Fu, Yuming; Hao, Yang

**Conference Paper**

# An Urban Accounting for Geographic Concentration of Skills and Welfare Inequality

55th Congress of the European Regional Science Association: "World Renaissance: Changing roles for people and places", 25-28 August 2015, Lisbon, Portugal

**Provided in Cooperation with:**
European Regional Science Association (ERSA)

*Suggested Citation:* Fu, Yuming; Hao, Yang (2015) : An Urban Accounting for Geographic Concentration of Skills and Welfare Inequality, 55th Congress of the European Regional Science Association: "World Renaissance: Changing roles for people and places", 25-28 August 2015, Lisbon, Portugal, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:
https://hdl.handle.net/10419/124703

# An Urban Accounting for Geographic Concentration of Skill and Welfare Inequality

*Yuming Fu*[#], *Yang Hao*[$]
*March 2015*
*JEL Code: E24, J24, R11, R13*
(Preliminary version, do not cite)

## Abstract

Using Jones (2014) generalized human capital accounting, we extend the urban accounting model of Desmet and Rossi-Hansberg (2013) to account for the geographic distribution of skills across US metropolitan areas. The methodology allows the productivity of high-skill workers to depend on location advantage and local skill mix; the latter also determines the productivity of low-skill workers. Urban friction, rising with population size, reduces worker consumption relative to their wage income. Amenities for high-skill and low-skill workers in each city are calibrated so that the utility for each skill type is equalized across cities. We examine counterfactual skill-mix distribution across cities and welfare gap between the skill groups by shutting down spatial heterogeneity in location advantage, amenity and excess friction respectively. We show that skill mix becomes more even across cities absent heterogeneity in location advantages or in excess friction but it becomes more dispersed absent amenity heterogeneity. The welfare gap widens when heterogeneity in any of the three factors is eliminated. The generalized urban accounting model can shed light on the causes of increased concentration of skilled workers in large cities in US highlighted by E. Moretti (2008) and Diamond (2012) and the implications for welfare gap between the skill groups.

**Key words**: urban accounting, generalized human capital accounting, skill distribution, welfare inequality

[#] Associate Professor, National University of Singapore, yuming.fu@nus.edu.sg. Corresponding at: Department of Real Estate, 4 Architecture Drive, Singapore 117566.

[$] PhD Candidate, National University of Singapore, hao.yang@nus.edu.sg. We thank all the participants of DRE Brown Bag lunch Seminar for their comments, especially critical reviews from Professor Deng Yongheng and Dr Li Qiang.

## *1. Introduction*

The United States and even all other countries in the world have been going through a tremendous structural transformation from manufacturing-oriented to knowledge-based economy. As with the intensified globalization and fast-upgrading new technologies, higher educated nations will continue to win in the competition for greater economic performance. So it is for the changing geography within America cities over the last thirty years. High-skill workers are largely demanded in more educated cities, meanwhile enhancing productivity of these skilled cities. As E. Moretti (2012) classifies, there exists "The Three Americas": Brain hubs, traditional industries, and the ones in the middle. Top-ten smartest cities gauged in terms of the percentage of college graduates of local labor force in America are Seattle (53%), San Francisco (50%), Raleigh (50%), Washington, D.C. (45%), Austin (44%), Minneapolis (43%), Atlanta (43%), Boston (41%), San Diego (40%) and Lexington (40%) (Neil Johnson, U.S. News University Directory, 2012). It is apparent that except college towns, Raleigh and Lexington, eight out of the ten are extremely large metropolises. In addition, high-skill workers are more and more tempted by desirable amenities cities can provide. For instance, natural amenities such as cooler summer, warm winter and shorter distance to coastal area, etc, and social amenities such as theater, café, museum, and even shopping centers, will pin down a picture of rise in locations equipped with them. These suggest that high-skill workers could be overrepresented in where urban productivity is soaring, quality of life is high, and costs of living are as well immensely large. These kinds of cities are called "Superstars" in Gyourko, Mayer, and Sinai (2006).

These cities, at the same time, are witnessed drastic wage inequality between the high-income and low-income groups. Take Atlanta for example, distinction of earnings between the top 5 percent and bottom 20 percent of households in Atlanta is as large as 18.8 times in 2012[1] (Alan Berube, Brookings paper, 2014). Wage premium is increasing associated with city size (Baum-Snow, Freedman, & Pavan, 2014; Baum-Snow & Pavan, 2012) and high-skill concentration (Davis & Dingel, 2012, 2013). Two main hypotheses document the reasons behind. Under demand-driven hypothesis, skill-biased technical change (Autor, Katz, & Kearney, 2004), complementarity between skilled workers (Behrens, Duranton, & Robert-Nicoud, 2010; Giannetti, 2003; Venables, 2011), and complementarity between cities and skills (Berry & Glaeser, 2005; Glaeser & Mare, 2001; Glaeser & Resseger, 2010) shapes the polarization of locations between high-skill workers and low-skill workers. Based upon supply-shift hypothesis, the young and educated are more likely to work in productive cities so as to gain learning-in-cities (Fu & Liao, 2012; Glaeser, 1999; Glaeser & Mare, 2001; Lucas, 2004); Moreover, recent studies eloquently document that cities better-off in amenity disproportionally attract more high-skill workers (Gyourko et al., 2006; Lee, 2010). Other studies pointing out that

[1] In 2012, Income for 20th-percentile household is $14,850 while income for 95th-percentile household is $279,827 on average.

industry-specific skill intensities as demand shift and low-skill protective labor market institutions such as labor unions are arguments for surging skill premium[2]. There has been no full accounting of skill concentration by all the three urban characters so far: efficiency, amenities, and excessive frictions.

Debates over welfare inequality have never reached a consensus. Moretti (2008) argues that if high-skill coming to cities because of large demand, relatively high standard of living is offset by higher housing costs, but if due to more amenities, high cost of housing reflects amenity consumption. By constructing heterogeneous CPIs for college graduates and high-school degree workers in different cities, Moretti is able to obtain real wage inequality, and he finds that demand-pull force is relatively more important than supply-push shock during 1980 and 2000; moreover, the over-exposure of college graduates in large as well as expensive cities leads to narrower well-being inequality than nominal wage gap. However, Diamond (2012) finds that welfare inequality increment is even more severe than nominal wage gap increment for 20 percent from 1980 to 2000. These two contradictory statements call into reconciliation. In this study, we attempt to provide an alternative view of the same welfare inequality question in the urban accounting framework with heterogeneous skills.

There is abundant literature about benefits of agglomeration economies to the high-skill group, the lower educated one however, are not drawn sufficiency attentions with respect to labor productivity. Moretti (Enrico Moretti, 2004a; E. Moretti, 2004b) finds earnings of the less educated are raised by the increasing supply of college graduates. He also accentuates that share of college graduates is positively correlated with wage of high-school graduates (E. Moretti, 2012). Glaeser's review of Moretti's book illustrates that taco stand worker earns quite different in Visalia and Menlo Park (Glaeser, 2013). The geographic pattern of the greatest minds as well as the lower educated both appealing to large cities like New York is expounded by Eeckhout, Pinheiro, and Schmidheiny (2010). Winters (2012) illustrates that low-skill workers benefit from locating near the high-skill by improving their labor force participation and employment. Pereira-Lopez and Soloaga (2013)find consistent evidence of imperfect substitution between different skills using Mexico metropolitan areas. Low-skill workers are more and more employed with higher payments in large cities, suggesting that low-skill must be more productive by nearing high-skill workers (Hao & Fu, 2014; Jones, 2014). Indeed, New York City is home to investment bankers and busboys, San Francisco to Internet entrepreneurs and grocery clerks, Boston to biomed engineers and the janitors who pick up their offices (Emily Badger, The Atlantic, 2014).

Productivity, skill concentration and welfare inequality are jointly determined in a spatial equilibrium. Jones (2014) generalized productivity accounting framework

[2] Katz and Autor (1999) and Goldin and Katz (2007) offer comprehensive surveys of literature on nominal wage inequality.

offers a brand new angle to view agglomeration economies, which is specialization instead of segregation. Quality of human capital investment, the service flow of human capital, is highly heterogeneous across cities, which are influenced by local degree of the division of labor. Institutions, market extent, and coordination costs all determine the division of labor. In an urban context, more location advantage and higher density in terms of city size and city skill-share positively lowers coordination costs and boost high-skill human capital service flow, i.e. high-skill productivity (Hao and Fu, 2014). Jones (2014) approach also allows imperfect substitution between skill groups. Low-skill productivity, i.e. low-skill wage when assuming a location-free human capital service flow, depends on local skill-mix. When the economy shifts to high-tech as the engine of growth, high-skill concentration generates a multiplier effect on low-skill workers (Moretti, 2012). As the elasticity of substitution changes, for example, lowers, the demand curve for high-skill is even more downward sloping; in this case, the relative output price between the two skill groups declines and low-skill workers are going to be demanded more. Therefore, alternation of the substitution elasticity plays a role of shaping distribution of skill-share. In addition, observed large supply of higher educated workers over the last three decades (1980-2010) for example, would have significant impacts on the distribution of skill-share as well as welfare inequality.

Distribution of urban amenities determines skill concentration and welfare inequality; moreover, they are simultaneously determined if amenities are endogenous. Besides Natural amenities, social amenities such as theaters rise with more concentration of workers or even of high-skill workers. Glaeser et.al (2001) find that high amenity cities have grown faster than low amenity city, and the demand for living in cities has risen for reasons beyond rising wages. Lee (2010) provides consumption-side explanation for rising wage premium, when high-skill workers have a bigger willingness-to-pay for urban amenities. Gyourko et.al (2006) depict “Superstar” cities that are both highly productive and desirable in amenities. High-skill workers living in these cities earn high wage, pay for more housing rents but largely due to consumption of amenities. The well-being inequality gap could be wider with the presence of “Superstar” cities, relative to an urban system without “Superstar” cities. The association between amenities and concentration motivates the study of agglomeration benefits from amenities.

Frictions deter people to be more productive and consume desirable amenities. Urban government inefficiency, labor loss per mile on road, and expanded urban size all act as barriers to come to the city. When a local government is less efficient, revenues collected from workers in the form of income tax for instance, will not be efficiently spent on public services. Living in a large city no doubt increases the distance and time on road. Hence, frictions, skill concentration and welfare inequality are determined at the same time in a spatial equilibrium. Due to the fact that amenities and frictions have direct influences on skill concentration and city size, they indirectly affect productivity according to Jones model.

Given location fundamentals and the presences of agglomeration economies, urban performances such as skill-mix, wages, and welfare of each skill group are mapped in a general equilibrium. People are attracted by high productivity and high amenities, but are repellant to frictions. Each group's welfare is gained by goods and services consumption using disposable earning, consumption of local amenities, but lost due to congestion-caused housing rents, commuting costs between residence and workplace, and local government inefficiency. We extend urban size accounting model of Desmet and Rosi-Hansberg (2013) to a two-skill accounting model, to decompose skill-mix, wage distribution, and welfare inequality into three location attributes (efficiency, amenities, and excessive frictions) and two types of agglomeration economies (productivity externalities and amenity externality). We use Jones' generalized division of labor approach to map location attributes to skill mix and productivity. Based on the win-win idea of specialization, both high-skill and low-skill workers are more specialized in productive cities. If productivity is positively correlated with amenities, that is to say, skill clusters are both productive and amenable; we could predict that welfare inequality gap is wider when any urban character is heterogeneous across cities, relative to the case of equalized urban character. The reason is that high-skill cohort has a stronger preference over urban amenities than low-skill do, wage premium is fixed and determined by local amenities (Hao and Fu, 2014), and both groups are exposed to the same frictions within a city. Henceforth, widening or narrowing welfare inequality gap lies in the association between productivity and amenities.

We meant to provide comparative statics using a two-city example to illustrate how welfare inequality is changed under different circumstances. (here not provided yet.)

Data of US Metropolitan areas in 2005 is adopted to perform counterfactual exercises. We find that large cities are more productive, not extremely amenable, and doing well in controlling excessive frictions. As for skill concentration, we show that skill-share distribution is more even in absence of heterogeneity in either efficiency or excessive friction, i.e. less skill concentration when either equalizing efficiency or excessive friction across cities, but more dispersed in absence of heterogeneity in urban amenities; these results imply that more educated cities are productive, not very amenable, and less frictional. As for wage distributions, they imply that higher-paid cities are productive, less desirable in amenities, and less congested for each skill group. Shutting down shocks in any of the three location attributes widens welfare inequality gap, which should be due to a negative association between productivity and amenities, which is revealed from either skill-share distribution or wage distributions. Efficiency contributes the greatest of narrowing welfare inequality.

As for the impacts of externalities, we find that externalities enhance the geographic concentration of high-skill workers if urban characters are in reality

heterogeneous across cities. Moreover, either productivity externalities or amenity externality contributes to narrowing welfare inequality gap.

The enlarged city size and intensified high-skill concentration is partly due to desirable amenities, so amenities have played an indirect role on productivity according to Jones (2014). This is what deviates from Moretti (2008) in which supply shock from amenities does not increase low-skill wage as well as high-skill wage. Although utilizing different methods to study welfare inequality, this paper documents that efficiency is the dominate character to account for welfare inequality, which is in line with Moretti (2008).

This paper studies changing welfare inequality between dispersed and even distributions of urban attributes, while Moretti (2008) and Diamond (2012) pay attention to the comparison between nominal wage gap and well-being inequality. In fact, our study would show that nominal wage gap is always narrower than well-being inequality as long as high-skill workers have stronger preferences over amenities. Our conjecture conforms to Diamond (2012) that, if urban wage premium keeps growing from 1980 to 2000, then welfare gap should expand, and the relative increments in welfare inequality must come from increasing divergent preferences to urban amenities between high-skill workers and low-skill workers.

Nonetheless, heterogeneous skill productivity is sensitive to results in Diamond (2012). She measures wage-based productivity with sectoral differences in local industrial employment composition, endogenous urban size, and endogenous skilled labor. From equation (25) and (26) in Diamond (2012), if elasticity of substitution is bigger than 1, productivity of respective high-skill and low-skill workers is declined in places of more concentration of either of the skill group. If elasticity of substitution is smaller than 1 but bigger than 0, larger urban size reduces both of the two skill productivity. These facts are contradictory to Hao and Fu (2014) in which productivity of either skill is raised by both city size and high-skill share. Moreover, complementarity between high-skill and low-skill workers is not conveyed in her productivity gauge. In addition, results of this paper show that large cities have low levels of amenity instead of high levels at least relative to its weighted average, therefore that low-skill are pushed out of high-amenity places as proposed by Diamond (2012) is logically not true. Instead, increasing welfare gap ascribes to the presence of “Superstar” cities, where both high-skill and low-skill workers are productive but amenities are more consumed by high-skill workers. One thing we should be aware of is that we do not take into account of heterogeneous housing costs for high-skill and low-skill workers. However, heterogeneity in housing costs consolidate our point of view because, low-skill workers are more likely to stay in expensive cities when they can pay a lower price for smaller units or lower quality housing services.

Our extended urban accounting framework is superior to Moretti (2008) and

Diamond (2012). It reconciles the conflict of welfare inequality between these two researches. We inherit the discussion in Moretti (2008), except demand-pull force works for both high-skill workers and low-skill workers, and supply-push force is especially stronger for the high-skill; moreover, we do conclude that nominal wage gap should be smaller than welfare inequality as documented in Diamond (2012). In addition, Moretti (2008) and Diamond (2012) hold a default consensus that clustering of high-skill group leads to enlarged welfare inequality, which is an assumption in urgent need of examination. This ad hoc assumption leaves Moretti and Diamond the choice to compare utility variation with nominal wage premium, but our framework allows not just comparing with nominal wage gap, but also decomposing skill concentration and welfare inequality simultaneously into location attributes and externalities. By doing so we could investigate whether skill concentration and welfare inequality respond to changes of these determinants in the same direction. As a result, welfare inequality gap is not always narrowed when high-skill share is dispersed, for example when amenity is equalized. Recall that wage premium is a function of local amenities level and preference differential to amenities between skills in Hao and Fu (2014). This implies that nominal wage premium virtually has little to do with welfare inequality[3].

