Analyzing Urban Systems: Have Mega-Cities Become Too Large?*

Klaus Desmet
Universidad Carlos III

Esteban Rossi-Hansberg
Princeton University

October 20, 2014

Abstract

With the trend toward greater urbanization and the rapid emergence of mega-cities continuing unabated, many policy makers ask themselves whether some cities are becoming too large, and whether policies should be aimed at stimulating the growth of intermediate-sized cities. In this chapter we use a simple model of a system of cities, together with some basic urban and aggregate data, to answer some of the following questions. Would there be any welfare gains from reducing the size of mega-cities and increasing that of intermediate-sized cities? Is there any sense in implementing policies that make cities more equal in terms of efficiency and amenities? Should infrastructure investments be focused on improving life in the most lagging cities? Is it advisable for policy to have a small city bias? As developing countries continue to grow, will their largest cities tend to become even larger or are we likely to see a more equal spatial distribution of people? We provide quantitative answers to these questions by using data from the U.S., as well as from two large developing countries, China and Mexico.

1. INTRODUCTION

The trend toward ever greater urbanization is continuing unabated across the globe. According to the United Nations, by 2025 close to 5 billion people will live in urbanized areas. Many cities, especially in the developing world, are set to explode in size. The Nigerian city of Lagos, for example, is expected to increase its population by 50%, to nearly 16 million, in the next decade and a half (UN-Habitat 2010). Naturally, there is an active debate on whether restricting the growth of mega-cities is desirable, and whether it can make residents of those cities and their countries better off.

Importantly, this debate is not so much about urbanization per se –whether people should move to cities or stay in the countryside– but rather about whether (some) of the world's mega-cities are

*We acknowledge the support of the International Growth Centre at LSE (Grant RA-2009-11-015).

creating mega-problems that could be avoided with suitable policies. People flock to cities in search of higher paying jobs and better amenities. Many of the world's large metropolises, such as Los Angeles and Mumbai, are highly productive and are located next to large bodies of water. As cities grow in size, however, they start suffering from increased congestion, higher crime rates, and air pollution. How fast the benefits of efficiency and amenities erode with population size because of increasing congestion costs depends on the quality of governance, responsible for the provision of road infrastructure, sewage systems, clean water, and security. Cities obviously differ in their efficiency, their amenities and the quality of their governance, and so there is no one answer to what their optimal size should be. We need analytical tools that can help us evaluate the desirability of policies that hinder or promote the growth of cities of different sizes. This will allow us to assess urban policies that depend on the size of the cities where they are implemented, namely scale-dependent policies.

When analyzing whether mega-cities have become "too large," policy makers often focus on an in-depth analysis of a particular city, such as Mexico City, Cairo, or Shanghai. However, no city is an island:[1] improving urban infrastructure in one city may attract immigration from other cities, and a negative shock in one location may be mitigated because people can move to another location. Considering the general equilibrium nature of any such urban policy is therefore key. That is, when deciding whether to make intermediate-sized cities more attractive, policy makers need to understand how that will affect both smaller and larger cities.[2]

There is thus a need for quantitative models of systems of cities that are complex enough to account for the general equilibrium nature of the problem but simple enough in terms of their structure and their data requirements to make them usable for policy makers. In Desmet and Rossi-Hansberg (2012) we develop such a framework. In this chapter we start by briefly sketching the main forces in that model, and discuss which data are needed to do policy analysis. (For the interested policymaker, the Technical Appendix provides a more detailed step-by-step practical guide on how to implement the methodology.) It then shows how this framework can be used to quantitatively analyze a number of important questions. Would there be any welfare gains from reducing the size of mega-cities and increasing that of intermediate-sized cities? Is there any sense in implementing policies that make cities more equal in terms of efficiency and amenities? As developing countries grow further, how

[1] As in John Dunne: *No man is an island entire of itself; every man is a piece of the continent, a part of the main.*

[2] The literature on firm dynamics and industrial policy has also focused on this problem. For example, is it desirable to subsidize small vs. large firms? (see, e.g., Restuccia and Rogerson, 2008)

will this affect their city-size distribution? Are cities in developing countries really too large? Should policy favor smaller cities or medium-large cities? Clearly, the answers to these questions will be country-specific. Different countries are at different levels of development, have different geographies and different histories. We will therefore compare the urban systems of three countries: the United States, which as the world's most advanced economy will serve as a benchmark; China, Asia's and the world's largest and fastest-growing country; and Mexico, one of Latin America's most important middle-income economies.

