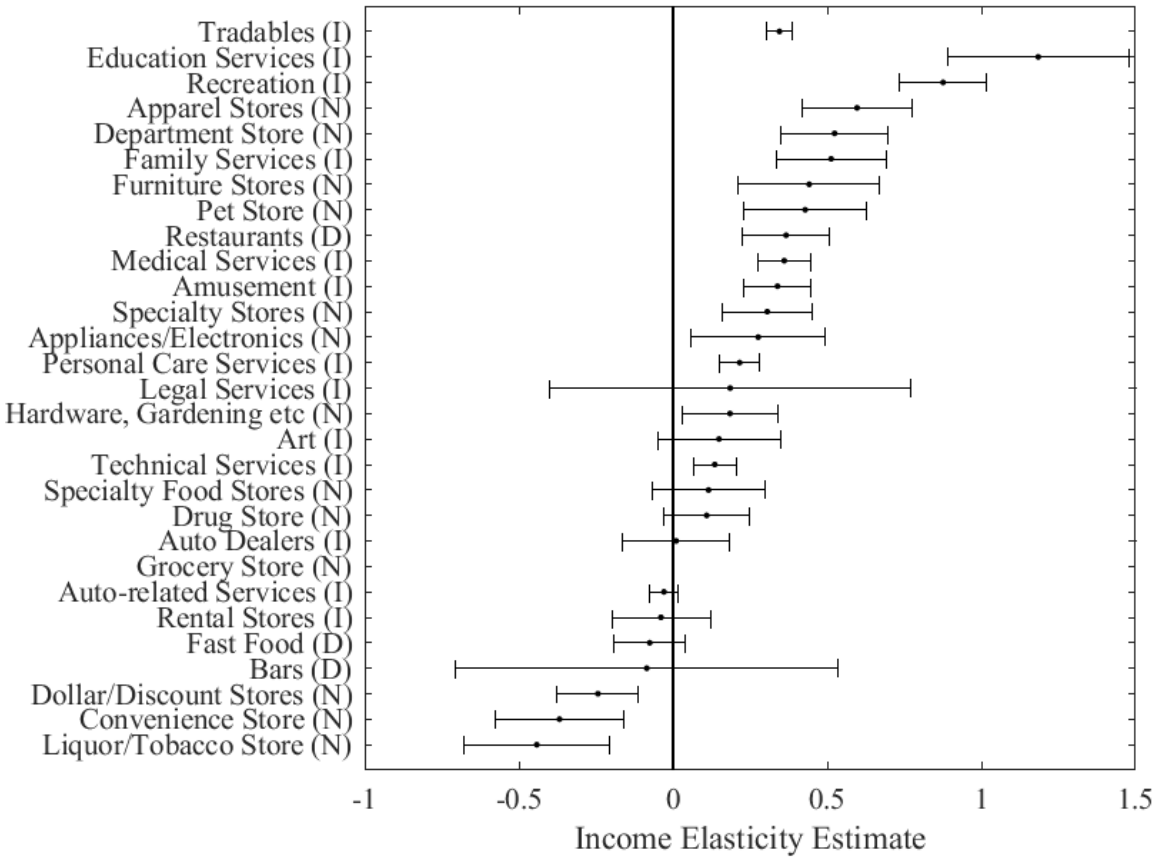
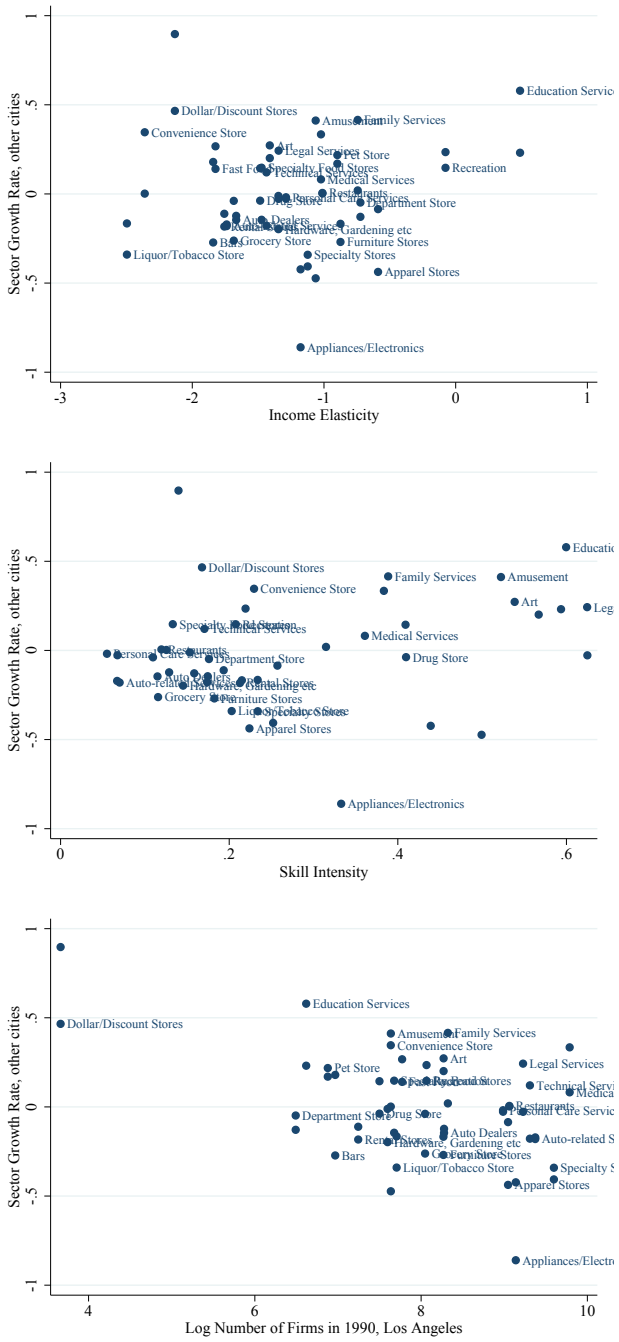


Figure 6: Income Elasticities by Sector



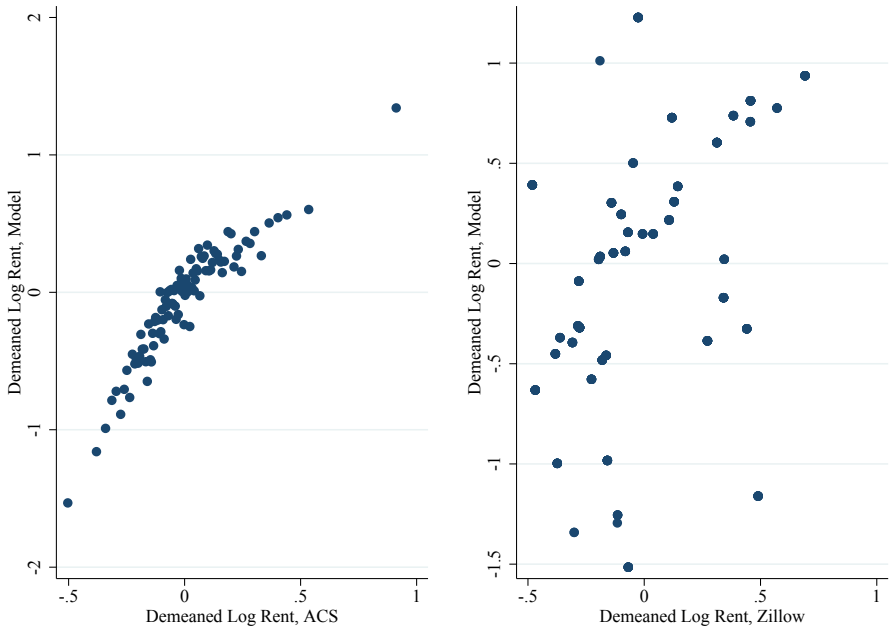
The figure plots estimated point estimates of income elasticities by goods sector using consumer expenditure data and 95% confidence intervals. Data source in parentheses (N=Nielsen, I=CEX Interview, D=CEX Diary).

Figure 7: Correlations of Local Sector Growth Rates in Price-Bartik



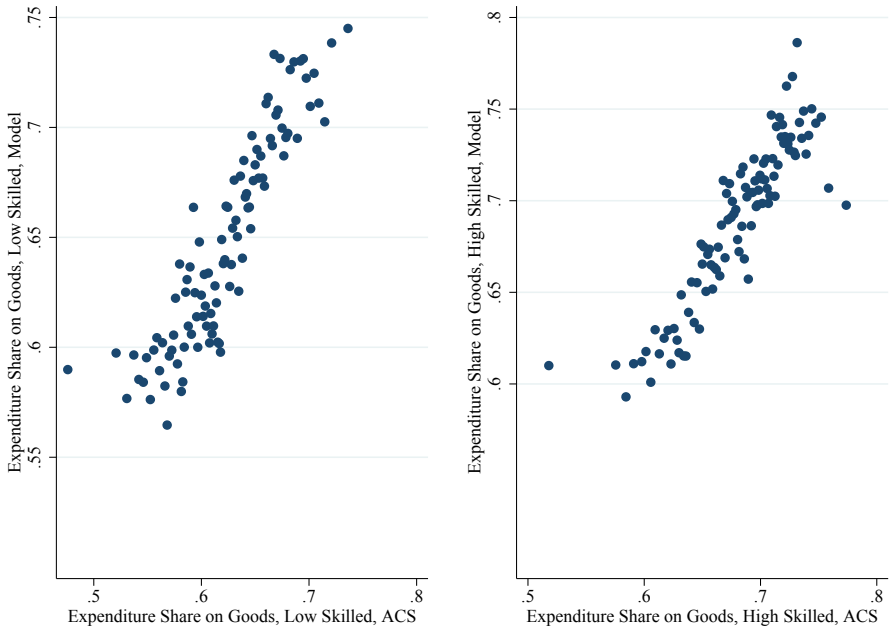
The figure plots sector growth rates in number of establishments from San Francisco Bay Area and San Diego for 1990-2000 and 2000-2014 and income elasticities (top), skill intensities by sector as skilled employment over total employment in sector nation-wide from Census 1990 and 2000 (middle) and initial log number of establishments in LA in 1990 (bottom).

