Intercity Trade and Housing Prices in US Cities[1]

Jeffrey P. Cohen[2] and **Yannis M. Ioannides**[3]

Version: May 1, 2023

JEL codes: R00, R11, R15, F40.

Keywords: Urban growth, economic geography, trade and growth, housing

[1]We acknowledge with thanks that we benefited from very insightful conversations with Costas Arkolakis, Federico Esposito and Vernon Henderson, from data shared by Albert Saiz, and from comments by Kevin Patrick Hutchinson, Zack Hawley and other participants at our presentations at the Urban Economics Association, the ASSA/AREUEA meetings, participants at C.R.E.T.E., Milos, Greece, especially Elias Dinopoulos, and at HULM (especially Satyajit Chatterjee and Morris Davis). We are especially grateful to the following individuals for lending us data: Roberto Cardarelli, Cletus Coughlin, Morris Davis, Jacob Haas, Lusine Lusinyan, Aaron Renn, and Steven Yamarik. We remain solely responsible for the content.

[2]Center for Real Estate and Urban Economic Studies, Department of Finance, School of Business, University of Connecticut, 2100 Hillside Road, Unit 1041RE Storrs, CT 06269-1041, USA. T: 860-486-1277. F: 860-486-0349. E: Jeffrey.Cohen@uconn.edu http://www.business.uconn.edu/person/jeffrey-cohen/

[3]Department of Economics, Tufts University, Medford, MA 02155, USA. T: 617-627-3294. F: 617-627-3917. E: yannis.ioannides@tufts.edu sites.tufts.edu/yioannides/

Intercity Trade and Housing Prices in US Cities

Jeffrey P. Cohen and Yannis M. Ioannides

Abstract

Estimating how much of household income is (and should be) spent on housing, and theoretical models of housing price growth determination, are important issues in urban and housing economics. Urban models often examine the consequences of domestic trade for city structure. We develop a theoretical model resting on a three-level hierarchy: the production of each city's traded final goods uses labor, physical capital, and locally produced intermediates interpreted as specialized labor. The traded goods are used to produce a composite that is used for consumption (of housing and non-housing goods) and investment.

We test whether this model of spatial equilibrium among cities with capital mobility between them exhibits empirical evidence on housing share expenditures that are consistent with the commonly assumed share of 30 percent. Using instrumental variables (IV), and a unique set of instruments including time-varying MSA-level military contract awards and MSA net migration flows, we identify MSA-level GDP growth impacts on MSA house price growth, based on an equation that follows from imposing spatial equilibrium across cities. In general, our empirical estimation of the theoretical model confirms a positive and significant relationship between GDP growth and house price growth. We also cannot reject the hypothesis that the housing consumption share is approximately 30 percent of household income. Our theoretical approach, synthesis of MSA-level data from numerous sources, and empirical analysis, are novel.

1 Introduction

The question of how much of their income households spend, and should spend, on housing is an issue that is of importance. For instance, guidance for individuals seeking mortgages and banks making mortgage decisions rely on rules of thumb for decision making. Landlords of potential tenants, and renters as well, depend on guidance for the share of household income that can and should be spent on housing. The popular press has long suggested that 30 percent of a household's income is a reasonable share to devote to housing costs.[4] The U.S. Department of Housing and Urban Development has determined that households spending more than 30 percent on housing costs are considered "cost-burdened".[5] Housing researchers who build utility-maximizing models of housing and non-housing (that is, traded goods) consumption should be able to explain the patterns of household decision making. Empirical assessments of the validity of the theory are important for understanding how well a housing theoretical model can approximate the market decision making process.

In building such theoretical models, a first step is to recognize that cities are vibrant hubs of economic activity and culture. They host a large and ever increasing share of the economy's population and output. For a city to function its economy must provide non-tradeable goods and services, which are required for each city's survival. Cities also typically produce tradeable goods, which are exported to the rest of the economy, thus allowing their economies to import goods that are consumed by their population and industries. Urban economic activity provides employment and is accommodated by each city's real estate sector. Real estate encompasses housing and non-housing structures. Housing prices and rents are key determinants of the cost of urban production and urban living. Urban economies are profoundly open to domestic competition.

Research on housing markets and prices typically looks either at the housing market alone, or at the housing and labor markets jointly. Research on the impact of trade on the urban economy, such as by Autor, Dorn, and Hanson (2013), considers the relationship

[4] See, among others, https://www.cnbc.com/2021/07/14/how-much-of-your-income-you-should-spend-on-housing.html.

[5] See https://www.huduser.gov/portal/sites/default/files/pdf/worst-case-housing-needs-2020.pdf

between trade and the labor markets. Our research reported here is motivated by a literature that links local housing markets and trade. It innovates by bringing into the analysis some additional but lesser known sources of data, which are critical for understanding urban economies as open economies.

First, our theoretical model rests on a three-level hierarchy: the production of each city's traded final goods uses labor, physical capital, and locally produced intermediates interpreted as specialized labor. The traded goods are used to produce a composite that is used for non-housing consumption and investment. Our theoretical model with housing builds and expands upon the basic structure of the model in Chapter 9 of Ioannides (2013), which does not include housing. We also incorporate state-of-the-art assumptions in urban-macro models. No variations of the model in Ioannides (2013) chapter 9 have been empirically estimated, so we offer some empirical insights here into the validity of the model and its assumptions.

Testing the hypothesis on the share of income spent on housing consumption, predicted by our theoretical model, is the primary interest of this paper. Therefore, the empirics are clearly connected to the theory, which in turn, are based on a notion of trade. Our empirical results indicate a positive and significant relationship between house price growth and GDP growth at the MSA level. They also support the hypothesis, based on the theory, that the share of housing consumption is approximately 0.3, in our theoretical overlapping generations model where there are two goods, housing and nonhousing consumption, in two time periods of the lifecycle (young and old).

This paper's empirical urban macro approach is innovative in that it synthesizes the city-level housing price data with MSA GDP and employment data, along with the unique city-level military contracts and net migration flows, and other instruments that we have developed. One data source is the Bureau of Economic Analysis (BEA) data on MSA GDP, which became available in 2002 and is reported annually for 381 US MSAs.[6] The BEA is an internationally-respected statistical agency that is a part of the U.S. Department of

[6]http://www.bea.gov/newsreleases/regional/gdp_metro/gdp_metro_newsrelease.htm

Commerce.[7] The BEA website describes GDP by metro area as "A comprehensive measure of the economies of counties, metropolitan statistical areas, and some other local areas. Gross domestic product estimates the value of the goods and services produced in an area. It can be used to compare the size and growth of county economies across the nation."[8] A second source that we rely on is little known data on federal military procurement contracts awarded to individual establishments, and the location of work performance, which we use to develop an instrument for MSA level GDP growth. We roll up each establishment's zip codes to the MSA level in order to obtain an estimate of the value of MSA-level military contract work. We use the growth in these MSA-level military contracts, and the level of these military contracts, as instruments for GDP growth and GDP, respectively. While Nakamura and Steinsson (2014) use national-level and state-level military procurement data as an instrument, our development of an MSA-level approach for estimating military procurement values is completely novel. Furthermore, we also incorporate data on annual net migration flows at the MSA level, developed from county-level migration flows (based on federal income tax returns data from the Internal Revenue Service), as an instrument for MSA level GDP.

The remainder of this paper is organized as follows. We begin with a brief survey of the literature on trade and house prices. Then we develop our theoretical model of spatial equilibrium with multiple city types. This model predicts a structural relationship of how GDP growth affects the growth of city-level house prices. It offers some structure for how one might empirically test the hypothesis that the share of consumption on housing is approximately 0.3. The empirical implications of spatial equilibrium have been tested before when analyzing interactions among US cities [c.f. Glaeser *et al.* (2014); Glaeser and Gyurko (2017)], yet ours is the first aggregative model of the urban economy that is estimated with MSA level data. In this regard, we then describe the data and discuss our empirical results. We conclude with some overall discussion and suggestions for future research.

[7] As with any data variables collected by official statistical agencies, there is always the potential for measurement error.

[8] See: https://www.bea.gov/data/gdp/gdp-county-metro-and-other-areas

2 Literature Review

There is relatively little literature that emphasizes empirically the structural implication of intercity trade, city output, and house price growth. Much of it pertains to either city trade or house prices, but less focus has centered on both together with gross domestic product.

