A Spatial Production Economy Explains Zipf's Law for Gross Metropolitan Product

Watanabe, Hiroki

5 November 2015

A Spatial Production Economy Explains Zipf's Law for Gross Metropolitan Product

Hiroki Watanabe*

November 5, 2015

Abstract

This paper provides a theoretical and empirical analysis of the distribution of GDP at city level (henceforth referred to as gross metropolitan product, GMP) with the aim of bridging the gap betwen the literature on agglomeration economies and the city-size distribution. We show that 1) it shares the same characteristics to the city-size counterpart: They are both fat-tailed, and 2) a 1% increase in employment leads to a 1.117% (or 1.180% in theory) increase in GMP. Free mobility of household forces a city to operate at the size where scale economies are present, or else, the city cannot offset the reduced housing consumption and increased congestion due to crowding set off by agglomeration economies, and loses its population and GMP to elsewhere. We establish a production economy model to break down the interplay above and derive the equilibrium GMP distribution, which tests well with the US data on GMP.

Keywords: Zipf's Law, Gibrat's Law, GDP by City, Production Economy
JEL classification: D51, E2, R12

1 Introduction

Why and how do cities exist? And if they do, then why and how do they come in different sizes? These questions have attracted a lot of interest among researchers both in economic and geographical fields. The answer to the first question is evaluated with GMP in *each* city, while the second question calls for the analysis of the *overall* distribution of city sizes. We now have a growing understanding of GMP and the size distribution of cities, separately. These questions are integrally related but the intersection of the two, the *distribution* of *GMP*, has never been analyzed to this date. The analysis is of significant importance for both the city-size distribution and agglomeration economy literature.

Four out of five people live in cities, and they do so for various reasons, i.e. better job prospects, decent wage, urban amenities, or family obligations. The resulting size distribution of cities has kept the rapt attention of researchers, creating a long line of work on the subject. Whereas a size distribution is something of note, the story does not end or complete there. The overriding research objective in the

*Department of Economics and Fiance, Lamar University (watanabe.wustl@gmail.com). I thank Kristian Behrens, Akin Buyukeren, Luiggi Donayre and Nicholas Sheared for their input. All remaining errors are mine.

literature is the welfare implication of the city-size distribution. No one moves in or out of a city just for the sake of making its size larger or smaller, nor does the city size itself feed its population. We move to a city because we think it is big (or small), but deep down, what we have in mind is not the size *itself* but what the size *has to offer*, i.e., other economic factors that differ city to city such as GMP, which correlates with the income and job opportunities in a city. Thus, the city-size distribution needs to be understood, and validated, in tandem with other spatial distributions.

As for agglomeration economies, the GMP of a city is not completely independent from other cities' GMP. GMP requires factors of production and some of them need to be brought in from somewhere else. And since inputs cannot simply fall off the sky, the distribution of GMP needs to be closely monitored when we study the agglomeration economies of any city.

Our major findings are as follows. First, two empirical regularities on the city-size distribution carry over to GMP. Most of GDP are generated in only a few cities, just as the city-size distribution, the regularity known as Zipf's law ([Gab99], [Dur07]). In fact only 20% of cities create as much as 79% of urban GDP.[1] GMP has a lower Pareto coefficient than its city-size counterpart. Its tail end is even heavier than the city-size distribution. Gibrat's law also extends to GMP. Urban economic growth rates are independent of the size of GMP. Second, GMP exhibits increasing returns to employment. New York's GMP is larger than any other city's, even size for size. This is consistent with our first finding that the GMP distribution has a heavier tail than the city-size distribution. We build a production economy model and establish that agglomeration economies are due to the trade-off between externalities and housing consumption. We prove that, as a direct result of free mobility, the equilibrium city size has to be such that an additional resident will reduce a housing lot size in the city but make up for it by raising citywide productivity.

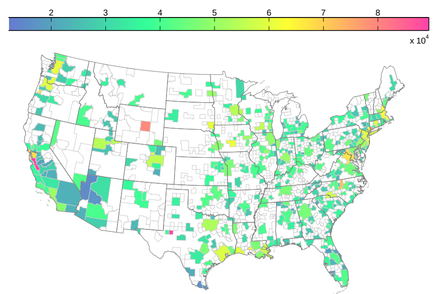


Figure 1. GMP per capita in 2010 (in 2005 USD).

In the existing city-size models, with the assumption of free mobility, consumers/workers will update their locations until they exhaust the locational arbitrage opportunities. Thus, regardless of the city size, cities become indifferent to consumers in equilibrium. This does not imply that workers are equally productive or that their income will be the same across the board. In practice, per capita GMP varies by location (cf. figure 1 and table 1). People enjoy the same util-

[1] This relation is known as a 20-80 rule: 20% of agents are accountable for 80% of the results, a typical sign that something of scale-free nature is at work.

ity level at the end of the day,[2] but what induces interurban migration depends on GMP. Overshadowed by the consuming interest in the city-size distribution, GMP is treated as a mere byproduct. Citywide aggregate production function is used as an intermediate step to reach the equilibrium city-size distribution. We should obtain the equilibrium production level in each city along with its size, but the predicted GMP has never been verified with empirical data.

On the other hand, the first question of why cities exist is of particular interest for urban economists or economic geographers because if they do not exist, then *urban* economics does not exist either after all. In the light of Starrett's spatial impossibility theorem [Sta78], we know that GMP must differ from city to city in the end, and in addition, GMP has to be large enough to offset all the negatives that come with urban life, or else cities will disappear. However, it is hard to figure out exactly where to look for the answer. Cities are the most complex socioeconomic organisms of all. Regardless of where we seek for the root cause, the traditional answer usually traces itself back to indivisibility. A favorable seaport cannot be taken apart and spread across the country for any practical use. Neither can urban amanities, local public goods or marketplaces [BK00], [BW93]. A firm benefits by sharing with other firms in its close proximity the labor pool [ABL07], a wide range of intermediate goods supply [FKV99] or supporting service industries [RB88], which become less effective or economically unfeasible if they are split apart. Positive externalities that firms in a crowded setting enjoy, such as knowledge spillovers [Mor04], decay rapidly with distance, and thus necessitate a significant level of dense economic activities within a narrow expanse of land to be of any use [Kru91]. Other motivating factors for agglomeration economies include matching [BRW06] and learning. See Duranton and Puga [DP04] for a full review on this matter. We shall focus on the increasing returns to scale at a city level as a driving force for the sake of deriving the GMP distribution. In any case, the existing agglomeration literature concentrates on why cities exist but has not considered how the resulting GDP is split.[3] In fact, some of them are not designed to address the distribution. The earlier New Economic Geography models such as [Kru91] do not offer much insight into the distribution of economic activities because there are only two regions in the model. There has been extensive empirical research on the subject [Sve75], [Seg76], [CS04]. Researchers typically try to isolate economies of localization (the boost in productivity from the nearby firms in the same industry) from economies of urbanization (the boost from the nearby firms in the multiple industries. See [Car87] for a review). The focus is the employment elasticity of GMP but not the distribution of GMP. Melo et al [MGN09] provide a comprehensive summary of the empirical studies.

Related to the empirical work above and our paper is a series of work that attempts to gauge the geographic concentration of industries [Kim95]. Our work differs from them in the following sense: The indices employed in these analyses are of the first [Kru91] or the second moment [EG94] of the distribution in na-

[2] Cities are put in equilibrium either by equating wage (e.g., [Dur07]) or utility level (e.g., [Gab99]).

[3] More recently, Caliendo et al. [CPRHS14] looked into the distribution of economic activities at the US state level to see how local level productivity changes affect the overall economy.

ture. As such, they are effective in addressing the degree of concentration for the observed distribution of economic activities, but they cannot examine the distribution itself: Low-order moments throw away lots of information on the distribution. Two distinct distributions, for example, one right-heavy and the other left-heavy, may end up with the same concentration index (we will show later that the GMP distribution is right heavy).

Be it elasticity or concentration, why has the agglomeration literature not paid attention to the distribution of GMP as much as the city-size distribution literature has to the distribution of city sizes? There seem to be at least two interrelated reasons. One of the reasons is the endogenous nature of the constraint that models of agglomeration economies have to follow. When the city size is in decline, it is obvious that there is a net outflow of people and those who left the city should be accounted for in the city they moved to. People simply *cannot* disappear. There is a simple one-to-one correspondence: One city's gain is another city's loss by the same exact number of people. Thus, it stands to reason that the researchers are keenly aware of the size distribution of cities as a necessary step to analyze individual city sizes. On the other hand, when the GMP is in decline, it is not as apparent or easy to locate where the lost *output* went as to track where *people* moved to. Unlike population, GMP does not have to add up to a fixed number and it *can* simply disappear in a flash. The sum of GMP (namely, GDP) depends on how the resources are allocated among the cities. For example, if we empty New York City and relocate all New Yorkers to other existing cities evenly, the total population of the US is still the same but GDP will be smaller. We tend to overlook the distribution of GMP in the absense of the GDP constraint, but that does not necessarily mean that a city can have an economic growth without affecting GMP of other cities. In fact, most models incorporate the substantial role that free mobility plays in characterizing the nature of agglomeration economies. For instance, Henderson [Hen86] states "On an aggregative level, resources are in some sense more productive in large cities — otherwise large cities with their high costs of living could not pay the high wage necessary to attract resident." (p.48). However, the notion has never been explicitly tested. The distribution of GMP is necessary to fully examine the nature of agglomeration economies as it describes how the arbitration of free mobility played out in the end.

The other reason is the nonexistence of the established empirical distribution of GMP itself. The city-size distribution is well-documented and its remarkable conformity to Zipf's law is both time- and country-invariant [So005]. Thus, a model of the city-size distribution will be immediately discarded if it leads to something other than Zipf's law, such as a degenerate distribution.[4] The same requirement would go for models of agglomeration economies if the empirical GMP distribution were proven to be robust. However, we did not know what the empirical GMP distribution looks like, and consequently, there has never been an immediate need to check a model's relevance to the reality in the distribution department.

[4]It needs to be stressed, however, that compliance to Zipf's law is merely a necessary condition for a model to be successful but *not* a sufficient condition. The same principle applies to modelling of the GMP distribution.

We intend to contribute to both agglomeration economy and city-size distribution literature in the following way: We will make it explicit how the free mobility of consumers endogenously gives rise to agglomeration economies. In addition (and as a consequence), we will derive the equilibrium GMP distribution, which provides an additional layer of empirical validity to test the models of the city-size distribution.