A first policy question we ask is how reducing spatial differences in city characteristics affects the city-size distribution. Many countries favor balanced spatial growth and try to reduce uneven development. Spatial concentration is often viewed with suspicious eyes, and many countries implement policies aimed at mitigating spatial differences. The hope is to create a less dispersed city-size distribution. When reducing differences in productivity across cities, our quantitative analysis shows different responses in different countries. In the United States the city-size distribution would become slightly less dispersed, with the larger cities becoming smaller and the smaller cities becoming larger. In Mexico, something similar happens: Mexico City declines, without any other city taking its place. In contrast, in China the opposite happens, with the city-size distribution becoming more dispersed, implying larger mega-cities. Although intuitively one would expect that reducing spatial differences should lead to a less dispersed city-size distribution, the opposite result in China illustrates the importance of doing model-based analysis. If not, policies might have unintended consequences: a policy aimed at reducing the dispersion in city sizes might actually increase it.

When mitigating spatial differences in city characteristics, we find limited welfare effects in the U.S. and Mexico, but very large effects in China. This reflects the important spatial disparities in productivity that plague China and the mobility restrictions under the *hukou* system. It suggests that as China develops further, there will be huge benefits from the spatial reallocation of people and economic activity. We do not find the same effects in Mexico, probably because it has a much longer history as a developing market economy, and thus suffers from less spatial distortions.

A second policy question we ask is whether policies that target lagging cities are more effective than policies that target cities of different sizes. Governments may choose to improve road infrastructure in remote areas or give incentives for firms to move to particular locations. A good example was the U.S. federal policy of the 1960s to assist the declining Appalachian region (Glaeser and Gottlieb, 2008). Another example are the European Union structural funds that subsidize investments in infrastructure and human capital in regions with an income per capita below a certain threshold.

Other policies favor cities of particular sizes. For example, China's urbanization policies in the 1980s and the 1990s had a strong small city bias, as it was based on "controlling the big cities, moderating development of medium-sized cities, encouraging growth of small cities" (Kamal-Chaoui et al., 2009). Our results suggest that policies that target small cities are less effective than those that target lagging cities. "Encouraging growth of small cities", as China used to do, may thus be unwarranted. Of course, to target lagging cities, there is a need for a methodology to rank cities on the basis of different characteristics. Our model provides a framework to do so.

We also analyze the spatial impact of country-wide shocks. As an example, we evaluate the effect of Mexico's 1994 crisis. Our results show that productivity differences across cities increased during this episode. This suggests that the country's largest cities, such as Mexico City, suffered less, in relative terms, compared to the smaller cities. So large cities seem better at absorbing economy-wide shocks.

The remainder of the chapter is organized as follows. Section 2 provides a brief summary of the theoretical framework, and discusses which data are needed. Section 3 implements the methodology for the benchmark case of the United States. Section 4 does the same for China and Mexico, and compares the findings. Section 5 concludes. A Technical Appendix guides the reader through a practical step-by-step discussion of how to do the analysis.

2. A SIMPLE FRAMEWORK OF A SYSTEM OF CITIES

To analyze how different policies affect a country's city size distribution, we need a framework that takes into account that each city is part of a system of cities, so that a shock to one city will affect other cities. For example, a policy that improves the amenities in a country's medium-sized cities will not only affect the attractiveness of those medium-sized cities, but will also have an impact on its larger cities. The framework should also capture that the size of a city depends on multiple determinants, such as its productivity, its amenities, or its governance. Since these different determinants are inter-related, improving a city along one dimension may not have the expected effect if the other dimensions are ignored. In Desmet and Rossi-Hansberg (2012) we develop such a model of a system of cities. In what follows we provide a brief overview of the forces at work in the model. For policymakers interested in implementing the methodology in other countries, the Technical Appendix provides a simple step-by-step guide on how to do so.