Figure 8: Model Fit: Rents



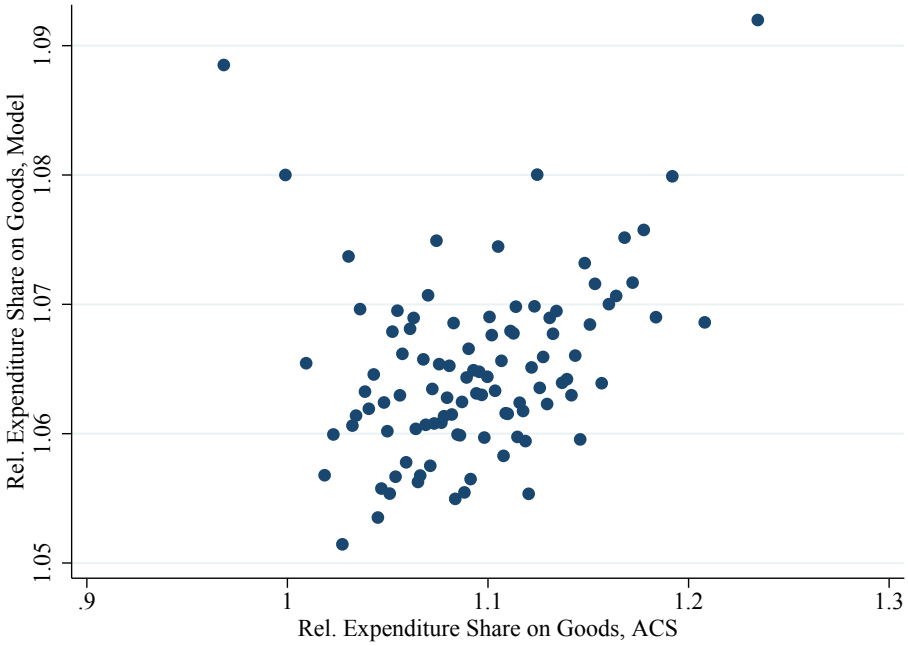
Left Panel: Log Rent by census tract in baseline model with price effects to data from the ACS 2014, binscatter with 100 bins, $Corr = .58$; Right Panel: Log Median Residential Rents from Zillow, $Corr = .43$

Figure 9: Model Fit: Expenditure Share on Goods



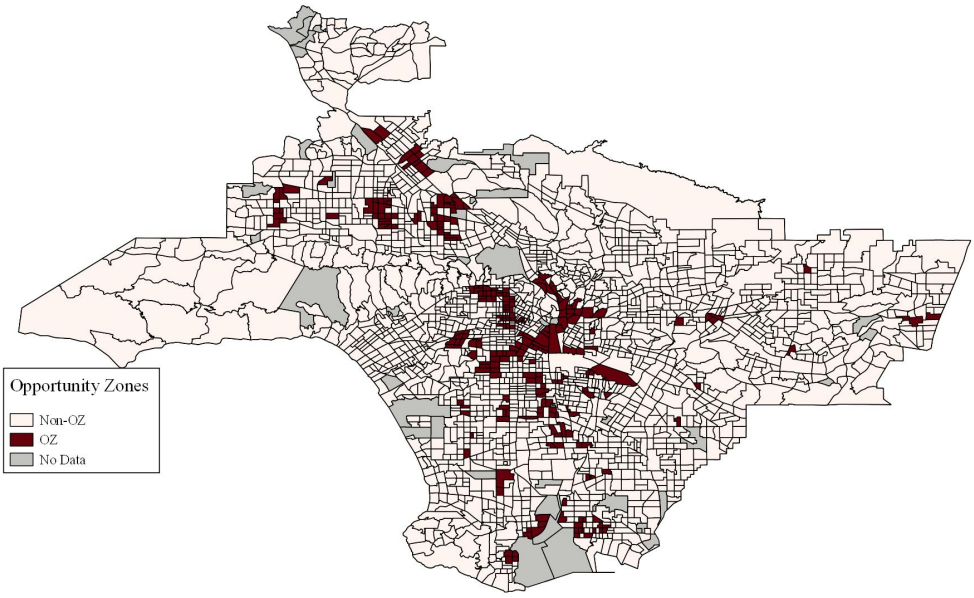
$Corr = .46$ in both, binscatter with 100 bins, ACS 2014 Data and Model-based.

Figure 10: Model Fit: Relative Expenditure Share on Goods



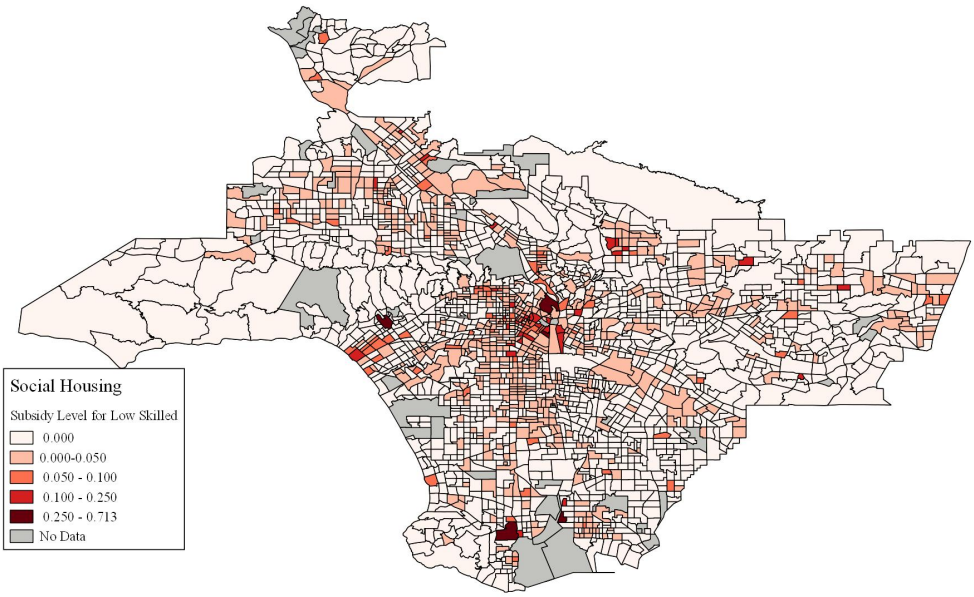
$Corr = .08$, binscatter with 100 bins, ACS 2014 Data and Model-based.

Figure 11: Opportunity Zones (OZ) in Los Angeles



The figure plots designated Opportunity Zones (257 Census Tracts).

Figure 12: Social Housing in Los Angeles, 2019



The figure plots the share of total housing costs of low skilled households that is covered by federal and state housing assistance in each Census Tract, Data from California Housing Partnership Preservation Database.