Contributions of this paper are significant. First, it is the first time to perform urban accounting of skill concentration, and different scenarios allow investigations of impacts of location attributes and agglomeration economies on welfare inequality. In addition, the paper elucidates how agglomeration economies work for skill concentration and welfare inequality. Second, this paper provides insight for future studies related to welfare inequality. Lindley and Machin (2014) document the labor market polarization. Their estimated implied relative demand shifts for college graduates vis-à-vis high school graduates are found widening over the period of 1980 to 2010. Utilizing this framework one could investigate how skill mix and welfare inequality gap change from 1980 to 2005 when cities becoming less alike tends to be more stylized. Third, for local policies intended to find a balance between economic development and welfare inequality, this paper offers a creative lens, that is, economic development needs to attract a large amount of high-skill workers, but when local government deliberately improve local amenities, welfare inequality between skill groups is wider.

Before we get to start the welfare inequality analyses, there are two issues should be dealt with in advance. One issue we should note is urban accounting framework of Desmet and Rossi-Hansberg (2013) upon which our model is based has some mistakes to be corrected. Doing so is to pave the foundation for comparison between

[3] Demand shock in Brinkman (2014) is decomposed into both industry-specific technology change and skill-specific technology change in his theoretical model; unsurprisingly, demand shock is still the dominant in determining concentration of high-skill in large cities conforming to Moretti (2008), although Brinkman does not touch welfare inequality question.

analyses of homogeneous skill and heterogeneous skill. Another issue needed to be coped with is productivity is not fully accounted for by "Solow residual" utilized in Desmet and Rossi-Hansberg (2013). In Solow's accounting framework, total wage is convex to Total Factor Productivity (TFP), as indicated by Appendix Figure 1. While under Jones human capital accounting framework, more even distribution of productivity yields higher total wages for high-skill and low-skill, since high-skill total wage and low-skill total wage is respectively concave to productivity, as shown by Appendix Figure 2. Therefore, welfare increments of either high-skill or low-skill when shutting down the difference of any city characteristic should be larger in heterogeneous skill case. We do find much higher levels of numerical utilities in the two-skill case than homogeneous-skill case in the results. so the next section is about this.

## 2. *The Homogeneous when Dispersed by City Characters*

*2.1 Benchmark model modification*

Urban outcome such as city size is shaped by positive effects from productivity, amenities, as well as costs and frictions arising from congestion. This section re-performs accounting exercises in Desmet and Rossi-Hansberg (2013) (i.e. D-RH) to pave the ground for heterogeneous workers analysis.

We need to set up an alternative utility function. The reason of doing so hinges on the mistake of the budget constraint (i.e. equation(3)) in D-RH (2013), which intends to show that disposable income subtracted by income loss in transportation is used to consume goods. The constraint, however, is

$$c_{it} = w_{it} h_{it} \left(1 - \tau_{it}\right) - R_{it} - T_{it}$$

where $\left(R_{it} + T_{it}\right)$ does not capitalize the loss due to distance to Central Business District (CBD) into income loss. Even if it equals to $\kappa(\frac{N_{it}}{\pi})^{\frac{1}{2}}$ in which $\kappa$ should have denoted commuting costs per mile in their point of view, we are not sure what is the measuring unit of $\kappa$, dollar or penny? Confusion caused by this problem would lead to different results with different measuring units of wage. The value of income loss should conform to earned wage income, but a parameter is unable to take up the role. Besides, inconsistency is found between above budget constraint and local government expenditure function in D-RH (2013), which is

$$G\left(h_{it} w_{it}, TC_{it}\right) = g_{it} h_{it} w_{it} \kappa TC_{it}$$

If $\kappa$ is commuting costs per mile, it is sufficient to denote income losses on road by solely $g_{it} \kappa TC_{it}$, which is the amount that urban government spends on building and maintaining urban infrastructure.

In fact, $\kappa$ captures labor loss per mile instead of labor income loss per mile on

road. And the correct form of budget constraint should have been

$$c_{it} = w_{it}h_{it}\left(1-\tau_{it}\right) - w_{it}h_{it}\left(R_{it}+T_{it}\right)$$

Then

$$c_{it} = \left(1-\theta\right)y_{it}\left(\left(1-\tau_{it}\right)-\left(R_{it}+T_{it}\right)\right) \tag{1}$$

And with the utility function $U_{it} = \log c_{it} + \varphi \log\left(1-h_{it}\right) + \gamma_{it}$, and First Order Conditions (FOC) with respect to consumption $[c_{it}]$ and working hour $[h_{it}]$, we get that

$$\varphi \frac{c_{it}}{1-h_{it}} = w_{it}\left(\left(1-\tau_{it}\right)-R_{it}-T_{it}\right) \tag{2}$$

Combing equation (1) with (2), we obtain that

$$\varphi \frac{\left(1-\theta\right)y_{it}\left(\left(1-\tau_{it}\right)-\left(R_{it}+T_{it}\right)\right)}{\left(1-\theta\right)y_{it}/h_{it}\left(\left(1-\tau_{it}\right)-R_{it}-T_{it}\right)} = 1-h_{it}$$

which essentially is

$$\varphi \frac{\left(1-\theta\right)y_{it}\left(\left(1-\tau_{it}\right)-\left(R_{it}+T_{it}\right)\right)}{\left(1-\theta\right)y_{it}\left(\left(1-\tau_{it}\right)-R_{it}-T_{it}\right)} = \frac{1-h_{it}}{h_{it}} \tag{3}$$

Equation (3) implies that working hour is a constant, and $h_{it} = \frac{1}{1+\varphi}$. In addition, this largely simplify the utility function to be

$$U_i = \log c_i + \gamma_i$$

$$\text{s.t.} \quad c_i = w_i\left(1-\tau_i-R_i-T_i\right)$$

Time dimension has no substantial effect either here or in D-RH (2013) except physical capital is rated universally in steady state, and we suppress the time subscript for convenience. Welfare losses are due to frictions. Frictions cause income losses, less income lead to less consumption, less consumption generates lower utility. Income loss is paid by the form of labor tax. We simplify disposable income as wage after paying labor tax at a new rate of $\hat{\tau}$, so the budget constraint is

$$\text{s.t.} \quad c_i = w_i e^{-\hat{\tau}_i}$$

From the equalization of above two budget constraints,

$$e^{-\hat{\tau}_i} \equiv 1-\tau_i-R_i-T_i$$

If $\tau_i + R_i + T_i$ is sufficiently small, then

$$\hat{\tau}_i = \tau_i + R_i + T_i \tag{4}$$

Deviated from D-RH (2013), the new $\hat{\tau}$ comprehensively captures the rate of income loss by frictions. The first component of frictions is the outcome of standard city size effect. All cities are monocentric that all residents within a city work in CBD and live in residential areas around the CBD. To make sure that a point of location within a city is indifferent from any other points within the city, a person lives closer to CBD should pay higher rent to bid for the location, because lower commuting costs occurred when travelling to CBD to work. Imagine you live at the fringe of a city paying no housing rent, as population increases and city size expands, it will cost more for you to commute since distance to CBD is farther. Imagine you live at the closest place to CBD, as population increases and city size expands, more people will compete with you for your location to avoid higher commuting costs, which results in higher bid rent for your place. Let the radius of a monocentric city be $\bar{d}_i$, and the per mile labor loss due to commuting is $\kappa$, then commuting-caused proportion of labor loss from a distance of d to CBD is

$$T(d) = \kappa d$$

The summation of commuting-caused proportion and renting-caused proportion of labor loss at distance d to CBD is

$$R_i(d) + T(d) = 0 + T(\bar{d}_i) = \kappa \bar{d}_i$$

Total population equals to the total housing units provided by the city,

$$N_i = \pi \bar{d}_i^2$$

then

$$R_i(\mathrm{d}) + T(\mathrm{d}) = \kappa \left( \frac{N_i}{\pi} \right)^{\frac{1}{2}} \quad \text{for} \quad \forall \quad \mathrm{d}. \tag{5}$$

The second component of frictions is caused by urban government inefficiency. Local government provides public services. A lot more effort by urban government must be taken in providing services when it is inefficient. Revenues come from the collected labor taxes in total. When government expenditure equals government revenue,

$$g_i w_i \kappa TC_i = \tau_i N_i w_i$$

where $g_i$ is the "excess friction" representing urban government inefficiency.

Total commuting is

$$TC_i = \int_0^{\bar{d}_i} (2\pi d^2) dd = \frac{2}{3} \pi^{-\frac{1}{2}} N_i^{\frac{3}{2}}$$

therefore labor wedge rate (labor income tax rate in general) is

$$\tau_i = g_i \kappa \frac{2}{3}\left(\frac{N_i}{\pi}\right)^{\frac{1}{2}} \tag{6}$$

From equation (4), (5), and (6), we get

$$\tau_i + R_i(\mathrm{d}) + T(\mathrm{d}) = g_i \kappa \frac{2}{3}\left(\frac{N_i}{\pi}\right)^{\frac{1}{2}} + \kappa\left(\frac{N_i}{\pi}\right)^{\frac{1}{2}}$$

and then a more comprehensive version of labor tax rate is

$$\hat{\tau}_i = g_i \kappa \frac{2}{3}\left(\frac{N_i}{\pi}\right)^{\frac{1}{2}} + \kappa\left(\frac{N_i}{\pi}\right)^{\frac{1}{2}} = \left(\frac{2}{3}g_i + 1\right)\kappa\left(\frac{N_i}{\pi}\right)^{\frac{1}{2}} = \hat{g}_i \kappa\left(\frac{N_i}{\pi}\right)^{\frac{1}{2}} \tag{7}$$

So far we get to see a clearer picture of labor tax rate. It captures the extent of frictions from two sources: one is the extent of standard friction caused by congestion, and it is embodied by $N_i$; the other is the extent of excess frictions coming from urban government inefficiency. The improvement based on D-RH (2013), allows us to clarify the relation between labor tax rate and the extent of urban inefficiency, and the relation between labor tax rate and the extent of congestion. Moreover, equation (7) has important implication for welfare analysis. A clearer picture between total friction and city size is drawn below, with the vertical axis be total friction, $\hat{\tau}_i N_i = \hat{g}_i \kappa \pi^{-\frac{1}{2}} N_i^{\frac{3}{2}}$, and horizontal axis be city size. Total friction is a convex function of city size $N_i$, which means

$$\hat{\tau}_i\left(\overline{N}_i\right) < \frac{\hat{\tau}_1\left(N_1\right) + \hat{\tau}_2\left(N_2\right)}{2}$$

Its curvature matters for welfare because when city size distribution is more dispersed, total friction tends to become larger, and more social welfare is forfeited.

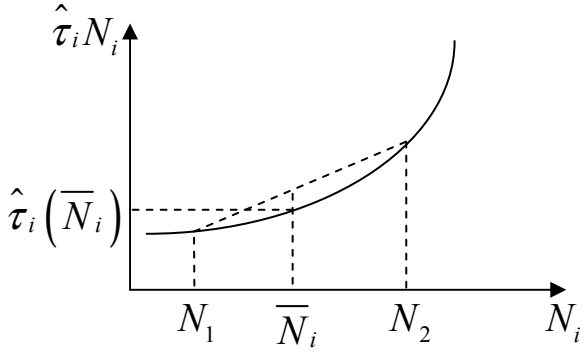



Figure: The convex relationship between total friction and city size

Take log of equation (7), and we get

$$\log\hat{\tau}_i - \frac{1}{2}\log N_i = cons\tan t + \log\hat{g}_i \tag{8}$$

where

$$cons\tan t = \log \kappa - \frac{1}{2}\log \pi$$

The deviation comes from excess frictions. After running regression on constant, error terms which are $\log \hat{g}_i$, and constant that could be used to derive $\kappa$, can be achieved.

$$\kappa = \exp\left( cons\tan t + \frac{1}{2}\log \pi \right)$$

From the budget constraint, rate of labor wedge is

$$\hat{\tau}_i = -\ln \frac{c_i}{w_i} = -\ln \frac{C_i / total\ pop_i}{w_i}$$

where $c_i$ is consumption per person, $C_i$ is aggregate consumption, $w_i$ is per working labor personal income.

Production is launched by productivity shifter, physical capital, and labor inputs since working hour is a constant,

$$Y_i = A_i K_i^{1-\alpha} N_i^{\alpha} \tag{9}$$

Wage is derived from the First Order Condition of output with respect to labor inputs, that is

$$w_i = A_i^{\frac{1}{\alpha}} \left( \frac{1-\alpha}{r} \right)^{\frac{1-\alpha}{\alpha}} \alpha \tag{10}$$

and productivity shifter is

$$A_i = y_i^{\alpha} \left( \frac{r}{1-\alpha} \right)^{1-\alpha} \tag{11}$$

where $y_i$ is per capita output, $r$ is the rental rate of physical capital. Even though D-RH (2013) do not adopt this measurement, they argue that this model-based calculation is highly correlated with empirical measure using actual data of physical capital stock at 0.9[4].

The indirect utility function that is indifferent across cities is

$$\overline{U} = \log\left( w_i e^{-\tau} \right) + \rho_i$$

In equilibrium, indirect utility is determined when labor market clears by

$$N = \sum_{i=1}^{I} N_i$$

[4] The correlation between productivity shifter obtained by equation (11) and the one measured in Hao and Fu (2013) is as high as 0.88. Hao and Fu (2013) adopts the same methodology as in D-RH (2013) to calculate the productivity shifter.

City size is now accounted for by three factors: productivity shifter $A_i$, amenity index $\rho_i$, and excess frictions $\hat{g}_i$, as below

$$N_i = \left( \frac{\frac{1}{\alpha}\log A_i + \log\left(\left(\frac{1-\alpha}{r}\right)^{\frac{1-\alpha}{\alpha}}\alpha\right) + \rho_i - \overline{U}}{\left(\frac{2}{3}g_i + 1\right)\kappa} \right)^2 \pi \tag{12}$$

Then amenity index is backed out from equation (8) in the equilibrium, which is

$$\rho_i = \left(\frac{N_i}{\pi}\right)^{\frac{1}{2}} \hat{g}_i\kappa - \frac{1}{\alpha}\log A_i - \log\left(\left(\frac{1-\alpha}{r}\right)^{\frac{1-\alpha}{\alpha}}\alpha\right) + \overline{U} \tag{13}$$

*2.2 Parameter calibration*

All parameters adopted in the following counterfactual analyses have the same values as in D-RH (2013), except the one representing labor losses per mile. To match the actual city size distribution based on the model, the value of $\kappa$=0.0015 instead of 0.02 as in D-RH (2013). The reason might be that data we use is from year 2005 instead of year 2005-year 2008. In the latter case, a larger number of large cities are exposed to severe urban congestions such that in equation (8), the constant term is escalated due to that rate of labor tax should be higher because of severer congestions. After all, intuitions would not change using either of the two samples. Other parameters follow D-RH (2013).

*2.3 Results*

Results and analyses of the cases without and with agglomeration economies are put into Appendix A.

## *3. Heterogeneous Skill when Dispersion by City Characters*

*3.1 Benchmark model*

We are intrigued by whether dispersion or convergence of any city characteristic contributes to reduction of welfare inequality gap between the high-skill and the low-skill. The understanding of skill heterogeneity is based on discussions in Hao and Fu (2014). High-skill workers and low-skill workers are more productive ascribing to local extent of the division of labor. Both high-skill and low-skill workers are more specialized when the division of labor is finer, implying that they are more productive in larger and more educated cities. Two channels we propose might contribute to narrower welfare inequality gap. One channel is low-skill workers in amenable but not large cities not only enjoy amenities but also may not be unproductive; the other channel is both high-skill and low-skill are more productive in large cities, and

excessive frictions do much more harm to the high-skill.

Based on above homogeneous skill model of urban accounting, we now present the case of heterogeneous skill. For low-skill workers, each of them get utility

$$U_{1i} = \log c_{1i} + \rho_{1i} \qquad \text{s.t.} \quad c_{1i} = w_{1i} e^{-\hat{\tau}_i}$$

For high-skill workers, each of them get utility

$$U_{2i} = \log c_{2i} + \rho_{2i} \qquad \text{s.t.} \quad c_{2i} = w_{2i} e^{-\hat{\tau}_i}$$

where $\rho_2 = \omega\rho,\ \rho_1 = \rho$; high-skill workers would like to pay more for the same urban amenities if $\omega > 1$.