In the framework of Desmet and Rossi-Hansberg (2012) a country has a given number of cities.

So the model is not suited to study city creation or destruction. Each city has three characteristics that affect its size:

- *Efficiency*

  A city's productivity refers to the efficiency with which a city is able to produce. One can either consider a city's productivity to be exogenously given, or one can assume that its productivity depends, at least partly, on its size, because of agglomeration economies. In other words, larger cities may be larger because they are more productive, but they may also be more productive because they are larger.

- *Amenities*

  A city's amenities refer to anything (positive or negative) that changes the attractiveness of a city as a place to reside, without directly affecting its productivity or distorting its labor market. This includes weather, natural features (being close to water), pollution, crime, cultural and sports activities, and any other factor that influences a city's quality of life.

- *Excessive frictions*

  As cities become larger, they become more congested. This restricts a city's population. But even among cities of the same size, not all are equal. Some cities are less efficient in dealing with congestion than others. This may reflect, for example, the fact that some city governments are less good in providing congestion-alleviating infrastructure than others. This would translate into higher local tax rates or other distortions. We refer to a city's excessive frictions as those distortions over and above what we would expect given the city's actual size.

Cities will be larger if they have higher efficiency, better amenities, or lower excessive frictions. Of course, what matters for an individual city is not just its own characteristics, but how those compare with the characteristics of all other cities. A highly productive city will have a smaller population if there are many other highly productive cities. A system of cities is in equilibrium if no one wants to (or can) move. This does not imply that wages should equalize across cities. Even if people can move at no cost, an individual may be willing to accept a lower wage in a city with better amenities. In many countries, such as China, there exist internal migration restrictions. In our model a city which restricts entry will show up as having worse amenities. We will return to the question of mobility when we discuss our findings for China.

An important advantage of the framework of Desmet and Rossi-Hansberg (2012) is that it can be implemented and used for policy analysis with relatively few data requirements. To identify *productivity*, we use information on income, hours worked and possibly (but not necessarily) capital. By using data on income, consumption and hours worked, we can easily get an estimate of a city's distortions. By comparing a city's distortions to the distortions it is expected to have given its size, we can identify a city's *excessive frictions*. Finally, by using data on population and matching a city's actual size to the one generated by the model, we can back out a city's implied *amenities*. The Technical Appendix provides further details on how to identify these three characteristics.

The bottomline is that having data on population, income, consumption, hours worked, and possibly capital at the level of cities suffices to identify the determinants of the city size distribution. Collecting such data is relatively straightforward. Table 1 gives a brief summary of how we did this for the U.S., China and Mexico. Some metropolitan data, such as GDP in the case of the U.S., and population in the case of China and Mexico, are directly provided by national statistical offices. Other metropolitan data, such as hours worked in the U.S. and Mexico, need to be constructed from micro-data. Even when not using micro-surveys, we often need to aggregate data up to the metropolitan level. For example, in the U.S. many data are provided at the county level (metropolitan statistical areas are a collection of counties), and in Mexico most variables are available at the municipal level (metropolitan areas are a collection of municipalities).

In order to make reasonable cross-country comparisons, another relevant point is to use geographic units which are comparable. In the U.S., metropolitan statistical areas (MSAs) are supposed to capture meaningful economic geographies, making them preferable to using either counties or places. In China, prefecture-level cities cover the entire area of the country, and should thus be understood as metropolitan areas with their rural hinterlands, making them hard to compare to MSAs in the U.S. That is why we rely, instead, on the *districts under prefecture-level cities*, which capture the urban part of the country, thus making them comparable to MSAs. In Mexico, we also use metropolitan areas, as defined by the national statistical institute (INEGI).