Figure 13: Opportunity Zones, % Change in Number of Firms, Baseline Calibration

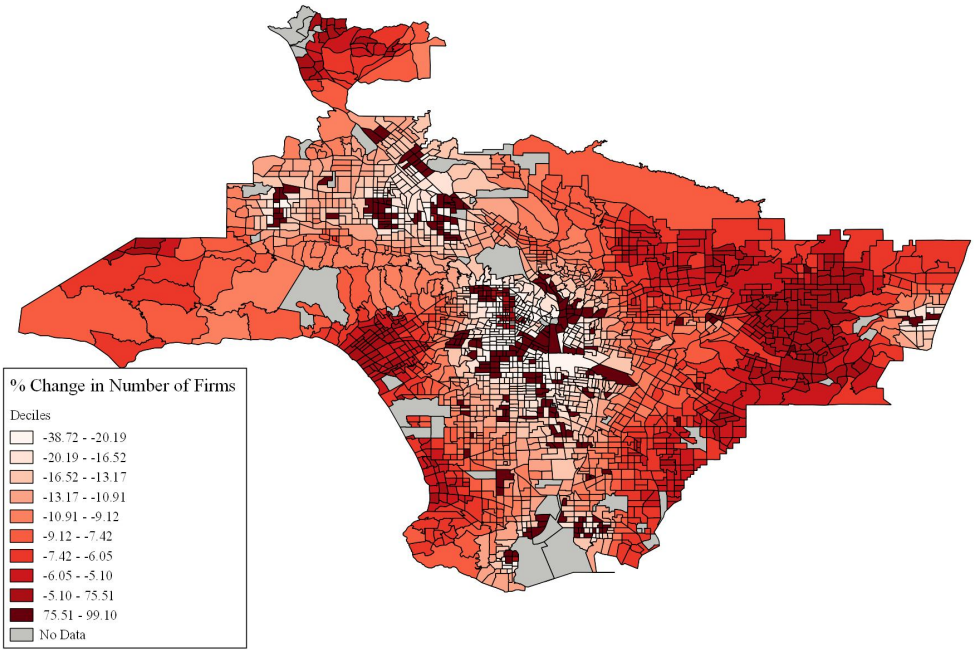


Figure 14: Opportunity Zones, % Change in Skill Ratio, Baseline Calibration

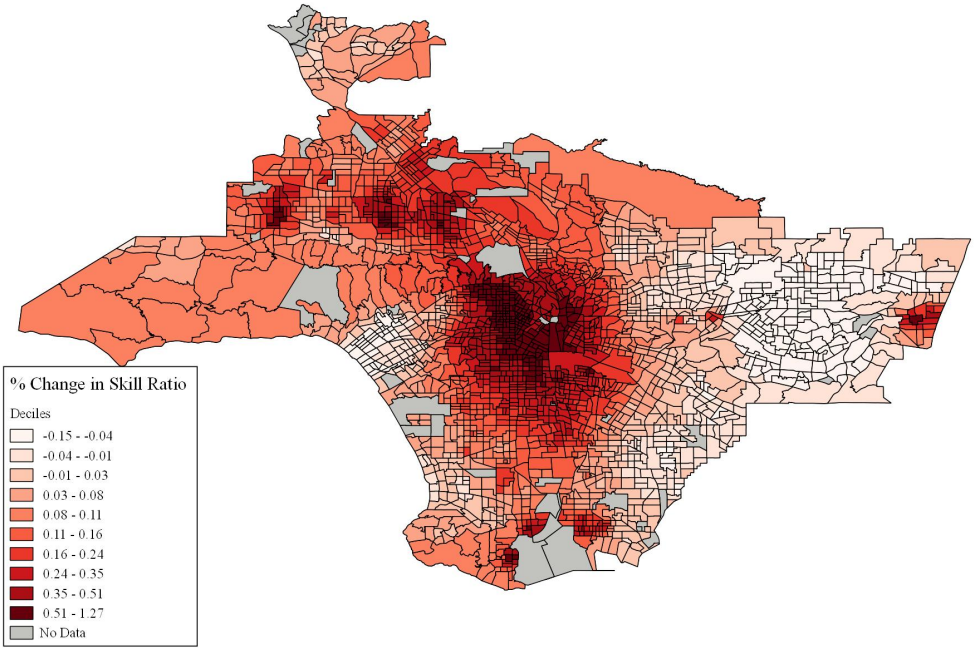


Figure 15: Social Housing, % Change in Number of Firms, Baseline Calibration

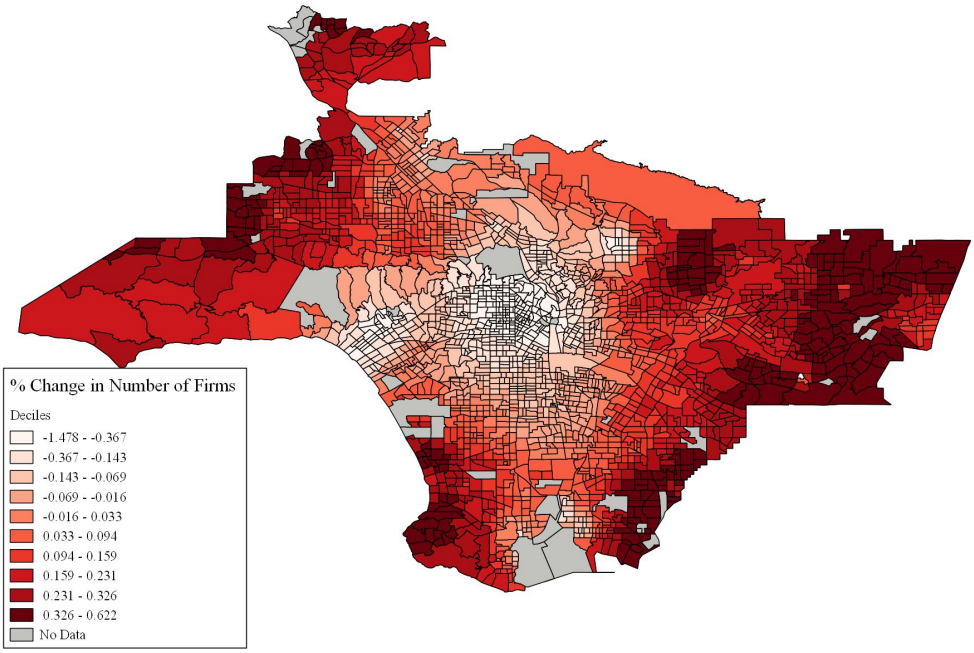


Figure 16: Social Housing, % Change in Skill Ratio, Baseline Calibration

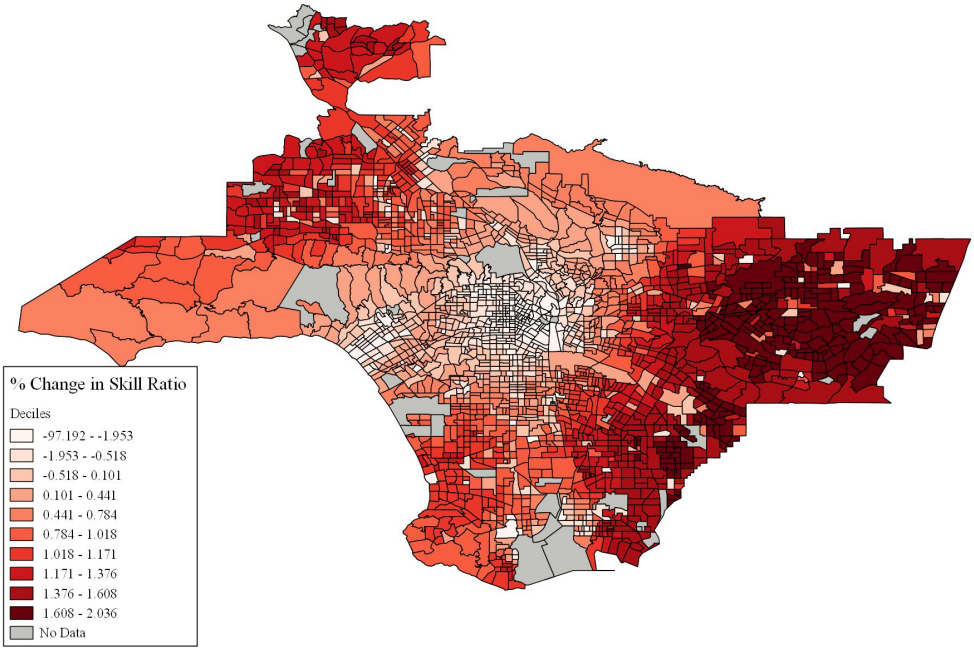
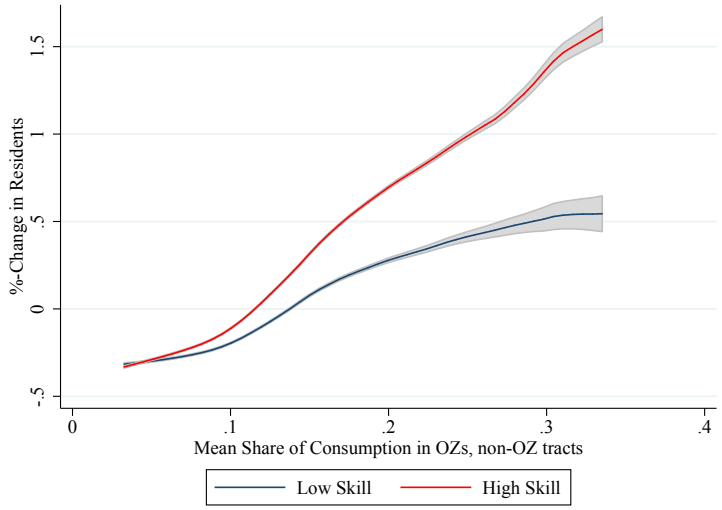
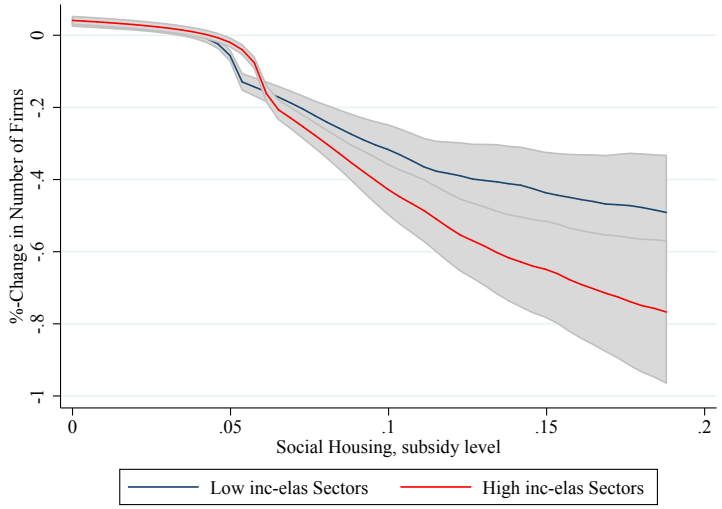


Figure 17: Opportunity Zones, Effect on Non-OZ Tracts, Baseline Calibration



The figure plots the % change in residents as a function of the mean share of consumption across sectors from non-OZ tracts in OZ tracts.

Figure 18: Social Housing, Mobility of Firms by Inc-Elasticity, Baseline Calibration



The figure plots the % change in the Number of Firms by Income Elasticity as a function fo Social Housing subsidy level.

Tables

Table 1: Number of Establishments in Recreation & Education vs Liquor & Convenience Stores, 2014

	(1)	(2)	(3)	(4)	(5)
	Income-elastic/ Inelastic	Recreation & Education	Liquor & Convenience	Recreation & Education	Liquor & Convenience
VARIABLES	Log Ratio	Log Number	Log Number	Number> 0	Number> 0
Log Skill Ratio	0.126***	0.320***	0.081***	0.181***	0.029***
	(0.010)	(0.021)	(0.017)	(0.008)	(0.010)
Observations	2,194	1,074	1,077	2,182	2,182
R-squared	0.070	0.180	0.022	0.149	0.004

Robust standard errors. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 2: Sector Income Elasticities and implied ν_j for Goods Sectors, 2014

	Sales Share	$(\widehat{\nu_j - \nu_{j^*}})\varepsilon$	**SE**	**Implied** ν_j
Liquor/Tobacco Store (N)	0.0062	-0.4419	0.1210	-2.4963
Convenience Store (N)	0.0153	-0.3682	0.1064	-2.3610
Dollar/Discount Stores (N)	0.0348	-0.2435	0.0672	-2.1322
Bars (D)	0.0027	-0.0851	0.3169	-1.8414
Fast Food (D)	0.0205	-0.0755	0.0587	-1.8237
Rental Stores (I)	0.0065	-0.0390	0.0813	-1.7568
Auto-related Services (I)	0.0352	-0.0293	0.0235	-1.7389
Grocery Store (N)	0.0612			-1.6852
Auto Dealers (I)	0.0866	0.0102	0.0890	-1.6666
Drug Store (N)	0.0338	0.1092	0.0712	-1.4847
Specialty Food Stores (N)	0.0110	0.1152	0.0928	-1.4737
Technical Services (I)	0.0310	0.1354	0.0349	-1.4367
Art (I)	0.0164	0.1491	0.1020	-1.4115
Hardware, Gardening etc (N)	0.0283	0.1849	0.0793	-1.3459
Legal Services (I)	0.0511	0.1856	0.2997	-1.3446
Personal Care Services (I)	0.0064	0.2162	0.0330	-1.2884
Appliances/Electronics (N)	0.0217	0.2765	0.1116	-1.1777
Specialty Stores (N)	0.0357	0.3059	0.0749	-1.1238
Amusement (I)	0.0129	0.3389	0.0561	-1.0631
Medical Services (I)	0.1452	0.3608	0.0444	-1.0229
Restaurants (D)	0.0442	0.3668	0.0716	-1.0119
Pet Store (N)	0.0046	0.4286	0.1019	-0.8985
Furniture Stores (N)	0.0142	0.4416	0.1167	-0.8747
Family Services (I)	0.0243	0.5128	0.0915	-0.7439
Department Store (N)	0.0170	0.5240	0.0891	-0.7234
Apparel Stores (N)	0.0363	0.5970	0.0907	-0.5893
Recreation (I)	0.0252	0.8760	0.0723	-0.0772
Education Services (I)	0.0044	1.1844	0.1503	0.4889
Tradables (I)	0.1675	0.3451	0.0218	-1.0517

Nominal income instrumented with dummy for high skill. All regressions include dummies for household size, age of householder and number of earners interacted with sector fixed effects, as well as Sector-MSA-Time fixed effects. All regressions are weighted by household weights in respective expenditure survey (N=Nielsen, I=CEX Interview, D=CEX Diary). Standard error are clustered at Household level.

Table 3: Correlations of Local Sector Growth Rates in Price-Bartik

	(1)	(2)	(3)
	Sector Growth Rate	Sector Growth Rate	Sector Growth Rate
VARIABLES	other cities	other cities	other cities
Income Elasticity	0.088		
	(0.075)		
Log Number of Firms, 1990		-0.113***	
		(0.029)	
Skill Intensity			0.369
			(0.248)
Observations	56	56	49
R-squared	0.050	0.232	0.068

Robust standard errors. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 4: Estimation of κ and δ_k, $\psi = -1.5$, 5km buffer

Panel A: First Stage Regression, $\frac{s_{kn,t-1}(g)}{s_{kn,t-1}(h)} \log \hat{s}_{kn,t}(g)$

	(1)	(2)
	1st	1st
VARIABLES	Price Effects	Price Effects
Avg Price IV	-0.037**	-0.075***
	(0.017)	(0.015)
Rel Price IV	0.502***	0.460***
	(0.063)	(0.070)
Rel Price IV X High	-0.136***	-0.122***
	(0.032)	(0.032)
Observations	8,362	8,362
R-squared	0.221	0.372
Controls	no	yes

Panel B: First Stage Regression, $\hat{\mathcal{L}}_{kn,t}$

	(1)	(2)	(3)	(4)
	1st	1st	1st	1st
VARIABLES	No Price Effects	Local	Price Effects	Price Effects
Avg Price IV	0.119***	0.116***	0.094***	0.116***
	(0.011)	(0.011)	(0.016)	(0.011)
Rel Price IV	0.306***	0.310***	0.284***	0.310***
	(0.061)	(0.061)	(0.063)	(0.061)
Rel Price IV X High	0.008***	0.002	0.000	0.002
	(0.002)	(0.001)	(.)	(0.001)
Observations	8,362	8,362	8,362	8,362
R-squared	0.867	0.867	0.827	0.867
Controls	yes	yes	no	yes