Then rate of labor tax from the two budget constraints is

$$\hat{\tau}_i = -\ln \frac{c_{1i}L_{1i} + c_{2i}L_{2i}}{w_{1i}L_{1i} + w_{2i}L_{2i}} = -\ln \frac{C_i}{total\ income_i} = -\ln \frac{c_i}{average\ income_i}$$

where $C_i$ is aggregate consumption; $c_i$ is per capita consumption. Therefore $\hat{\tau}_i$ is the same as in the homogenous case. The proportion of losses on wage due to commuting and renting (i.e. labor tax rate) is identical for either of the two skilled groups' residents who live at the same distance to the CBD. But amount of income losses occurs more for higher paid workers even though higher paid workers live at the same distance to CBD as those who earn less.

All workers live within the city circle. Imagine high-skill workers all live in the arc area that is $L_{2i}/N_i$ proportional to the whole circle, because each of them occupies 1 housing unit. Similarly for low-skill workers, their occupied arc area has the proportion of $L_{1i}/N_i$ to the whole circle. City size or area size of the circle is the sum of both two skilled groups, $N_i = L_{1i} + L_{2i}$.

Then total commuting of low-skill workers is

$$TC_{1i} = \int_0^{\bar{d}_i} (2\pi d^2 \frac{L_{1i}}{N_i}) dd = \frac{2}{3} \pi^{-\frac{1}{2}} N_i^{\frac{1}{2}} L_{1i}$$

Then total commuting of high-skill workers is

$$TC_{2i} = \int_0^{\bar{d}_i} (2\pi d^2 \frac{L_{2i}}{N_i}) dd = \frac{2}{3} \pi^{-\frac{1}{2}} N_i^{\frac{1}{2}} L_{2i}$$

Local government face that government expenditure equals government revenue, that is

$$g_i \left( w_{1i} \kappa TC_{1i} + w_{2i} \kappa TC_{2i} \right) = \tau_i \left( L_{1i} w_{1i} + L_{2i} w_{2i} \right) \tag{16}$$

where $\kappa TC_{1i}$ is labor losses for total commuting of low-skill workers, $w_{1i}\kappa TC_{1i}$ is

income losses for total commuting of low-skill workers. Then rate of labor wedge is

$$\tau_i = g_i \kappa \frac{2}{3}\left(\frac{N_i}{\pi}\right)^{\frac{1}{2}}$$

At any point within the city, labor losses due to commuting and renting is the same for high-skill and low-skill workers,

$$R_i(\mathrm{d}) + T(\mathrm{d}) = \kappa(\frac{N_i}{\pi})^{\frac{1}{2}}$$

Same as the improvement we make in our homogenous case, labor tax rate $\hat{\tau}_i$ is defined as

$$e^{-\hat{\tau}_i} \equiv 1 - \tau_i - R_i - T_i$$

so $\hat{\tau}_i = \tau_i + R_i + T_i$ if $\tau_i + R_i + T_i$ is sufficiently small.

Therefore,

$$\hat{\tau}_i = \hat{g}_i \kappa(\frac{N_i}{\pi})^{\frac{1}{2}}$$

Analogously, we could retrieve the estimations of excess friction $\hat{g}_i$ and $\kappa$.

For both low-skill workers and high-skill workers, their indirect utility functions are respectively

$$\overline{U}_1 = \log\left(w_1 e^{-\hat{\tau}}\right) + \rho_1 \tag{17}$$

$$\overline{U}_2 = \log\left(w_2 e^{-\hat{\tau}}\right) + \rho_2 \tag{18}$$

It includes complete factors that account for utility: wage that is promoted by more extensive division of labor, labor tax rate that represents the extent of income losses from standard frictions and excess frictions, and skill-specific preferences for urban amenities.

Production function is defined as in Hao and Fu (2014),

$$Y_i = K_i^{1-\alpha} H_i^{\alpha} \tag{19}$$

Human capital stock is

$$H_i = \left[\left(h_1 L_{1i}\right)^{\frac{\varepsilon-1}{\varepsilon}} + \left(h_{2i} L_{2i}\right)^{\frac{\varepsilon-1}{\varepsilon}}\right]^{\frac{\varepsilon}{\varepsilon-1}} \tag{20}$$

First Order Conditions are then

$$w_1 = \frac{\partial Y}{\partial L_1} = \frac{\partial Y}{\partial H}\frac{\partial H}{\partial H_1}\frac{\partial H_1}{\partial L_1} = K^{1-\alpha}\alpha H^{\alpha-1} G_1 \mathrm{h}_1 \tag{21}$$

$$w_2 = \frac{\partial Y}{\partial L_2} = \frac{\partial Y}{\partial H}\frac{\partial H}{\partial H_2}\frac{\partial H_2}{\partial L_2} = K^{1-\alpha}\alpha H^{\alpha-1} G_2 \mathrm{h}_2 \tag{22}$$

$$r = \frac{\partial Y}{\partial K} = (1-\alpha) K^{-\alpha} H^{\alpha} \tag{23}$$

Derivations of equations (21)-(23) are

$$K = \left(\frac{1-\alpha}{r}\right)^{\frac{1}{\alpha}} H$$

$$\frac{w_2}{w_1} = \frac{G_2 \mathrm{h}_2}{G_1 \mathrm{h}_1}$$

Together with

$$H = G_1 h_1 \tilde{L} = \left(\frac{H}{H_1}\right)^{\frac{1}{\varepsilon}} h_1 \tilde{L} = h_1 L_1^{\frac{1}{1-\varepsilon}} \left(L_1 + \frac{w_2}{w_1} L_2\right)^{\frac{\varepsilon}{\varepsilon-1}}$$

$$\frac{G_2}{G_1} = \left(\frac{H_1}{H_2}\right)^{\frac{1}{\varepsilon}} = \left(\frac{h_1 L_1}{h_2 L_2}\right)^{\frac{1}{\varepsilon}}$$

We get that

$$H = h_1 L_1 \left(1 + \left(\frac{\mathrm{h}_2}{\mathrm{h}_1}\right)^{1-\frac{1}{\varepsilon}} \left(\frac{L_1}{L_2}\right)^{\frac{1}{\varepsilon}} \frac{L_2}{L_1}\right)^{\frac{\varepsilon}{\varepsilon-1}}$$

Then low-skill wage is

$$w_1 = \alpha\left(\frac{1-\alpha}{r}\right)^{\frac{1-\alpha}{\alpha}} h_1^{1-\frac{1}{\varepsilon}} L_1^{-\frac{1}{\varepsilon}} H^{\frac{1}{\varepsilon}} = \alpha\left(\frac{1-\alpha}{r}\right)^{\frac{1-\alpha}{\alpha}} h_1 \left(1 + \left(\frac{\mathrm{h}_2}{\mathrm{h}_1}\right)^{1-\frac{1}{\varepsilon}} \left(\frac{L_2}{L_1}\right)^{1-\frac{1}{\varepsilon}}\right)^{\frac{1}{\varepsilon-1}} \tag{24}$$

High-skill wage is

$$w_2 = \alpha\left(\frac{1-\alpha}{r}\right)^{\frac{1-\alpha}{\alpha}} h_2^{1-\frac{1}{\varepsilon}} L_2^{-\frac{1}{\varepsilon}} H^{\frac{1}{\varepsilon}} = \alpha\left(\frac{1-\alpha}{r}\right)^{\frac{1-\alpha}{\alpha}} h_2^{1-\frac{1}{\varepsilon}} h_1^{\frac{1}{\varepsilon}} L_1^{\frac{1}{\varepsilon}} L_2^{-\frac{1}{\varepsilon}} \left(1 + \left(\frac{\mathrm{h}_2}{\mathrm{h}_1}\right)^{1-\frac{1}{\varepsilon}} \left(\frac{L_2}{L_1}\right)^{1-\frac{1}{\varepsilon}}\right)^{\frac{1}{\varepsilon-1}} \tag{25}$$

High-skill quality is

$$h_2 = h_1 \left(\frac{w_2}{w_1}\right)^{\frac{\varepsilon}{\varepsilon-1}} \left(\frac{L_2}{L_1}\right)^{\frac{1}{\varepsilon-1}} \tag{26}$$

High-skill productivity is as well the urban productivity since low-skill quality, $h_1$, is assumed to be identical across locations.

Since low skill productivity could be written as

$$w_1 = \alpha\left(\frac{1-\alpha}{r}\right)^{\frac{1-\alpha}{\alpha}} h_1\left(1+\left(\frac{h_2}{h_1}\frac{w_1}{w_2}\right)^{\varepsilon-1}\right)^{\frac{1}{\varepsilon-1}}$$

also from both utility functions of the two skills, we obtain

$$\frac{w_1}{w_2} = e^{\frac{\overline{U}_1}{\overline{U}_2}}\frac{e^{\rho_2}}{e^{\rho_1}} \quad \text{and} \quad \frac{e^{\rho_2}}{e^{\rho_1}} = \frac{e^{\omega\rho}}{e^{\rho}} = e^{\rho(\omega-1)}$$

Substitute these two equation into (24) and (25), we get

$$w_1 = \alpha\left(\frac{1-\alpha}{r}\right)^{\frac{1-\alpha}{\alpha}} h_1\left(1+\left(\frac{\exp\left(\frac{\overline{U}_1}{\overline{U}_2}+\rho(\omega-1)\right)}{h_1}h_2\right)^{\varepsilon-1}\right)^{\frac{1}{\varepsilon-1}} \tag{27}$$

$$w_2 = \frac{\alpha\left(\frac{1-\alpha}{r}\right)^{\frac{1-\alpha}{\alpha}} h_1}{\exp\left(\frac{\overline{U}_1}{\overline{U}_2}+\rho(\omega-1)\right)}\left(1+\left(\frac{\exp\left(\frac{\overline{U}_1}{\overline{U}_2}+\rho(\omega-1)\right)}{h_1}h_2\right)^{\varepsilon-1}\right)^{\frac{1}{\varepsilon-1}} \tag{28}$$

If elasticity of substitution $\varepsilon < 2$ (in fact, $\varepsilon = 1.6$ in estimations of Hao and Fu (2014)), this equation says that low-skill wage $w_1$ is a concave function of urban productivity $h_2$; moreover due to high-skill wage is proportional to low-skill wage, high-skill wage $w_2$ is also a concave function of urban productivity $h_2$. From the data, we find that $\exp\left(\frac{\overline{U}_1}{\overline{U}_2}+\rho(\omega-1)\right) \in (0,1)$, so high-skill wage should have a steeper slope with respect to productivity, shown as below

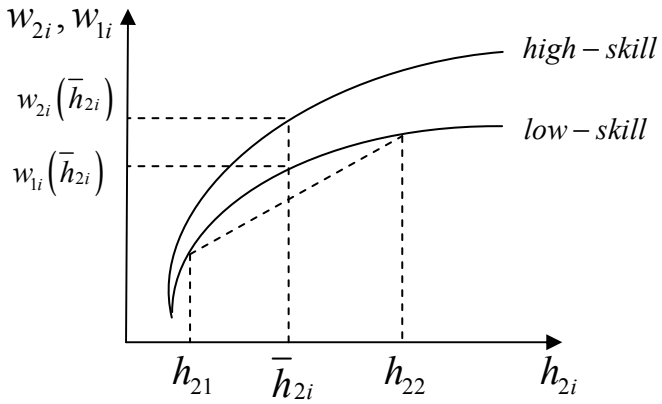



Figure: The concave relationship between wage and urban productivity

Concavity means

$$w_{2i}\left(\overline{h}_{2i}\right) > \frac{w_{21}\left(h_{21}\right)+w_{22}\left(h_{22}\right)}{2} \quad \text{and} \quad w_{1i}\left(\overline{h}_{2i}\right) > \frac{w_{11}\left(h_{21}\right)+w_{12}\left(h_{22}\right)}{2}$$

These inequalities suggest that more convergence in productivity across cities results in higher high-skill wage and low-skill wage. A more converged productivity distribution predicts a more uniform distribution of city size controlling for other factors. Note that the relationship between the rate of labor loss and city size is still a convex curvature. Therefore when city size distribution is more converged, total frictions tend to become smaller and less welfare is lost. So far we have distinguished two approaches of productivity dispersion which influence welfare, by lowering welfare as well as raising welfare.

In equilibrium, suppose there are I cities in a city system, i=1, 2, 3,…., I. Then

$$\overline{U}_1 = U_{1i} \quad for\ \forall city\ i,$$

$$\overline{U}_2 = U_{2i} \quad for\ \forall city\ i.$$

The economy-wide indirect utility $\overline{U}_1$ is determined when low-skill labor market clears,

$$\sum_{i=1}^{I} L_{1i} = \overline{L}_1 \tag{29}$$

The economy-wide indirect utility $\overline{U}_2$ is determined when high-skill labor market clears,

$$\sum_{i=1}^{I} L_{2i} = \overline{L}_2 \tag{30}$$

and $\overline{L}_1 + \overline{L}_2 = N$ .

In equilibrium,

$$\overline{U}_1 = \log\left(\alpha\left(\frac{1-\alpha}{r}\right)^{\frac{1-\alpha}{\alpha}} h_1\left(1+\left(\frac{\mathrm{h}_2}{\mathrm{h}_1}\right)^{1-\frac{1}{\varepsilon}}\left(\frac{L_2}{L_1}\right)^{1-\frac{1}{\varepsilon}}\right)^{\frac{1}{\varepsilon-1}} e^{-\hat{g}_i\kappa\left(\frac{N_i}{\pi}\right)^{\frac{1}{2}}}\right)+\rho_1 \tag{31}$$

$$\overline{U}_2 = \log\left(\alpha\left(\frac{1-\alpha}{r}\right)^{\frac{1-\alpha}{\alpha}} h_2^{1-\frac{1}{\varepsilon}} h_1^{\frac{1}{\varepsilon}} L_1^{\frac{1}{\varepsilon}} L_2^{-\frac{1}{\varepsilon}}\left(1+\left(\frac{\mathrm{h}_2}{\mathrm{h}_1}\right)^{1-\frac{1}{\varepsilon}}\left(\frac{L_2}{L_1}\right)^{1-\frac{1}{\varepsilon}}\right)^{\frac{1}{\varepsilon-1}} e^{-\hat{g}_i\kappa\left(\frac{N_i}{\pi}\right)^{\frac{1}{2}}}\right)+\rho_2 \tag{32}$$

Now we can back out amenity index $\rho_2$ and $\rho_1$ from above two equations

when normalizing reservation utilities as: $\overline{U}_1 = 10$, $\overline{U}_2 = 20$. If we assume that $\rho_2$ and $\rho_1$ has linear relationship b/w each other such as $\rho_2 = \omega\rho_1$, then when shutting down the difference of low-skill preferred amenities to be $\overline{\rho}_1$, equalized amenities for high-skill workers is $\tilde{\rho}_2 = cons\tan t + \omega\overline{\rho}_1 + error$.

*3.2 The calibration of parameters*

We set and calibrate values for the parameters in the model. $\alpha$=0.6666 is human capital input share in the production function, $r$=0.02 is interest rate for renting physical capital, $\kappa$=0.0015 is labor losses per mile when commuting and estimated by 2005 CPS-MORG files data, $\omega$=1.125 refers to preference differential to urban amenities between high-skill and low-skill workers, $\varepsilon$ =1.6 is elasticity of substitution, $h_1$=33610.32 is low-skill quality of human capital and we proxy for it by median schooling years of the low-skill group, $\eta$=0.1 is elasticity of city size in affecting local productivity, $\beta$=0.4 is elasticity of share of high-skill workers in affecting local productivity, $\zeta$ =0.02 is elasticity of city size in determining amenity index, $\overline{U}_1$=10 is the low-skill indirect utility and $\overline{U}_2$=20 is the high-skill indirect utility.