Once we have measures of these three determinants for each city in a country, we can do counterfactual policy exercises. For example, we ask ourselves what would happen if efficiency differences across Chinese cities would be reduced to the level found in the United States. Would this reduce or increase the size of China's mega-cities? Would this improve the well-being of Chinese citizens? Which cities would gain and which cities would lose? Given that technological diffusion is likely to mitigate spatial differences in productivity in China over the next decades, these are relevant

Table 1. Variables

Variable	Source
1. United States	
Unit of observations	Metropolitan Statistical Areas 2005-2008
Population	Bureau of Economic Analysis
Income	Bureau of Economic Analysis
Consumption	Bureau of Economic Analysis, American Community Survey
	(constructed, Desmet and Rossi-Hansberg, 2012)
Hours worked	Current Population Survey
	(constructed, Desmet and Rossi-Hansberg, 2012)
Capital	Bureau of Economic Analysis, American Community Survey
	(constructed, Desmet and Rossi-Hansberg, 2012)
2. China	
Unit of observations	Districts under prefecture-level cities 2005
Population	China city statistics
Income	China city statistics
Consumption	China city statistics
	(constructed, Desmet and Rossi-Hansberg, 2012)
Hours worked	2005 1% Population Survey
	(constructed, Desmet and Rossi-Hansberg, 2012)
3. Mexico	
Unit of observations	Metropolitan Areas, 1989, 1994, 2000 and 2005
Population	Mexican census
Income	Encuesta Nacional de Ingresos y Gastos de los Hogares
	(microdata, constructed)
Consumption	Encuesta Nacional de Ingresos y Gastos de los Hogares
	(microdata, constructed)
Hours worked	Encuesta Nacional de Ingresos y Gastos de los Hogares
	(microdata, constructed)

questions for policymakers.

3. THE UNITED STATES AS A BENCHMARK

Large cities must either be highly productive, have attractive amenities or be efficiently run. If not, they would never have grown to the size they are. But because there are congestion costs associated with becoming large, policy often aims to reduce heterogeneity across cities by revamping backward locations. This is done by making productive investments (increasing efficiency), improving their attractiveness as a place to live (increasing amenities), or strengthening local governance (lowering excessive frictions). By analogy with business cycle policy, which aims to smooth shocks over time, regional policy tries to smooth differences across space.

To analyze the impact of such policies, we start by presenting results for the United States. This example will also provide a useful benchmark to compare China and Mexico with. To present an upper bound of the potential effects of spatial smoothing, our counter-factual policy analysis consists of completely shutting down spatial differences in each of the three city characteristics (efficiency, amenities, and excessive frictions) by setting their values to the population weighted average. That is, we want to see what happens to the city-size distribution, to the fate of individual cities, and to overall welfare when we completely eliminate spatial differences in any of the three dimensions. While it is unlikely that any policy is actually able to completely smooth out differences across space, it provides a useful upper bound to what policy could actually achieve.

Figure 1 shows the results. The upper-left panel shows the actual city-size distribution. Each of the other three panels presents the actual and the counter-factual distributions of city sizes when we shut down the spatial variation in one of the city characteristics. In all figures the horizontal axis shows the log of population size and the vertical axis the log of the probability of cities being larger than that size. For a given city size on the horizontal axis it shows on the vertical axis the share of cities that are larger than that size. A steeper slope implies a less dispersed city size distribution, with the smaller cities being larger and the larger cities being smaller. This is a common way of depicting size distributions of cities since it emphasizes the upper tail of the distribution and, perhaps more important, a distribution exhibiting Zipf's law would show up as a straight line with slope of -1 (similar to the one for the U.S. in the upper-left corner of Figure 1).