Panel C: First Stage Regression, $\hat{\mathcal{L}}_{kn,t} \times High$

	(1)	(2)	(3)	(4)
	1st	1st	1st	1st
VARIABLES	No Price Effects	Local	Price Effects	Price Effects
Avg Price IV	0.060***	0.058***	0.047***	0.058***
	(0.006)	(0.006)	(0.008)	(0.006)
Rel Price IV	-0.244***	-0.244***	-0.262***	-0.244***
	(0.029)	(0.029)	(0.027)	(0.029)
Rel Price IV X High	0.811***	0.807***	0.807***	0.807***
	(0.031)	(0.031)	(0.031)	(0.031)
Observations	8,362	8,362	8,362	8,362
R-squared	0.906	0.906	0.894	0.906
Controls	yes	yes	no	yes

Panel D: Second Stage Regressions

	(1)	(2)	(3)	(4)	(5)
		IV	IV	IV	IV
VARIABLES	OLS	No Price Effects	Local	Price Effects	Price Effects
$-\frac{\hat{\kappa}}{1-\eta}$	0.020			-5.717***	-4.695***
	(0.048)			(1.007)	(0.768)
$\hat{\delta}_{low}$	-0.514***	0.385	0.260	-0.920	-0.883
	(0.065)	(0.511)	(0.517)	(1.014)	(0.736)
$\hat{\delta}_{high} - \hat{\delta}_{low}$	2.675***	1.728***	1.714***	0.769**	1.019***
	(0.052)	(0.090)	(0.088)	(0.306)	(0.247)
Observations	8,362	8,362	8,362	8,362	8,362
Controls	yes	yes	yes	no	yes
1st Stage F-Stat		46	44.03	8.80	19.19

All specifications include the sum of shares interacted with year-dummies and skill-year FX. Controls include changes in household income, changes in residential rents, log distance to city center (City of Los Angeles City Hall), log population density in 1990 and log average slope. Standard errors clustered at level of Census Tract. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 5: Estimation of θ

Panel A: First Stage Regression

	(1)	(2)	(3)	(4)
VARIABLES	1st	1st	1st	1st alt. IV
Log Avg Slope X ν_j	0.116***	0.061***	0.185***	
	(0.010)	(0.013)	(0.015)	
Log Avg Slope X rank(ν_j)				0.008***
				(0.001)
Observations	152,323	79,550	58,783	152,323
Sample	Chains	Retail	Services	Chains
Tract-Year FX	yes	yes	yes	yes
Sector-Year FX	yes	yes	yes	yes
Chain-Year FX	yes	yes	yes	yes
Number of clusters	6113	5285	5601	6113

Panel B: Second Stage Regressions

	(1)	(2)	(3)	(4)
VARIABLES	IV	IV	IV	alt. IV
$\frac{\sigma-1}{\theta}$	0.251***	0.225	0.317***	0.164***
	(0.066)	(0.146)	(0.098)	(0.062)
Observations	152,323	79,550	58,783	152,323
Sample	Chains	Retail	Services	Chains
Tract-Year FX	yes	yes	yes	yes
Sector-Year FX	yes	yes	yes	yes
Chain-Year FX	yes	yes	yes	yes
Number of clusters	6113	5285	5601	6113
1st Stage F-Stat	144.2	20.96	162.3	154.7

Standard errors clustered at level of Year-Zipcode. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 6: Main Calibration of Model

Parameter	Description	Value	Source
Preferences			
κ	Resident Supply Elasticity	2.4	Estimated
ϵ_h	Income Elasticity of Housing	.507	Assumed
ϵ_g	Income Elasticity of Goods	.507	Assumed
ν_j	Sector Income Elasticities	see table 2	Estimated
η	EoS between Housing and Goods	.493	Assumed
γ	EoS across Sectors	2	Assumed
σ	Eos within Sector	5	Assumed
Firm Supply			
θ	Firm Supply Elasticity	16	Estimated
Shopping Frictions			
ϕ^1	Distance Elasticity, w/ Price Effects	-1.5	Assumed
ϕ^2	Distance Elasticity, w/o Price Effects	0	Assumed
Spillover Elasticities			
δ^1_{low}	Low Skilled, w/ Price Effects	0	Estimated & Scaled
δ^2_{low}	Low Skilled, w/o Price Effects	0	Estimated & Scaled
δ^1_{high}	High Skilled, w/ Price Effects	1	Estimated & Scaled
δ^2_{high}	High Skilled, w/o Price Effects	1.25	Estimated & Scaled
Skill Premium			
ρ	Rel. Labor Endowment of High Skilled	1.7	IPUMS 2014

Table 7: Opportunity Zones, Baseline Calibration, % Changes

	Firms All local	HHs All	HHs Low Skill	HHs High Skill	HHs Skill Ratio
Opportunity Zone	79.98	0.73	0.59	1.18	0.59
No Opportunity Zone	-10.39	-0.09	-0.09	-0.09	0.09
R-squared	0.95	0.44	0.46	0.43	0.47

Table 8: Opportunity Zones, Calibration without Price Effects, % Changes

	Firms All local	HHs All	HHs Low Skill	HHs High Skill	HHs Skill Ratio
Opportunity Zone	90.86	0.02	0.02	0.03	0.00
No Opportunity Zone	-12.93	-0.00	-0.00	-0.00	0.01
R-squared	0.99	0.07	0.08	0.05	0.71

Table 9: Opportunity Zones, % Changes in local Skill Ratio and Welfare by Skill

	w/ Price Effects	w/o Price Effects	w/ Price Effects No Spillover	w/o Price Effects No Spillover	Homothetic Preferences	Homothetic Preferences No Spillover
Skill Ratio						
OZ	0.59	0.00	0.39	0.01	0.21	0.21
Non-OZ	0.09	0.01	0.12	0.01	0.21	0.21
Welfare						
Low Skill	-0.09	-0.14	-0.09	-0.14	-0.19	-0.19
High Skill	-0.08	-0.16	-0.16	-0.16	-0.19	-0.28

Table 10: Social Housing, Baseline Calibration, % Changes

	Firms	HHs	HHs	HHs	HHs
	All local	All	Low Skill	High Skill	Skill Ratio
Subsidy Level	-4.54	2.36	29.73	-50.77	-107.66
Constant	0.05	-0.02	-0.29	0.45	0.80
R-squared	0.23	0.09	0.78	0.61	0.85

Table 11: Social Housing, Calibration without Price Effects, % Changes

	Firms	HHs	HHs	HHs	HHs
	All local	All	Low Skill	High Skill	Skill Ratio
Subsidy Level	0.01	-0.10	32.64	-63.86	-115.84
Constant	0.00	0.00	-0.32	0.59	0.71
R-squared	0.00	0.00	0.63	0.34	0.64

Table 12: Social Housing, % Changes in local Skill Ratio and Welfare by Skill

	w/ Price Effects	w/o Price Effects	w/ Price Effects No Spillover	w/o Price Effects No Spillover	Homothetic Preferences	Homothetic Preferences No Spillover
Skill Ratio						
Subsidy Level	-107.66	-115.84	-95.04	-92.92	-162.47	-143.81
Constant	0.80	0.71	0.76	0.75	1.25	1.19
Welfare						
Low Skill	0.10	0.08	0.10	0.10	0.16	0.17
High Skill	-0.23	-0.35	-0.11	-0.10	-0.38	-0.18

References

Agarwal, Sumit, Jensen, J Bradford, & Monte, Ferdinando. 2018. The Geography of Consumption.

Aguiar, Mark, & Bils, Mark. 2015. Has consumption inequality mirrored income inequality? *American Economic Review*, **105**(9), 2725–56.

Ahlfeldt, Gabriel M, Redding, Stephen J, Sturm, Daniel M, & Wolf, Nikolaus. 2015. The economics of density: Evidence from the Berlin Wall. *Econometrica*, **83**(6), 2127–2189.

Albouy, David, Ehrlich, Gabriel, & Liu, Yingyi. 2016. Housing demand and expenditures: How rising rent levels affect behavior and cost-of-living over space and time.

Allen, Treb, Arkolakis, Costas, & Li, Xiangliang. 2015. Optimal city structure. *Yale University, mimeograph*.

Athey, Susan. 2002. Monotone comparative statics under uncertainty. *The Quarterly Journal of Economics*, **117**(1), 187–223.

Atkin, David, Faber, Benjamin, & Gonzalez-Navarro, Marco. 2018. Retail globalization and household welfare: Evidence from mexico. *Journal of Political Economy*, **126**(1), 1–73.

Barnatchez, Keith, Crane, Leland Dod, & Decker, Ryan. 2017. An assessment of the national establishment time series (nets) database.

Baum-Snow, Nathaniel, & Hartley, Daniel A. 2016. Accounting for central neighborhood change, 1980-2010.

Behrens, Kristian, Duranton, Gilles, & Robert-Nicoud, Frédéric. 2014. Productive cities: Sorting, selection, and agglomeration. *Journal of Political Economy*, **122**(3), 507–553.

Borusyak, Kirill, & Jaravel, Xavier. 2018. *The Distributional Effects of Trade: Theory and Evidence from the United States*.

Borusyak, Kirill, Hull, Peter, & Jaravel, Xavier. 2018. *Quasi-experimental shift-share research designs*. Tech. rept. National Bureau of Economic Research.

Brinkman, Jeffrey, Coen-Pirani, Daniele, & Sieg, Holger. 2015. Firm dynamics in an urban economy. *International Economic Review*, **56**(4), 1135–1164.

Brueckner, Jan K, Thisse, Jacques-Francois, & Zenou, Yves. 1999. Why is central Paris rich and downtown Detroit poor?: An amenity-based theory. *European economic review*, **43**(1), 91–107.

Busso, Matias, Gregory, Jesse, & Kline, Patrick. 2013. Assessing the incidence and efficiency of a prominent place based policy. *American Economic Review*, **103**(2), 897–947.

Comin, Diego, Lashkari, Danial, & Mestieri, Martı. 2018. Structural Change with Long-run Income and Price Effects.

Couture, Victor. 2016. Valuing the consumption benefits of urban density. *University of California, Berkeley. Processed.*

Couture, Victor, & Handbury, Jessie. 2017. *Urban revival in America, 2000 to 2010.* Tech. rept. National Bureau of Economic Research.

Couture, Victor, Gaubert, Cecile, Handbury, Jessie, & Hurst, Erik. 2019. Income Growth and the Distributional Effects of Urban Spatial Sorting.

Davis, Donald R, Dingel, Jonathan I, Monras, Joan, & Morales, Eduardo. 2019. How segregated is urban consumption? *Journal of Political Economy*, **127**(4), 000–000.

Davis, Morris A, Gregory, Jesse, & Hartley, Daniel A. 2018. The Long Run Effects of Low Income Housing on Neighborhood Composition.

DellaVigna, Stefano, & Gentzkow, Matthew. 2019. Uniform pricing in us retail chains. *The Quarterly Journal of Economics*, **134**(4), 2011–2084.

Diamond, Rebecca. 2016. The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000. *American Economic Review*, **106**(3), 479–524.

Diamond, Rebecca, & McQuade, Tim. 2019. Who Wants Affordable Housing in Their Backyard? An Equilibrium Analysis of Low-Income Property Development. *Journal of Political Economy*, **127**(3), 000–000.