The remainder of the paper is organized as follows: Section 2 investigates into the nature of GMP distribution and provides descriptive statistics on GMP along with city size. In section 3, we introduce a spatial production economy model to explain the findings in section 2 before we empirically evaluate our model's performance in section 4. Section 5 concludes our study.

2 GMP Actualities

We start off with a casual observation of GMP and then establish Zipf's and Gibrat's law for GMP. After that, we will identify the relationship between GMP and city size.

2.1 Data Description

The US Bureau of Economic Analysis reports annual GDP by metropolitan statistical area (MSA) along with the US GDP and estimated employment. Descriptive statistics for the data employed are in table 1.

The data set we use is more inclusive than previous studies. There are 366 cities with accompanying GMP figures. The largest sample size used so far to test GMP is 30 by Mion and Naticchioni [MN05] (See [MGN09] for a full list of papers in this field).[5] For example, MSA's like Beaumont-Port Author MSA, TX (population 388,745 as of 2010) are too small to be included in the data set in [Seg76]. At the time of writing, the GMP of a city smaller than the 366th largest city (Palm Coast, FL with the GMP of $1,132 million) is not reported.

Let us start with the distribution of GMP. Figure 2 (in color) is a map of the United States with MSA's colored according to their population density and GMP in 2010.[6] Figure 2(a) comes with no surprise. It is well documented that the city-size distribution is tail heavy. What is newsworthy is figure 2(b). GMP shares the same pattern to city size in terms of distribution. Figure 3 represents the probability density function (PDF) and rank-size plot of GMP in 2010. New York accounts for the lion's share of GDP, followed by Los Angeles, and there are lots of mid-sized cities that are dwarfed by the high-ranked cities.

[5] The aforementioned study [Seg76] has 58 locations but output is limited to the manufacturing sector rather than GMP as a whole. These studies often quote census for manufacturers alone.

[6] Population data also include micropolitan statistical area along with MSA. For definition of MSA, see http://www.census.gov/population/metro/about/.

Data	Employment	GMP	GMP per capita	GMP growth
Unit	MSA	Million 2005 USD	2005 USD	Annual rate
Year	2010	2010	2010	2010 over 2005
Coverage (MSA/USA)	84.90%	87.60%	87.60%	87.60%
Largest city	#001 New York	#001 New York	#001 Midland, TX	#001 Midland, TX
Largest size	18,919,787	1,147,160	89,350	11.60%
73rd largest city	#073 Akron, OH	#073 Worcester, MA	#073 Waterloo, IA	#073 Dallas, TX
Size share of #1-73	71.80%	78.75%	29.40%	N/A
City near arithmetic mean	#072 North Port, FL	#063 Madison, WI	#150 Greenville, SC	#171 Topeka, KS
Arithmetic mean	707,308	31,763	36,002	.49%
Median city	#183 Laredo, TX	#183 Bellingham, WA	#183 Gainesville, FL	#183 Monroe, LA
Median size	251,539	8,414	34,048	.35%
Smallest city	#366 Carson City, NV	#366 Palm Coast, FL	#366 Palm Coast, FL	#366 Lake Charles, LA
Smallest size	55,212	1,132	11,793	-5.98%
Standard deviation	1,582,442	86,824	11,257	.0212
Skewness	6.669	7.500	1.409	.9088
City near geometric mean	#149 Naples, FL	#154 Kalamazoo, MI	#178 Mobile, AL	N/A
Geometric mean	321,416	11,078	34,460	N/A
Mean of log value	12.68	23.13	10.45	N/A
Standard deviation of log value	1.062	1.216	.2920	N/A
Skewness of log value	1.109	1.103	.2251	N/A

Table 1. Descriptive Statistics. The statistics above the line (shaded in blue) are related to a linear scale and below the line (shaded in green) are related to a log scale. The mean of log value is same as the log of geometric mean. The first 73 cities make up for the upper 20% of the total number of cities.

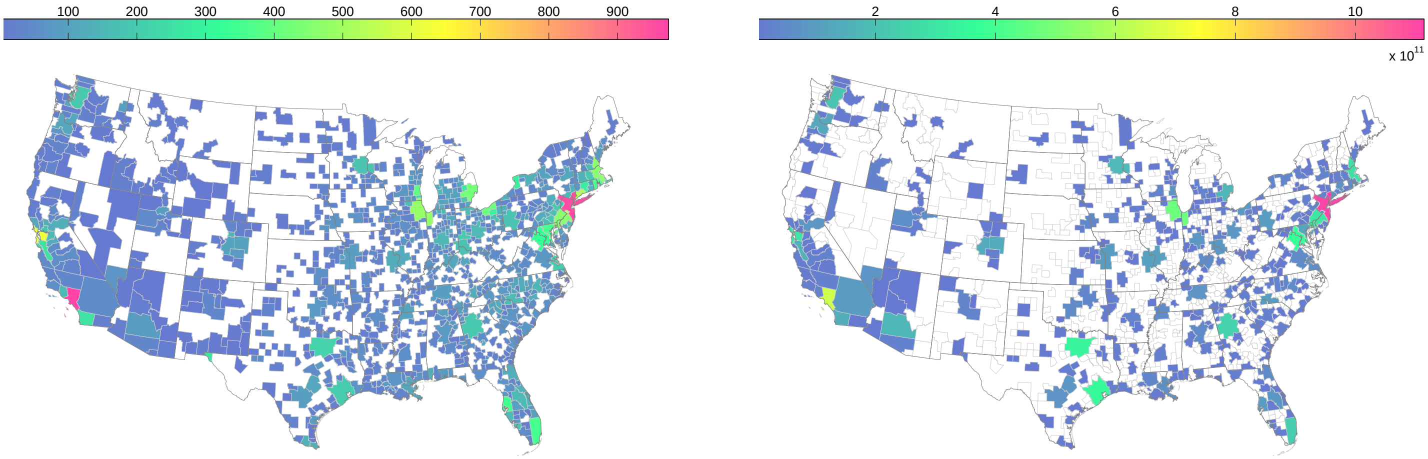
2.2 Gibrat's Law for GMP

Next, we look into the dynamics: Does a large GMP make a city grow fast? The answer: No. Gibrat's law implies that the size of a city does not have any bearing on its growth rate. The city-size distribution is known to follow it well ([IO03]). It turns out that GMP does the same. We carried out both parametric (below) and non-parametric estimations (in appendix A.1) following [Eec04] to examine the relationship between GMP and GMP growth rate.

First, we regress GMP growth rate on GMP. We gathered data from 2005 and 2010 to compute growth rates. GMP is defined by the geometric mean $Y := \sqrt{Y_{05}Y_{10}}$, assuming exponential growth. Estimates are reported in table 2. Figure 4(a) seems to indicated that the regression line is pulled upwards partly because of New York.[7] To counteract this sensitivity to predominantly large cities, we regressed GMP on the log of GMP as well (figure 4(b)).

The null is not rejected at the 5% level of confidence on GMP or on the log thereof. The estimates seem to agree with Gibrat's law. Non-parametric estimation

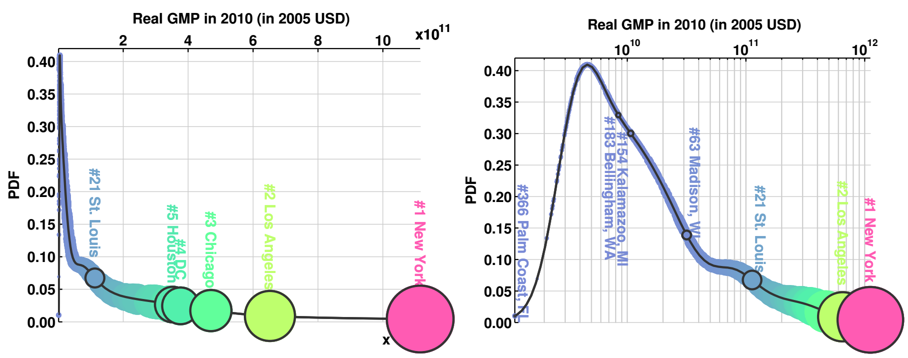
[7] The coefficient on GMP may well have been negative had New York's growth rate been negative. The estimates' dependence on New York is not all that welcoming because, while it is large, New York is still just one observation as much as Beaumont, TX is.



(a) Population Density in 2010 (persons/km^2).

(b) GMP in 2010 (in 2005 USD).

Figure 2.



(a) PDF of GMP. Dots are size proportionate to GMP.

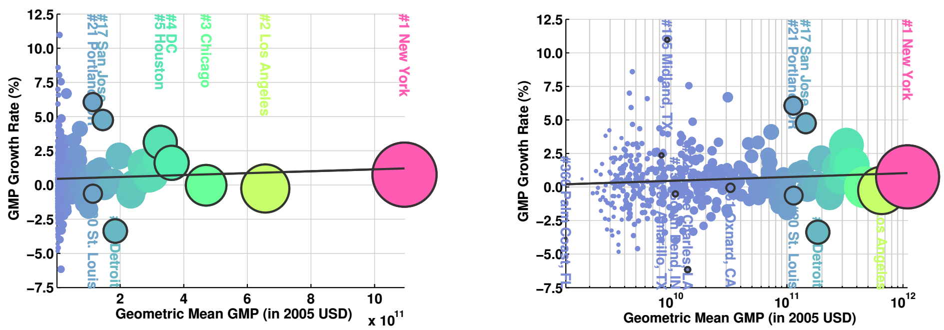
(b) PDF of GMP in log scale.

Figure 3. PDF plots of GMP. See table 1 for the explanation of the selected cities above.

essentially shadows the result above. See appendix A.1 for details and analysis on variance.

	Regressor			R^2	Figure
	Intercept	GMP	log(GMP)		
Coefficient	4.482e-03	6.981e-15		8.077e-04	4(a)
t-statistic	3.84	.54			
Coefficient	-2.373e-02		1.230e-03	5.044e-03	4(b)
t-statistic	-1.13		1.36		

Table 2. Ordinary least squared (OLS) estimate of growth rate.



(a) OLS over GMP.

(b) OLS over log of GMP.

Figure 4.