*3.3 The results*

*3.3.1 The case without agglomeration economies*

We adopt data of year 2005 from NBER version of Current Population Survey (CPS) as well for the heterogeneous study of skills. We define skill-share of high-skill population divided by low-skill population[5]. A cutoff of skilled population to determine whether a city exists is defined as the total labor multiplied by the share of that skill group. The threshold of total labor is 2600, which is 3000 used in D-RH (2013) multiplied by the mean of cities' labor force share[6].

In Figure 1, amenity being equalized is the weighted average of low-skill preferred amenities. By documenting the significantly linear relationship between high-skill preferred amenities and low-skill preferred amenities, we are able to eliminate amenities difference while keeping skill-heterogeneous preferences over amenities. High-skill population distribution is less dispersed in both counterfactual

[5] Regressions of population and total wages on urban characters could be found in respectively Appendix Table 1 and Appendix Table 2.

[6] Mean of labor force share = 1 – mean of old people share.

efficiency and counterfactual excessive frictions, while more mixed in counterfactual amenity case, even though high-skill worker reallocation is the largest of the three. High-skill labor abundant cities are productive, less congestible, but not necessarily amenable except in mega cities.

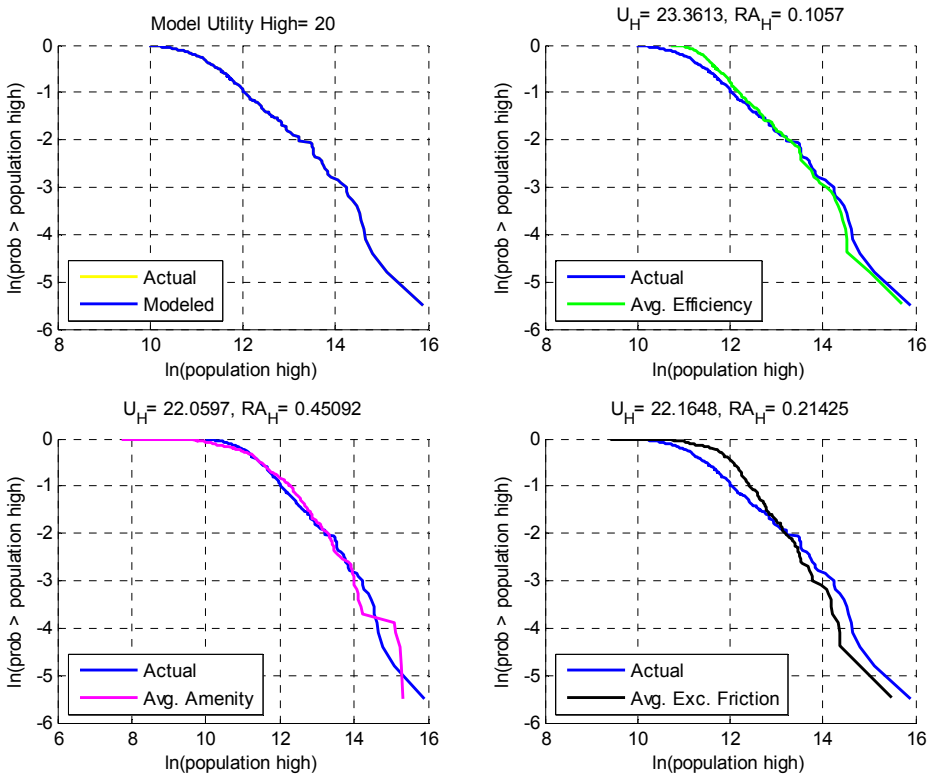



Figure 1: Counterfactuals of High-skill Population Size Without One Shock, $\kappa = 0.0015$, $\eta = 0$, $\beta = 0$, $\zeta = 0$

Relative to homogeneous skill in D-RH(2013) and our improved model, proportional utility increments are fairly massive while reallocations are relatively modest, for example in counterfactual efficiency, 10 percent of high-skill worker mobile to cities other than origins. The relatively large increase in utility could be the result of small reallocation but shutting down of distortions; moreover, concavity of total wage on population generates more welfare gains with less dispersed size distribution, and convexity of aggregate friction lowers utility losses.

Figure 2 shows similar to high-skill counterfactual population distribution. The three urban characters as well are important to Low-skill workers. Counterfactual utility is respectively higher than in homogenous exercises. So, counterfactual city size distribution is analogous to counterfactual high-skill and low-skill population distribution.

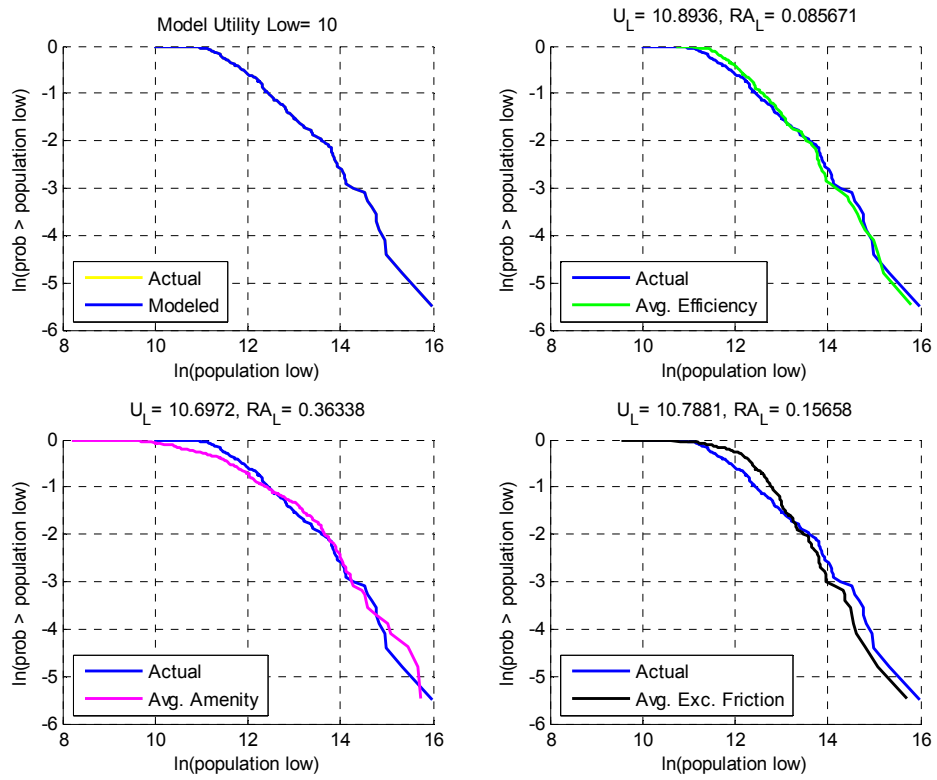



Figure 2: Counterfactuals of Low-skill Population Size Without One Shock, $\kappa = 0.0015$, $\eta = 0$, $\beta = 0$, $\zeta = 0$

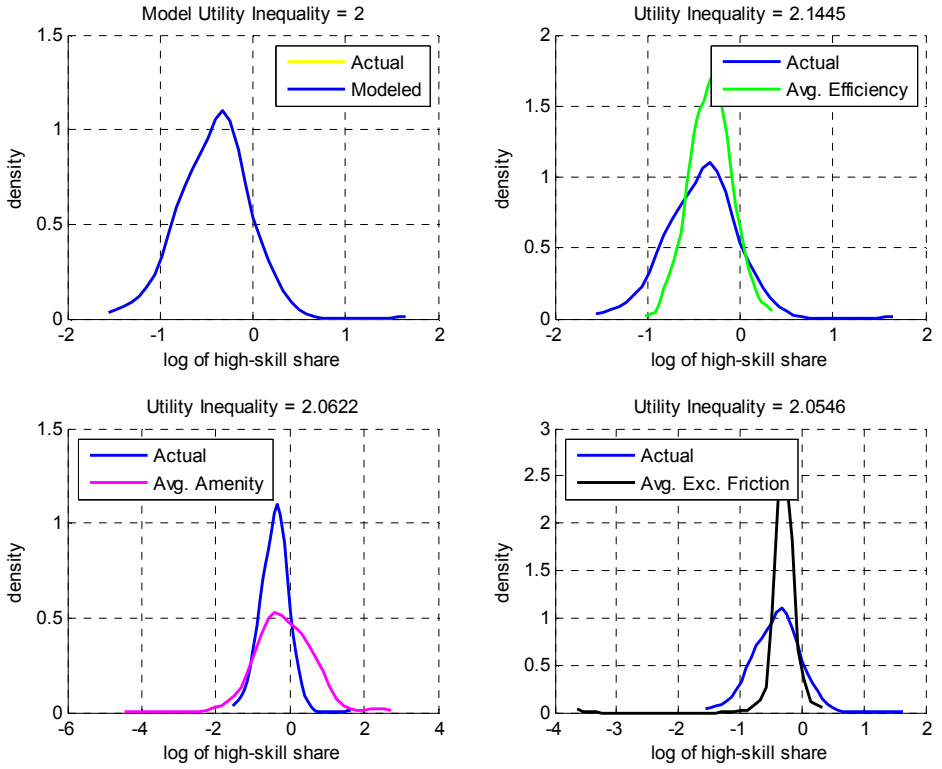



Figure 3: Counterfactuals of City Education Level Without One Shock, $\kappa = 0.0015$, $\eta = 0$, $\beta = 0$, $\zeta = 0$

Our primary concern is accounting for the relative geographic concentration of high-skill workers, and its welfare inequality implication. In Figure 3, cities that are

too small to be existed are not presented. Dispersed distribution of either efficiency or excessive frictions generates geographic concentration of high-skill workers relative to that of low-skill workers. Instead, a less even distribution of amenities over cities causes no geographic clustering of high-skill workers relative to that of low-skill. But we should keep in mind that the combined effects of all the three spatial factors determine whether high-skill workers relative to their counterparts, low-skill, concentrate in certain cities. In counterfactual efficiency and excessive friction cases, skill-share distribution is more contracted, while eliminating amenities variation across cities expands skill share distribution. These imply that more educated cities are more productive, less congested, while not very amenable.

In any of the three counterfactual cases, welfare inequality gap is wider than in actual. It implies that more dispersed distribution of any urban character expands welfare inequality. Equalizing efficiency plays the most significant role of widening welfare inequality. For instance, welfare inequality gap is raised to be 2.14 in counterfactual efficiency case, rather than 2.06 in counterfactual amenity case and 2.05 in counterfactual excessive frictions case. That is to say, dispersed distribution of efficiency narrows welfare inequality gap for about 14 percentage points, compared with the case of shutting down efficiency variation by its weighted average.

Eliminating efficiency generates a much more shrunken high-skill wage distribution in Figure 4, implying that higher-paid cities for high-skill workers are more productive (This is the demand pull force(increase in standard of living is offset by higher cost of living) in Moretti (2008)).

By the similar token, higher-paid cities are relatively worse-off in high-skill preferred amenities. Once equalizing amenities for all cities, higher-paid cities attract even more high-skill workers because they prefer more of amenities, facilitating rise in productivity and thus even higher payment. Analogously, lower-paid cities are relatively better-off in high-skill preferred amenities; losing advantage in amenities means losing population especially high-skill population, so wage becomes even lower. (This is the supply push force(higher cost of housing reflects consumption of desirable local amenities) discussed in Moretti (2008)).

Due to the fact that dispersed distribution of efficiency contributes the most to narrowing welfare inequality gap, so demand pull force outweighs supply shift force; in addition, welfare inequality is reduced when either efficiency or excessive frictions are more alike, therefore welfare inequality must be more shrunken when all sorts of urban factors are less identical.

In the case of excessive frictions, lower-paid cities pay even less while high-skill wage in higher-paid cities are barely changed. This possibly is due to that small cities are not crowded, and after mitigating the excessive frictions variation, high-skill workers reallocate to large and productive cities and leave the small cities even more

unproductive, so lower-paid cities can only offer lower wage; moreover, large and productive cities are more productive ascribing to the migration of high-skill workers, however the congestion becomes even worse, these two opposing forces leave wage quantitatively unchanged.

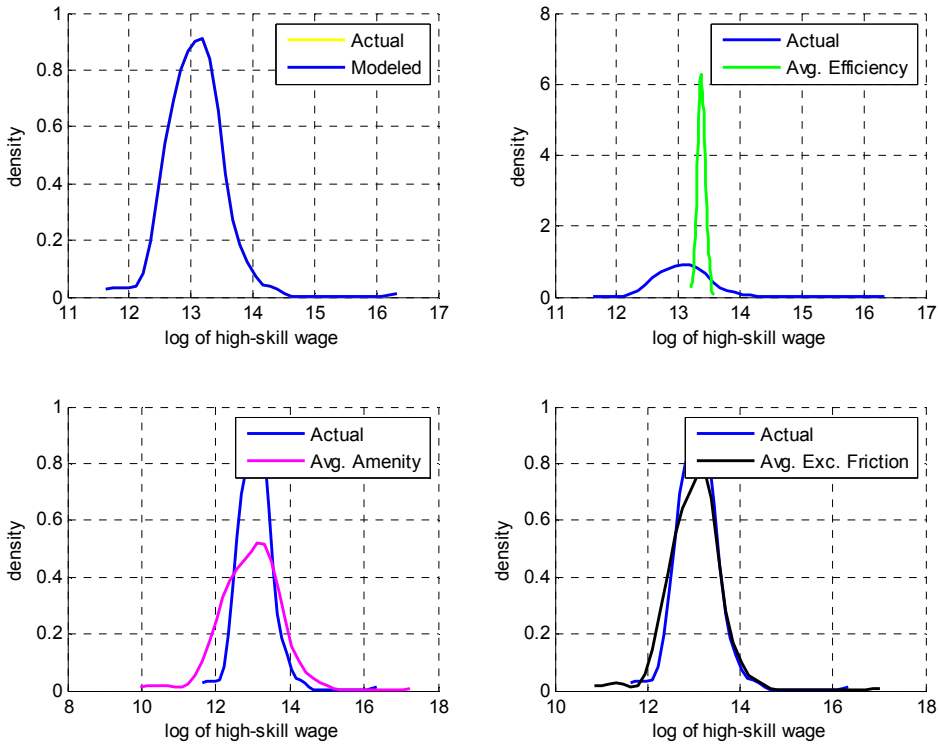



Figure 4: Counterfactuals of High-skill Wage Without One Shock, $\kappa = 0.0015$, $\eta = 0$, $\beta = 0$, $\zeta = 0$

As can be seen from Figure 5, eliminating efficiency generates a much more shrunken low-skill wage distribution, implying that higher-paid cities for low-skill workers are more productive.

Similar to high-skill wage, higher-paid cities are also worse-off in amenities to low-skill workers. Equalizing amenities makes those originally higher-paid cities more attractive to low-skill workers, large concentration of low-skill workers do not enhance productivity; besides, compensation for amenities is higher in amenable places, earnings for low-skill workers would largely decline. When there are a large number of high-skill workers as well as low-skill workers present in this kind of cities, productive and less amenable, high-skill welfare should be smaller than low-skill welfare, thus welfare inequality is narrowed.

The different patterns in counterfactual amenities between the two skill cohorts might result from: Denser high-skill cities are more productive while low-skill concentration is incapable of enhancing agglomeration economies. So high-skill workers get to choose their ideal locations and low-skill workers act as followers, of course both groups get their respective indifferent utilities.