Diamond, Rebecca, McQuade, Timothy, & Qian, Franklin. 2018. *The effects of rent control expansion on tenants, landlords, and inequality: Evidence from San Francisco.* Tech. rept. National Bureau of Economic Research.

Dolfen, Paul, Einav, Liran, Klenow, Peter J, Klopack, Benjamin, Levin, Jonathan D, Levin, Laurence, & Best, Wayne. 2019. *Assessing the Gains from E-commerce.* Tech. rept. National Bureau of Economic Research.

Fajgelbaum, Pablo, & Gaubert, Cecile. 2019. *Optimal spatial policies, geography and sorting.* Tech. rept. National Bureau of Economic Research.

Gaubert, Cecile. 2018. Firm sorting and agglomeration. *American Economic Review*, **108**(11), 3117–53.

Glaeser, Edward L, Kim, Hyunjin, & Luca, Michael. 2018. Nowcasting gentrification: using yelp data to quantify neighborhood change. *Pages 77–82 of: AEA Papers and Proceedings*, vol. 108.

Guerrieri, Veronica, Hartley, Daniel, & Hurst, Erik. 2013. Endogenous gentrification and housing price dynamics. *Journal of Public Economics*, **100**, 45–60.

Handbury, Jessie. 2013. Are poor cities cheap for everyone? non-homotheticity and the cost of living across us cities. *The wharton school research paper*, 1–054.

Hanoch, Giora. 1975. Production and demand models with direct or indirect implicit additivity. *Econometrica (pre-1986)*, **43**(3), 395.

Helliwell, John F, & Verdier, Genevieve. 2001. Measuring internal trade distances: a new method applied to estimate provincial border effects in Canada. *Canadian Journal of Economics*, 1024–1041.

Hottman, Colin, & Monarch, Ryan. 2018. Estimating Unequal Gains across US Consumers with Supplier Trade Data. *FRB International Finance Discussion Paper*.

Hubmer, Joachim. 2018. *The Race Between Preferences and Technology*. Tech. rept.

Krugman, Paul. 1991. Increasing returns and economic geography. *Journal of political economy*, **99**(3), 483–499.

Lee, Sanghoon, & Lin, Jeffrey. 2017. Natural amenities, neighbourhood dynamics, and persistence in the spatial distribution of income. *The Review of Economic Studies*, **85**(1), 663–694.

Matsuyama, Kiminori. 2019. Engel's Law in the Global Economy: Demand-Induced Patterns of Structural Change, Innovation, and Trade. *Econometrica*, **87**(2), 497–528.

Monte, Ferdinando, Redding, Stephen J, & Rossi-Hansberg, Esteban. 2018. Commuting, migration, and local employment elasticities. *American Economic Review*, **108**(12), 3855–90.

Neumark, David, Zhang, Junfu, & Wall, Brandon. 2005. Employment Dynamics and Business Relocation: New Evidence from the National Establishment Time Series.

Redding, Stephen, & Weinstein, David. 2017. *Aggregating from micro to macro patterns of trade*. Tech. rept. National Bureau of Economic Research.

Redding, Stephen J, & Weinstein, David E. 2019. Measuring Aggregate Price Indexes with Taste Shocks: Theory and Evidence for CES Preferences.

Reynolds, C Lockwood, & Rohlin, Shawn M. 2015. The effects of location-based tax policies on the distribution of household income: evidence from the federal Empowerment Zone program. *Journal of Urban Economics*, **88**, 1–15.

Schiff, Nathan. 2014. Cities and product variety: evidence from restaurants. *Journal of Economic Geography*, **15**(6), 1085–1123.

Su, Yichen. 2018a. Measuring the Value of Urban Consumption Amenities: A Time-Use Approach.

Su, Yichen. 2018b. The Rising Value of Time and the Origin of Urban Gentrification.

Theodos, Brett, Meixell, Brady, & Hedman, Carl. 2018. Did States Maximize Their Opportunity Zone Selections? *Retrieved from Urban Institute: https://www. urban. org/sites/default/files/publication/98445/did_states_maximize_their_ opportunity_zone_selections_1. pdf.*

Tsivanidis, Nick. 2018. The Aggregate and Distributional Effects of Urban Transit Infrastructure: Evidence from Bogotás TransMilenio.

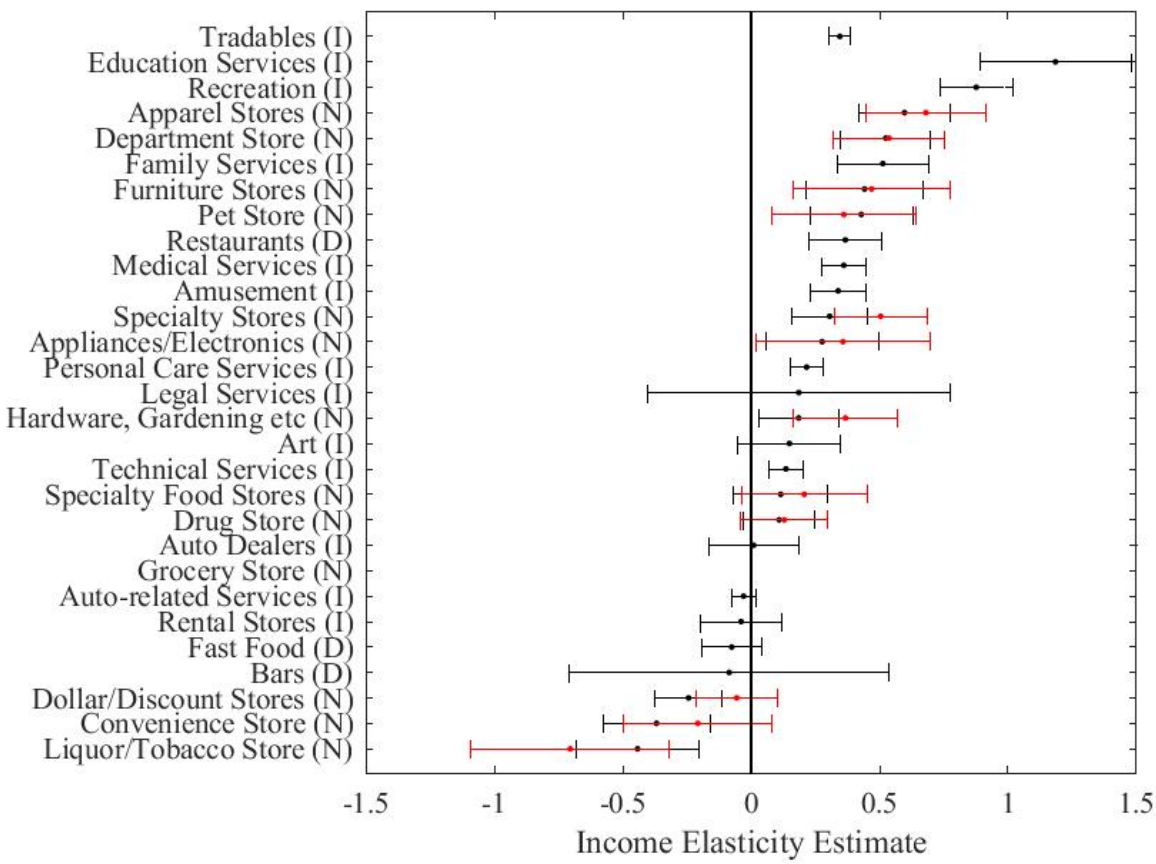
Waldfogel, Joel. 2008. The median voter and the median consumer: Local private goods and population composition. *Journal of Urban Economics*, **63**(2), 567–582.

Ziv, Oren. Productivity, Density, and Sorting.