2.3 Zipf's Law for GMP

As we have seen in figure 3, GMP seems to be well traced by a power law. OLS estimation confirms the power-law behavior of GMP, as documented in table 3 and figure 5. The Pareto exponent is -.9003 on employment,[8] whereas we have

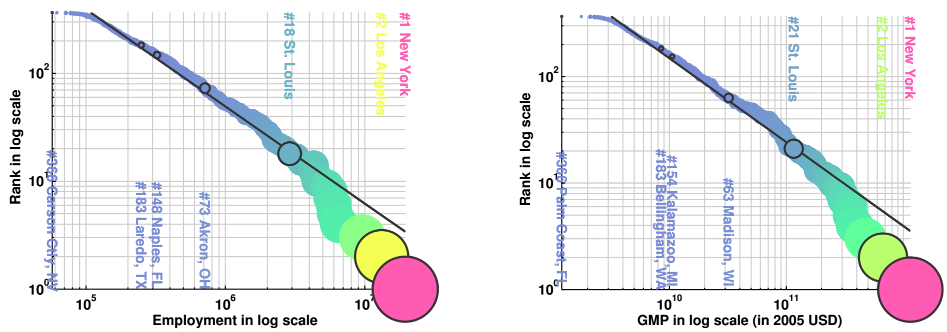
	Regressor			R^2	Figure
	Intercept	log(Employment)	log(GMP)		
Coefficient	16.34	-.9003		.9763	5(a)
t-statistic	174.57	-122.50			
Coefficient	23.13		-.7875	.9756	5(b)
t-statistic	152.89		-120.58		

Table 3. Rank-size and rank-GMP regression

-.7878 on GMP. This is indicative of the fact that the GMP distribution is even more skewed than the corresponding city-size distribution. This is to be theoretically verified with proposition 3.2.

We include OLS just for illustration, with the caveat that it would not work had we had the extensive data. As pointed out by Gabaix and Ioannides [GI04], the city-size distribution does not sit well with the assumptions on errors in OLS

[8] Employment data are based on population estimates that the Bureau of Economic Analysis uses to compute per capita GMP.



(a) Rank-size regression

(b) Rank-GMP regression

Figure 5.

estimation. The same criticism applies to the GMP distribution as well. In addition, due to the limited data range, it is likely that Zipf's law applies only to the upper tier and that the untruncated GMP distribution deviates from Zipf's law for small cities (cf. [Eec04]). In this case, a distribution other than a Pareto distribution, such as a lognormal or double Pareto lognormal [GZS10], is an apt choice to describe the data. Unfortunately, an exhaustive data set is not available for GMP. In the absence of the lower end of the distribution, Zipf's law can be used to describe the remaining mid to upper end of the distribution. Nevertheless, we will explain the GMP distribution both theoretically (in section 3.2.1) and empirically (in section 4.1) without OLS. The case in point is not whether Zipf's law describes the upper end of the distribution in particular but that the GMP distribution has a fat tail.

2.4 City Size and GMP

Figure 6 shows the relationship between working population and the aggregate product in a city. There seems to be a log-linear relationship between them with coefficient slightly but statistically significantly greater than one, indicating that increasing returns to scale are at work between city size and GMP. Table 4 reports the results with figure 6.[9] Our numbers are not too far off from the findings from the ones found in the literature. For example Shefer [She73] finds that a 1% rise in input will results in a 1.12% increase in output (note, however, that this is just for the primary metal industry, whereas our numbers are for GMP).

[9]Note that

$$\log(Y/L) = \gamma_0 + \gamma_1 \log L \quad \Rightarrow \quad \log Y = \gamma_0 + (\gamma_1 + 1) \log L$$

on a per-capita basis. On aggregate level, $\log Y = \beta_0 + \beta_1 \log L$ so that $\gamma_1 = \beta_1 - 1$, as can be seen in table 4.

Regressand	Parameter	Value	Intercept	Employment	log(Employment)	R^2	$\bar{R}^2$	Fig.
GMP	Actual	Coefficient	-6.492e+09	5.409e+04		.9717	.9716	6(a)
		t-value	-7.75	111.81				
log(GMP)	Actual	Coefficient	8.960		1.117	.9528	.9527	6(b)
		t-value	54.02		85.73			
		t-statistic	4.79		-4.81			
log(GMP)	Theoretical	Coefficient	8.165		1.180	.9498	.9497	6(b)
		t-value	47.74		87.79			
GMP per capita	Actual	Coefficient	3.422e+04	2.522e-03		.1257	.1233	6(c)
		t-value	56.69	7.23				
log(GMP per capita)	Actual	Coefficient	8.960		.1173	.1821	.1799	6(d)
		t-value	54.02		9.00			
		t-statistic	4.79		-4.81			
log(GMP per capita)	Theoretical	Coefficient	8.165		.1800	.1302	.1278	6(d)
		t-value	47.74		13.39			
Housing	Actual	Coefficient	-1.736e+09	8489		.9571	.9569	6(e)
		t-value	-9.88	84.60				
log(Housing)	Actual	Coefficient	4.149		1.305	.8769	.8765	6(f)
		t-value	11.98		47.81			
		t-statistic	-4.57		4.59			
log(Housing)	Theoretical	Coefficient	5.732		1.180	.8688	.8684	6(f)
		t-value	16.03		41.88			

Table 4. $\bar{R}^2$ is an adjusted value of R^2. For t-value, the null is coefficient equals zero, whereas for t-statistic, the null is coefficient equals theoretical value. Section 4.2 explains theoretical value.

3 Model

3.1 Spatial Production Economy

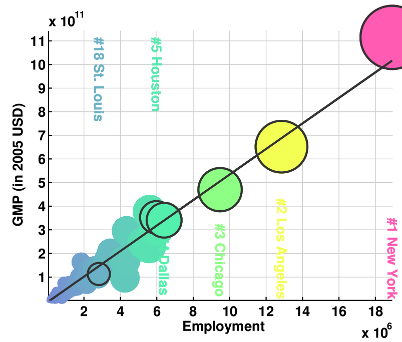
We construct an intercity general equilibrium model to seek a comprehensive explanation for all the empirical findings in section 2. In particular we develop a production economy with three commodities: composite goods, housing and leisure,[10] and two types of agents: worker/consumer and landlord.

There are I cities in the economy. S^i residents live in city i, totalling $S = \sum_{i=1}^{I} S^i$ of urban population nationwide. Each city has a demographic composition similar to the Alonso model (cf. Berliant and Fujita [BF92]). See figure 7 for one example representation of agents involved in this production economy. In each city lives a landlady who owns all the area H in city i. She is retired and lives off her rental income $r^i H$, where r^i marks the city's rental rate (think of her as the first settler in town or a developer). She is an immobile[11] landlady and assumed to consume only composite goods and leisure out of one unit of time she is endowed with.[12] The remainder of the urban population are mobile, active and identical

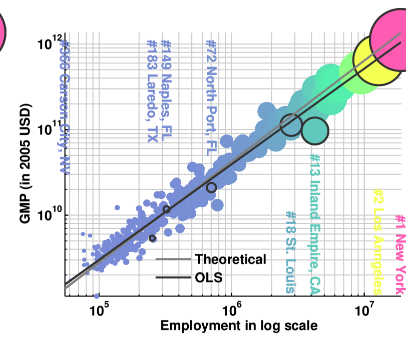
[10] Alternatively, we can include capital goods but due to the lack of data, we limit ourselves to three goods in this economy. See appendix A.2 for how capital stock is dealt in the literature.

[11] We assume that she cannot change her city of residence so that we can count the rental income toward GMP where it is collected. Otherwise, city i's rental income may show up in city j. However, we will not count her toward S^i for notational ease. We will return to the role of her location choice in section 4.2.

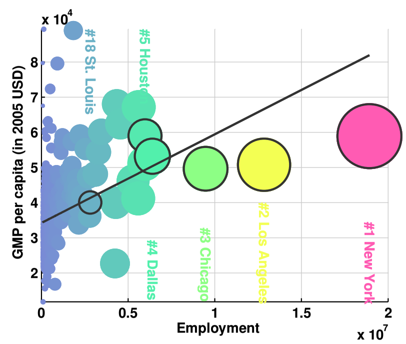
[12] She lives in a special lot designated for a landlady within the city but outside H to keep our analysis tractable. She cannot be an absentee landlady because of footnote 11.



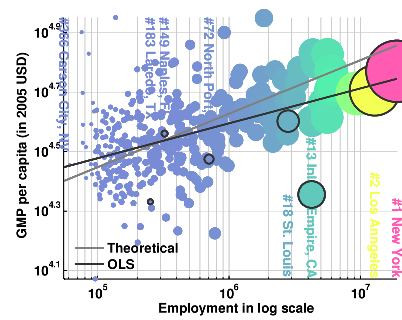
(a) OLS



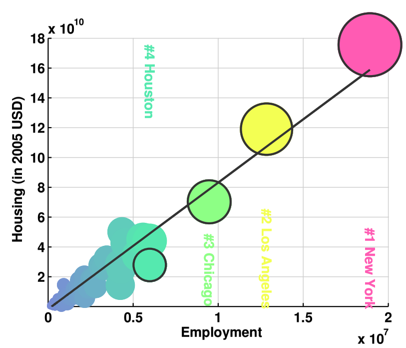
(b) OLS (log)



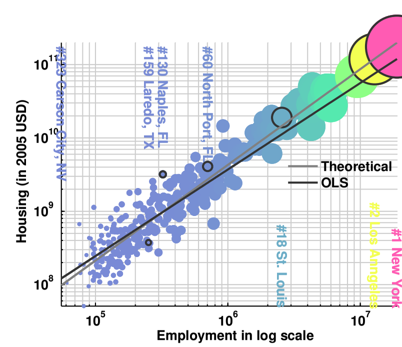
(c) OLS GMP per capita



(d) OLS GMP per capita (log)



(e) OLS Housing



(f) OLS Housing (log)

Figure 6.

workers/consumers who supply labor l_R^i out of one unit of time they are endowed with to produce a basket of goods c_R^i that includes all the goods and services other than housing h_R^i and leisure $(1 - l_R^i)$. Their consumption bundle x^i and endowment

e^i are given by

$$x_R^i = \begin{pmatrix} c_R^i \\ h_R^i \\ 1 - l_R^i \end{pmatrix}, \quad x_L^i = \begin{pmatrix} c_L^i \\ h_L^i \\ 1 - l_L^i \end{pmatrix}, \quad e_R^i = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad e_L^i = \begin{pmatrix} 0 \\ H \\ 1 \end{pmatrix},$$

where subscript R denotes a representative working *R*esident and L denotes the *L*andlady.