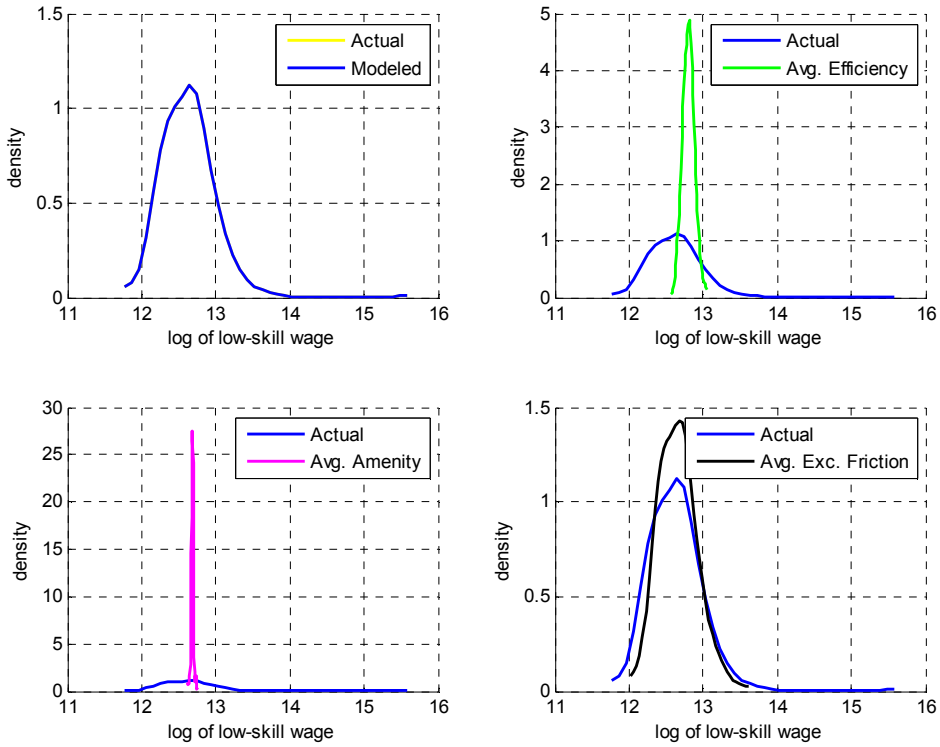



Figure 5: Counterfactuals of Low-skill Wage Without One Shock, $\kappa$ = 0.0015, $\eta$ = 0, $\beta$ = 0, $\zeta$ = 0

Following what has been examined in Hao and Fu (2014), we could derive location advantage shifter $A = a^{\eta_1}$ as below

$$h_2 = f(a, N, S) = a^{\eta_1} N^{\eta} S^{\beta} \tag{33}$$

where high-skill share $S = L_2 / L_1$, and city size $N = L_2 + L_1$. We set the elasticity of productivity to city size at 0.1, and high-skill share at 0.4.

In a larger city, sorts of leisure such as theaters, shopping malls and cafes, etc, are widely provided since there is a larger market for them; on the other round, these facilities are appealing to workers who have strong preferences to them. Same as D-RH (2013), urban amenities could be expressed as

$$\rho_i = \overline{\rho}_i N_i^{\zeta} \tag{34}$$

and we set the externality parameter at 0.02.

In Figure 6, Dispersed distribution of efficiency is a compounding outcome of exogenous location advantage and endogenous agglomeration economies. According to Fu and Hao (2014), agglomeration economies is composed of both effects from city size and from urban skill-mix; larger city size and denser concentration of high-skill workers facilitates more extensive division of labor and enhance urban productivity.

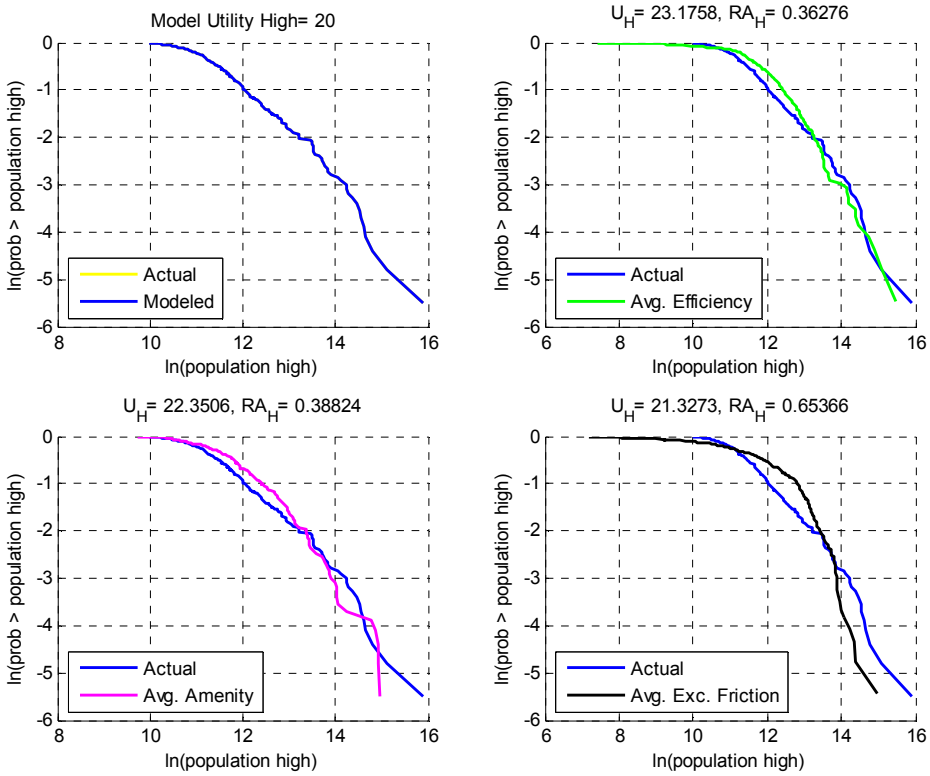


Figure 6: Counterfactuals of High-skill Population Size Without One Shock, $\kappa = 0.0015$, $\eta = 0.1$, $\beta = 0.4$, $\zeta = 0.02$

Remarked in Figure 5, urban amenities could be decomposed as exogenous physical amenities and endogenous effects from urban size. In presence of agglomeration economies, cross-city variations in efficiency and amenities are largely reduced. Counterfactual efficiency and excessive frictions utilities are both lowered a little bit, while counterfactual amenity utility is rising. Reasons: in counterfactual amenity case, some cities with large quantity of high-skill workers have worse physical amenities so that equalization of it could generate a small increase in utility; in counterfactual efficiency and excessive frictions, the condition of no exploitation of two types of externality is exacerbated. But by and large, the change in utilities are virtually tiny, since there are three forces that are counteractive to each other: on one hand with externalities, underlying differences across cities are reduced so that welfare gains is increased due to concave total wages and convex total friction; on the other hand with the existence of externalities, more alike cities occur welfare losses since they are unable to exploit externalities.

In Figure 7, it is similar to Figure 6 high-skill population counterfactual analyses. Relative to Figure 1(Figure 2), reallocations of high-skill(low-skill) workers in Figure 6(Figure 7) are pretty drastic. Reason: the effects of agglomeration economies get compounded. Relative to Figure 1(Figure 2), city selection effects in Figure 6(Figure 7) are stronger (i.e. many more cities disappear or become extremely small) with the presence of externalities; the reason is compounded effects of agglomeration economies. But in no-externality case, 1 city disappears respectively in counterfactual efficiency and excess friction case. In all-externality case, 4 cities disappear in

counterfactual efficiency, 7 cities exit in counterfactual excess friction. No cities exit when equalizing amenities. So city selection effects are much weaker than in homogenous skill analyses.

From counterfactual amenity exercises in Figure 1, 2, 6 and 7, the accordance in size patterns we observed gives some implications: first, extremely small cities are amenable because of natural advantage in amenity; second, that medium-sized cities are insufficiently amenable at least relative to its weighted average is due to low natural amenity and due to small population; third, mega cities are amenable is definitely ascribing to both ample natural amenities and massive social amenities.

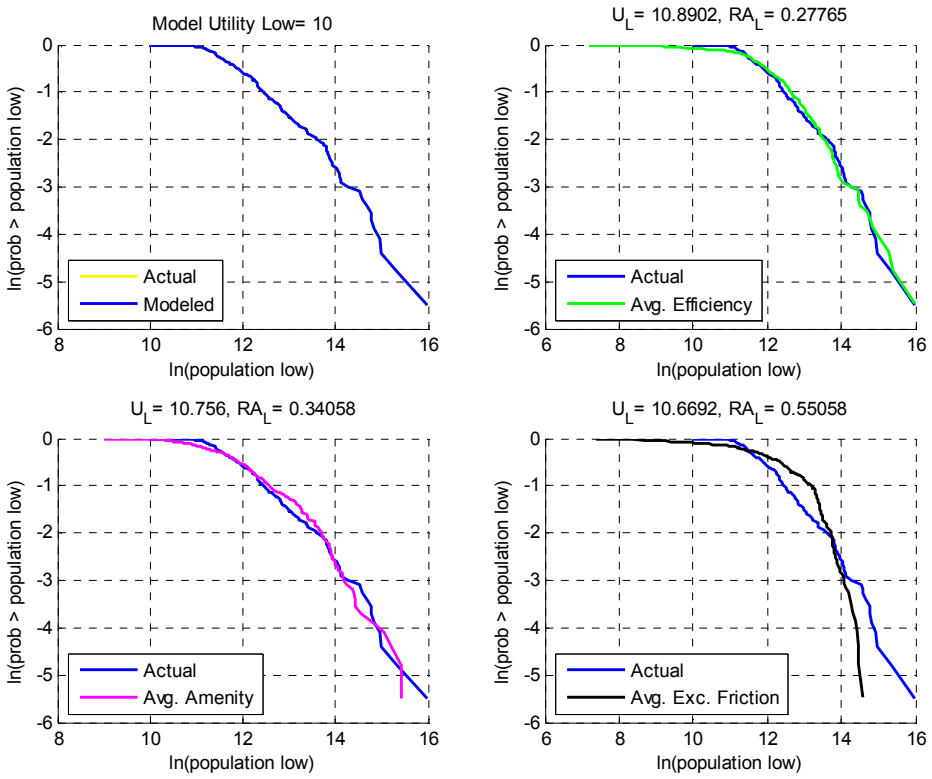



Figure 7: Counterfactuals of Low-skill Population Size Without One Shock, $\kappa = 0.0015$, $\eta = 0.1$, $\beta = 0.4$, $\zeta = 0.02$

Reallocations of counterfactual amenity and excessive frictions are much bigger than that of counterfactual efficiency. Reason: amenities have direct impacts on city size and skill-share, and those directly influence productivity; excessive frictions have direct impacts on city size which affect productivity, and productivity affects the skill composition, and skill composition affects productivity. We are unsure about counterfactual amenity's large reallocation because it is unclear to conclude about the counterfactual size distribution(Figure 6 and 7) and skill-share distribution (Figure 8); however, equalizing excessive frictions generates more even city size distribution and skill-share distribution, less dispersed distributions of these two determinants of urban productivity imply a more even efficiency distribution, which together with equalized excessive frictions, induce a larger population reallocation.

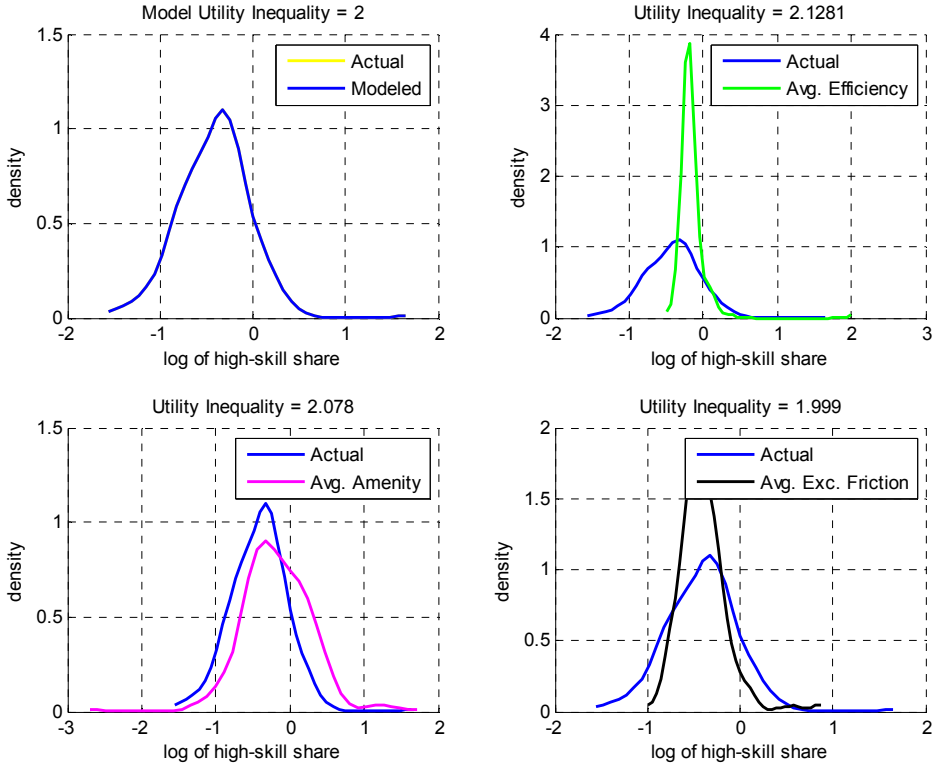



Figure 8: Counterfactuals of City Education Level Without One Shock, $\kappa = 0.0015$, $\eta = 0.1$, $\beta = 0.4$, $\zeta = 0.02$

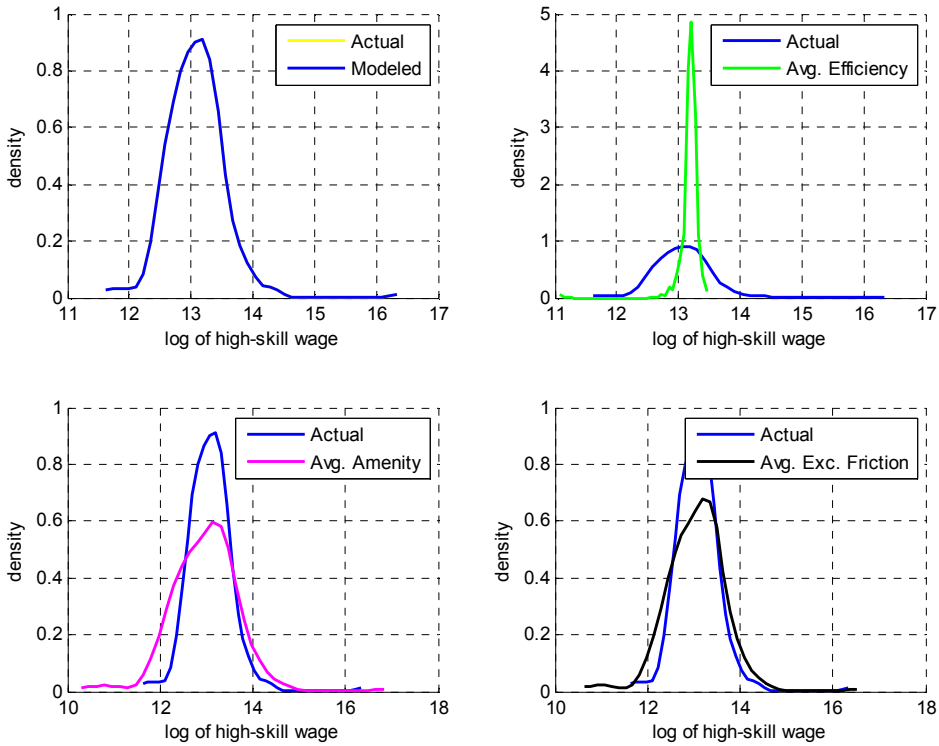



Figure 9: Counterfactuals of High-skill Wage Without One Shock, $\kappa = 0.0015$, $\eta = 0.1$, $\beta = 0.4$, $\zeta = 0.02$

In Figure 8, as for efficiency, it is the most contributable factor to welfare inequality for about 13 percentage points. These imply that more educated cities are

more productive in natural advantage, less frictional, while not very naturally amenable. Notice that in counterfactual efficiency case, some more cities gain higher skill-share. These happen to those college towns which are small but educated, such as Madison (WI), Boulder (CO), Ann Arbor (MI), etc. Once the natural advantage of them is averaged, they lose more college graduates and lose even more low-skill workers because of no jobs created for the low-skill. And the fact is they are mostly become too small to be a city because productivity is their major comparative advantage.

Figure 9 shows that higher-paid cities are more productive in natural advantage, less frictional, and not very desirable in natural amenities. Compared with the case without externalities, lower-paid cities provide even lower high-skill wages. This might be account of the fact that these less productive cities happen to be very amenable cities to high-skill workers, compensation has to be paid to get indifferent utilities in equilibrium.