Several applications have been published of approaches to estimate the external shocks to a city's economy, which is traced back to an exports price index first developed by Pennington-Cross (1997). But much of the related literature pertains to international trade rather than intercity trade. Other subsequent applications include Hollar (2011), which is a study on central cities and suburbs; Larson (2013), which considers housing and labor markets in growing versus declining cities; and Carruthers *et al.* (2006) on convergence. A more recent paper by Li (2017) uses a rudimentary empirical analysis to motivate a theoretical model and simulations for US cities that describe the relationship between house prices and comparative advantage. The theory is the primary focus of that paper. Our understanding of export-oriented cities would benefit from further analytical and empirical attention, together with fewer limitations of some of the other exports data sources. This is in view of the sparseness of published research integrating theoretical underpinnings with rigorous empirical modelling on house prices, GDP, and mobility of goods and residents at the MSA level.

A more promising approach would build on Rosen (1979) and Roback (1982). They develop models that have been more recently used to describe spatial equilibrium in housing markets, as in Glaeser *et al.* (2014), which tests a MSA-level model, and Glaeser and Gyurko (2017), which explores the implicatins of spatial equilibrium by means of numerical calibrations. Both those articles aim at a number of key stylized facts of urban housing markets, that is, positive serial correlation of price changes at one-year frequencies and mean reversion over longer periods, strong persistence in construction, and highly volatile prices and construction levels within markets over time. The spatial equilibrium approach typically assumes perfect labor mobility and fixed land. Spatial equilibrium within each city implies that land rent is function of distance to the city center, and transport costs within the city.[9]

In the real world, cities are not autarkic and typically interact with other cities. Ioannides (2013), Ch. 7 and 9, develops standard approaches for modeling interactions among systems of cities, relying on the notion of spatial equilibrium.[10] Duranton and Puga (2014) offer a comprehensive treatment of growth in a system of cities that also links theoretical predictions with empirical specifications.

In spatial equilibrium models, estimating intercity shipping costs has become an issue of interest. The Donaldson and Hornbeck (2016) pioneering work in this area uses intercity distances to proxy for shipping costs and emphasizes market access [Fujita *et al.* (1999)]. The concept of access to highway and/or road infrastructure as has been explored in a number of studies after Donaldson and Hornbeck (2016) systemitizes the role of geography. In this large follow-up literature on access, highway access for a given residential property is defined as the drive time from that property to the nearest highway exit, with the implication that the highway provides access to various employment and goods markets. In the European context, Hoogendoorn *et al.* (2019) consider a new highway tunnel that was opened in

[9] The spatial equilibrium condition, which expresses arbitrage, turns out to have major implications for urban growth equations in the context of economic integration. These consequences have been emphasized recently by Hsieh and Moretti (2019). They show empirically that spatial equilibrium introduces dependence among city growth rates, which makes the contribution of a particular city to aggregate growth differ significantly from what one might naively infer from the growth of the city's GDP by means of a standard growth-accounting exercise. They show that the divergence can be dramatic. E.g., despite some of the strongest rate of local growth, New York, San Francisco and San Jose were only responsible for a small fraction of U.S. growth during their study period. By contrast, almost half of aggregate US growth was driven by growth of cities in the South. This divergence is due to the fact that spatial equilibrium imposes restrictions on city-specific TFP growth rates. Future work might condsider both international and intercity trade in the Hsieh and Moretti (2017) context.

[10] In a related literature originating with Henderson (1974), city types differ according to the number and types of final goods produced, or whether or not they produce only intermediate goods and import all final goods. Ioannides (2013), Chapter 7, develops a variety of rich urban structures in a static context and *ibid.,* Chapter 9, in a dynamic one. Both approaches impose intracity and intercity spatial equilibrium. In the case of the static model, manufactured goods may be either produced locally or imported from other cities. Manufactured goods are produced using raw labor and intermediate goods interpreted as specialized labor, which are themselves produced from raw labor, using increasing returns to scale (IRS) technologies. In the case of the dynamic model, manufactured goods are produced using raw labor and intermediate goods interpreted as specialized labor, which are themselves produced from raw labor, using increasing returns to scale (IRS) technologies, and physical capital. In either case, those goods are combined locally to produce a final good that may be used for either consumption or investment. Urban functional specialization, rather than sectoral, as articulated by Duranton and Puga (2005), also leads to structural differences. In other words, certain economic functions, like management, research and development and corporate headquarters may be located in different places than manufacturing.

2003 in the Netherlands, and how access by road associated with the tunnel impacted house prices. They consider housing accessibility to a highway improvement in terms of travel time savings, along with a variety of distance decay functions, and find an access elasticity of 0.8. [11] House prices are significantly impacted by better access. Using a difference-in-differences estimation approach, Levkovich *et al.* (2016) study the house price impacts of new highways in the Netherlands. Levkovich *et al.* (2016) also develop an access index, based on the population density at various locations, and find that the elasticity of house prices with respect to access is roughly 1.76. In other words, improved market access has a greater than proportional increase in house prices. In our context, we empirically proxy for market access with one of our instrumental variables (IV), that is, the number of planned highway miles, based on data available at the MSA level in Baum-Snow (2007). This is our instrument for the spatial complexity term that we define in more detail in the theory section below.

In light of one of our goals of examining the determinants of MSA-level house price growth, another challenge is identifying the causal relationships between house price growth and GDP growth at the city level. We propose a novel instrument, which is the growth rate of MSA-level military contracts. The idea of military growth contracts as an instrument has been used in other contexts, such as Nakamura and Steinsson (2014), although their approach is to use more readily-available state-level or national-level military contracts data. They note that others have argued national-level military expenditure is exogenous to the business cycle, and state-level expenditures are similarly exogenous to state-level output because the military does not engage in a build-up in response to states' business cycle conditions. We further build upon the Nakamura and Steinsson (2014) approach with our method of rolling up project-level data using zip code data of the work performance locations of establishments to which these federal awards were granted. Specifically, Nakamura and Steinsson (2014) interact national-level procurement data with state-level dummies as one approach to instrumentation for national-level GDP; and a second IV approach following the general idea of Bartik (1991), with which Nakamura and Steinsson (2014) scale national spending for each state by the ratio of state-level military spending at a point in time to

[11] Their access index definition is based on a function of travel time by car to a particular destination and the number of employment opportunities available at that destination.

the state's average output in the first 5 years of their sample. Our problem is more complex with several endogenous variables and multiple instruments, but our MSA-level estimates of military expenditures allow for a MSA-level GDP instrument that has not been previously explored.

But there are other factors expected to be correlated with GDP growth. Specifically, migration is another important piece of this puzzle. Card (2009) uses immigration as an IV for regional wages and inequality. This literature motivates our exploration of growth of annual MSA net migration flows as an additional IV for the MSA GDP growth variable in our empirical model.

3 Trade and the Housing Market

Drawing on standard approaches for modeling interactions among systems of cities, the present paper works with an aggregate model of an economy as being made up of cities of different types. Types differ according to the number and types of final goods produced, or whether or not they produce only intermediate goods and import all final goods. This literature originated by Henderson (1974). The present paper draws from Ioannides (2013), Chapter 7, which develops a variety of rich urban structures in a static context and *ibid.*, Chapter 9, which obtains its dynamic counterpart.[12] Although cities are different because of their specialization, individuals are indifferent as to where they locate, that is spatial equilibrium holds. Each city specializes in the production of a different tradeable manufactured good,[13] which uses raw labor, physical capital, and intermediate goods interpreted as specialized labor, which are themselves produced from raw labor, using increasing returns to scale (IRS) technologies.[14] The tradeable manufactured goods are combined in each city to produce a final composite good that may be used for either consumption or investment.

An overview of our theoretical approach is as follows. The economy consists of many cities

[12]This main model of Ioannides, Ch. 9, constitutes an original adaptation of Ventura (2005)'s model of global growth to the urban structure of a national economy by building on key features of Ioannides (2013), Ch. 7.

[13]The specialization may be functional rather than sectoral, as articulated by Duranton and Puga (2005).