On the production side, there are many firms in a city who employ one worker each and produce the identical immobile commodity in a perfectly competitive environment. The production plan y^i of a representative firm in city i is given by

$$y^i = \begin{pmatrix} f(l_F^i, h_F^i\,;\, S^i) \\ -h_F^i \\ -l_F^i \end{pmatrix},$$

where l_F^i denotes labor demand and h_F^i denotes land input used as a production site. We let the production function $f(\cdot)$ depend on the city size S^i to allow for externalities within the city such as knowledge spillover effects (as evidenced by [Hen03] for example) or congestion to have an impact on productivity of individual firms operating in the same city.[13]

Intracity production economy in city i is identified by

$$\mathcal{P}^i := \left\{ \left(X_N^i, \succsim_N, e_N^i \right)_{N \in \{R, L\}}, Y_F^i \right\},$$

where $X_N^i := \mathbb{R}_+^3$ is a consumption set of a representative worker or the landlady, $\succsim_N$ is a complete preorder over consumption set X_N^i, and Y_F^i is a production set of a representative firm given by

$$Y_F^i := \left\{ y^i = (c_F^i, h_F^i, l_F^i)' \in \mathbb{R}_+^3 : c_F^i \leq f(l_F^i, h_F^i; S^i) \right\}.$$

A feasible allocation in $\mathcal{P}^i$ is defined as follows:

Definition 3.1: Feasible Allocation

For given $S^i \in [0, S]$, an allocation $(x_R^i, x_L^i, y^i) \in X_R^i \times X_L^i \times Y_F^i$ in the intracity production economy $\mathcal{P}^i$ is feasible iff

$$x_R^i S^i + x_L^i = y^i S^i + e_R^i S^i + e_L^i. \tag{1}$$

To find GMP we need to compute the *value* of each commodity. Let $p^i := \left(1, r^i, w^i\right)'$ be the price on a composite good, housing lot and leisure. We take composite goods as a numéraire.[14] There are two equivalent ways to define GMP. From the production point of view, GMP $\mathcal{Y}^i$ is defined by the total value of all the final goods and services produced in the city, $\mathcal{Y}^i = p^i \cdot \left(y^i + e_R^i S^i + e_L^i\right)$. From the consumers' end, GMP is the sum of all the expenditures on goods and services, $\mathcal{Y}^i = p^i \cdot \left(x_R^i S^i + x_L^i\right)$. They come out to the same number due to Walras' law.

[13] Since we bundle all the goods in a single basket, there is no distinction between localization economies and urbanization economies in our model. An empirical study by Henderson [Hen86] supports the former whereas Carlino [Car79] supports the latter. Melo et al [MGN09] conclude that it all depends on the nature of the data employed and the type of the industry studied.

[14] Note that none of the commodities are tradable beyond the city border in this economy.

Definition 3.2: GMP

GMP in the intercity production economy $\mathcal{P}^i$ of size S^i is identified by

$$\mathcal{Y}^i := p^i \cdot \left(y^i + e^i_R S^i + e^i_L\right) = p^i \cdot \left(x^i_R S^i + x^i_L\right). \tag{2}$$

In application, GDP does not count leisure time. We consume leisure for the price of the opportunity cost (namely, lost wage), but in practice there is no explicit/accounting trace of market transactions for the consumption of leisure to track down the leisure portion of GDP. In particular we produce and consume $w^i\left\{\left(1 - l^i_R\right)S^i + \left(1 - l^i_L\right)\right\}$ worth of leisure, but this part is excluded from recorded GDP and by extension, from GMP as well.[15] Thus, we shall redefine $\mathcal{Y}^i$ with only the first two entries and take out the last entry (leisure)

$$Y^i := \begin{pmatrix}1\\ r^i\end{pmatrix} \cdot \left[\begin{pmatrix}c^i_R\\ h^i_R\end{pmatrix} S^i + \begin{pmatrix}c^i_L\\ h^i_L\end{pmatrix}\right] = \begin{pmatrix}1\\ r^i\end{pmatrix} \cdot \left[\begin{pmatrix}f(l^i_F, h^i_F; S^i)\\ -h^i_F\end{pmatrix} S^i + \begin{pmatrix}0\\ 0\end{pmatrix} S^i + \begin{pmatrix}0\\ H\end{pmatrix}\right] \tag{3}$$

for statistical purposes.

As we understand from our empirical findings in section 2, Y^i exhibits increasing returns to scale S^i. Then according to (3), individual profit $f\left(\cdot; S^i\right) - r^i h^i_F$ needs to be increasing in S^i so that Y^i is convex in S^i.

To find the equilibrium price vector, first define $\theta^i := (\theta^i_R, \theta^i_L)$ as a vector of a representative resident and landlady's share of profit (θ^i_R, $\theta^i_L \in [0, 1]$ and $\theta^i_R S^i + \theta^i_L = 1$).

Definition 3.3: Intracity Equilibrium

For a given θ^i and e^i, an intracity equilibrium in city i with $S^i > 0$ is a feasible allocation $\left(x^{i\,*}_R, x^{i\,*}_L, y^{i*}\right)$ and price vector p^{i*} such that

1. *For $N = R$ and L*
$$p^{i^*} \cdot x^{i\,*}_N \leq p^{i^*} \cdot e^i_N + \theta^i_N p^{i^*} \cdot y^{i^*} S^i. \tag{4}$$
2. *For $N = R$ and L*
$$p^{i^*} \cdot x^i_N \leq p^{i^*} \cdot e^i_N + \theta^i_N p^{i^*} \cdot y^{i^*} S^i \quad \Rightarrow \quad x^{i^*} \succeq_N x^i, \tag{5}$$
for any $x^i_N \in X^i$.
3. *For any $y_i \in Y^i_F$,*
$$p^{i^*} \cdot y^{i^*} \geq p^{i^*} \cdot y^i. \tag{6}$$

To identify the equilibrium city size, let the intercity production economy $\mathcal{P} := \left\{\left(\mathcal{P}^i, S^i\right)_{i=1}^I, S\right\}$ and define

Definition 3.4: Intercity Equilibrium

For a given ownership matrix $\left(\theta^i\right)_{i=1}^I \in [0, 1]^{2I}$ and endowment matrix $\left(e^i\right)_{i=1}^I \in \prod_i \left(X^i_R \times X^i_L\right)$, an intercity equilibrium in the production economy $\mathcal{P}$ is a list of a feasible allocation matrix $\left(x^{i\,*}_R, x^{i\,*}_L, y^{i*}\right)_{i=1}^I \in \prod_i \left(X^i_R \times X^i_L \times Y^i_F\right)$, price matrix $\left(p^{i*}\right)_{i=1}^I \in \mathbb{R}^{3I}_+$, and size distribution $\left(S^i\right)_{i=1}^I \in [0, S]^I$ such that for any i and j with $S^i > 0$ and $S^j > 0$,

[15] We use the data provided by the US Bureau of Economic Analysis, and their figures are based on tax reports. We do not pay tax on leisure.

1. $\left(x_R^{i\,*}, x_L^{i\,*}, y^{i*}, p^{i*}\right)$ *is an intracity equilibrium*
2. $$x_R^{i\,*} \sim_R x_R^{j\,*}. \tag{7}$$
3. *Urban population adds up to* $$\sum_i S^i = S. \tag{8}$$

The second item (7) is due to free mobility of workers. This does not apply to landladies, who are locked in their place of residence to keep the housing portion of GMP where it is generated.

The equilibrium city-size distribution is the size component of an equilibrium in $\mathcal{P}$ and the GMP distribution is (3) computed with an equilibrium in $\mathcal{P}$.

3.2 Application

To derive the exact distribution of GMP for empirical testing, consider an application of the spatial production model developed in section 3.1 with production function and labor market in the style of Eeckhout [Eec04] with the explicit presence of landladies (see figure 7 for a schematic representation of the agents and commodities involved in this example). We will find the analytical solution to the intercity equilibrium, from which we obtain the equilibrium GMP distribution.

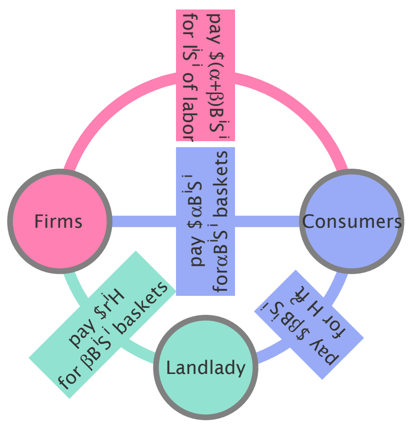


Figure 7. Commodity flow. Leisure is excluded in accordance with the practical definition of GDP adopted by the US Bureau of Economic Analysis.

3.2.1 Intracity Equilibrium

To start off, pick any city i and consider its intracity equilibrium. Firm's production plan is specified by

$$f\left(l_F^i, h_F^i; S^i\right) = A^i a_+\left(S^i\right) a_-\left(S^i\right) l_F^i, \tag{9}$$

where A^i is a stochastic citywide productivity parameter, $a_+(\cdot)(> 0)$ measures the positive externality shared among the firms operating within the same city, and $a_-(\cdot)(\in (0, 1))$ measures congestion externality. On the plus side, city size is assumed to raise the productivity of all the firms operating in the city. Positive externality enhances with size ($a'_+(\cdot) > 0$). On the downside, whereas each consumer supplies l_R^i units of gross labor, congestion externality adversely affects effective labor. The fraction $1 - a_-(S^i)$ of labor hours will be spent on commuting rather than on production. The level of reduction in effective labor aggravates with the size of a city ($a'_-(\cdot) < 0$).[16] Firms do not pay for the time lost in commuting and workers assume responsibility for the time cost of commuting. That is, firms will pay (ostensible) wages at the rate of ω^i only for the fraction of l^i_R when their worker is present at work, i.e., only for $a_-\left(S^i\right) l^i_R$ hours out of l^i_R. On an hourly basis, (effective) wage is knocked down to $w^i\left(S^i\right) := \omega^i a_-\left(S^i\right)$ for each hour devoted for work, inclusive of commuting time.[17] We will discuss the role of landlady's labor supply later. We assume that firms do not require land as input in accordance with [Eec04] for simplicity, but land can readily be incorporated into our production economy in section 3.1 as a factor of production.

Profit (6) turns into

$$p^i \cdot y^i = A^i a_+\left(S^i\right) a_-\left(S^i\right) l^i_F - \omega^i a_-\left(S^i\right) l^i_F = \left[B^i\left(S^i\right) - \omega^i a_-\left(S^i_t\right)\right] l^i_F, \tag{10}$$

where $B^i(S^i) := A^i a_+\left(S^i\right) a_-\left(S^i\right)$. Since production function (9) exhibits constant returns to scale in l^i_F,

$$p^i \cdot y^{i^*} = 0 \tag{11}$$

in equilibrium (otherwise y^{i^*} violates profit maximiation condition (6)). Hence, if $l^i_F > 0$, it must follow that

$$B^i(S^i) = \omega^i a_-\left(S^i\right)(= w^i\left(S^i\right)) \tag{12}$$

in equilibrium.