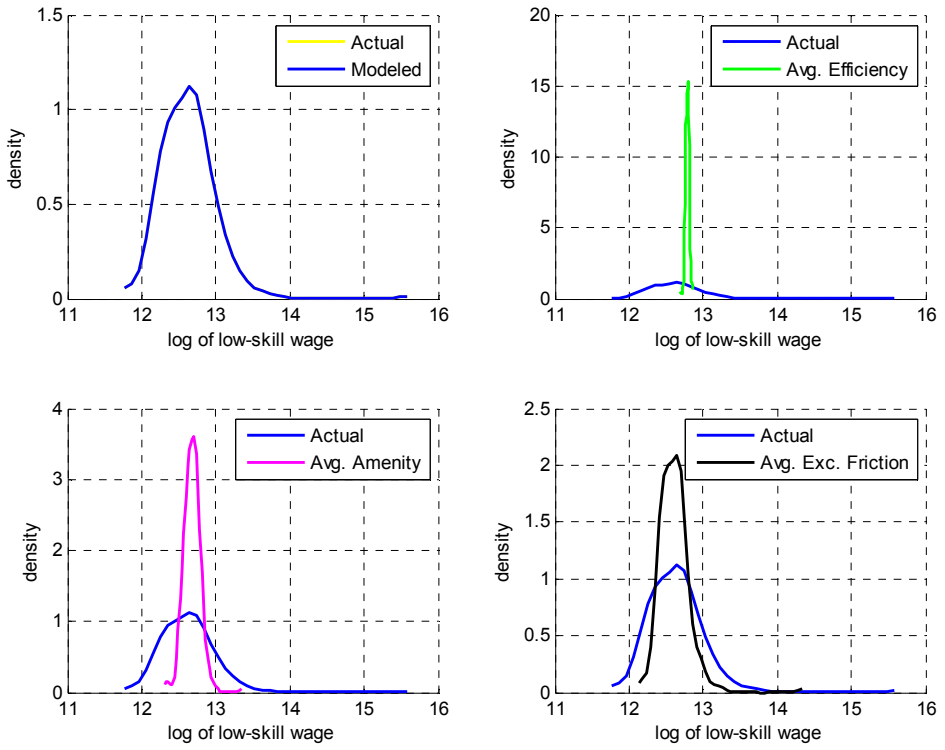



Figure 10: Counterfactuals of Low-skill Wage Without One Shock, $\kappa = 0.0015$, $\eta = 0.1$, $\beta = 0.4$, $\zeta = 0.02$

Figure 10 demonstrates that distribution of low-skill wage in counterfactual amenity case is less even than in Figure 5. When externalities are controlled, low-skill preferred amenities are less varied, equalizing which would bring about reaction not as fierce as in Figure 5. Location fundamentals and agglomeration economies together and simultaneously determine skill composition and wages; they are respectively unable to be accounted for by neither one of them alone.

Table 1 reports utility levels of high-skill and low-skill workers in counterfactual exercises. Roles played by location fundamentals and agglomeration economies are displayed in this table. Dispersed distribution of any of the three urban characters leads to a narrower welfare inequality gap between the high-skill and low-skill groups. Productivity is the greatest contributor. Comparison between 2.1445 and 2.0955 implies that role of externalities on welfare inequality is to shift counterfactual scenarios closer to reality. If in reality welfare inequality is narrower, then externalities narrow inequality. Comparison between 2.0825 and 2.0780 suggests analogously that amenity externality plays the same role as productivity externalities.

**Table 1:**

***Skilled utility in counterfactual scenarios***

| Scenarios / Skills | Averaged urban characters | $\eta=0$ $\beta=0$ $\zeta=0$ | $\eta=0.1$ $\beta=0$ $\zeta=0$ | $\eta=0$ $\beta=0.4$ $\zeta=0$ | $\eta=0.1$ $\beta=0.4$ $\zeta=0$ | $\eta=0.2$ $\beta=0.4$ $\zeta=0$ | $\eta=0.1$ $\beta=0.8$ $\zeta=0$ | $\eta=0.1$ $\beta=0.4$ $\zeta=.02$ |
|---|---|---|---|---|---|---|---|---|
| High-skill | Productivity | 23.3613 | 22.7806 | 23.0880 | 22.5443 | 22.1412 | 22.4081 | 23.1758 |
| | Amenity | 22.0597 | 22.2528 | 22.2388 | 22.4295 | 22.6196 | 22.6057 | 22.3506 |
| | Excessive frictions | 22.1648 | 22.2031 | 22.3205 | 22.3678 | 22.3984 | 22.5913 | 21.3273 |
| Low-skill | Productivity | 10.8936 | 10.8081 | 10.8645 | 10.7584 | 10.6778 | 10.7285 | 10.8902 |
| | Amenity | 10.6972 | 10.7355 | 10.7327 | 10.7704 | 10.8078 | 10.8051 | 10.7560 |
| | Excessive frictions | 10.7881 | 10.7961 | 10.8192 | 10.8288 | 10.8353 | 10.8739 | 10.6692 |
| Welfare inequality | Productivity | 2.1445 | 2.1078 | 2.1251 | 2.0955 | 2.0736 | 2.0887 | 2.1281 |
| | Amenity | 2.0622 | 2.0728 | 2.0721 | 2.0825 | 2.0929 | 2.0921 | 2.0780 |
| | Excessive frictions | 2.0546 | 2.0566 | 2.0630 | 2.0656 | 2.0672 | 2.0776 | 1.9990 |

**Table 2:**

***Statistics in selected counterfactual scenarios***

| S=L2/L1 N=L2+L1 | Scenarios / Items | 2005 Actual | $\eta=0,\ \beta=0,\ \zeta=0$ | | | $\eta=0.1,\ \beta=0.4,\ \zeta=.02$ | | |
|---|---|---|---|---|---|---|---|---|
| | | | Avg. productivity | Avg. amenity | Avg. exc. friction | Avg. productivity | Avg. amenity | Avg. exc. friction |
| Statistics of skill-share S | mean | 0.7175 | 0.7454 | 1.2252 | 0.7614 | 0.8779 | 0.9479 | 0.6966 |
| | Std. Dev. | 0.3866 | 0.1727 | 1.5755 | 0.1422 | 0.4399 | 0.5669 | 0.2372 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | median | 0.6747 | 0.7279 | 0.8386 | 0.7503 | 0.8286 | 0.8204 | 0.6533 |
| | Skewness | 6.7264 | 0.7936 | 5.7063 | 0.5158 | 13.7397 | 3.8262 | 3.7914 |
| | Kurtosis | 76.7596 | 4.1234 | 43.3111 | 8.5677 | 204.3486 | 25.8559 | 23.6435 |
| Correlations | corr(N,S) | 0.0901 | 0.0881 | -0.1019 | -0.0052 | -0.1565 | -0.1694 | 0.4470 |
| | corr(W2,S) | 0.8679 | -0.9701 | -0.1430 | -0.2849 | -0.6976 | -0.2633 | -0.2328 |
| | corr(W1,S) | 0.9059 | 0.9968 | -0.3831 | 0.4101 | 0.2574 | -0.7285 | -0.3461 |
| | corr(W2,N) | 0.0287 | -0.0785 | -0.0137 | -0.0361 | 0.5918 | -0.0049 | -0.2339 |
| | corr(W1,N) | 0.0345 | 0.0867 | 0.0915 | 0.0495 | -0.0451 | 0.0438 | -.3955 |
| Wage premium (weight = size) | w.a.W2/w.a.W1 | 1.3124 | 1.8100 | 0.8551 | 1.1581 | 1.6276 | 0.9229 | 1.1499 |
| | w.a.(W2/W1) | 1.3124 | 1.8100 | 0.8551 | 1.1581 | 1.6276 | 0.9229 | 1.1499 |

Note: 1. Skewness quantifies how symmetrical the distribution is (Above are all positively skewed: mean>median>mode). Kurtosis quantifies whether the shape of the data distribution matches the Gaussian distribution. 2. Weighted wage premium is used to investigate wage variation source (weight= size). 3. In no-externality case, 1 city disappear respectively in counterfactual efficiency and excess friction case. In all-externality case, 4 cities disappear in counterfactual efficiency, 7 cities exit in counterfactual excess friction. No cities exit when equalizing amenities. 4. Table 2 reports information of all cities including those should have been disappeared.

The counterfactual cases without externalities and with all three externalities are reported in Table 2. All counterfactual distributions of skill-share are similar to the reality skill share that they are all positively skewed, with mean of skill is bigger than its median. There are a few extensively educated cities, but the number of more educated cities is smaller than those with smaller than average skill-share.

Take the correlation between high-skill wage and skill-share in the counterfactual amenity case as an example, the coefficient is -0.1430. It conforms to above analyses, in which more educated cities are not very amenable and higher-paid cities are more desirable in amenities. Wage premium is essentially determined by amenities and skill-specific differential of willingness to pay for amenities. We investigate the sources of its variation across cities via comparison between weighted average of high-skill wage divided by weighted average of low-skill wage, and weighted average of wage premium.

When comparing 0.1727 with 0.4399, we find that 0.4399 is closer to mean, 0.3866. It implies that presences of externalities shift skill-share into its actual distribution shape. So is the implication from comparison between 1.5755 and 0.2372 in counterfactual amenities exercises. If skill is more concentrated in certain cities that leads to dispersed skill-share distribution (or heterogeneous skill-mix), then externalities play the role of enhancing skill concentration.

From Appendix Table 3, we find that reallocations of high-skill workers and low-skill workers in all counterfactual excessive frictions scenarios are larger than

those in counterfactual efficiency cases. Reasons are discussed above: indirect mechanism of excessive frictions determining urban productivity.

## *4. Conclusions*

Welfare inequality, skill-mix, urban size, skill-mix, and wage distributions are in virtue the mapping outcomes of exogenous spatial characteristics and externalities. This paper aims to account for geographic concentration of skills, skilled wage distributions and understands how heterogeneous urban attributes together with agglomeration economies contribute to welfare inequality gap. We first modify the theoretical model that Desmet and Ross-Hansberg (2013) has built. Since labor income losses instead of labor losses should have been accounted for, working hours then are invariant over cities. Based on our improvements made, we find analogous outcomes with analogous intuitions to explain these outcomes, even though counterfactual distribution patterns are much clearer after modification.

Three urban characteristics: efficiency, amenities, and excessive frictions, are found to be important to determine skill-share distribution. Dispersion of any of the three urban characters leads to narrower welfare inequality gap. Equalizing efficiency and excessive frictions give more equalized skill-share distributions, while shutting down differences in urban amenities generates a more dispersed skill-share distribution, i.e. the high-skill workers are concentrating into some cities. We also find urban efficiency has the most tremendous contribution to reducing welfare inequality. In general, dispersed efficiency narrows welfare inequality gap for about 14 percentage points compared with the case of shutting down efficiency variation by its weighted average.

This paper demonstrates the core of agglomeration economies is about specialization, of both two skill groups, instead of segregation between the two skill groups when they are imperfectly substitutable. More empirical and theoretical research on welfare inequality gap between skill groups in the future could benefit from this study.

# Appendix:

## Appendix A: Homogeneous skills analyses

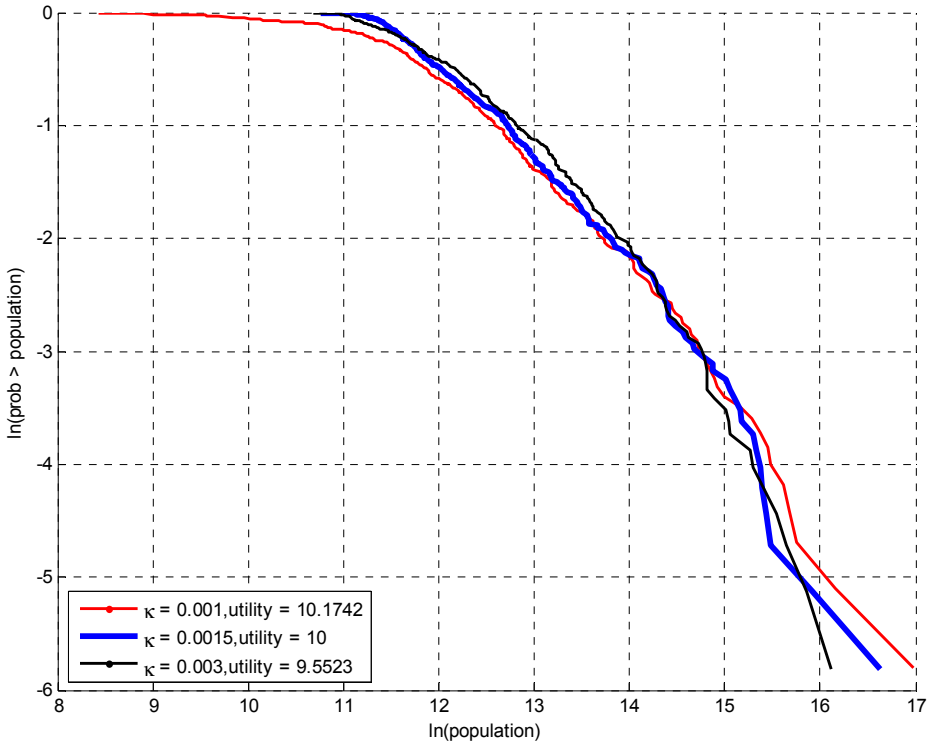



Figure 1: The City Size Distribution for Different Levels of Commuting-caused Labor Loss

Notes:

- More labor loss per mile due to commuting leads to more utility loss; more labor loss per mile leads to a more even city size distribution, b/c larger cities that are already congested are losing population and smaller cities are gaining population. Results are contrary to the case when $\kappa$ is half of it.
- Aggregate friction is a convex function of city size, so more even city size distribution lowers utility losses. So utility loss amount of enlarging $\kappa$ is composed of two parts: one is the loss due to workers moving out of large as well as productive cities, the other one is shrunken loss ascribing to a more even city size distribution. And utility gain amount of lowering $\kappa$ incorporates two parts: one is the gain resulting from more firms and workers moving into large as well as productive cities, the other one is augmented loss on account of a more uneven city size distribution.
- A large quantity of small cities disappears (log(8)*oldshare), b/c their comparative advantages is the extremely low congestions.
- Small cities turn to be even smaller instead of larger when $\kappa$ is doubled. Reason could be: firms and workers leave mega cities for comparatively smaller yet relatively productive cities, the second-tier ones. Besides, a more even city size distribution generates a level of total friction and smaller utility loss. So workers previously in small cities also would move to these second-tier cities without being exposed to as severe congestions as in mega cities. And we do observe expanded population size of second-tier cities in Figure 1($\kappa$=0.03).