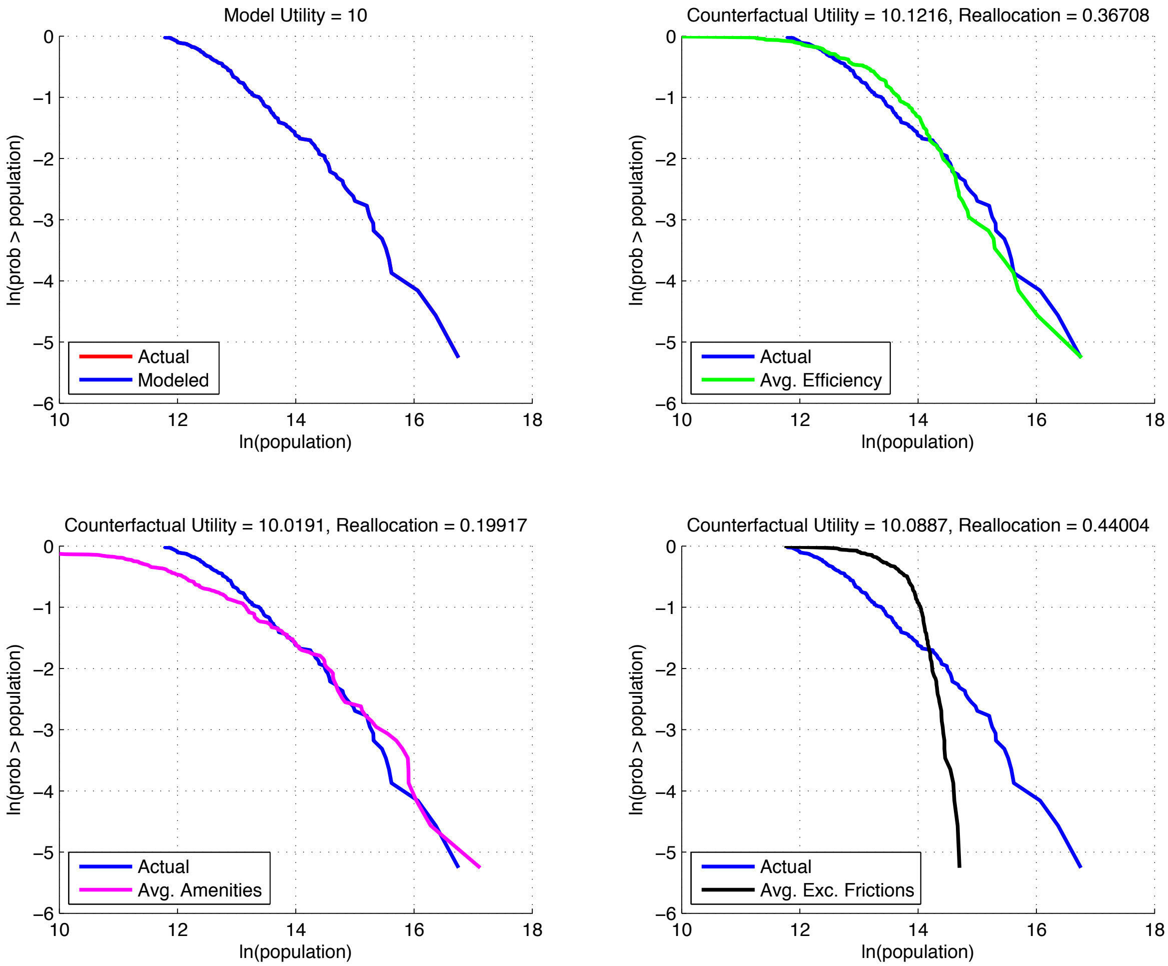

Figure 1: Counter-factuals Without Differences in One City Characteristic (United States)

By comparing the actual distribution with the counter-factual distributions, we notice that efficiency and amenity differences have a limited effect on the city-size distribution, whereas differences in excessive frictions play a much more important role. Indeed, the counter-factual distributions when differences in efficiency or amenities are eliminated hardly change the city-size distribution. That is, if all cities had the same level of efficiency or amenities, the city-size distribution would not change much (except for some cities becoming extremely small). In contrast, if all cities had the same level of excessive frictions, the dispersion in city sizes would become much smaller, making cities much more similar in size than in reality. In the same way that growth accounting decomposes the relative role played by different growth determinants, this exercise of urban accounting can be interpreted as decomposing the relative role played by different city characteristics. What this decomposition shows is that the city-size distribution in the U.S. is mostly due to differences in excessive frictions.