Note here that aggregate production may exhibit agglomeration economies due to positive externality $a_+(\cdot)$, but internal scale economies are still absent because individual production function is linear in l^i_F. For more on a dialectic between increasing and constant returns to scale, see Rossi-Hansberg and Wright [RHW07].

Next order of business is the consumers. Represent $\succeq_R$ in $\mathcal{P}^i$ by

$$u_R\left(c^i_R, h^i_R, l^i_R\right) = \alpha \log c^i_R + \beta \log h^i_R + \gamma \log\left(1 - l^i_R\right), \tag{13}$$

where α, β, $\gamma > 0$, and assume $\alpha + \beta + \gamma = 1$ without loss of generality. According to the feasibility condition (1) household income is given by $p^i \cdot e^i_R + \theta^i_R p^i \cdot y^i S^i$. Since firms earn zero profit (11), household income simplifies to labor income $p^i \cdot e^i_R = w^i\left(S^i\right) \cdot 1$ alone, with which to buy composite goods c^i_R, housing h^i_R and leisure $\left(1 - l^i_R\right)$ at the price of $p^i = \left(1, r^i, w^i\right)'$. Marshallian demand is

$$x^i_R\left(p^i, w^i\right) = \begin{pmatrix} c^i_R\left(p^i, w^i\right) \\ h^i_R\left(p^i, w^i\right) \\ 1 - l^i_R\left(p^i, w^i\right) \end{pmatrix} = \begin{pmatrix} \alpha w^i\left(S^i\right) \\ \beta w^i\left(S^i\right)/r^i \\ \gamma \end{pmatrix} = \begin{pmatrix} \alpha B^i\left(S^i\right) \\ \beta B^i\left(S^i\right)/r^i \\ \gamma \end{pmatrix} \tag{14}$$

The second equality holds as a result of profit maximization (6) and its consequence (12). Labor supply l^i_R of a typical household will be $1 - \gamma = \alpha + \beta$. Material balance

[16] Commuting may be embedded through a specific intracity structure such as a monocentric city model [Alo64]. Since we just need it to work as a dispersion force in each city, we summarized commuting in a single function $a_-(\cdot)$ for our purpose.

[17] Hence, the opportunity cost of leisure is $w^i\left(S^i\right)$ rather than ω^i.

(1) requires that $\left(1-l_R^i\right)S^i+1-l_L^i=-l_F^iS^i+1\cdot S^i+1$. Since utility maximization for the retired landlady (5) results in ${l_L^i}^*=0$ in this economy (see (17) below), ${l_R^i}^*={l_F^i}^*$, which furthermore implies that the equilibrium production plan will be

$$y^{i^*}=\begin{pmatrix} f\left(l_F^i,\,h_F^i;\,S^i\right) \\ 0 \\ -l_F^i \end{pmatrix}=\begin{pmatrix} (\alpha+\beta)B^i\left(S^i\right) \\ 0 \\ -(\alpha+\beta) \end{pmatrix}. \tag{15}$$

Turning to the landlady, represent $\succeq_L$ in $\mathcal{P}^i$ by

$$u_L\left(c_L^i,\,h_L^i,\,l_L^i\right)=c_L^i\mathbf{1}_{\{l_L^i=0\}}\left(l_L^i\right), \tag{16}$$

where $\mathbf{1}_{\{l_L^i=0\}}(\cdot)$ is an indicator function that takes the value of one when $l_L^i=0$ and zero otherwise. Since she is retired, any hour of labor $l_L^i>0$ will instantly push her utility level down to zero regardless of an increment in her utility level from an increased consumption of composite goods financed through her labor income.[18] That is, her utility level is nonnegative over the plane $l_L^i=0$ in $\mathbb{R}_+^3$ and zero elsewhere. Once again, since the share of zero profit (11) earns her nothing, the budget constraint (4) implies that the landlady's income is $p^i\cdot e_L^i=r^iH+w^i\left(S^i\right)$. Her Marshallian demand is

$$x_L^i\left(p^i,w^i\right)=\begin{pmatrix} c_L^i\left(p^i,w^i\right) \\ h_L^i\left(p^i,w^i\right) \\ 1-l_L^i\left(p^i,w^i\right) \end{pmatrix}=\begin{pmatrix} r^iH \\ 0 \\ 1 \end{pmatrix}. \tag{17}$$

Then residential utility maximization (14), profit maximization (15) and landlady's utility maximiation (17) rewrite feasibility condition (1) as

$$\begin{pmatrix} \alpha B^i\left(S_R^i\right) \\ \beta B^i\left(S_R^i\right)/r^i \\ 1-(\alpha+\beta) \end{pmatrix}S^i+\begin{pmatrix} r^iH \\ 0 \\ 1 \end{pmatrix}=\begin{pmatrix} (\alpha+\beta)B^i\left(S^i\right) \\ 0 \\ -(\alpha+\beta) \end{pmatrix}S^i+\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}S^i+\begin{pmatrix} 0 \\ H \\ 1 \end{pmatrix}. \tag{18}$$

from which, along with the first order condition (12), we can find the equilibrium price vector in city i as

$$p^{i^*}=\begin{pmatrix} 1 \\ r^i \\ w^i \end{pmatrix}=\begin{pmatrix} 1 \\ \beta B^i\left(S^i\right)S^i/H \\ B^i\left(S^i\right) \end{pmatrix} \tag{19}$$

(see figure 7 to trace the commodity flow in equilibrium over $\mathcal{P}^i$). Observe that the rent r^i goes up if 1) city i draws a good technological shock A^i, 2) positive externality $a_+\left(S^i\right)$ intensifies, or 3) city i becomes less crowded ($a_-\left(S^i\right)$). Likewise the leisure becomes expensive for the same reasons. Reasons 2) and 3) are triggered by urban growth, whereas reason 1) is independent of S^i.

[18] Alternatively, we could model her as an active worker, which complicates our notations without much gain in insight.

It is worth pointing out that our economy makes a judicious use of limited labor. The working population is capped at S. The intracity equilibrium and more noticeably, the intercity equilibrium, allocate more people to a city with a good production environment and pull back labor from a city of low productivity. Residential indirect utility is increasing in $B^i\left(S^i\right)$ in equilibrium (see section 3.2.2 below). Migration dynamics are such that there is an inflow when $B^i\left(S^i\right)$ is above the national average and an outflow if it is below. The economy has an auto-correcting mechanism built into it: If population allocation ever deviates from the equilibrium, it creates an economic incentive for people to move to a productive city.[19]

With feasibility condition (18) and the equilibrium price (19) we obtain GMP (2) in equilibrium as follows:

$$\begin{aligned} \mathcal{Y}^i &= p^{i*} \cdot \left(x_R^{i\,*} + x_L^{i\,*}\right) \\ &= \underbrace{(\alpha+\beta)B\left(S^i\right)S^i + \beta B\left(S^i\right)S^i}_{\text{Reported portion of GMP } Y^i \ (3)} + B\left(S^i\right)\left(\gamma S^i + 1\right) \\ &= (1+\beta)B\left(S^i\right)S^i + B\left(S^i\right). \end{aligned} \tag{20}$$

On the second line in (20) are the values of composite goods, housing and leisure for each. Only the first two are included in the reported GMP.

Equilibrium GMP (20) leads to the following:

PROPOSITION 3.1: CITYWIDE SCALE ECONOMIES IN INTRACITY ECONOMY

Consider the equilibrium in an intracity economy $\mathcal{P}^i$. The reported portion of GMP Y^i exhibits increasing returns to scale S^i iff $B^i\left(S^i\right)S^i$ exhibits increasing returns to scale S^i.

Proof. Apparent from (20). □

In reference to section 2.4, observed data seem to suggest that positive externality does outstrip negative externality. However, as we will see in proposition 3.2, citywide scale economies in *inter*city equilibrium will be positive without specifically assumed increasing returns to scale on $B^i\left(S^i\right)S^i$.

3.2.2 INTERCITY EQUILIBRIUM

To find the intercity equilibrium in definition 3.4, rewrite indifference principle (7) in terms of a utility function (13) so that indirect utility $u\left(x_R^{i\,*}\right) = u\left(x_R^{j\,*}\right)$ for any i and j with $S^i, S^j > 0$. This leads to

$$B^i\left(S^i\right)\left(S^i\right)^{\frac{-\beta}{\alpha}} = B^j\left(S^j\right)\left(S^j\right)^{\frac{-\beta}{\alpha}} =: K, \tag{21}$$

where K is a location-invariant constant. According to (20), GMP is $\mathcal{Y}^i = (1+\beta)K\left(S^i\right)^{\frac{\beta}{\alpha}+1} + K\left(S^i\right)^{\frac{\beta}{\alpha}}$ and the reported portion of GMP will be

$$Y^i = (\alpha+2\beta)K\left(S^i\right)^{\frac{\beta}{\alpha}+1} \tag{22}$$

[19] The allocation of labor is still not efficient though, due to externalities.

in the neighborhood of the equilibrium size, which breaks down into composite goods production/consumption of $(\alpha+\beta)K\left(S^i\right)^{\frac{\beta}{\alpha}+1}$ and housing production/consumption of $\beta K\left(S^i\right)^{\frac{\beta}{\alpha}+1}$. Note that productivity parameter A^i no longer makes an explicit appearance in (22) but GMP is still positively related to it: A high value of A^i is reflected in Y^i through increased S^i. Now we have

Proposition 3.2: Citywide Scale Economies in Intercity Economy

If an intercity economy $\mathcal{P}$ is in equilibrium, the reported portion of GMP Y^i exhibits increasing returns to size in the neighborhood of equilibrium size S^i.

Proof. Immediate from (22). □

In comparison to proposition 3.1 it is curious that we have to assume $B^i\left(S^i\right)S^i$ to be increasing returns to scale only in $\mathcal{P}^i$ but not in $\mathcal{P}$. The short answer to this enigma is that free mobility puts the city size where scale economies are at work. For illustrative purposes, assume that $a_+\left(S^i\right)a_-\left(S^i\right)$ takes the form $\left(S^i\right)^{\delta\left(S^i\right)}$. Then proposition 3.1 specifically requires $\delta\left(S^i\right) > 0$ in $\mathcal{P}^i$ but proposition 3.2 does not because perfect mobility will bring $\delta\left(S^i\right)$ above zero anyway. Now assume that $\mathcal{P}$ is in equilibrium. Suppose that in some city i, size S^i rose by one (call this new resident Axel). In this case, housing consumption will be reduced in the city because residents have to make room for Axel's house out of a fixed supply H of land, and he also exacerbates congestion in the city. However, since $\mathcal{P}$ is in equilibrium, reduction in utility level from curtailed housing consumption needs to be offset by either c^i_R or l^i_R in compliance with utility equalization (7). Since l^i_R is independent of size (i.e., leisure consumption does not and cannot accommodate the change to the city residents introduced by Axel), compensation must be made through increased consumption in c^i_R alone. Then the question becomes: Can he produce enough composite goods to leave everyone in city i on the same indifference curve?