[14]See Appendix A for details.

which are differentiated in terms of production. Each city type produces a different tradeable manufactured good, which uses raw labor, physical capital and a range of intermediates interpreted as specialized labor. There is free mobility of capital and mobility. As Fujita and Thisse (2009), p. 113, emphasize, while the mobility of capital is driven by differences in nominal returns, workers move when there is a positive difference in utility (real wages). In other words, differences in living costs matter to workers but not to owners of capital. The economy is populated by two overlapping generations of individuals, who maximize their lifetime utility by choosing their lifetime consumption of rental housing and non-housing consumption of a composite good that is produced in each city by combining all tradeable manufactured goods. This is a fully micro-founded model whose implications under spatial equilibrium constitute the basis of our empirical approach. This article is the first empirical estimation of a fully micro-founded life cycle consistent urban macro growth model.

3.1 A Model of Urban Economic Integration and Specialization

The exposition that follows extends the main model in Ioannides (2013), Chapter 9, in order to allow for housing.[15] It is a dynamic model that allows for differences across cities of local congestion parameters κ_i , although for empirical implementation it makes sense to assume no heterogeneity in κ_i. It assumes that individuals are free to move within and across cities, that is, spatial equilibrium is imposed in terms of individuals' lifetime utilities. This implies in turn conditions on intercity housing price and income growth patterns. The paper also assumes perfect capital mobility, which implies that nominal returns are equalized across all cities. Indeed, this assumption has a major implication, namely that the growth process in all cities depend on national physical capital accumulation.

This section aims at obtaining a more general expression for spatial equilibrium and its implications for the growth rate of the price of housing. A number of individuals $\bar{N}_t$ are born every period and live for two periods. The economy has the demographic structure of the overlapping generations model, the workhorse of modern macroeconomics. We assume that

[15] While that model is explicated for simplicity in terms of two types of cities, it may be restated in terms of a continuum of goods.

individuals born at time t work when young, consume non-housing consumption in the form of the composite aggregate that is produced locally with all traded manufactured goods and of rental housing, $(C_{1,t}, G_{1,t})$, respectively, using their labor income net of their savings in period 1. They consume again when they are old in period 2, $(C_{2,t+1}, G_{2,t+1})$ respectively, out of their total resources. They leave no bequests. We assume Cobb-Douglas preferences over first- and second-period consumption bundles for the typical individual,

$$U_t = S^*[C_{1,t}^{1-\beta} G_{1,t}^{\beta}]^{1-S}[C_{2,t+1}^{1-\beta} G_{2,t+1}^{\beta}]^{S},\ 0 < S < 1, \tag{1}$$

where $S^* \equiv [S^{1-\beta}\beta^{\beta}]^{-S}[(1-S)^{1-\beta}\beta^{\beta}]^{-(1-S)}$, and parameter S satisfies $0 < S < 1$. Note that this formulation subsumes time preference, as it is not critical for our analysis.[16]

Labor, net of commuting time, supplied by the young generation in a particular city at t is given by

$$H_t = N_t\left(1 - \kappa N_t^{\frac{1}{2}}\right),$$

with N_t the number of the members of the young generation at t, $\kappa \equiv \frac{2}{3}\pi^{-\frac{1}{2}}\kappa'$, and κ' the time cost per unit of distance traveled. It follows that employment per capita is given as $\left(1 - \kappa N_t^{\frac{1}{2}}\right)$.

Let W_t denote the wage rate per unit of time. Spatial equilibrium within the city obtains when labor income net of land rent is independent of location. This along with the assumption that the opportunity cost of land is 0, and therefore the land rent at the fringe of the city is also equal to 0, yields an equilibrium land rental function; see Chapter 7, Ioannides (2013). It declines linearly as a function of distance from the CBD and is proportional to the contemporaneous wage rate, W_t. It is convenient to close the model of a single city and to express all magnitudes in terms of city population, N_t, to be referred to as size, too. We again assume that all land rents in a given city are redistributed to its residents when they are young, in which case total rental income may be written in terms of the number of young residents as $\frac{1}{2}\kappa W N_t^{\frac{3}{2}}$. This yields first period net labor income per young resi-

[16] An additively separable version of (1) readily follows by taking logs. However, it is equivalent to the levels-version. Furthermore, uncertainty could be dealt with by taking logs.

dent, after redistributed land rentals are added and net of individual commuting costs, of $\left(1-\kappa N_t^{\frac{1}{2}}\right) W_t$. With a given wage rate, individual income declines with city size, entirely because of congestion. But, there are benefits to urban production which are reflected on the wage rate.

The canonical model of an urban economy assumes that capital is free to move. Thus, nominal returns to capital are equalized across all cities. The model assumes that young individuals are free to move, which in the context of our two-overlapping generations requires that lifetime utility is equalized across all cities. By using these conditions simultaneously, we obtain a relationship between housing prices, consumption good prices and nominal incomes across cities, which may be taken to the data.

With free capital mobility, an economy-wide wide nominal return to physical capital, K_{t+1}, is established. This nominal return to phyisical capital is R_{t+1}, each time period $t+1$. All capital is held by the members of young generation at time t. The price vector $(P_{i,t}, P_{i,G,t})$ is composed of the price of composite consumption good, and the (rental) price of housing in city i at time t, respectively. The indirect utility function corresponding to (1) is:

$$V_{i,t} = R_{t+1}^{S(1-\beta)} P_{i,t}^{-(1-S)(1-\beta)} P_{i,G,t}^{-(1-S)\beta} P_{i,t+1}^{-(1-S)(1-\beta)} P_{i,G,t+1}^{-S\beta} \left(1-\kappa N_t^{\frac{1}{2}}\right) W_t. \tag{2}$$

We assume that capital depreciates fully in one period. The young maximize utility by saving a fraction S of their net labor income. The productive capital stock in period $t+1$, K_{t+1}, is equal to the total savings of the young at time t. Therefore, the law of motion may be written as: $K_{t+1} = SN_t \left(1-\kappa N_t^{\frac{1}{2}}\right) W_t$.

We refer to the case where capital and labor are free to move as *economic integration*. With economic integration, industries will locate where industry productivities, the industry-specific TFP functions Ξ_{Jt}'s,[17] are the most advantageous, and capital will seek to locate so as to maximize its return. Unlike the consequences of economic integration as examined by Ventura, *op. cit.,* where aggregate productivity is equal to the most favorable possible in the economy, here urban congestion may prevent industry from locating so as to take

[17] See Appendix A for details on the specification of the urban production structure and clarification of the role of the industry-specific TFP functions Ξ_{Jt}'s.

greatest advantage of locational factors alone. Put differently, free entry of cities into the most advantageous locations may be impeded by competing uses of land as alternative urban sites, at the national level. However, utilities enjoyed by city residents at equilibrium do depend on city populations, and therefore, spatial equilibrium implies restrictions on the location of individuals. We simplify the exposition by assuming that all cities have equal unit commuting costs κ.

3.1.1 Spatial Equilibrium

We derive next the implications of spatial equilibrium in our dynamic economy. It is simplest to work with two types of cities and therefore two traded goods. Let Q_{Xit}, Q_{Yjt} denote the total quantities of the traded goods X, Y produced by cities i, j, that specialize in their production, respectively. The formulation is symmetrical for the two city types, and therefore, we work with a city of type X. This derivation may be extended to many city types and goods. We suppress redundant subscripts and write for the nominal wage and the nominal gross rate of return in a type$-X$ city:

$$W_{Xt} = (1 - \phi_X)\frac{P_X Q_X}{H_X},\ R_{Xt} = \phi_X \frac{P_X Q_X}{K_X}, \tag{3}$$

where P_X denotes the local price of traded good X, which is expressed in terms of the local price index, the numeraire, which is equal to one in all cities. We also assume initially that there are no intercity shipping costs for traded goods. With economic integration, the gross nominal rate of return is equalized across all city types, that is:

$$R_t = R_{Xt} = R_{Yt}.$$

Spatial equilibrium for individuals requires that indirect utility, (2), be equalized across all cities. In view of free capital mobility, spatial equilibrium across cities of different types requires that:

$$P_{X,t}^{-(1-S)(1-\beta)} P_{X,G,t}^{-(1-S)\beta} P_{X,t+1}^{-S(1-\beta)} P_{X,G,t+1}^{-S\beta} \left(1 - \kappa N_{Xt}^{\frac{1}{2}}\right) W_{Xt}$$

$$= P_{Y,t}^{-(1-S)(1-\beta)} P_{Y,G,t}^{-(1-S)\beta} P_{Y,t+1}^{-S(1-\beta)} P_{Y,G,t+1}^{-S\beta} \left(1 - \kappa N_{Yt}^{\frac{1}{2}}\right) W_{Yt} \tag{4}$$