Appendix Figures and Tables

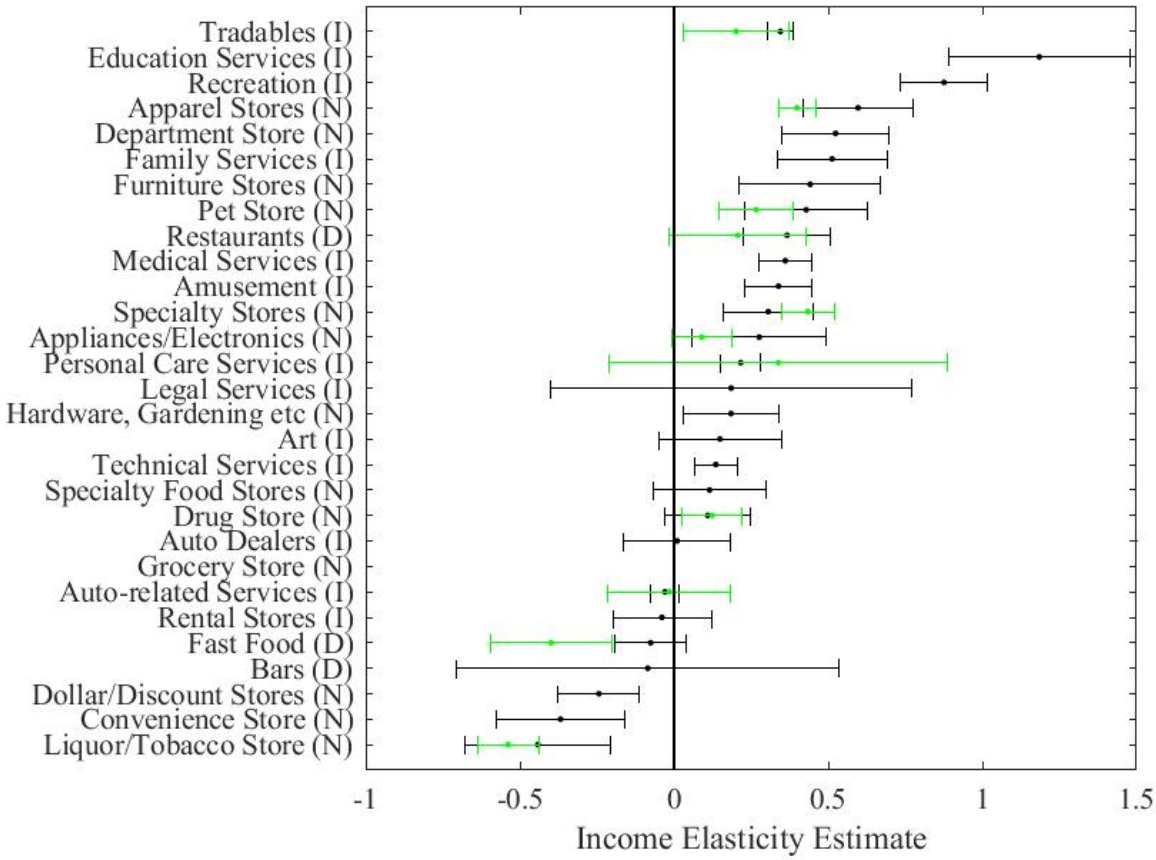
A.1 Additional Figures

Figure A.1: Income Elasticities by Sector with zip code fixed effects



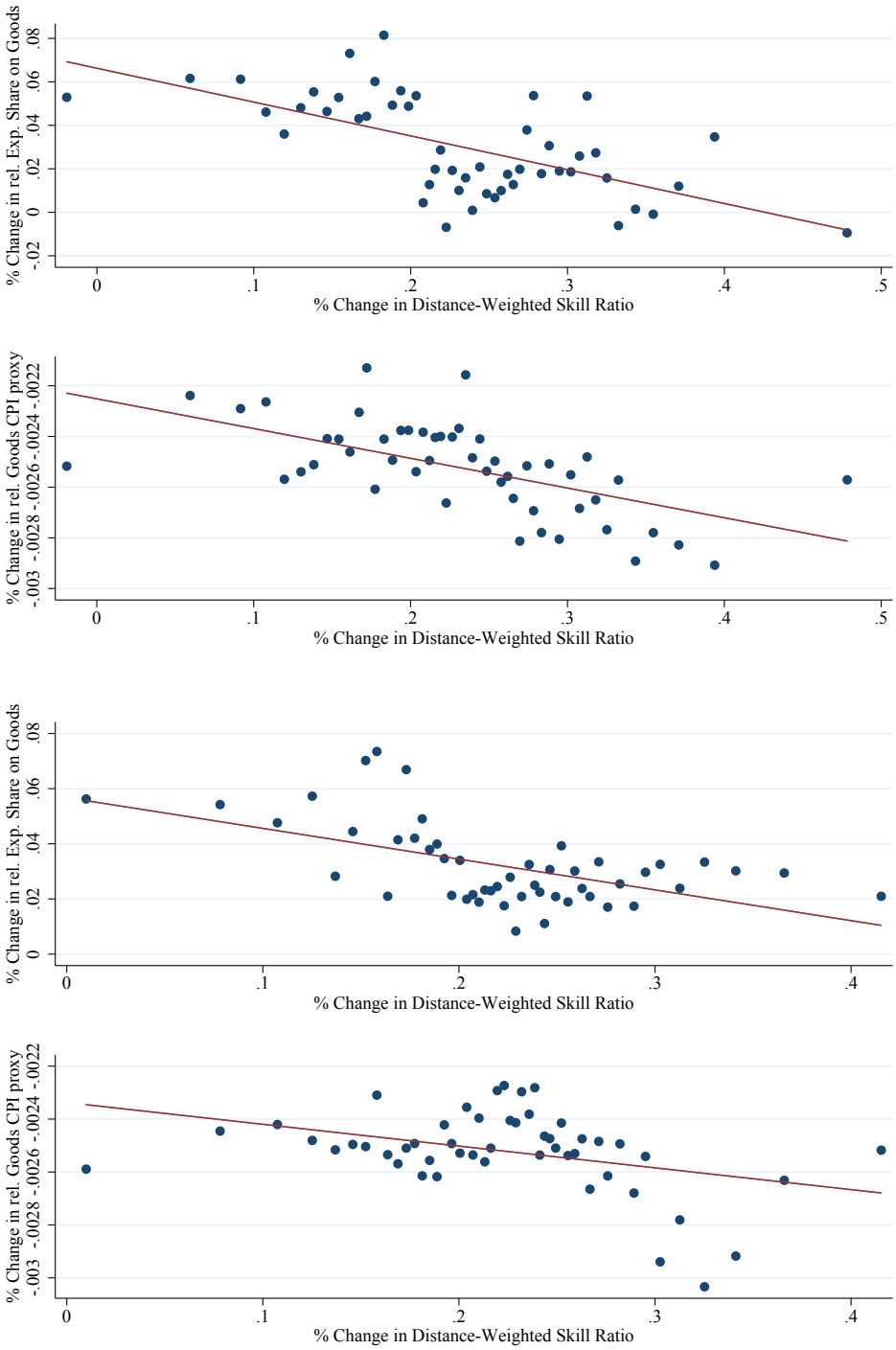
The figure plots estimated point estimates of income elasticities by goods sector using consumer expenditure data and 95% confidence intervals. Estimates from Nielsen with Zip-Code fixed effects in red. Data source in parentheses (N=Nielsen, I=CEX Interview, D=CEX Diary).

Figure A.2: Income Elasticities by Sector with sector definition from alternative source



The figure plots estimated point estimates of income elasticities by goods sector using consumer expenditure data and 95% confidence intervals. Estimates from data alternative source covering approximately the same sector in green. Data source in parentheses (N=Nielsen, I=CEX Interview, D=CEX Diary).

Figure A.3: Tract-level Changes in Relative Expenditure Share and Local Price Index Proxy



The figure plots % change in ratio of expenditures shares on goods of high over low skilled HHs (upper panel of top graph), and % change in ratio of price index proxy on goods of high over low skilled HHs (lower panel of top graph) as a function of % changes in the distance-weighted skill ratio (measure of spillovers) without controls. Bottom graph shows the same relationships with regression controls. I construct the price index proxy as in 30, but I use observed changes in the number of varieties in each tract instead of sector growth rates from other MSAs. I multiply the tract-sector-specific growth rates with $\frac{\theta}{(1-\sigma)} - 1$, based on a first-order approximation of the model expression for the price index in 12. Binscatter with 50 bins, data from ACS 2014, NETS.

Figure A.4: Opportunity Zones, % Change in Price Index for High Skilled

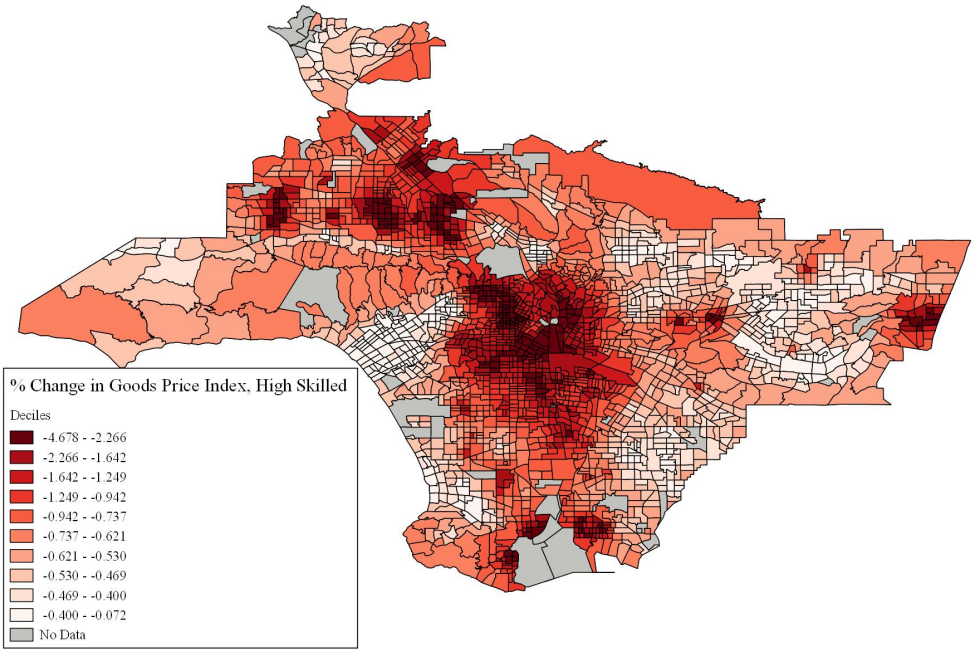


Figure A.5: Opportunity Zones, % Change in Spillovers

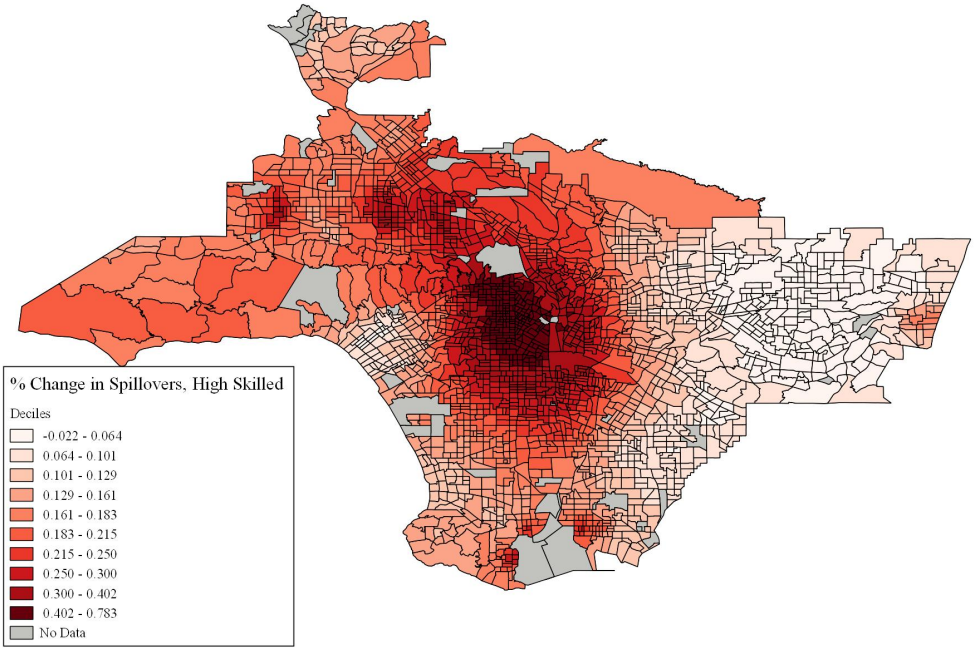


Figure A.6: Social Housing, Relative % Change in Number of Firms by Income Elasticity