It is the answer to this question that yields endogenous agglomerative force in $\mathcal{P}$ in equilibrium. The marginal rate of substitution between c^i_R and h^i_R is $\frac{\beta}{\alpha}\frac{c^i_R}{h^i_R}$ baskets of composite goods for each ft^2. Now, Axel carves $\frac{\partial h^i_R}{\partial S^i}=\frac{h^i_R}{S^i}$ ft^2 from every resident's lot in the city. Thus, each city resident needs to have $\left(\frac{\beta}{\alpha}\frac{c^i_R}{h^i_R}\right)\frac{h^i_R}{S^i}=\frac{\beta}{\alpha}c^i_R\frac{1}{S^i}$ more baskets to keep to the countrywide utility level (or else the current allocation will not be an equilibrium). Then the addition of Axel into the city needs to raise the individual production of composite goods as follows:

$$\begin{aligned}
\frac{\partial f\left(\cdot\,;S^i\right)}{\partial S^i} &= \frac{\beta}{\alpha}f\left(\cdot\,;S^i\right)\frac{1}{S^i}\\
\Rightarrow\quad \delta\left(S^i\right)A^i\left(S^i\right)^{\delta\left(S^i\right)-1} &= \frac{\beta}{\alpha}A^i\left(S^i\right)^{\delta\left(S^i\right)-1}\\
\Rightarrow\quad \delta\left(S^i\right) &= \frac{\beta}{\alpha}(>0).
\end{aligned}\tag{23}$$

[20] If not, for example, if $\delta\left(S^i\right) < \frac{\beta}{\alpha}$, then Axel cannot make up for the lost individual housing unit by producing more composite goods through enhanced pooled production externality net of the congestion externality. The knowledge spillover effect he brings in (less the congestion externality he exerts) is not enough to render the dwindled housing consumption tolerable for the current residents. In this case, city i is better off bumping him out, i.e., it should reduce S^i, contradicting the fact that $\mathcal{P}$ is in equilibrium. And vice versa, city i should be larger if $\delta\left(S^i\right) > \frac{\beta}{\alpha}$. Everyone welcomes Axel and wants more residents to move in in this case. Thus, free mobility arbitrages the gap between externality component $\delta\left(S^i\right)$ and countrywide constant $\frac{\beta}{\alpha}$ and *forces* the city to operate in the domain where scale economies are present. Note that utility equalization (7) only applies to cities with $S^i > 0$. If city's aggregate production function does *not* exhibit increasing returns to scale *anywhere* over $0 < S^i \leq S$, intercity migration will leave the city trailing behind others and eventually turn it rural. At such a location, $\delta\left(S^i\right) < \frac{\beta}{\alpha}$ and all the residents will be drained off to other cities until S^i shrinks to zero. Thus, increasing returns to scale at the aggregate level are an eligibility requirement to be listed under MSA. See appendix A.3 for further discussion on naturally occurring scale economies in $\mathcal{P}$ as opposed to presumed scale economies in $\mathcal{P}^i$.

Note that the preceding argument refers only to the positive aspect of intercity resource allocation. The equilibrium $\mathcal{P}$ guarantees that utility level is only *equalized* but not necessarily *maximized*. Free mobility simply assigns workers to each city according to exogenously drawn productivity parameter A^i so as to even out the utility level across the country. For instance, Boston may use *more* people to raise its intracity utility level while New York may use *fewer* people to raise its intracity utility level. Nevertheless, the city-size and GMP distribution may stay put as long as the utility level is equalized among cities. A worker makes a location choice based on her own utility level without reference to the boon and scar that she leaves on existing residents' welfare. A social planner may impose an optimal allocation of workers to maximize, for example, a population-weighted utility level over the country, but he will have to forgo utility equalization (7) to do so, which may not be sustainable in practice.[21]

It is crucial that we unbundle housing consumption from the composite good. Location-variant housing rent plays a vital role in the economy as we discuss in appendix A.3. If we include housing as part of a composite good, per capita consumption level, and consequently the individual production level, will be the same across the cities because the consumption of leisure is the same everywhere and free mobility guarantees an equal utility level. Then aggregate production level becomes directly proportional to the city size. Alternatively, we can unbundle the composite good and create markets for many commodities. In that case we may have increasing returns to scale in *production* but the price of individual commodi-

[20] We took the landlady out of equation because her marginal rate of substitution between composite goods and housing is zero. Cutbacks in housing lot do not affect her at all due to her preferences (16).

[21] There have been several attempts to identify the optimal city-size *distribution*. There is no agreement on the objective function to use in the literature. It can be a social utility level as we mentioned in the main text or as in [Fuj13]. It can also be a cost of public good provision [SSW74], GMP [Suh91], or benefit from local interaction [Tab82].

ties tend to negate the variations in output levels and *GMP* will be only proportional to the city size. A positive technological shock enhances the production, which reduces the equilibrium price in a perfectly competitive market. Thus, the *value* of the output will exhibit constant returns to scale, which is not compatible with our findings in section 2.4. We will have to forgo the assumption of perfectly competitive market in this case.

3.3 Distribution of GMP

Eeckhout [Eec04] has shown that S^i follows the lognormal distribution due to the central limit theorem. The equilibrium size of a city can be written historically as a sum of the log of error terms over time. *The city size depends on the cumulative effect of multiplicative nature* (cf. footnote 1) rather than of additive nature [LSA01], leading to the lognormal distribution as a result of proportional growth (see appendix A.4 for details). In particular $\log\left(S^i\right) \sim N\left(\mu_S, \sigma_S^2\right)$. In conjunction with (22) we obtain the following:

Proposition 3.3: GMP Distribution
The reported portion of GMP follows a lognormal distribution:

$$\log Y \sim N\left(\left(\frac{\beta}{\alpha}+1\right)\mu_s + \log(\alpha+2\beta)K, \left\{\left(\frac{\beta}{\alpha}+1\right)\sigma_s\right\}^2\right) \tag{24}$$

There is a log-linear relationship between GMP and city size (22) and city size follows a lognormal distribution. Naturally, GMP also follows a lognormal distribution by extension. The variance of $\log(Y)$ is inflated by $\frac{\beta}{\alpha}+1$ due to citywide scale economies (proposition 3.2). This observation is consistent with our findings in section 2.3. GMP (in log scale) spreads further than its city-size counterpart in $\mathcal{P}$.

Eeckhout [Eec04] also establishes Gibrat's law $\frac{d\log S_t^i}{dt} \approx \epsilon$ (t denotes time. See appendix A.4). Then from (22),

$$\frac{d\log Y_t^i}{dt} = \left(\frac{\beta}{\alpha}+1\right)\frac{d\log S_t^i}{dt} \approx \left(\frac{\beta}{\alpha}+1\right)\epsilon_t^i,$$

where ϵ_t^i is an i.i.d. random variable. Thus GMP also follows Gibrat's law. Note, once again, that variation in Y^i is inflated by $\frac{\beta}{\alpha}+1$ and this is coherent with our empirical findings in section 4.1.

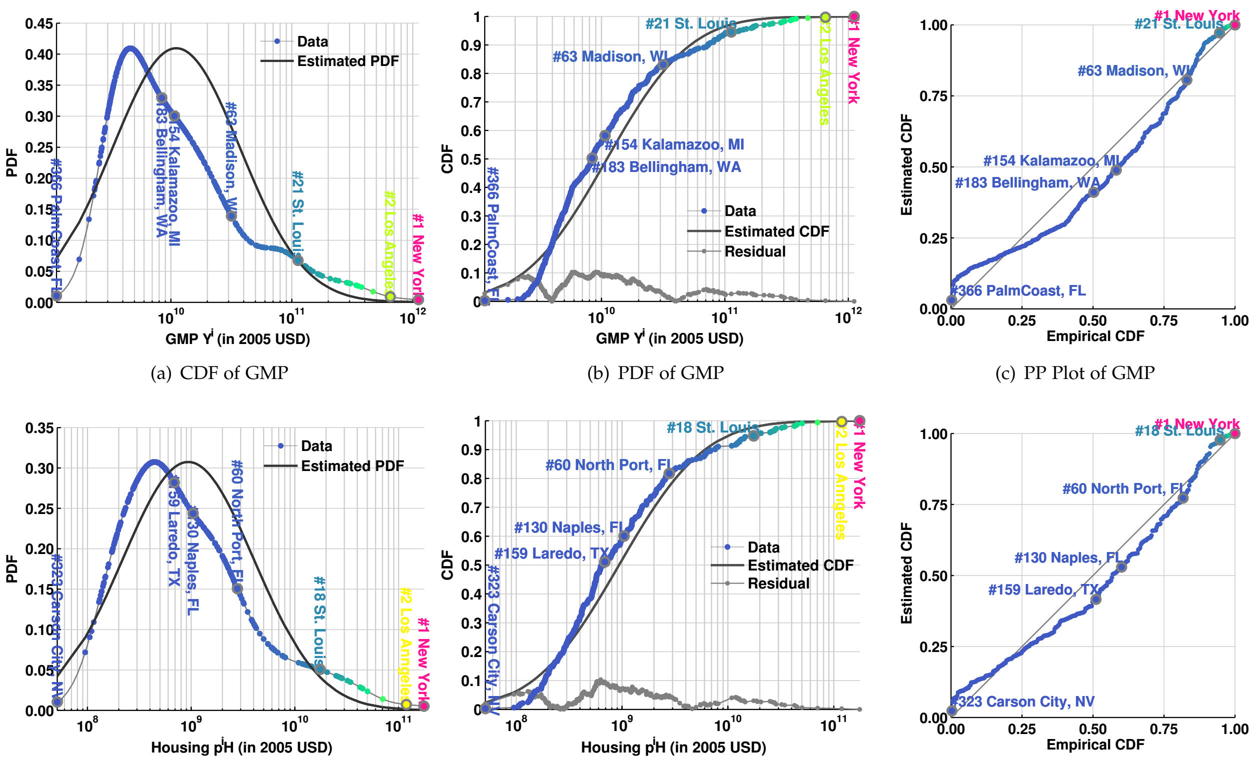
4 Empirical Implementation

4.1 Distribution of GMP

We will put proposition 3.3 to an empirical test in this section. First, rewrite the GMP distribution (24) as $\log Y \sim N(\mu, \sigma^2)$. The maximum likelihood estimator of (24) is $\hat{\mu} = \frac{\sum \log Y^i}{N}$ and $\hat{\sigma}^2 = \frac{\sum\left(\log Y^i - \hat{\mu}\right)^2}{N}$. We report our estimations in table 5 with

	Employment	GMP	Housing
Data Size I	366	366	323
Censored (Unreported Cities)	0%	0%	11.75%
Censored (Unreported Value)	0%	0%	13.87%
Estimated Mean $\hat{\mu}_S$, $\hat{\mu}$	12.68	23.13	20.66
Estimated Variance $\hat{\sigma}_S^2$, $\hat{\sigma}^2$	1.128	1.478	2.133
Theoretical Variance $\left\{\left(\frac{\hat{\beta}}{\hat{\alpha}}+1\right)\hat{\sigma}_S\right\}^2$	–	1.571	1.571
Log Likelihood/I	.2709	.2370	.1968
Kolomogorov-Smirnov Statistic	.1122	.1032	.1031

Table 5.