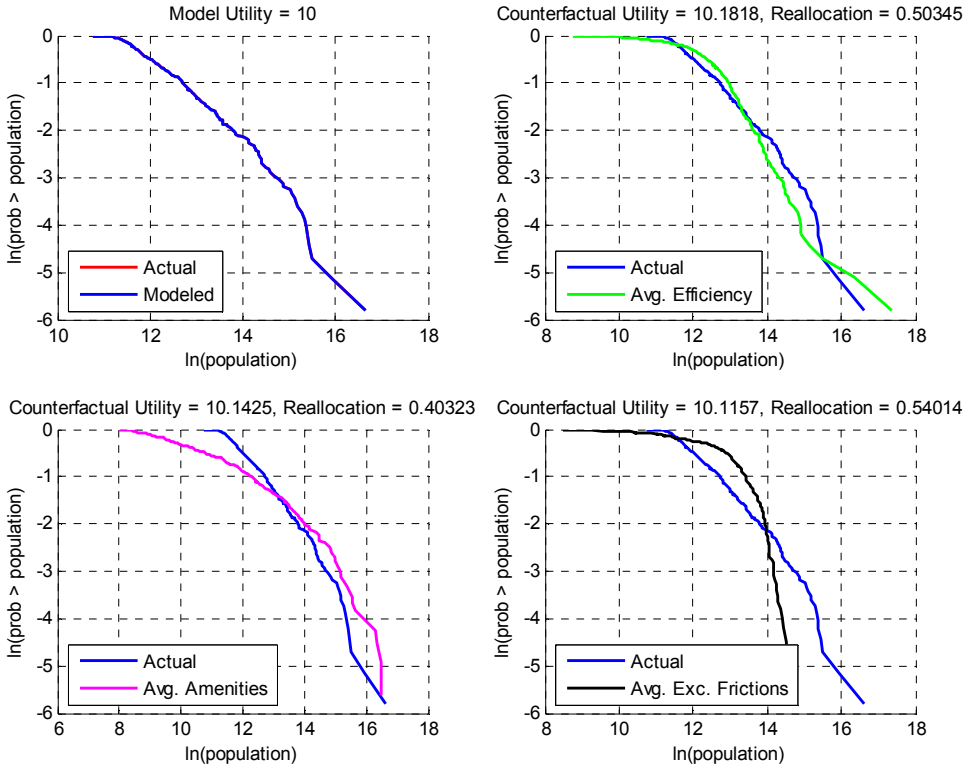



Figure 2: Counterfactuals Without One Shock, $\kappa = 0.0015$ , $\omega = 0$ , $\zeta = 0$

Notes:

- When respective distortion from any of the 3 urban characters is shut down, welfare gain is modest but population reallocation is large.
- Compared to D-RH(2013), utility levels are all higher, this is due to even more inefficient cities exit. So the city selection effect mentioned in D-RH(2013) is stronger here(city selection effect: namely is that many cities exit or become extremely small).
- Two exceptions in the case of equalizing productivity: one is mega cities gain population, the other one is extremely small cities lose population and even cannot survive. Reason could be: 1) some small cities are virtually productive such as college towns, and productivity is one major advantage; 2) certain mega cities are essentially not that productive. Implication is: calling for research on skill-mix.
- Equalizing amenity: city size distribution is more dispersed. It implies that small cities could be amenable, while large cities might be not as attractive in amenities as small cities.

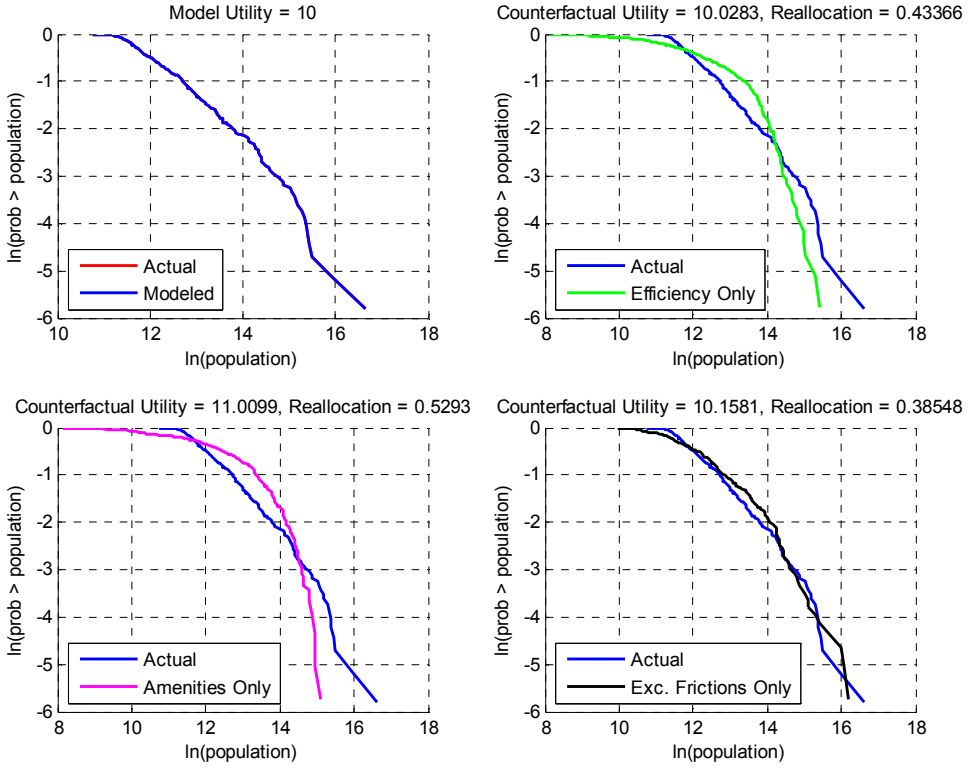



Figure 3: Counterfactuals With One Shock, $\kappa = 0.0015$ , $\omega = 0$ , $\zeta = 0$

Notes:

- Two forces city size distribution are at work to: one is dispersion force of differentiated urban factor, the other one is even force of the two equalized urban factors. In efficiency only and amenity only cases, city size distribution is less dispersed, suggesting that the even force of equalized amenities and excessive frictions(efficiency and excessive frictions) outperforms the dispersion force coming from efficiency(amenities). In excessive frictions case, the picture is more mixed and it is hard to say which force dominates.
- Either one or two of the three urban characters are able to document the city size distribution.

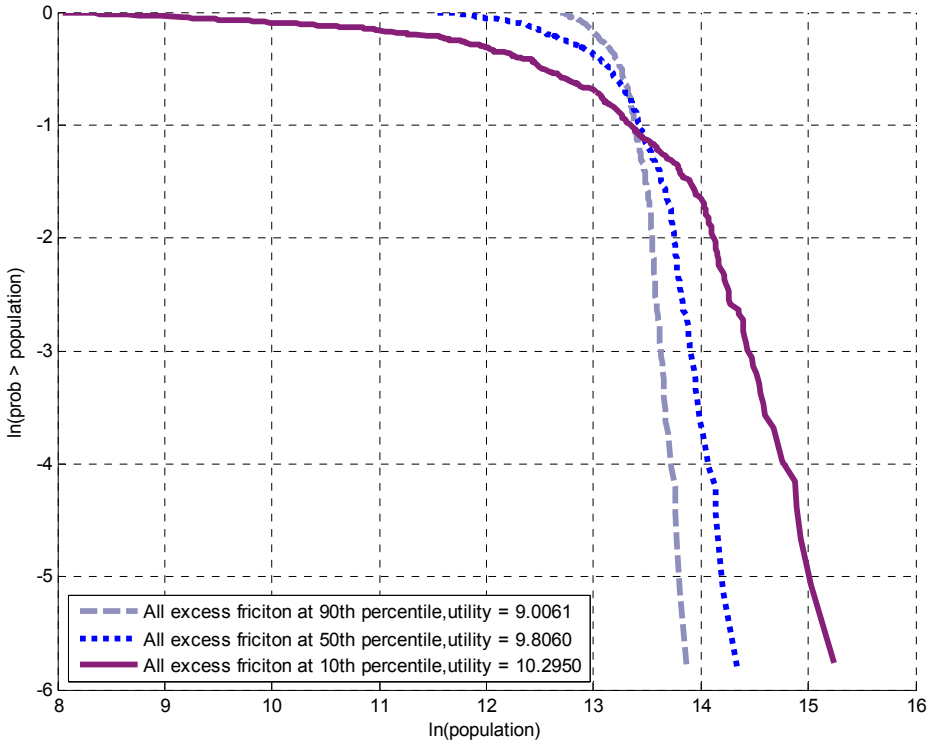



Figure 4: The City Size Distribution with Alternative Excessive Frictions

Notes:

- 90th percentile utility is smaller than in D-RH(2013), and 10th percentile utility is bigger than in D-RH(2013).
- 50th percentile is different from weighted average in above counterfactuals. With smaller counterfactual utility in 50th percentile excessive frictions, we could predict that median of excessive frictions is bigger than its weighted average.
- At 10th percentile excessive frictions, many more inefficient cities disappear.

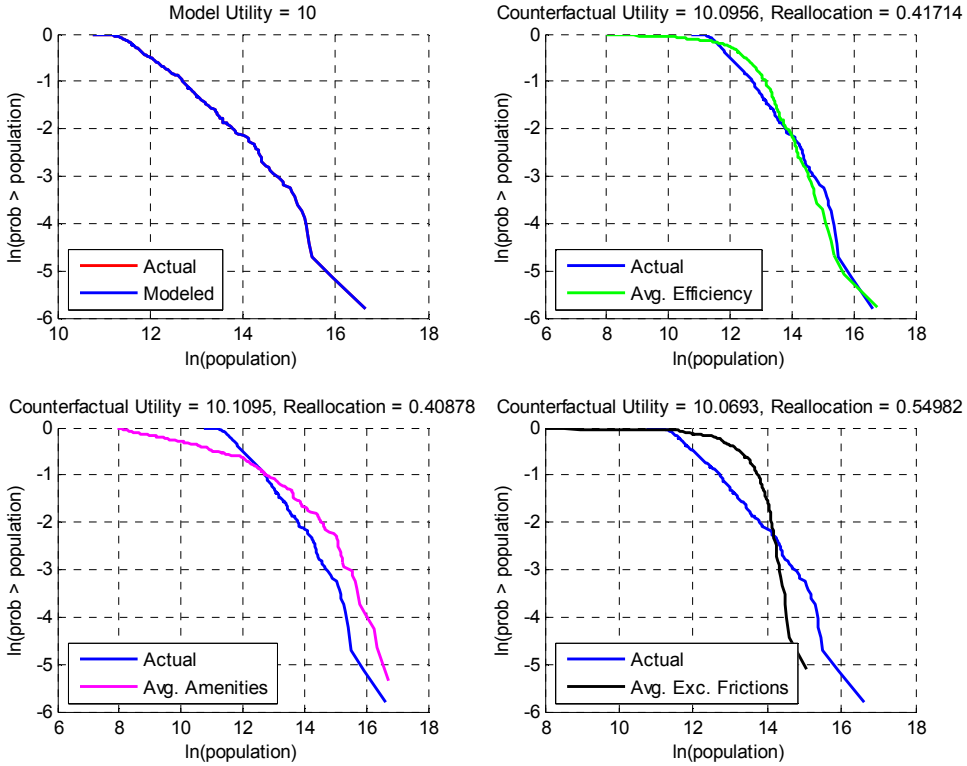



Figure 5: Counterfactuals Without One Shock, $\kappa = 0.0015$ , $\omega = 0.02$ , $\zeta = 0$

Notes:

- Relative to Figure 2 without agglomeration economies, two more counteractive forces are working: underlying variations across cities are largely reduced in the case of agglomeration economies, so welfare loss is reduced due to less distortion and due to convex total frictions; when equalizing exogenous location advantage, cities become more alike and then there will be no exploitation of externality, so welfare gain is reduced.
- Utility levels in the three counterfactual cases are lowered than in Figure 2, suggesting that reduced amount of welfare gain is bigger than that of welfare loss.
- A clearer trend of city size distribution in counterfactual efficiency of Figure 5 is shown, relative to that of Figure 2. Mega cities lose population, and this implies that mega cities are productive due to their high levels of exogenous location advantage.
- With respect to counterfactual efficiency and counterfactual amenity, our Figure 5 is much clearer in trends than those in D-RH(2013). It is a piece of evidence of our improved model.

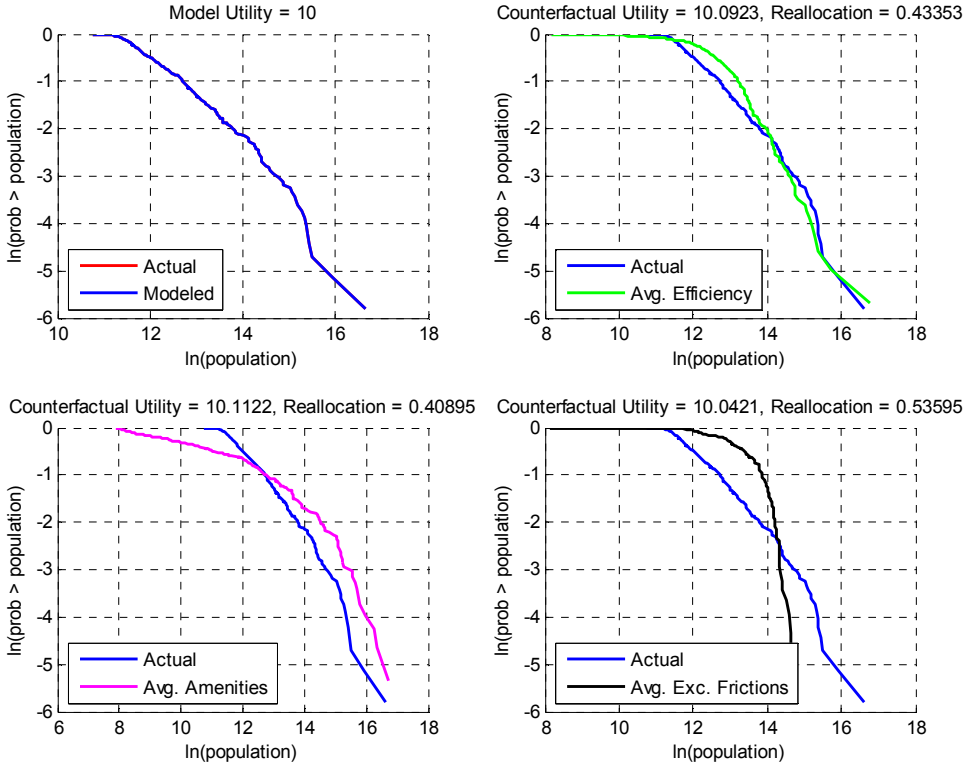



Figure 6: Counterfactuals Without One Shock, $\kappa = 0.0015$ , $\omega = 0.02$ , $\zeta = 0.02$

Notes:

- Compared with Figure 5, no qualitatively substantial change in patterns of city size distribution.
- In line with D-RH(2013), utility in counterfactual amenity is slightly higher than in Figure 5, while utilities in counterfactual efficiency and counterfactual excessive frictions respectively go down a little bit. Reasons: in counterfactual amenity, some large cities have worse physical amenities so that equalization of it could generate a small increase in utility; in counterfactual efficiency and excessive frictions, the condition of no exploitation of two types of externality is exacerbated.
- More small cities exit, b/c loss is compounded in the presence of externality.