By taking logs we have:

$$-(1-S)(1-\beta)\ln P_{X,t} - (1-S)\beta \ln P_{X,G,t} - S(1-\beta)\ln P_{X,t+1} - S\beta \ln P_{X,G,t+1} + \ln\left(1 - \kappa_X N_{Xt}^{\frac{1}{2}}\right) + \ln W_{Xt}$$

$$= -(1-S)(1-\beta)\ln P_{Y,t} - (1-S)\beta \ln P_{Y,G,t} - S(1-\beta)\ln P_{Y,t+1} - S\beta \ln P_{Y,G,t+1} + \ln\left(1 - \kappa_Y N_{Yt}^{\frac{1}{2}}\right) + \ln W_{Yt}. \tag{5}$$

By rearranging this equation we obtain a condition for spatial equilibrium for each city, relative to cities of other types. Without loss of generality, we refer to the other city generically as the average city, n. In other words, spatial equilibrium is expressed for city i, relative to the urban economy:

$$GR_{t+1,t}(P_{i,G}) - GR_{t+1,t}(P_{G,\mathrm{n}}) = -\frac{1-\beta}{\beta}\left[GR_{t+1,t}(P_i) - GR_{t+1,t}(P_{\mathrm{u}})\right] + \frac{1}{S\beta}\left[GR_{t+1,t}\Upsilon_i - GR_{t+1,t}(\Upsilon_{\mathrm{n}})\right]$$

$$-\frac{1}{S}\ln\left[\frac{P_{i,G,t}}{P_{n,G,t}}\right] - \frac{1-\beta}{S\beta}\ln\left[\frac{P_{i,t}}{P_{n,t}}\right] - \left[\ln\left(1 - \kappa_i N_{i,t}^{\frac{1}{2}}\right) - \ln\left(1 - \kappa_{\mathrm{n}} N_{\mathrm{n},t}^{\frac{1}{2}}\right)\right], \tag{6}$$

where $GR_{t+1,t}$ stands for the growth rate between period t and t+1, Υ_i is income (or GDP) per capita in city i and Υ_{n} is national income per capita, $P_{i,G}$ is the price of housing, G, in city i, and $P_{G,\mathrm{n}}$ is the national average price of housing; $P_{i,t}$ is the goods price in city i, and $P_{n,t}$ is the national average of the goods price; and the last term, $\left[\ln\left(1 - \kappa_i N_{i,t}^{\frac{1}{2}}\right) - \ln\left(1 - \kappa_{\mathrm{n}} N_{\mathrm{n},t}^{\frac{1}{2}}\right)\right]$, is the difference between the local and national spatial complexity terms. Clearly, broadly similar empirical models, obtained from simpler behavioral models, may be nested within (6). In particular, the coefficient of $GR_{t+1,t}(P_j) - GR_{t+1,t}(P_{\mathrm{j,u}})$, the growth rate of the city price index relative to a national average, is predicted to be positive and greater than 1; the coefficient of $GR_{t+1,t}\Upsilon_i - GR_{t+1,t}(\Upsilon_{\mathrm{n}})$, the growth rate of GDP relative to a national average, is predicted to be positive and greater than 1.

We impose a simplifying assumption on the exponents S and β in the utility function in (1) above. Recall that S is share of lifetime income spent on all consumption when old, and β is share of income in a given time period spent on housing relative to consumption of other goods. We assume they are equal.[18] This simplifies our (6) dramatically (and also alleviates multicolinearity in the empirical specification) as follows:

$$
\begin{aligned}
&GR_{t+1,t}(P_{i,G}) - GR_{t+1,t}(P_{G,\mathrm{n}}) = \\
&\frac{1}{\beta}\left[\ln\left[\frac{P_{i,t}}{P_{n,t}}\right] - [GR_{t+1,t}(P_i) - GR_{t+1,t}(P_{\mathrm{u}})] - \ln\left[\frac{P_{i,G,t}}{P_{n,G,t}}\right]\right] + \\
&\frac{1}{\beta^2}\left[[GR_{t+1,t}\Upsilon_i - GR_{t+1,t}(\Upsilon_{\mathrm{n}})] - \ln\left[\frac{P_{i,t}}{P_{n,t}}\right]\right] + \\
&[GR_{t+1,t}(P_i) - GR_{t+1,t}(P_{\mathrm{u}})] + \left[\ln\left(1-\kappa_i N_{i,t}^{\frac{1}{2}}\right) - \ln\left(1-\kappa_{\mathrm{n}} N_{\mathrm{n},t}^{\frac{1}{2}}\right)\right].
\end{aligned}
\tag{7}
$$

Since the theory implies the coefficient equals 1 on the third term in equation (7), and therefore it does not involve any parameter to be estimated, we rewrite (7) as follows:

$$
\begin{aligned}
&GR_{t+1,t}(P_{i,G}) - GR_{t+1,t}(P_{G,\mathrm{n}}) - [GR_{t+1,t}(P_i) - GR_{t+1,t}(P_{\mathrm{u}})] = \\
&\frac{1}{\beta}\left[\ln\left[\frac{P_{i,t}}{P_{n,t}}\right] - [GR_{t+1,t}(P_i) - GR_{t+1,t}(P_{\mathrm{u}})] - \ln\left[\frac{P_{i,G,t}}{P_{n,G,t}}\right]\right] \\
&+\frac{1}{\beta^2}\left[[GR_{t+1,t}\Upsilon_i - GR_{t+1,t}(\Upsilon_{\mathrm{n}})] - \ln\left[\frac{P_{i,t}}{P_{n,t}}\right]\right] \\
&+\left[\ln\left(1-\kappa_i N_{i,t}^{\frac{1}{2}}\right) - \ln\left(1-\kappa_{\mathrm{n}} N_{\mathrm{n},t}^{\frac{1}{2}}\right)\right].
\end{aligned}
\tag{8}
$$

The last term in brackets in the right hand side above proxies for spatial complexity, regulation, and housing supply factors. It may be approximated as:

$$
\kappa\left[\ln\left[1-N_{i,t}^{\frac{1}{2}}\right] - \ln\left[1-N_{\mathrm{n},t}^{\frac{1}{2}}\right]\right].
$$

Our final estimation equation is given by Eq. (8) above, but with the simlification of the spatial complexity term described immediately above. We anticipate the coefficients to be positive on the first two terms on the right side of this equation. Clearly, GDP growth, the

[18]This assumption is reasonable given that they are both likely around 0.3 in the U.S.

ratio of house prices in levels, the ratio of goods prices in levels, and spatial complexity all are anticipated to be endogenous. We discuss our IV approach in the data section below. Last, we note that the above framework was chosen for ease of exposition but may be expanded to allow for a large number of goods associated with a large number of cities. In that case, spatial equilibrium is defined in terms of the entire range of goods traded.

Finally, many urban models factor in amenities, such as schools, parks, and environmental quality, all of which vary across cities. We address this in our empirical model by including fixed effects for MSAs, which absorb the amenity effects on house price growth.[19] Since our main focus is on estimating the share of housing expenditures rather than the impacts of amenities, this is a reasonable approach to simplify the theory in a relatively tractable manner. Also, a helpful referee pointed out that including year fixed effects, as we do in our empirical analysis, implies any elasticities are short-run estimates, which is sensible in our context because we are estimating our model over a finite period of time.

4 Overview of Data

We have assembled data from a variety of sources, which we use as comprehensively as possible to investigate the parameter β in the theoretical model. We describe these data to provide an overall view of the empirical resources we bring to our approach.

4.1 Data Sources

In this section we describe the major sources of data, which we merge and append into one large dataset that we use for the estimation of Eq. (8). The combining of MSA data from these sources lead us to an unbalanced panel dataset, comprised of 197 MSAs, with annual data for the period 2003-2017. Since the dataset is unbalanced, we end up with 2,587 observations in the OLS estimation and 1,020 observations in the IV estimation. One important note is that since we combine MSA-level data from many different sources, several

[19]See, for instance, Glaeser et al. (2005).

of which use their unique definitions of the cities included in the respective MSAs, it is not always possible, without imposing some stringent and unreasonable assumptions, for us to merge data from all sources for all MSAs. This is particularly problematic for the largest MSAs that include many surrounding cities/suburbs. Due to these difficulties related to data merging, we drop some of the largest MSAs from our sample (such as New York City, Boston, Los Angeles, and Chicago).[20] Therefore, our analysis primarily focuses on mid-sized and small MSAs.