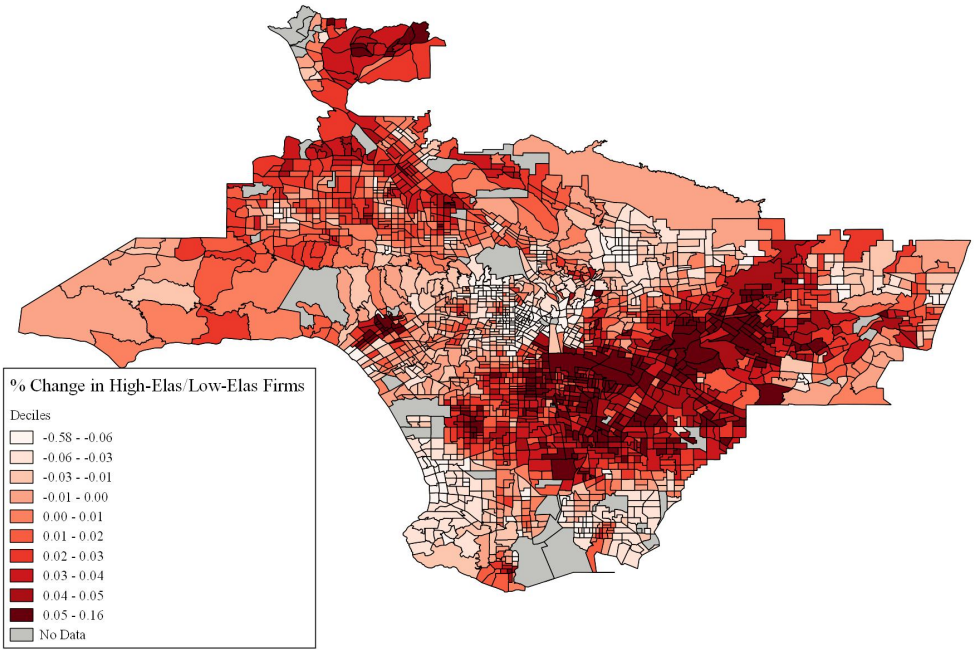
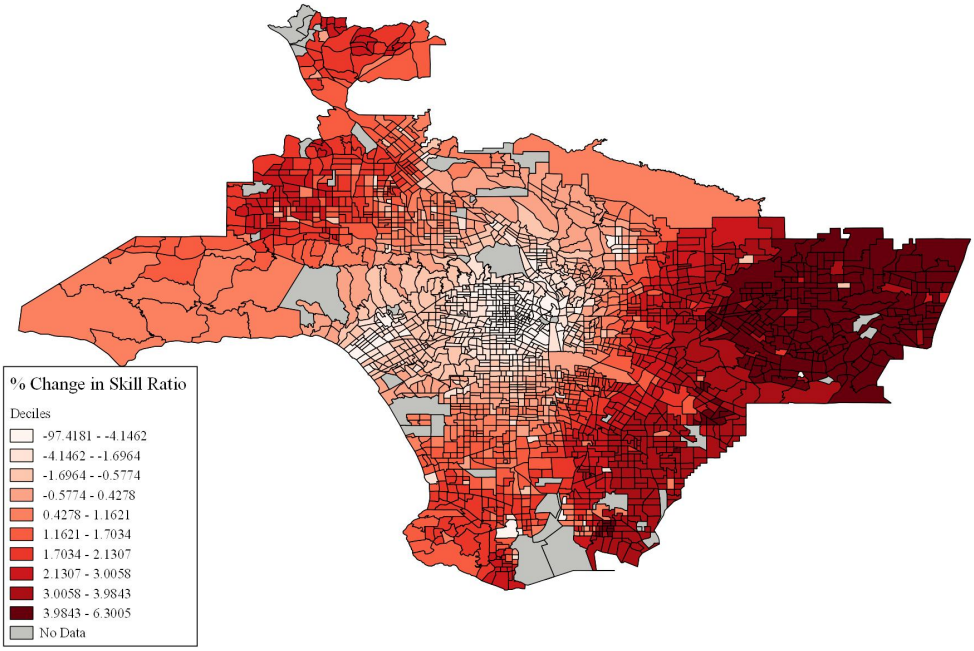


Figure A.7: Social Housing, % Change in Skill Ratio, Calibration w/o Price Effects



A.2 Additional Tables

Table A.1: Skill Premium, ACS 2014

VARIABLES	(1) All US	(2) LA Sample
High Skilled HH Head	0.639***	0.705***
	(0.001)	(0.006)
Observations	4,242,708	137,063
R-squared	0.284	0.254

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table A.2: Expenditure Share on Housing by Skill

VARIABLES	(1) All US ACS 2014	(2) LA Sample ACS 2014	(3) All US CEX 12-16	(4) LA Sample CEX 12-16	(5) LA Tracts ACS 2014	(6) LA Tracts Model
High Skilled HH Head	-0.0449***	-0.0535***	-0.0168***	-0.0072*	-0.0578***	-0.0398***
	(0.0002)	(0.0015)	(0.0015)	(0.0043)	(0.0006)	(0.0002)
Observations	4,078,372	127,523	40,868	5,578	4,388	4,388
R-squared	0.1257	0.0700	0.1438	0.1444	0.9573	0.9984

(1) and (2) include Puma-year FX and dummies for sex and age of HH head, HH size and home ownership. (3) and (4) include MSA-year FX and same dummies as above. (5) and (6) use tract level data and model outcomes and include tract FX. Robust Standard errors. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.3: Sector Income Elasticities, Robustness

	Baseline		Zipcode Fx		Alternative Source	
	$\widehat{(\nu_j - \nu_{j*})}\varepsilon$	SE	$\widehat{(\nu_j - \nu_{j*})}\varepsilon$	SE	$\widehat{(\nu_j - \nu_{j*})}\varepsilon$	SE
Liquor/Tobacco Store (N)	-0.4419	0.1210	-0.7041	0.1971	-0.5380	0.0508
Convenience Store (N)	-0.3682	0.1064	-0.2072	0.1473		
Dollar/Discount Stores (N)	-0.2435	0.0672	-0.0558	0.0814		
Bars (D)	-0.0851	0.3169				
Fast Food (D)	-0.0755	0.0587			-0.3986	0.1009
Rental Stores (I)	-0.0390	0.0813				
Auto-related Services (I)	-0.0293	0.0235			-0.0149	0.1013
Grocery Store (N)						
Auto Dealers (I)	0.0102	0.0890				
Drug Store (N)	0.1092	0.0712	0.1292	0.0865	0.1239	0.0493
Specialty Food Stores (N)	0.1152	0.0928	0.2069	0.1252		
Technical Services (I)	0.1354	0.0349				
Art (I)	0.1491	0.1020				
Hardware, Gardening etc (N)	0.1849	0.0793	0.3675	0.1027		
Legal Services (I)	0.1856	0.2997				
Personal Care Services (I)	0.2162	0.0330			0.3382	0.2796
Appliances/Electronics (N)	0.2765	0.1116	0.3574	0.1733	0.0899	0.0493
Specialty Stores (N)	0.3059	0.0749	0.5045	0.0926	0.4340	0.0434
Amusement (I)	0.3389	0.0561				
Medical Services (I)	0.3608	0.0444				
Restaurants (D)	0.3668	0.0716			0.2061	0.1135
Pet Store (N)	0.4286	0.1019	0.3610	0.1429	0.2665	0.0616
Furniture Stores (N)	0.4416	0.1167	0.4692	0.1549		
Family Services (I)	0.5128	0.0915				
Department Store (N)	0.5240	0.0891	0.5369	0.1102		
Apparel Stores (N)	0.5970	0.0907	0.6805	0.1190	0.3991	0.0306
Recreation (I)	0.8760	0.0723				
Education Services (I)	1.1844	0.1503				
Tradables (I)	0.3451	0.0218			0.2013	0.0877

As in Table 2. Zipcode Fx refers to specification in Nielsen where the Sector-MSA-Time Fx is replaced by Sector-Zipcode-Time Fx. Alternative Source refers to estimates from other samples covering the approximately same sector.

Table A.4: Estimation of κ and δ_k, Robustness

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	10km Buffer	10km Buffer	$\psi = 3$	$\psi = 3$	Pop-based	Pop-based	Weighted ^{IV}	Weighted ^{IV}	Rent Share	Rent Share
VARIABLES	Local	Price Effects	Local	Price Effects	Local	Price Effects	Local	Price Effects	Local	Price Effects
$-\frac{\hat{\kappa}}{1-\eta}$		-4.596***		-4.853***		-4.508***		-4.628***		-7.713***
		(0.920)		(0.848)		(0.824)		(0.725)		(2.948)
$\hat{\delta}_{low}$	-1.042***	-1.273***	0.581***	-0.492	1.101***	-0.538	-0.146	-0.927*	0.260	-1.752
	(0.189)	(0.291)	(0.220)	(0.387)	(0.308)	(0.564)	(0.272)	(0.485)	(0.517)	(1.491)
$\hat{\delta}_{high} - \hat{\delta}_{low}$	1.605***	0.693***	1.001***	0.589***	1.763***	1.007***	1.715***	1.029***	1.714***	1.848***
	(0.074)	(0.266)	(0.029)	(0.147)	(0.171)	(0.321)	(0.088)	(0.250)	(0.088)	(0.390)
Observations	8,362	8,362	8,362	8,362	8,358	8,358	8,362	8,362	8,362	8,362
Controls	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
1st Stage F-Stat	210.6	14.28	17.39	11.31	64.69	13.86	67.35	18.48	44.03	2.442

All specifications include the sum of shares interacted with year-dummies (Borusyak *et al.* (2018)) and skill-year FX. Controls include log distance to city center (City of Los Angeles City Hall), log population density in 1990 and log average slope. Standard errors clustered at level of Census Tract. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.5: Opportunity Zones, Summary Statistics

	OZ	Non-OZ	Difference
Log Number of HHs	6.827	6.907	-0.0795*
Log HH Income	10.74	11.22	-0.484***
Log Skill Ratio	-1.423	-0.692	-0.731***
Log Number of Firms	4.264	4.380	-0.116*
Log Number of Firms, low Inc-Elas	3.491	3.539	-0.0471
Log Number of Firms, high Inc-Elas	3.571	3.765	-0.195**
Log Rent per sqft	0.341	0.330	0.0111
Observations	257	1937	

Table A.6: Opportunity Zones, Decomposing the HH Response, Baseline Model

	U	U	Rent	P(g)	P(g)	CPI(g)	CPI(g)	B
	Low Skill	High Skill	All	Low Skill	High Skill	Low Skill	High Skill	High Skill
Opportunity Zone	0.16	0.25	0.61	-2.75	-2.85	-2.94	-3.15	0.38
No Opportunity Zone	-0.13	-0.19	-1.47	-0.90	-0.72	-0.73	-0.48	0.18
R-squared	0.61	0.74	0.79	0.81	0.80	0.74	0.69	0.74

Table A.7: Opportunity Zones, Decomposing the HH Response, Model w/o Price Effects

	U	U	Rent	P(g)	P(g)	CPI(g)	CPI(g)	B
	Low Skill	High Skill	All	Low Skill	High Skill	Low Skill	High Skill	High Skill
Opportunity Zone	-0.13	-0.15	-1.43	-1.02	-1.02	-0.85	-0.83	0.01
No Opportunity Zone	-0.14	-0.16	-1.49	-1.03	-1.03	-0.84	-0.82	0.01
R-squared	0.99	0.99	1.00	1.00	1.00	1.00	1.00	0.84

Table A.8: Opportunity Zones, Decomposing the Firm Response

	$M_n(j)$	$M_n(j)$	$M_n(j)$	$M_n(j)$
	w/ Price Effects	w/ Price Effects	w/o Price Effects	w/o Price Effects
Opportunity Zone	91.30		104.95	
Opportunity Zone X ν_j		1.06		0.00
R-squared	0.96	1.00	1.00	1.00
Sector FX	yes	yes	yes	yes
Tract FX	no	yes	no	yes

Table A.9: Social Housing, Decomposing the HH Response, Baseline Model

	U	U	Rent	P(g)	P(g)	CPI(g)	CPI(g)	B
	Low Skill	High Skill	All	Low Skill	High Skill	Low Skill	High Skill	High Skill
Subsidy Level	12.39	-12.81	62.43	16.62	-15.88	0.30	0.35	-20.03
Constant	-0.02	0.02	-0.04	-0.04	0.02	-0.00	-0.00	-0.14
R-squared	0.78	0.59	0.78	0.77	0.55	0.11	0.10	0.20

Table A.10: Social Housing, Decomposing the HH Response, Model w/o Price Effects

	U	U	Rent	P(g)	P(g)	CPI(g)	CPI(g)	B
	Low Skill	High Skill	All	Low Skill	High Skill	Low Skill	High Skill	High Skill
Subsidy Level	13.60	-10.78	56.01	17.99	-13.76	0.00	0.00	-37.99
Constant	-0.05	0.02	0.01	-0.07	0.02	0.00	-0.00	-0.29
R-squared	0.63	0.37	0.49	0.63	0.37	0.04	0.10	0.11