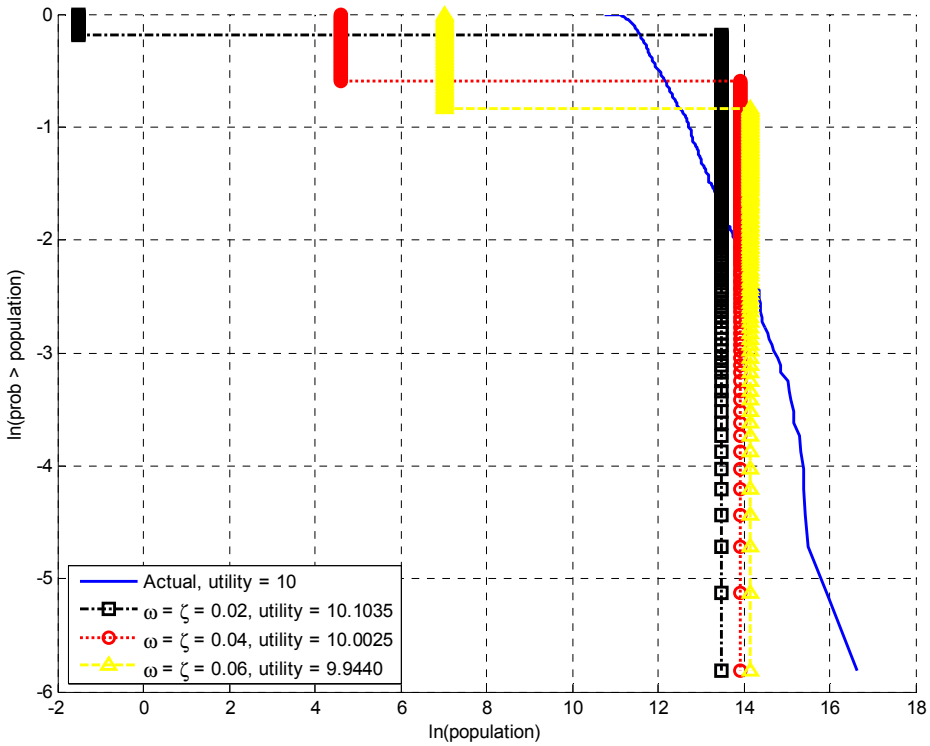
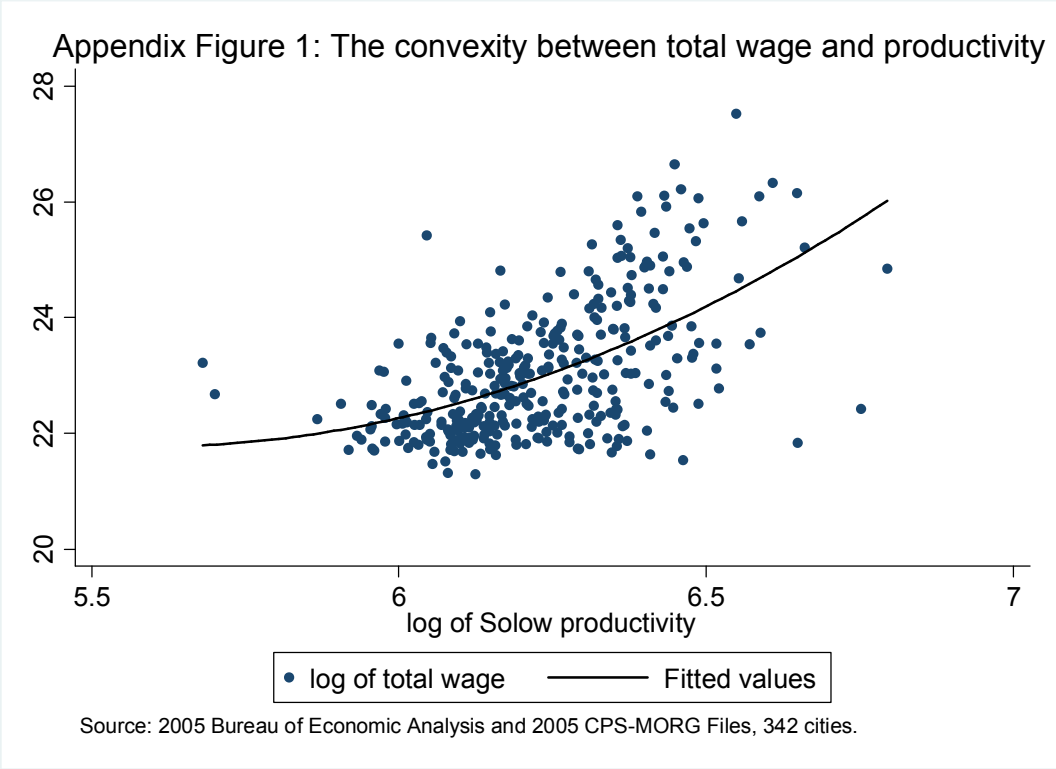



Figure 7: Counterfactuals Without All Three Shocks, $\kappa = 0.0015$

Notes:

- Eliminating all shocks leaves only two sizes of cities, one large and one small. Utility is slightly up in the 0.02 and 0.04 cases instead of down in D-RH(2013). Reason: lots of worker in large and productive cities are able to exploit productive externalities.
- As effects of agglomeration ascend, large and small cities both become larger, number of small cities increases and number of large cities declines.
- Utility is descended while increasing externalities. Reason: no variation in shocks dampens the condition of no exploitation of externalities, so larger externalities lead to more loss.

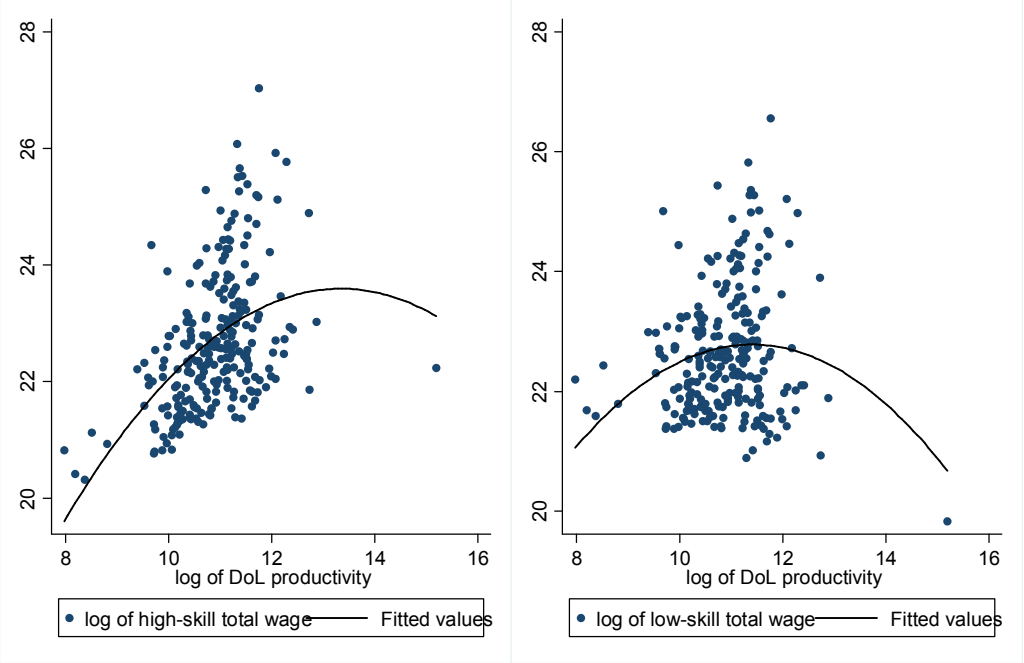
## Appendix B: Figures and Tables




Notes:

- Data provides evidence of convex relation between total wage and Solow residual productivity. So uneven productivity distribution boosts total wage and welfare.

Appendix Figure 2: The concavity between skilled total wage and productivity




Source: 2005 Bureau of Economic Analysis and 2005 CPS-MORG Files, 342 cities.

Notes:

- Data offers evidence of concave relation between total wage and Jones Division of Labor(DoL) productivity, respectively for high-skill and low-skill. So even productivity distribution increases total wage and welfare. This is one opposing component in our urban accounting.

**Appendix Table 1:**

***Iterated SUR of log skilled population and urban characteristics, 2005***

| | (1) | (2) |
|---|---|---|
| VARIABLES | logL2 | logL1 |
| Log of productivity | 1.013*** | 0.818*** |
| | (0.20) | (0.13) |
| High-skill amenity | 0.859*** | |
| | (0.18) | |
| Excessive frictions | -0.816*** | -0.877*** |
| | (0.06) | (0.05) |
| Low-skill amenity | | 1.609*** |
| | | (0.13) |
| Constant | 1.899 | 12.833*** |
| | (1.97) | (0.41) |
| Observations | 241 | 241 |
| R-squared | 0.389 | 0.517 |

Standard errors in parentheses

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

Notes:

- All urban characters are location fundamentals.
- We perform regressions of skilled population on location fundamentals of our calculated indexes of efficiency, amenities and excessive frictions. Efficiency is location advantage calculated by Hao and Fu (2014) in case of endogenous concerns.
- Estimations are performed by Ordinary Least square method in a system of Seemingly Unrelated Regression.
- Efficiency has a larger impact on high-skill population as well as low-skill population, implying that low-skill workers benefit in employment from locating nearing high-skill workers because more high-skill workers generate higher urban productivity.

**Appendix Table 2:**

***Iterated SUR of skilled total wage and urban characteristics, 2005***

| | (1) | (2) |
|---|---|---|
| VARIABLES | logwL2 | logwL1 |
| Log of productivity | 1.568*** | 0.784*** |
| | (0.22) | (0.14) |
| High-skill amenity | 1.276*** | |
| | (0.20) | |
| Excessive frictions | -0.894*** | -0.913*** |
| | (0.07) | (0.05) |
| Low-skill amenity | | 1.863*** |
| | | (0.15) |
| Constant | 7.466*** | 23.820*** |
| | (2.18) | (0.47) |
| Observations | 241 | 241 |
| R-squared | 0.333 | 0.441 |

Standard errors in parentheses

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

Notes:

- All urban characters are location fundamentals.
- We perform regressions of skilled total wages on location fundamentals of our calculated indexes of efficiency, amenities and excessive frictions. Efficiency is location advantage in case of endogenous concerns. Estimations are performed by Ordinary Least square method in a system of Seemingly Unrelated Regression.
- Efficiency and urban amenities have significant positive impacts on total wages of both high-skill and low-skill workers, while excessive frictions have significant negative impacts.

**Appendix Table 3:**

***Skill Reallocations in counterfactual scenarios***

| Scenarios | Averaged | $\eta = 0$ | $\eta = 0.1$ | $\eta = 0$ | $\eta = 0.1$ | $\eta = 0.2$ | $\eta = 0.1$ | $\eta = 0.1$ |
|---|---|---|---|---|---|---|---|---|

| Skills | urban characters | $\beta = 0$ $\zeta = 0$ | $\beta = 0$ $\zeta = 0$ | $\beta = 0.4$ $\zeta = 0$ | $\beta = 0.4$ $\zeta = 0$ | $\beta = 0.4$ $\zeta = 0$ | $\beta = 0.8$ $\zeta = 0$ | $\beta = 0.4$ $\zeta = .02$ |
|---|---|---|---|---|---|---|---|---|
| High-skill | Productivity | 0.1057 | 0.0673 | 0.0996 | 0.0795 | 0.0769 | 0.0878 | 0.3628 |
| | Amenity | 0.4509 | 0.4300 | 0.4235 | 0.4075 | 0.3916 | 0.3955 | 0.3883 |
| | Excessive frictions | 0.2143 | 0.1936 | 0.2368 | 0.2144 | 0.1882 | 0.2605 | 0.6537 |
| Low-skill | Productivity | 0.0857 | 0.0584 | 0.0836 | 0.0678 | 0.0663 | 0.0745 | 0.2777 |
| | Amenity | 0.3634 | 0.3796 | 0.3589 | 0.3713 | 0.3837 | 0.3681 | 0.3406 |
| | Excessive frictions | 0.1566 | 0.1574 | 0.1929 | 0.1889 | 0.1797 | 0.2382 | 0.5506 |

## Appendix C: Data

1980 MSA has been converted to 2003 standard of MSA definition by OMB.

- Aggregate consumption in 1980 and 2005 are retrieved from Bureau of Economic Analysis, Regional Economic Accounts. Table of Personal income and earnings and Table of National Income and Product Accounts. Method of calculating consumption of each city is in D-RH (2013), instead, I use Market based PCE in addenda and exclude durable goods consumption.
- Total populations in 1980 and 2005 are obtained from Bureau of Economic Analysis, Regional Economic Accounts.
- Urban employment is total population excluding retired people in 1980 (2005). We use share of people aged 65 and above in year 1980 (2005).

$$N_i = total\ population_i * (1 - oldshare80)$$

and old people share in 1980 is calculated based on IPUMS 5% state sample census in 1980, the share in 2005 is calculated based on census 2000 and census 2005 assuming aged population share grows with the same exponential speed during 2000 and 2010.

- Aggregate personal income is from Bureau of Economic Analysis, Regional Economic Accounts.
- Skilled wages and populations in 2005 are from Current Population Survey,

MORG Files in NBER, 1980 information is from IPUMS 5% state sample census 1980.

## References


Autor, D., Katz, L., & Kearney, M. S. (2004). *Trends in U.S. Wage Inequality: Re-Assessing the Revisionists*. 1-89.

Baum-Snow, N., Freedman, M., & Pavan, R. (2014). *Why Has Urban Inequality Increased?* Manuscript.

Baum-Snow, N., & Pavan, R. (2012). Understanding the City Size Wage Gap. *Review of Economic and Statistics, 79*.

Behrens, K., Duranton, G., & Robert-Nicoud, F. (2010). *Productive cities: Sorting, Selection, and Agglomeration*.

Berry, C. R., & Glaeser, E. L. (2005). The Divergence of Human Capital Levels across Cities. *Papers in Regional Science, 84*(3), 407-444

Brinkman, J. (2014). *The Supply and Demand of Skilled Workers in Cities and the Role of Industry Composition*. Federal Reserve Bank of Philadelphia. 1-40.

Davis, D. R., & Dingel, J. I. (2012). *A Spatial Knowledge Economy*. NBERWorking Paper No. 18188.

Davis, D. R., & Dingel, J. I. (2013). *The Comparative Advantage of Cities*.

Desmet, K., & Rossi-Hansberg, E. (2013). Urban Accounting and Welfare. *The American Economic Review, 103*, 2296-2327.

Diamond, R. (2012). *The Determinants and Welfare Implications of US Workers Diverging Location Choices by Skill: 1980-2000*. Job Market Paper. 1-89.

Eeckhout, J., Pinheiro, R., & Schmidheiny, K. (2010). *Spatial sorting: Why New York, Los Angeles and Detroit attract the greatest minds as well as the unskilled*.

Fu, Y. M., & Liao, W. C. (2012). *What Drive the Geographic Concentration of College Graduates in US? Evidence of Learning in Cities from Internal Migration*.

Giannetti, M. (2003). On the Mechanics of Migration Decisions: Skill Complementaries and Endogenous Price Differential. *Journal of Development Economics, 71*, 329-349.

Glaeser, E. L. (1999). Learning in Cities. *Journal of Urban Economics, 46*(2), 254-277.

Glaeser, E. L. (2013). A Review of Enrico Moretti's The New Geography of Jobs. *Journal of Economic Literature, 51*(3), 825-837.

Glaeser, E. L., & Mare, D. C. (2001). Cities and Skills. *Journal of Labor Economics, 19*(2), 316-342.

Glaeser, E. L., & Resseger, M. G. (2010). The Complementarity between Cities and Skills. *Journal of Regional Science, 50*(1), 221-244. doi: 10.1111/j.1467-9787.2009.00635.x

Goldin, C., & Katz, L. (2007). *Long-Run Changes in the U.S. Wage Structure: Narrowing, Widening, Polarizing*. NBER Working Papers 13568. National Bureau of Economic Research, Inc.

Gyourko, J., Mayer, C. J., & Sinai, T. (2006). *Superstar Cities*.

Hao, Y., & Fu, Y. M. (2014). *Productivity Accounting across Cities: When Skill is Heterogeneous*. National University of Singapore. 1-48.

Jones, B. F. (2014). The Human Capital Stock: A Generalized Approach. *The American Economic Review, 104*(11), 3752-3777.

Katz, L., & Autor, D. (1999). Changes in the Wage Structure and Earnings Inequality. *Handbook of Labor*

*Economics, Ashenfelter and Crad, Eds, Elsevier*.

Lee, S. H. (2010). Ability Sorting and Consumer City. *Journal of Urban Economics, 68*, 20-33.

Lindley, J., & Machin, S. (2014). Spatial Changes in Labour Market Inequality. *Journal of Urban Economics, 79*, 121-138.

Lucas, R. E. (2004). Life Earnings and Rural-Urban Migration *Journal of Political Economy, 112*(1), S29-S59.

Moretti, E. (2004a). Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data. *Journal of Econometrics, 121*(1-2), 175-212. doi: 10.1016/j.jeconom.2003.10.015

Moretti, E. (2004b). Human Capital Externalities in Cities. *Handbook of Regional and Urban Economics, 4*, 2243-2291.

Moretti, E. (2008). *Real Wage Inequality*. IZA discussion papers, No. 3706. 1-49.

Moretti, E. (2012). *The New Geography of Jobs*. Houghton Mifflin Harcourt: Mariner Books; Reprint edition (March 19, 2013).

Pereira-Lopez, M., & Soloaga, I. (2013). *External Returns to Higher Education in Mexico 2000-2010*. Documento de trabajo No.3.

Venables, A. J. (2011). Productivity in Cities: Self-selection and Sorting. *Journal of Economic Geography, 11*, 241-251.

Winters, J. V. (2012). *Human Capital Externalities and Employment Differences across Metropolitan Areas of the U.S*.