Specifically, we draw our data for the primary variables in Eq. (8) from the following sources:

- House Price Index, ($P_{i,G}$): The Freddie Mac HPI data for each MSA is used as our house price index measure. In Tables 2 and 3 below, "House Price Growth" stands for the growth rate of the House Price Index, $GR_{t+1,t}(P_{i,G})$, for a given MSA i at a point in time (that is, year), t.
- Payroll Employment, $\ln\left[1 - N_{i,t}^{\frac{1}{2}}\right]$: We use annual payroll employment data, as reported by the Bureau of Labor Statistics, for each MSA. We also construct an average employment variable for all MSAs. In the tables below, "MSA employment" stands for the MSA employment level at a given year; "MSA employment" stands for $\ln\left[1 - N_{i,t}^{\frac{1}{2}}\right]$, and "national average employment" stands for $\ln\left[1 - N_{\mathrm{n},t}^{\frac{1}{2}}\right]$.
- Gross Domestic Product per capita in city i, Υ_i: Annual (nominal) GDP for MSAs,

[20] We acknowledge it would strengthen the paper if we could include all MSAs in the U.S. We have spent a significant amount of time and effort exploring this, and there are some issues that we explain here that hopefully will make the challenges more clear. If all the data were collected and obtained at the zip code or Census tract levels, then it would be straightforward to do some re-coding and roll-up the estimates to one common MSA definition. In fact, this is what is done for the Internal Revenue Service migration data, and the military contracts data from the federal government. But that is not how the data are reported everywhere. Some of the different federal agencies and organizations (e.g., the Census Bureau; the Bureau of Economic Analysis; Freddie Mac; the Bureau of Labor Statistics; the highways data from Baum-Snow (2007); and the unavailable land area from Saiz (2010)) use different definitions of the cities included in the largest MSAs, and also the MSA definitions themselves vary over time for some agencies. The smaller MSAs tend not to change definitions much if at all, with perhaps the exception of a handful of MSAs (i.e., at most 3 or so) across the various data vintages. Without obtaining the disaggregated data (e.g., zip code or tract level) for each variable from every source, it is not straightforward to do a comprehensive data merge in a reliable manner.

obtained from the BEA, in millions of \$.[21] In Tables 1 and 2 below, "MSA GDP growth" stands for $GR_{t+1,t}\Upsilon_i$.

- Goods Prices (P_i) : There are no MSA level goods prices for a comprehensive set of U.S. cities. Therefore, we use regional goods price indexes from the four regions of the U.S. Bureau of Labor Statistics, to proxy for MSA level goods prices. One advantage is that for reasonably small MSAs, a regional price is likely taken as given by the MSA, thereby alleviating some of the potential endogeneity concerns of the goods prices. "MSA cpi" in Tables 2 and 3 represent the CPI estimate for the region in which a particular MSA is located; and "cpi national average" is the average of the regional CPIs for all regions, weighted by the number of MSAs in each region.

- Prices Term: In Tables 2 and 3, the first regressor, denoted as "Prices Term", stands for:

$$\left[\ln\left[\frac{P_{i,t}}{P_{n,t}}\right] - [GR_{t+1,t}(P_i) - GR_{t+1,t}(P\mathrm{u})] - \ln\left[\frac{P_{i,G,t}}{P_{n,G,t}}\right]\right]$$

- GDP Term: In Tables 2 and 3, the second regressor, denoted as "GDP Term", stands for:

$$\left[[GR_{t+1,t}\Upsilon_i - GR_{t+1,t}(\Upsilon\mathrm{n})] - \ln\left[\frac{P_{i,t}}{P_{n,t}}\right]\right]$$

- Spatial Complexity Term: In Tables 2 and 3, the third regressor, denoted as "Spatial Complexity Term", stands for:

$$\left[\ln\left[N_{i,t}^{\frac{1}{2}}\right] - \ln\left[N_{\mathrm{n},t}^{\frac{1}{2}}\right]\right].$$

Data for our instruments are as follows:

- Military Contracts Awarded: As described in the literature review section above, other researchers, including Nakamura and Steinsson (2014), have used state-level military contracts data as an instrument for GDP. Their identifying assumption relies on the fact that "the U.S. does not embark on a military build-up because states that receive

[21] Missing MSAs were included manually from https://apps.bea.gov/iTable/index_regional.cfm

a disproportionate amount of military spending are doing poorly relative to other states." Our approach is a significant innovation in that we have obtained data on the exact locations (including zip codes) of the contracting entities for all awarded military contracts, from the years 2003-2017, obtained from the Federal Procurement Data System. [22] We aggregate the individual contract award amounts by zip codes in each year, and then roll up the zip code level data to the MSA level. This leads us to an estimate of the value of military contracts awarded by MSA in each year. We rely on an argument similar to that of Nakamura and Steinsson (2014), since the military does not choose to engage in build-ups because of some MSAs doing "worse" or "better" than others in terms of their economic activities (MSA-level GDP).

- Planned Highway Miles: We use the number of planned highway miles, normalized by population (in millions) in each year, at the MSA level, from Baum-Snow (2007).[23] We use the Baum-Snow (2007) "leccna" variable, as our estimate for planned highway miles in all years of our sample, which we subsequently normalize by the population (millions) in each year. Cities that ship more goods domestically are expected to rely heavily on the highway network (see, for example, Duranton et al., 2014), which was planned many years ago. For this reason, the size of the planned highway network (per person, which varies over time) in a particular city is used as an instrument for that city's spatial complexity term; [24] a larger network in the city should lead to higher complexity, better access, and in turn, be a valid instrument for the spatial complexity term (which is essentially an employment estimate, following the exposition in the theory section above). This variable also represents the congestion and/or roads quality and accessibility within each city. As described above in the literature review section, highway access is an important issue in the urban economics literature, and

[22](https://www.fpds.gov/fpdsng_cms/index.php/en/, accessed in January 2020). Note that some entries are negative because the contract awards include funds that were returned to the federal government for non-performance or other reasons, therefore there is substantial variation in our MSA-level estimates. We drop outliers by eliminating observations for MSAs with military expenditures year over year growth rates above 250 percent.

[23]Details on how this variable is constructed can be found in Baum-Snow (2007), pages 802-803.

[24]While we use the estimate without year-to-year variation in the number of planned highway miles by MSA from Baum-Snow (2007), normalizing by population leads to annual variation and also offers a more precise instrument for the GDP measure that is in per-capita terms.

this IV approach is how we tie access into the model. This instrument is expected to be uncorrelated with shocks to city-level spatial complexity because they pertain to past plans for highway rays and miles that were in the original plan (from 1947). Shocks to employment around 60 years later should be uncorrelated with the original plans. Our focus on highways that were in the original plan enables us to avoid the complications of new plans for highway construction, which more likely would be considered to be correlated with "shocks" to employment. For instance, while a new decision to build another highway would be expected to be correlated with a city's domestic employment, it also can be considered a shock to a city's current employment if the new plan is unexpected. Therefore, focusing on highways that were in the original plan from the 1940's (as opposed to more recent plans) leads to a credible instrument for current employment (that is, spatial complexity).

- Net Migration: We use annual MSA to MSA net migration data as an instrument.[25] The net migration between MSA i and all other MSAs is calculated for each MSA i, and this calculation is repeated for each MSA in each year of the sample. These net migration estimates for each MSA are used as an instrument for the "GDP Growth Term".

- Undevelopable Land Area: We utilize the Saiz (2010) data on MSA unavailable land area, and we normalize this by the MSA resident population (in millions) in each year (from the Census Bureau). This allows for time-varying intensities of the undevelopable land, as land that is not usable in a relatively densely populated area is expected to be worth more than in a rural area. The ratio of the city to the national average developable land per capita is an instrument for the ratio of city to national house prices in the "Prices Term", of the final form of our model.

[25] Aaron Renn provided the migration inflows and outflows data aggregated up to the MSA level, which we use to calculate the net migration flows in each year. The inflows and outflows data originated from the U.S. Internal Revenue Service (IRS)

4.2 Descriptive Statistics

Table 1 presents descriptive stats for the data used for the period of 2003-2017 in the regression for Eq. (8) above. The first column of Table 1 reports the national averages for all MSAs. The average GDP growth rate was 3.73 percent, the average house price index growth rate was approximately 2.76 percent. Military contracts grew at an average of 9.25 percent per annum.[26] There were 42.5 planned highway miles per million population in the average MSA in our sample. The average unavailable land index was 0.71 percent per million population. The average number of payroll employment over all MSAs in all years was approximately 450,000 workers.