Table A.11: Social Housing, Decomposing the Firm Response

	$M_n(j)$	$M_n(j)$	$M_n(j)$	$M_n(j)$
	w/ Price Effects	w/ Price Effects	w/o Price Effects	w/o Price Effects
Subsidy Level	-4.56		0.00	
Subsidy Level X ν_j		-1.94		-0.00
R-squared	0.23	0.97	1.00	1.00
Sector FX	yes	yes	yes	yes
Tract FX	no	yes	no	yes

Appendix A

[COMING SOON]

A.3 Geography of Los Angeles

A.4 Census and ACS Data

A.5 Sector Definitions for Local Consumption

A.6 NETS Data

A.7 CEX Expenditure Data

A.8 Nielsen Expenditure Data

A.9 Housing Data

A.10 Amenity Controls

A.11 Social Housing Data

Appendix B

B.1 Properties of Household Preferences

The following analysis is under the assumptions of given prices and from the view of an individual household of any type. To save on notation I omit location and type subscripts
First, let's look at the price index of goods responds to changes in real consumption

$$\frac{U\partial P(g)}{P(g)\partial U} = \frac{1}{1-\gamma}\sum_{j=1}^{J}\tilde{s}(j)\nu_j = \frac{\bar{\nu}}{1-\gamma}$$

where $\bar{\nu}$ is the expenditure share weighted income elasticity of demand parameter across sectors inside the goods sector. Second, I can compute the expenditure elasticity with respect to real consumption

$$\begin{aligned}\frac{U\partial E}{E\partial U} &= \frac{1}{1-\eta}I^{\eta-1}\left(a_h r^{1-\eta}U^{\epsilon_h}\epsilon_h + a_g P(g)^{1-\eta}U^{\epsilon_g}\epsilon_g + a_g P(g)^{1-\eta}U^{\epsilon_g}(1-\eta)\frac{U\partial P(g)}{P(g)\partial U}\right)\\ &= \frac{1}{1-\eta}\left(s(h)\epsilon_h + s(g)\left(\epsilon_g + \frac{1-\eta}{1-\gamma}\bar{\nu}\right)\right) = \frac{\bar{\epsilon}}{1-\eta}\end{aligned}$$

where $\bar{\epsilon}$ is the expenditure share weighted average income elasticity of demand parameter across housing and goods.
With the above result, I can compute the expenditure elasticity of demand for housing

$$\frac{\partial \log C(h)}{\partial \log E} = \eta + \epsilon_h\frac{\partial \log U}{\partial \log E} = \eta + (1-\eta)\frac{\epsilon_h}{\bar{\epsilon}}$$

and goods,

$$\frac{\partial \log C(g)}{\partial \log E} = \eta + (1-\eta)\left(\frac{\epsilon_h - \frac{\eta}{1-\gamma}\bar{\nu}}{\bar{\epsilon}}\right).$$

For the expenditure elasticity of demand for a particular sector in the service industry, it holds that

$$\frac{\partial \log C(j)}{\partial \log E} = (\gamma-\eta)\frac{\log P(g)}{\log E} + \eta + (\nu_j+\epsilon_g)\frac{\partial \log U}{\partial \log E} = \eta + (1-\eta)\frac{\epsilon_g+\nu_j}{\bar{\epsilon}} + (1-\eta)\frac{(\gamma-\eta)}{(1-\gamma)}\frac{\bar{\nu}}{\bar{\epsilon}}$$

Next, we can compute the mobility elasticity with respect to income. Recall

$$\lambda_n = \frac{B_n U_n^{\kappa}}{\sum_{n'} B_{n'}U_{n'}^{\kappa}}$$

So,

$$\frac{E\partial\lambda_n}{\lambda\partial E} = \kappa\frac{E\partial U}{U\partial E}\Big|_n - \frac{E\partial\Phi}{\Phi\partial E} = \kappa(1-\eta)\left(\frac{1}{\bar{\epsilon}_n} - \sum_{n'}\lambda_{n'}\frac{1}{\bar{\epsilon}_{n'}}\right)$$

These elasticities imply the following:

- Engel aggregation: $s(h)\frac{\partial \log C(h)}{\partial \log E} + s(g)\sum_j \tilde{s}(j)\frac{\partial \log C(j)}{\partial \log E} = 1$
- Conditional on prices income elasticities of demand parameters $\epsilon_g, \upsilon(q)$ are defined up to scale. Consumption choices are not affected by scaling the parameters by a constant factor. Furthermore, if κ is scaled by the same factor agents mobility choices are unaffected.
- As a result of the above I can normalize one income elasticity parameter and one taste shifter without affecting the economic choices of agents.
- Sufficient: If $0 < \eta < 1$ and $\gamma > 1$ then $\epsilon_i > 0, \forall i \in \{h, g\}$ and $\bar{\nu} < 0$ such that utility is increasing in expenditure and the inner price index is increasing in expenditure.
-
  1. $\epsilon_i = 1 - \eta, \forall i$ and $\nu_j = 0, \forall j$: preferences are homothetic nested CES, many trade models
  2. $\epsilon_i = 1-\eta, \forall i$ and $\exists \nu_j \neq 0$: upper nest is homothetic and within sectors non-homothetic, Borusyak & Jaravel (2018)
  3. $\epsilon_i \neq 1-\eta, \forall i$ and $\nu_j = 0, \forall j$, upper nest is non-homothetic and lower nest homothetic, Comin et al. (2018), Matsuyama (2018)
- In the case of homothetic upper nest ($\epsilon_g = 1 - \eta$): $\frac{U\partial E}{E\partial U} = 1 + s(g)\frac{\bar{\nu}}{1-\gamma}$

B.2 Proofs

9.2.1 Proof of Proposition 1

Proof. The proof is straightforward and can be found in a similar form in Matsuyama (2019). Recall the expression for the expenditure share on goods from sector j in location n' by household k in n and taking logs

$$\log s_{knn'}(j) = \log a_g + \log \alpha_j + (\gamma - \eta)\log P_{kn}(j) + (\eta - 1)\log I_k + (\epsilon_g + \nu_j)\log U_{kn} + (1-\gamma)\log P_n(j)$$

Taking prices and nominal income as given, I take the derivative with respect to U_{kn}

$$\frac{\partial \log s_{knn'}(j)}{\partial U_{kn}} = \frac{1}{U_{kn}}\left(\epsilon_g + \nu_j + \frac{\gamma - \eta}{1-\gamma}\bar{\nu}_{kn}\right)$$

Note that $s_{knn'}(j)$ is increasing in real consumption if $\epsilon_g + \nu_j > \frac{\gamma-\eta}{1-\gamma}\bar{\nu}_{kn}$ which captures the property that as household get richer they allocate more spending to sector with higher income elasticity. For any $\nu_1 > \nu_2$,

$$\frac{\partial \log s_{knn'}(1)}{\partial U_{kn}} - \frac{\partial \log s_{knn'}(2)}{\partial U_{kn}} = \frac{1}{U_{kn}}(\nu_1 - \nu_2) > 0.$$

This establishes log-supermodularity of $s_{knn'}(j)$ in U_{kn} and ν_j. The result holds by the same logic for $\tilde{s}_{kn}(j)$. ■

9.2.2 Proof of Corollary 1

Proof. Given $U_{high,n} > U_{low,n}$ Proposition 1 implies for any $\nu_1 > \nu_2$

$$\frac{s_{high,n}(1)}{s_{high,n}(2)} > \frac{s_{low,n}(1)}{s_{low,n}(2)}.$$

With $I_k > 0, \forall k$

$$\frac{\pi_{high,n}(1)}{\pi_{high,n}(2)} > \frac{\pi_{low,n}(1)}{\pi_{low,n}(2)}$$

where $\pi_{kn}(j) = s_{kn}(j)I_k$. We want to show for any $x_n > x'_n$ and $\nu_1 > \nu_2$

$$\frac{\pi_{high,n}(1)x_n + \pi_{low,n}(1)(1-x_n)}{\pi_{high,n}(1)x'_n + \pi_{low,n}(1)(1-x'_n)} > \frac{\pi_{high,n}(2)x_n + \pi_{low,n}(2)(1-x_n)}{\pi_{high,n}(2)x'_n + \pi_{low,n}(2)(1-x'_n)}$$

Note that the left hand side is increasing in $\pi_{high,n}(1)$ since $x_n > x'_n$. Aplying the log-spm of $\pi_{kn}(j)$ we can write

$$\frac{\pi_{high,n}(1)x_n + \pi_{low,n}(1)(1-x_n)}{\pi_{high,n}(1)x'_n + \pi_{low,n}(1)(1-x'_n)} > \frac{\frac{\pi_{low,n}(1)\pi_{high_n}(2)}{\pi_{low,n}(2)}x_n + \pi_{low,n}(1)(1-x_n)}{\frac{\pi_{low,n}(1)\pi_{high_n}(2)}{\pi_{low,n}(2)}x'_n + \pi_{low,n}(1)(1-x'_n)} = \frac{\pi_{high,n}(2)x_n + \pi_{low,n}(2)(1-x_n)}{\pi_{high,n}(2)x'_n + \pi_{low,n}(2)(1-x'_n)}$$

This completes the proof. ■

B.3 Proof of Corollary 2

Proof. The proof uses results from Athey (2002) on monotone comparative statistics of sums of log-spm functions. I can write the goods price index as

$$P(U_{kn}, x_n)^{1-\gamma} = \sum_j \underbrace{\alpha_j U_{kn}^{\nu_j}}_{=f(U_{kn},\nu_j)} \underbrace{P_n(j, x_n)^{1-\gamma}}_{=u(\nu_j,x_n)}$$

Theorem 1 in Athey (2002) states that iff $f(U_{kn}, \nu_j)$ is log-spm in U_{kn} and ν_j a.e. and $u(x_n, \nu_j)$ is log-spm in x_n and ν_j a.e. then $P(U_{kn}, x_n)^{1-\gamma}$ is log-spm in U_{kn} and x_n a.e. To show log-spm of $u(x_n, \nu_j)$ I start with equation 14 implies

$$\frac{M_n(j)}{L_n} = \frac{\pi_n(j)}{L_n f^e(j)}.$$

By corollary 1 $\frac{\pi_n(j)}{L_n f^e(j)}$ is log-spm in x_n and ν_j, hence $\frac{M_n(j)}{L_n}$ is log-spm in x_n and ν_j.[1] Specifically for $x_n > x'_n$ and $\nu_1 > \nu_2$,

$$\frac{M_n(1;x_n)}{M_n(1;x'_n)} > \frac{M_n(2;x_n)}{M_n(2;x'_n)}$$

Applying equation 12 under the assumption that shopping frictions are infinite outside n we can directly see that

$$\frac{P_n(1;x_n)^{1-\gamma}}{P_n(1;x'_n)^{1-\gamma}} > \frac{P_n(2;x_n)^{1-\gamma}}{P_n(2;x'_n)^{1-\gamma}}$$

and conclude that $P_n(j;x_n)^{1-\gamma}$ is log-spm in x_n and ν_j.

Lastly, log-supermodularity of $f(U_{kn},\nu_j)$ is given by proposition 1. Hence, I can apply therorem 1 in Athey (2002) and conclude that $P(U_{kn},x_n)^{1-\gamma}$ is log-spm in x_n and U_{kn}. ■

[1] Dividing by a positive constant $f^e(j)$ does not affect log-supermodularity.