(a) CDF of GMP (b) PDF of GMP (c) PP Plot of GMP

(d) CDF of Housing (e) PDF of Housing (f) PP Plot of Housing

Figure 8.

supporting density plots in figure 8. Housing portion of GMP, $\beta K\left(S^i\right)^{\frac{\beta}{\alpha}+1}$, is also available and they are expected to follow the lognormal distribution as well.

The overall fit is not too far off. The maximum discrepancy between the empirical and estimated CDF (Kolomogorov-Smirnov statistic) is .1032. We are more than certain that the fit would improve if we used an inclusive data set. Due to truncation to the left of the distribution, the tail end of the distribution does not extend as far as the theory predicts. GMP is not available for smaller cities.

As we will see in section 4.2, the ratio between β and α is .1800. Following the GMP distribution (24), the theoretically expected value of σ^2 is $(1.180\hat{\sigma}_S^2)^2 = 1.571$ (cf. theoretical variance in table 5). Our expected variance in GMP computed from the expected expenditure shares α and β and estimated variance σ^2 in employment is very close to the actual variance in GMP (we missed the actual value only by 5.86%). This confirms the validity of the form (24) along with the selection of the utility and production functions in section 3.2. On the other hand, the variance on housing is larger than the theoretical value by 35.86%. We will explore the cause of the large gap in housing in section 4.2.

4.2 Scale Economies

According to (22), the ratio between GMP and housing is $\alpha + 2\beta$ to β. The actual ratio is \$1.162e+13 to \$1.603e+12 among the MSA's, indicating that the expenditure share β of the housing sector is 13.23%.[22] Hence, the expected ratio $\frac{\hat{\beta}}{\hat{\alpha}} = .1800$. Taking a log of (22),

$$\log Y^i = \log(\alpha + 2\beta)K + \left(\frac{\beta}{\alpha} + 1\right)\log S^i = \log(\alpha + 2\beta)K + 1.180\log S^i.$$

The actual value of the coefficient is 1.117 in table 4 rather than 1.180, meaning that our model overshoots the coefficient by only 5.64% (or, the economy is off where it is supposed to be by 5.64%). The housing portion of GMP $\log \beta K \left(S^i\right)^{\frac{\beta}{\alpha}+1}$ also shares the same coefficient in (22). Here the predicted value comes *short* of the actual value by 10.59%. The large discrepancy may be because of the censored data,[23] or the landlords' or developers' registered addresses, which may be different from the city where they real estate is located. The fact that imputed rent is excluded and that houses usually last longer than the duration of a fiscal year exacerbates the deviation even further.

5 Conclusion and Extensions

There are two sets of research on the nature of geographic concentration of economic activities: one on agglomeration economies and the other on the city-size distribution. Despite the fact that they address the closely related questions of the existence and size of a city, a line of communication between the two has never

[22] Note that GDP includes real estate sales and excludes imputed rents. Thus, the figure is related to, but does not necessarily represent the expenditure share in a particular year.

[23] Housing output values are not reported for all MSA's and the missed portion is not negligible. For example, housing sector in Dallas (rank #4 in employment) is censored out of the data. However, these censored values are included in the citywide and nationwide aggregates.

been established. We filled in the gap with the aim of making explicit the indispensable role that free mobility plays in characterizing the nature of agglomeration economies and providing an extra checking device for the models of the city-size distribution.[24]

We have discovered that the GMP distribution follows the same pattern to the city-size distribution, and we sought a systematic illustration of how our local economies are related to their employment and output levels on a national scale. Proposition 3.2 further revealed that GMP grows more than 1% against a 1% growth in employment. Large cities make up for an exceeding share of GDP and they do so more than their city size alone can account for. Consequently, due to agglomeration economies of GMP in employment, the GMP distribution is even more skewed than the city-size distribution.

We constructed a production economy model that endogenously gives rise to agglomeration economies in equilibrium. The interplay between externalities and housing consumption drives the cities to operate at the size where increasing returns to scale are present. Empirical testing verifies our model's prediction; however, the result could have been better. Due to data truncation, our predicted distribution does not trace the lower end of the distribution well. Ideally, we would like to test our prediction with a more exhaustive data set, which, for the moment, does not exist.

Our objective was to explain the GMP distribution in a consistent manner. As far as we know this is the first attempt to analyze GMP as a distribution. Along the way, we have left several prospects for future work. We assumed a single-input production function. Local output may well be affected by capital (cf. appendix A.2), educational attainment, location of the city, access to a rich labor pool, or urban infrastructure, which, obviously vary from city to city. We also packed the consumption goods other than housing and leisure into a single basket. In reality a city comes with various industries. Some of them may exhibit increasing returns to scale and some may not both within and across the industries. Cross-sectional GMP analysis is called for to decode the internal workings of local economies that, on aggregate, exhibit increasing returns to scale and result in a fat-tail distribution. Lastly, we assumed that all goods are immobile. Data reported by the Bureau of Economic Analysis are based on tax filed in each MSA. As such, the scope of GMP matches the range of production in MSA, but it does not necessarily match the range of consumption in the city. The openness of a spatial economy may be addressed by adding shipping firms to the production economy. One way to do so is to assume that a shipping firm takes goods in city i as input and "produces" the same good in city j as output, in less than a one-to-one ratio to reflect shipping charges of iceberg form.

[24] For example, Eeckhout's model [Eec04] that we based our model on passes the test, and thus it does not only explain the city-size distribution but also the GMP distribution.

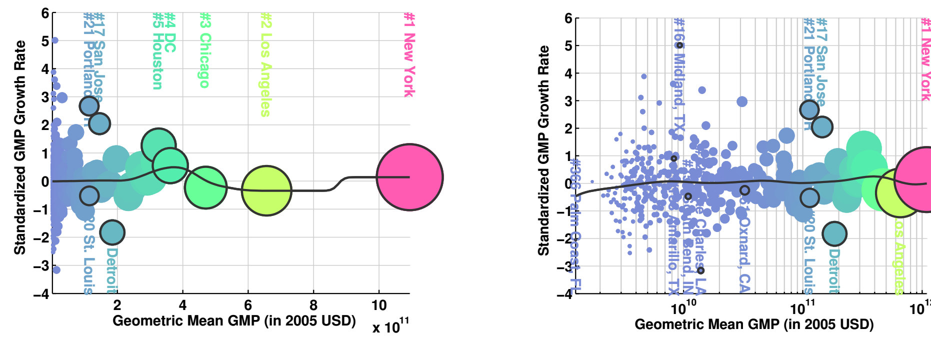
A Appendix

A.1 Non-Parametric Estimation of Growth Rate

We estimate the conditional expectation of GMP growth rate $E[g|Y] = m(Y)$ with a Nadaraya-Watson kernel estimator [Wat64]

$$\hat{m}(Y) = \sum_{i=1}^{I} g^i \frac{K_h(Y - Y^i)}{\sum_{j=1}^{I} K_h(Y - Y^j)}.$$

Y denotes GMP and g denotes its growth rate. Sample size is $I = 366$ with each city indexed by a superscript i. $K_h(\cdot)$ is a scaled kernel with a bandwidth h. For non-parametric estimation we standardize the growth rate to take out the nation-wide growth rate, which is captured by the intercept in parametric estimation in section 2.2. Figure 9 plots the growth rate and its kernel estimation. We tried to



(a) Kernel regression on GMP with bandwidth $h = 6.137e + 10$.

(b) Kernel regression on the log of GMP with bandwidth $h = .3932$.

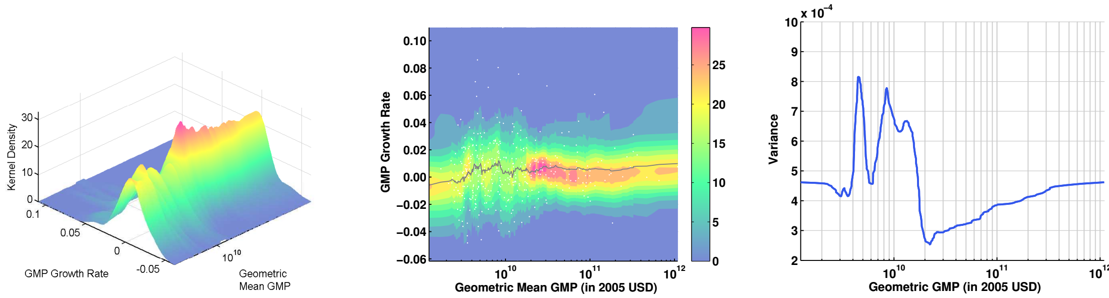
Figure 9.

estimate $\hat{m}(Y)$ first (figure 9(a)). The disperse spread of GMP towards the upper end swings the estimate from side to side and makes it hard to interpret the relationship. The log of GMP in figure 9(b) seems to exhibit a more discernible pattern. There is a slight inclination to the left and right tails, probably because of a smaller number of observations to the both ends than in the mid range. Other than that, our estimate seems to be in support of Gibrat's law for GMP.

Figure 10 presents the kernel estimate of GMP growth rate. Aside from an increase in variance in the lower mid range and decrease in the mid range, there does not seem to be a systematic correlation between GMP and its variance in growth rate.

A.2 Capital Stock

The two lines of work mentioned in section 1 take different approaches to theorizing about their respective target objective. The city-size distribution models based



(a) Kernel Density Estimation

(b) Kernel Density Estimation. Gray line tracks the mean.