Table 1 also reports descriptive statistics for 4 regions of the U.S. First, it is noteworthy that cities in the West had the highest average GDP growth rate and house price growth rate, both at slightly over 4 percent, while MSAs in the South had the second highest average GDP growth rate and house price growth rate, with about 4 percent and 3 percent, respectively. The MSAs in the Midwest had the lowest average rate of growth of military contracts, at about 11 percent. The CPI also grew fastest, on average, in cities in the West. Cities in the East had the highest average employment, at about 770,000 on average, while employment was lowest in cities in the Midwest where it was roughly half the number of the East, on average.

5 Estimation Results

The spatial equilibrium, Eq. (8), dictates our choice of variables in the empirical analysis. Our objective is to determine an empirical estimate of the size and sign of the parameter β. For the estimation of Eq. (8), the first term of the dependent variable is the difference between the housing price growth rate in MSA i and the national housing price growth rate. The second term in the dependent variable is the difference between the MSA goods price

[26] Note that there are large fluctuations in the military contracts awarded and in some cases, when awarded funds are not expended in a given year, a portion of those funds are returned in a subsequent year (and counted as negative expenditures). This leads to a large amount of variation in the data, and for this reason we omitted some extreme outliers.

growth rate in city i and the national goods price growth rate. This essentially leads to a measure of the difference in local to national house price growth, in real terms.

The first independent variable, which we refer to as the "Prices Term", is the first term in brackets on the right side of Eq. (8). Similarly, the second term in brackets on the right side of Eq. (8) is referred to in the tables below as the "GDP Term". The third term on the right side of Eq. (8) comprises an additional regressor that is proxy for spatial complexity, regulation, and housing supply factors at the MSA-level and national level. This parsimonious specification helps minimize the potential multicolinearity that is present with a more general specification that would need to be estimated without the assumption that $\beta = S$. This version of Eq. (8) is first estimated by OLS with fixed effects for each of the years and MSAs in the sample, and the results are presented in column 1 of Table 3. In these regression results, some of the regressors have the incorrect sign. But the OLS specification ignores potential endogeneity, which might explain why these results are not intuitive based on the predictions of Eq. (8). In other words, it is possible that endogeneity of the underlying variables in these regressors could be leading to biased estimates, so endogeneity bias might lead us to reject the theory underlying Eq. (8) in terms of the signs and significance of the parameter estimates. To investigate this possibility, below we explore an IV approach.

As a starting point for the IV model, we explore the first stage regressions for each endogenous variable (i.e., the "Prices Term" with parameter β^{-1} and the "GDP Term" with parameter β^{-2}). The "Spatial Complexity Term" is also endogenous and therefore we instrument for that term as well, although our hypothesis on β are not directly dependent on this spatial complexity estimate. Table 2 presents the first stage regressions for each of the endogenous regressors. In the first column (for the "Prices Term"), the instrument for the "Prices Term", given as

$$\ln(\text{MSA cpi/cpi national average}) - (\text{MSA cpi growth} - \text{national average of cpi growth})$$

$$-(\text{MSA unavailable land per capita/national average of unavailable land per capita}),$$

is statistically significant. The "GDP Term" instrument, which is given as

$$\text{(MSA military contracts growth} - \text{national average of military contracts growth)}$$

$$+\text{(MSA net migration growth} - \text{national average of net migration growth)}$$

$$-\ln\text{(MSA cpi/cpi national average)},$$

is statistically insignificant in this first stage regression for the "Prices Term". The planned highway miles per capita is significant, as is the instrument for the "Spatial Complexity Term". Importantly, the F-statistic implies that all of the variables in the model are highly jointly significant. Also, the R^2 for this first stage regression of the "Prices Term" is quite high, and equal to 0.933. All of this offers strong evidence that the instruments for the "Prices Term" are valid.

For the "GDP Term" first stage regression, shown in column 2 of Table 2, none of the instruments are individually significant. While at first glance this could be concerning, the F-statistic $p-$value is very small (equal to 0.0008), implying joint significance of all of the instruments for the "GDP Term". Witten et. al. (2013) note this type of regression result is not uncommon when there is multicolinearity among some of the instruments. In fact, we find that the net migration instrument and the military contracts instrument have a correlation of approximately 0.28, while the planned highway miles instrument and the military contracts instrument have a correlation of approximately -0.27. This could be leading to multicolinearity that inflates the standard errors in this first stage regression, but nevertheless the instruments are likely valid, as evidenced by the low $p-$value on the F-statistic.

Finally, for the "Spatial Complexity Term", the net migration instrument is a significant determinant of this "Spatial Complexity Term", the F-statistic has a very small $p-$value (0.006), and the R^2 is very high.

We also examine the $J-$statistic, to test the hypothesis that the overidentification restriction is valid here. We find that the $J-$statistic has a $p-$value of 0.4391. Since this is higher than the critical $p-$value of 0.05, we cannot reject the hypothesis that the overidentification restriction is valid. This is further evidence of the effectiveness of the chosen instruments. Also, given the joint significance of all of these instruments in the first stage regressions, we conclude that the instruments are reasonable and we proceeded to use their predicted values in the second stage regressions.

Now that we demonstrated these first stage regression results support the correlation between the instruments and each of the endogenous variables, we take the predicted values for the "Prices Term", the "GDP Term" and the "Spatial Complexity Term" based on the first stage regressions, and then move to present the second stage results for the IV estimation. Moving to the IV parameter estimates for the predicted value of the "Prices Term" and the predicted value of the "GDP Term", in column 2 of Table 3, they both have the anticipated sign (i.e., positive) and are statistically significant with $p-$value < 0.05. The MSA-level predicted "Spatial Complexity Term" is insignificant, with $p-$value 0.829, however this variable is not a primary focus of our analysis since we are concerned with estimating the sign, significance, and magnitude of β. The R^2 is 0.21 for this second stage IV estimation.

Importantly, we can interpret the estimates as follows. First, since the IV estimate of β^{-1} is 3.75 and β^{-2} is approximately 4.80, this implies it is not unreasonable to infer that our empirical estimate of β is approximately equal to 0.3. The 95 percent confidence intervals on the coefficient estimates in the second stage imply that we cannot reject the hypothesis that β is equal to approximately 0.3., since the range of this confidence interval for the predicted "Prices Term" is between 0.122 and 7.379.

Finally, a helpful referee suggested we test whether we can reject the hypothesis of the restriction of the coefficients β are the same, for the parameter estimates for β^{-1} in the predicted "Prices Term" and β^{-2} in the predicted "GDP Term". To accomplish this, we use the "testnl" command in Stata, and find that the $p-$value is 0.4635. Since this is larger than the critical $p-$value of 0.05, we cannot reject the hypothesis that the coefficients β are the same for the second stage parameter estimates on the predicted "Prices Term" and the predicted "GDP Term".

6 Conclusions

This is the first paper estimating an equilibrium urban macro model that links a city's presence in domestic trade to its house price growth rate performance. We estimate a spatial equilibrium equation based on our urban macro model. Our primary empirical findings confirm the comparative statics implications of our theoretical model. In the spatial equilibrium equation, we have controlled for endogeneity with IV and GMM approaches.

Our findings offer evidence that approximately 30 percent of a typical consumer's income is spent on housing and 70 percent on other goods. Furthermore, the empirical estimate of β of approximately 0.3, together with our assumption that β equals S, implies that the typical consumer spends 70 percent of their income when young and 30 percent when old. The bottom line is that the data seems to support the theory we have developed that led to Eq. (8).

An implication of these estimates is that the behavioral model helps in addressing another issue. If we were to interpret the price of housing as the user cost of housing, then expected capital gains on housing (from increases in the "GDP Term") reduce its user cost. For spatial equilibrium, this is consistent with a lower growth rate of per capita real income in the same city. In other words, and without making a causal claim (but see Glaeser and Gyourko (2017) and Hsieh and Moretti (2019)), expected capital gains in housing are associated with lower real income growth. Additionally, when GDP grows faster in a particular MSA than the nationwide average, one expects faster growth in house prices relative to national house prices, after filtering out changes in the growth rate of local goods prices relative to national goods prices.

In addition to our development of the theoretical equilibrium urban macro model in this context, one of our other contributions is our merging and applying a set of novel data for testing the theory. Our instruments, including our use of both military contracts data aggregated to the MSA level and net migration at the MSA level, are novel.

It would be interesting to explore the potential of the model to explain housing price dynamics and economic growth in a number of truly global cities, such as New York, San

Francisco, Vancouver, London, Singapore, Hong Kong, etc. In those cities and many others, it is not only trade but also foreign investment in housing and real estate that plays an important but not well understood role. These issues clearly deserve attention in future research.[27] At least in the U.S. context, inclusion of such larger MSAs that are comprised of many smaller cities would require the ability to merge data on many different variables with MSA definitions that are consistent across variables. Given that our data for the largest cities are drawn from sources relying on several different definitions, a first step in improving the data availability for the largest cities might be to devote resources for facilitating a broad and collaborative approach to data collection and merging across federal and local government agencies.

Future work might also consider incorporating the effects of contagion and diffusion across MSAs, and tie these forces closely to cross-city migration, both in the theory and empirics.[28] Exploring convergence or divergence could also be an interesting aspect to consider in the type of framework that we have developed.[29]

[27]Robustness of our findings to the exclusion of major markets is unknown but is of interest for future work.

[28]e.g., Sinai and Souleles, 2013; Schubert 2021; Defusco et al, 2013).

[29]Van Nieuwerburgh and Weill (2010) find evidence of divergence.

7 References

Autor, David H., David Dorn, and Gordon H. Hanson. 2013. “The China Syndrome: Local Labor Market Effects of Import Competition in the United States.” *American Economic Review.* 103(3):2121-2168.

Bartik, Timothy. 2002. “Instrumental variable estimates of the labor market spillover effects of welfare reform.” *WE Upjohn Institute Staff Working Paper,* (02-78).

Baum-Snow, Nathaniel. 2007. “Did Highways Cause Suburbanization?” *Quarterly Journal of Economics.* 122(2):775-805.

Cardarelli, Roberto, and Lusine Lusinyan. 2015. “U.S. Total Factor Productivity Slowdown: Evidence from the U.S. States.” IMF working paper WP/15/116. May.

Carruthers, John I., Michael K. Hollar, and Gordon F. Mulligan. 2006. “Growth and Convergence in the Space Economy: Evidence from the United States. Presented at the ERSA, Volos, Greece.

Davis, and Michael G. Palumbo. 2008. “The Price of Residential Land in Large U.S. Cities. *Journal of Urban Economics.* 63(1):352–384.

DeFusco, A., Ding, W., Ferreira, F. V., and Gyourko, J. 2013. “The role of contagion in the last American housing cycle.” Unpublished manuscript.

Doherty, Conor. 2017. “Why Falling Home Prices Could Be a Good Thing.” The Upshot. *New York Times.* February 10. https://www.nytimes.com/2017/02/10/upshot/popping-the-housing-bubbles-in-the-american-mind.html

Donaldson, Dave, and Richard Hornbeck. 2016. “Railroads and American Economic Growth.” *Quarterly Journal of Economics.* 799–858.

Duranton, Gilles, and Diego Puga. 2005. “From Sectoral to Functional Urban Specialization.” *Journal of Urban Economics.* 57(2):343–370.

Duranton, Gilles, and and Diego Puga. 2014. "The Growth of Cities." pp. 781-853. In Steven N. Durlauf and Philippe Aghion (eds.) *Handbook of Economic Growth,* Vol. 2A. Amsterdam: North-Holland.

Duranton, Gilles, Peter M. Morrow, and Matthew A. Turner. 2014. "Roads and Trade: Evidence from the US." *Review of Economic Studies.* 81(2):681–724. doi:10.1093/restud/rdt039

Fujita, Masahisa, Paul Krugman and Anthony J. Venables. 1999. *The Spatial Economy: Cities, Regions, and International Trade.* Cambridge, MA: MIT Press.

Glaeser, Edward L., and Joseph Gyourko. 2017. "Housing Dynamics." Harvard Institute of Economic Research Discussion Paper Number 2137.

Glaeser, E. L., Gyourko, J., Saks, R. E. 2005. "Why have housing prices gone up?" *American Economic Review,* 95(2), 329-333.

Glaeser, Edward L., Joseph Gyourko, Eduardo Morales, and Charles G. Nathanson. 2014. "Housing Dynamics: An Urban Approach." *Journal of Urban Economics*. 81:45–56.

Gruber, Jonathan. 1997. "The Consumption Smoothing Benefits of Unemployment Insurance." *The American Economic Review.* 87(1):192-205.

Gyourko, Joseph, Christopher Mayer, and Todd Sinai. 2013. "Superstar Cities" *American Economic Journal: Economic Policy*. 5(4):167–199.

Henderson, J. Vernon. 1974. "The Sizes and Types of Cities." *American Economic Review*. 64:640-656.

Hollar, Michael K. 2011. "Central Cities and Suburbs: Economic Rivals or Allies?" *Journal of Regional Science*. 51(2):231–252.

Hoogendoorn, S., van Gemeren, J., Verstraten, P., Folmer, K. 2019. House Prices and Accessibility: Evidence from a Quasi-Experiment in Transport Infrastructure. *Journal of Economic Geography*. 19(1):57-87.

Hornbeck, Richard, and Enrico Moretti. 2015. "Who Benefits From Productivity Growth? The Local and Aggregate Impacts of Local TFP Shocks on Wages, Rents, and Inequality." working paper.

http://faculty.chicagobooth.edu/richard.hornbeck/research/papers/tfp_hm_may2015.pdf

Hsieh, Chiang-Tai, and Enrico Moretti. 2019. "Housing Constraints and Spatial Misallocation." *American Economic Journal: Macroeconomics.* 11(2):1-39.

Ioannides, Yannis M. 2013. *From Neighborhoods to Nations: the Economics of Social Interactions.* Princeton, NJ: Princeton University Press.

Larson, William D. 2013. "Housing and Labor Market Dynamics in Growing versus Declining Cities." Bureau of Economic Analysis working paper. November.

Levkovich, O., Rouwendal, J., Van Marwijk, R. 2016. "The Effects of Highway Development on Housing Prices." *Transportation.* 43(2):379-405.

Li, Jinyue. 2017. "Housing Prices and the Comparative Advantage of Cities." Working Paper, Chinese University of Hong Kong.

Nakamura, E., and Steinsson, J. 2014. "Fiscal stimulus in a monetary union: Evidence from US regions." *American Economic Review.* 104(3), 753-92.

Pennington-Cross, Anthony. 1997. "Measuring External Shocks to the City Economy: An Index of Export Prices." *Real Estate Economics.* 25(1):105–128.

Saiz, Albert. 2010. "The Geographic Determinants of Housing Supply." *Quarterly Journal of Economics.* 1253–1296.

Schubert, G. 2021. "House price contagion and us city migration networks." Job Market Paper. Cambridge, Mass.: Joint Center for Housing Studies, Harvard University.

Sinai, T., and Souleles, N. 2013. "Can owning a home hedge the risk of moving?" *American Economic Journal: Economic Policy*, 5(2), 282-312.

United States Bureau of the Census. *Annual Survey of Manufactures.*

https://www.census.gov/programs-surveys/asm.html

United States Bureau of the Census. *State Government Finances.*

https://www.census.gov/govs/state/historical_data.html

Van Nieuwerburgh, S., Weill, P. O. 2010. "Why has house price dispersion gone up?" *The Review of Economic Studies*, 77(4), 1567-1606.

Ventura, Jaume. 2005. "A Global View of Economic Growth." In *Handbook of Economic Growth*, eds. Aghion, Philippe, and Steven N. Durlauf, Ch. 22, 1419-1497. Amsterdam: Elsevier.