(c) Kernel Estimate of Variance

Figure 10.

on general equilibrium typically do not include capital stock as part of the production function ([Dur07], [Eec04] for example);[25] whereas most agglomeration models do. Labor alone serves its purpose to explain the actual city-size distribution without involvement of capital stock. It is easy to measure a city size, but measuring citywide capital stock is not as straightforward as a head count. In fact, there is no data on the level of capital stock at city level in the Untied States. Those studies that quote capital stock use the estimated level based on factors related to capital such as local public goods, housing and state roads, mixed in with predetermined weights [Seg76], or estimated retrospectively *from* the pair of labor and GDP per capita at city level [Sve75]. Capital stock is known to be correlated with city size, which causes a multi-collinearity problem. According to [Seg76], capital stock's contribution to GMP is .116 as opposed to labor's .891. We did not test our model on capital stock for lack of its direct measurement. Capital stock can be readily incorporated into our model in way similar to land.

A.3 Unconditional Scale Economies

To further understand the reasoning behind proposition 3.2, conduct comparative statics on $\frac{\beta}{\alpha}$. Imagine that β goes down (or equivalently, α goes up). In intracity equilibrium in $\mathcal{P}^i$, $B^i\left(S^i\right)$ can take any value. In intercity equilibrium in $\mathcal{P}$, $B^i\left(S^i\right)$ is subject to utility equalization condition (21). As the expenditure share of housing β decreases, we can pack lots of people in a city (because they do not care about the lot size much) and produce lots of composite goods (which they do care about). The story ends here in $\mathcal{P}^i$. In $\mathcal{P}$, it goes further. Increased city size makes the city appealing because people are willing to swap a large parcel of land for only a few composite goods to squeeze Axel in when β is small. People view a large city with small houses more favorably than a small city with large houses. To offset

[25]There are some exceptions. For example, Rossi-Hansberg and Wright [RHW07] address city-size distribution with capital stock incorporated into the model. Even then, actual capital stock level is not used for empirical testing.

the rush of people into a large city, the effective wage rate (19) in the city goes down in equilibrium to meet the shared utility level (7). $\mathcal{P}^i$ does not factor in the levelling effect of the wage across the cities as much as $\mathcal{P}$ does. Notice that as β becomes smaller, scale economies also weaken because the residents do not care about housing, and they become more willing to give up their land in exchange for even a limited increase in composite goods. The housing market is indispensable in this sense to observe endogenously induced agglomeration economies in $\mathcal{P}$.

This observation compares to a closed and open monocentric city model (cf. [Bru87]). In a closed monocentric city, size is exogenous but utility level is endogenous just as in $\mathcal{P}^i$. In an open monocentric city, size is endogenous but utility level has to match the national level as in $\mathcal{P}$. Since the wage rate depends on city size, $\mathcal{P}$ picks the levelling effect of wage but $\mathcal{P}^i$ does not, and therefore, we have to throw in an additional assumption on production function in $\mathcal{P}^i$.

A.4 The City-Size Distribution and Gibrat's Law

Denote discrete time by subscript t and define $\Lambda\left(S_t^i\right) := a_+\left(S_t^i\right)a_-\left(S_t^i\right)\left(S_t^i\right)^{\frac{-\beta}{\alpha}}$ and suppose $\Lambda(\cdot)$ is invertible in the neighborhood of equilibrium S^i. Then $A_t^i\Lambda\left(S_t^i\right) = K$ from (21) so that $S_t^i = \Lambda^{-1}(K/A_t^i)$. With the law of motion $A_t^i = (1+\sigma_t^i)A_{t-1}^i$, we have

$$S_t^i = (1+\epsilon_t^i)S_{t-1}^i, \tag{25}$$

where $1+\epsilon_t^i := 1/\Lambda^{-1}(1+\sigma_t^i)$. Then

$$\frac{d\log S_t^i}{dt} \approx \epsilon_t^i,$$

for a small ϵ, leading to Gibrat's law. Computing the city size recursively, (25) also implies

$$\log S_t^i \approx \log S_0^i + \sum_{\tau=1}^{t}\epsilon_t^i,$$

which leads to the lognormal city-size distribution as $t\to\infty$ by the central limit theorem. See [Eec04] for details.

References

[ABL07] Fredrik Andersson, Simon Burgess, and Julia I Lane, *Cities, matching and the productivity gains of agglomeration*, Journal of Urban Economics **61** (2007), no. 1, 112–128.

[Alo64] William Alonso, *Location and land use. Toward a general theory of land rent*, Harvard University Press, 1964.

[BF92] Marcus Berliant and Masahisa Fujita, *Alonso's discrete population model of land use: Efficient allocations and competitive equilibria*, International Economic Review (1992), 535–566.

[BK00] Marcus Berliant and Hideo Konishi, *The endogenous formation of a city: Population agglomeration and marketplaces in a location-specific production economy*, Regional Science and Urban Economics **30** (2000), no. 3, 289–324.

[Bru87] Jan K Brueckner, *The structure of urban equilibria: A unified treatment of the Muth-Mills model*, Handbook of Regional and Urban Economics **2** (1987), 821–845.

[BRW06] Marcus Berliant, Robert R Reed, and Ping Wang, *Knowledge exchange, matching, and agglomeration*, Journal of Urban Economics **60** (2006), no. 1, 69–95.

[BW93] Marcus Berliant and Ping Wang, *Endogenous formation of a city without agglomerative externalities or market imperfections: Marketplaces in a regional economy*, Regional Science and Urban Economics **23** (1993), no. 1, 121–144.

[Car79] Gerald A Carlino, *Increasing returns to scale in metropolitan manufacturing*, Journal of Regional Science **19** (1979), no. 3, 363–372.

[Car87] ______, *Productivity in cities: Does city size matter?*, Federal Reserve Bank of Philadelphia Business Review (1987), 3–12.

[CPRHS14] Lorenzo Caliendo, Fernando Parro, Esteban Rossi-Hansberg, and Pierre-Daniel Sarte, *The impact of regional and sectoral productivity changes on the US economy*, Tech. report, National Bureau of Economic Research, 2014.

[CS04] Federico Cingano and Fabiano Schivardi, *Identifying the sources of local productivity growth*, Journal of the European Economic Association **2** (2004), no. 4, 720–742.

[DP04] Gilles Duranton and Diego Puga, *Micro-foundations of urban agglomeration economies*, Handbook of regional and urban economics **4** (2004), 2063–2117.

[Dur07] Gilles Duranton, *Urban evolutions: The fast, the slow, and the still*, The American Economic Review (2007), 197–221.

[Eec04] Jan Eeckhout, *Gibrat's law for (all) cities*, American Economic Review (2004), 1429–1451.

[EG94] Glenn Ellison and Edward L Glaeser, *Geographic concentration in us manufacturing industries: A dartboard approach*, Tech. report, National Bureau of Economic Research, 1994.

[FKV99] Masahisa Fujita, Paul Krugman, and Anthony J. Venables, *The spatial economy*, The MIT Press, 1999.

[Fuj13] Shota Fujishima, *Evolutionary implementation of optimal city size distributions*, Regional Science and Urban Economics **43** (2013), no. 2, 404–410.

[Gab99] Xavier Gabaix, *Zipf's law for cities: an explanation*, Quarterly Journal of Economics (1999), 739–767.

[GI04] Xavier Gabaix and Yannis M Ioannides, *The evolution of city size distributions*, Handbook of Regional and Urban Economics **4** (2004), 2341–2378.

[GZS10] Kristian Giesen, Arndt Zimmermann, and Jens Suedekum, *The size distribution across all cities – double Pareto lognormal strikes*, Journal of Urban Economics **68** (2010), no. 2, 129–137.

[Hen86] J Vernon Henderson, *Efficiency of resource usage and city size*, Journal of Urban economics **19** (1986), no. 1, 47–70.

[Hen03] ______, *Marshall's scale economies*, Journal of Urban Economics **53** (2003), no. 1, 1–28.

[IO03] Yannis M Ioannides and Henry G Overman, *Zipfs law for cities: an empirical examination*, Regional Science and Urban Economics **33** (2003), no. 2, 127–137.

[Kim95] Sukkoo Kim, *Expansion of markets and the geographic distribution of economic activities: The trends in US regional manufacturing structure, 1860-1987*, The Quarterly Journal of Economics (1995), 881–908.

[Kru91] Paul Krugman, *Increasing returns and economic geography*, Journal of Political Economy **99** (1991), no. 3, 483–499.

[LSA01] Eckhard Limpert, Werner A Stahel, and Markus Abbt, *Log-normal distributions across the sciences: Keys and clues*, BioScience **51** (2001), no. 5, 341–352.

[MGN09] Patricia C Melo, Daniel J Graham, and Robert B Noland, *A meta-analysis of estimates of urban agglomeration economies*, Regional Science and Urban Economics **39** (2009), no. 3, 332–342.

[MN05] Giordano Mion and Paolo Naticchioni, *Urbanization externalities, market potential and spatial sorting of skills and firms*, CEPR Discussion Papers (2005).

[Mor04] Enrico Moretti, *Human capital externalities in cities*, Handbook of regional and urban economics **4** (2004), 2243–2291.

[RB88] Francisco L Rivera-Batiz, *Increasing returns, monopolistic competition, and agglomeration economies in consumption and production*, Regional Science and Urban Economics **18** (1988), no. 1, 125–153.

[RHW07] Esteban Rossi-Hansberg and Mark LJ Wright, *Urban structure and growth*, The Review of Economic Studies **74** (2007), no. 2, 597–624.

[Seg76] David Segal, *Are there returns to scale in city size?*, The Review of Economics and Statistics (1976), 339–350.

[She73] Daniel Shefer, *Localization economies in SMSA's: A production function analysis*, Journal of Regional Science **13** (1973), no. 1, 55–64.

[Soo05] Kwok Tong Soo, *Zipf's law for cities: A cross-country investigation*, Regional science and urban Economics **35** (2005), no. 3, 239–263.

[SSW74] Joseph A Swanson, Kenneth R Smith, and Jeffrey G Williamson, *The size distribution of cities and optimal city size*, Journal of Urban Economics **1** (1974), no. 4, 395–409.

[Sta78] David Starrett, *Market allocations of location choice in a model with free mobility*, Journal of Economic Theory **17** (1978), no. 1, 21–37.

[Suh91] Seoung Hwan Suh, *The optimal size distribution of cities*, Journal of Urban Economics **30** (1991), no. 2, 182–191.

[Sve75] Leo Sveikauskas, *The productivity of cities*, The Quarterly Journal of Economics (1975), 393–413.

[Tab82] Takatoshi Tabuchi, *Optimal distribution of city sizes in a region*, Environment and Planning A **14** (1982), no. 1, 21–32.

[Wat64] Geoffrey S Watson, *Smooth regression analysis*, Sankhyā: The Indian Journal of Statistics, Series A (1964), 359–372.