8 Appendix A (Not for Publication): The Urban Production Structure

We develop first the case where all cities are autarkic, that is no intercity trade, and cities produce both manufactured tradeable goods, and use them in turn to produce the composite used for consumption and investment. Each of the manufactured tradeable goods, $j = X, Y$, is produced by a Cobb-Douglas production function, with constant returns to scale, using a composite of raw labor and physical capital, with elasticities $1-\phi_J$, and ϕ_J, respectively, and a composite made of intermediates. The shares of the two composites are $u_J, 1-u_J$ respectively. There exists an industry $J-$specific total factor productivity, Ξ_{Jt}. Production conditions for each of two industries J are specified via their respective total cost functions:

$$B_{Jt}(Q_{Jt}) = \left[\frac{1}{\Xi_{Jt}}\left(\frac{W_t}{1-\phi_J}\right)^{1-\phi_J}\left(\frac{R_t}{\phi_J}\right)^{\phi_J}\right]^{u_J}\left[\sum_m P_{Zt}(m)^{1-\sigma}\right]^{\frac{1-u_J}{1-\sigma}} Q_{Jt}, \tag{9}$$

where Q_{Jt} is the total output of good $J = X, Y$, P_{Zt} is the price of the typical intermediate, elasticity parameters u_J, ϕ_J satisfy $0 < u_J, \phi_J < 1$, and the elasticity of substitution in the intermediates composite σ is greater than 1. The TFP term Ξ_{Jt}, summarizes the effect on industry productivity of geography, institutions and other factors that are exogenous to the analysis.

Each of the varieties of intermediates used by industry J are produced according to a linear production function with fixed costs (which imply increasing returns to scale), with fixed and variable costs incurred in the same composite of physical capital and raw labor that is used in the production of manufactured goods X and Y. The shares of the productive factor inputs used are the same as, ϕ_J and $1-\phi_J$, $J = X, Y$, respectively.[30] The respective

[30] This may be generalized to allow for input-output linkages by requiring (see also Fujita, *et al.* (1999), Ch. 14), that each intermediate good industry use its own composite as an input. This is accomplished by introducing as an additional term $\left[\int_0^{M_{it}} p_{it}^{1-\epsilon_i}\right]$ on the r.h.s. of the cost function $b_{it}(Z_{Jt})$.

total cost function is

$$b_{it}(Z_{Jt}(m)) = \frac{f + cZ_{Jt}(m)}{\Xi_{Jt}} \left[\left(\frac{W_t}{1-\phi_J} \right)^{1-\phi_J} \left(\frac{R_t}{\phi_J} \right)^{\phi_J} \right],$$

and $Z_{Jt}(m)$, the quantity of the input variety m used by industry $J = X, Y$. Its price is determined in the usual way from the monopolistic price setting problem [Dixit and Stiglitz (1977)] and it is equal to marginal cost, marked up by $\frac{\sigma}{\sigma-1}$:

$$P_{Z,J,t} = \frac{\sigma}{\sigma-1} \frac{c}{\Xi_{Jt}} \left(\frac{W_t}{1-\phi_J} \right)^{1-\phi_J} \left(\frac{R_t}{\phi_J} \right)^{\phi_J}.$$

At the monopolistically competitive equilibrium with free entry, each of the intermediates is supplied at quantity $(\sigma - 1)\frac{f}{c}$, and costs $\frac{\sigma f}{\Xi_{Jt}} \left(\frac{W_t}{1-\phi_J} \right)^{1-\phi_J} \left(\frac{R_t}{\phi_J} \right)^{\phi_J}$ per unit to produce. Its producer earns zero profits.

TABLES

Table 1: Mean and Standard Deviation of Data, Key Variables and Instruments

	(1) All MSAs	(2) Northeast	(3) Midwest	(4) West	(5) South
GDP Growth	3.73	3.00	3.32	4.33	3.92
	(4.059)	(2.785)	(3.667)	(4.910)	(4.075)
House Price Growth	2.76	2.50	1.72	4.23	2.81
	(7.126)	(4.272)	(4.291)	(10.77)	(6.855)
CPI Growth	2.01	2.03	1.79	2.16	2.08
	(1.275)	(1.267)	(1.276)	(1.085)	(1.350)
Military Contracts Growth	9.25	3.42	11.45	9.60	9.04
	(74.04)	(32.15)	(89.82)	(81.04)	(64.81)
Planned Highway Miles Per Capita	42.49	29.93	35.08	57.50	43.45
	(80.62)	(52.80)	(54.36)	(127.7)	(68.47)
Unavailable Land Per Capita	0.71	0.77	0.42	1.12	0.70
	(0.959)	(0.918)	(0.695)	(1.272)	(0.856)
Employment	4.5e+05	7.7e+05	3.7e+05	4.0e+05	4.4e+05
	(867366.0)	(1896743.6)	(688568.9)	(544492.8)	(611531.4)
Net Migration Growth		11.91	11.85	11.61	11.36
		(43.32)	(43.33)	(43.33)	(43.31)
N	3184	349	929	672	1234

Notes: Overall national and regional mean values for each variable; standard deviations in parentheses. Time period is annual from 2003-2017, for 197 MSAs (but regressions that follow in subsequent tables are based on unbalanced panels). Per capita data are in terms of millions of population.

Table 2: First Stage Regressions

Dependent Variable Listed in Column Headings

	(1) Prices Term	(2) GDP Term	(3) Spatial Complexity Term
Instrument 1 (see notes)	1.032**	-0.909	-0.00103
	(0.000)	(0.105)	(0.851)
Instrument 2 (see notes)	-0.0000703	0.00111	0.00000812
	(0.196)	(0.482)	(0.145)
Planned Highway Miles Per-Capita	0.0104*	0.0185	-0.00103
	(0.021)	(0.612)	(0.481)
Migration Instrument (see notes)	-0.0380*	0.411	0.00392*
	(0.038)	(0.138)	(0.029)
Constant	-0.418	-7.053	-1.116**
	(0.138)	(0.132)	(0.003)
r2	0.933	0.197	0.999
F statistic p-value	0.00000335	0.000824	0.00580
N	1020	1020	1020

Notes: p-values in parentheses

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

First Stage Regressions with Fixed Effects for MSA and Year (Time period is annual from 2003-2017; unbalanced panel of MSAs). Robust Standard Errors Clustered by Region (Northeast, Midwest, West, South).

Regressors:

Prices Term defined as ln(MSA cpi/cpi national average) – (MSA cpi growth – national average of cpi growth) – ln(MSA house price index/national average of house price index).

GDP Term defined as (MSA GDP growth – national average GDP growth) – ln(MSA cpi/cpi national average)

Spatial Complexity Term defined as ln(MSA employment – national average employment)

Instruments:

The instrument for the Prices Term is labelled above as Instrument 1, which is: ln(MSA cpi/cpi national average) – (MSA cpi growth – national average of cpi growth) – (MSA unavailable land per capita/national average of unavailable land per capita).

The instrument for the GDP Term is labelled above as Instrument 2, which is: (MSA military contracts growth – national average of military contracts growth) + (MSA net migration growth – national average of net migration growth) - ln(MSA cpi/cpi national average).

Planned Highway Miles Per-Capita (MSA) is the instrument for the Spatial Complexity Term.

Migration Instrument is defined as ln(MSA net migration/national average of net migration), used for overidentification.

Table 3: OLS and 2SLS (Second Stage) Regression Results

Dependent Variable:
(MSA House Price Growth - National Average House Price Growth) - (MSA CPI Growth - National Average of CPI Growth)

	(1) OLS	(2) 2SLS
Prices Term	-1.288^{**}	3.751^{*}
	(0.006)	(0.043)
GDP Term	0.369^{**}	4.802^{**}
	(0.004)	(0.007)
Spatial Complexity Term	25.47	-22.86
	(0.230)	(0.829)
Constant	24.56	-4.453
	(0.247)	(0.972)
R-squared	0.157	0.214
J Statistic p-value	-	0.439
F-Statistic p-value	-	0.00141
N	2587	1020

Notes: *p*-values in parentheses
$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$
Regressions include Fixed Effects for MSA and Year (Time period is annual from 2003-2017; unbalanced panel of MSAs). Robust Standard Errors Clustered by Region (Northeast, Midwest, West, South).

For 2SLS estimates, the Prices Term, GDP Term, and Spatial Complexity Term are the predicted values of those terms from the first stage regressions, obtained using the estimates in Table 2.

Regressors:
Prices Term defined as ln(MSA cpi/cpi national average) – (MSA cpi growth – national average of cpi growth) – ln(MSA house price index/national average of house price index).

GDP Term defined as (MSA GDP growth – national average GDP growth) – ln(MSA cpi/cpi national average)

Spatial Complexity Term defined as ln(MSA employment – national average employment)