Eliminating differences in city characteristics amounts to smoothing differences across space. In the same way that the business cycle literature analyzes the welfare benefits of smoothing temporal shocks, we can analyze the welfare effects of smoothing spatial shocks. As the numbers at the top of the different panels in Figure 1 indicate, those welfare effects are modest: welfare would increase by 1.2% if all cities had the same efficiency, by 0.2% if all cities had the same amenities, and by 0.9% if all cities had the same excessive frictions. In terms of consumption equivalence, the corresponding figures would be, respectively, 12%, 2% and 9%. Finding positive welfare effects, though modest, is not surprising. Eliminating differences across space tends to spread people more equally across locations. Given that congestion costs increase with size in a convex way, this leads to welfare gains. As mentioned before, from a policy point of view, the modest welfare gains are, if anything, upper bounds. Indeed, completely eliminating differences is probably impossible, given that some of a city's characteristics may be given by nature or geography and thus difficult to change or amend. In addition, our counter-factual exercises assume people can move across locations at no cost.

Since city characteristics may be correlated with size, spatial smoothing can often be reinterpreted as a size-dependent policy, affecting the fate of mega-cities relative to medium-sized and smaller cities. For example, Figure 1 shows that equalizing efficiency makes the city-size distribution slightly less disperse: the larger cities become smaller and the smaller cities become larger. Los Angeles would lose 29% of its population. The respective figures for New York and Chicago would be losses of 77% and 46%. In that sense, a policy that smooths efficiency differences amounts to a size-dependent policy that benefits smaller cities relative to larger cities. The same size-dependency is present when equalizing differences in excessive frictions. In contrast, amenities are less strongly correlated with size. For example, when equalizing amenities across locations, some of the larger cities would lose (Los Angeles and San Diego would lose 8% and 42% of their populations), whereas others would gain (New York and Philadelphia would increase their populations by 44% and 39%).

Although smoothing differences in efficiency and amenities has a limited effect on the city-size distribution (the counter-factual distribution and the actual distribution do not look all that different), the examples of the individual cities mentioned above illustrate that the ordering of cities changes substantially. For instance, when equalizing efficiency across all cities, New York is no longer the country's largest city and drops to 6th position, overtaken by cities such as Riverside, Los Angeles, Chicago and Phoenix. The country's largest city is now Riverside, with a population very similar to

New York's actual population. This reordering of cities implies that behind the veil of an apparently stable city-size distribution, there would be substantial reallocation of population. When calculating reallocation following the methodology of Davis and Haltiwanger (1992) by adding the number of new workers in expanding cities as a proportion of total population, we find a reallocation of 37% when we eliminate differences in efficiency. The corresponding figures when eliminating differences in amenities and excessive frictions are, respectively, 20% and 44%. (These reallocation numbers are given at the top of each panel in all figures.)

Once again it is important to recall that the welfare differences from smoothing are computed under the assumption of free mobility. If one were to take into account the costs of moving and the magnitude of reallocation, it is likely that any modest positive welfare gain from smoothing would vanish and become negative. Indeed, many cities would undergo large shocks, giving people an incentive to move, even when incurring moving costs. This could easily turn the small welfare gains into large welfare losses. From that point of view, regional policies aimed at reducing spatial differences may very well be counterproductive: they would force many people to move, with possibly important welfare costs.

4. COMPARING WITH CHINA AND MEXICO

4.1. China

Reducing spatial dispersion to the level of the U.S.

In the United States we found that even if we were to be able to eliminate all spatial differences in efficiency (or any other determinant of city sizes), the gains in terms of well being would be limited. We now turn our attention to China, a country that has been rapidly urbanizing over the last couple of decades. As is well known, China is characterized by much larger spatial differences than the United States. For example, the productivity of the city at the 80th percentile in China is 71% higher than the city at the 20th percentile. The corresponding figure for the U.S. is a much lower 32%. As China grows further and develops, it is likely that these spatial differences will tend to converge to the ones observed in more mature economies such as the United States. For policy makers who need to make long-term decisions about, for instance, how much and where to invest in infrastructure, it is essential to understand how China's future growth will affect the spatial distribution of its people and its economy. Therefore, rather than analyzing what would happen in China if we completely smoothed out spatial differences, it may be more relevant from a policy point of view to ask ourselves how the Chinese city-size distribution would change if it had the same dispersion in the different city characteristics as the United States.

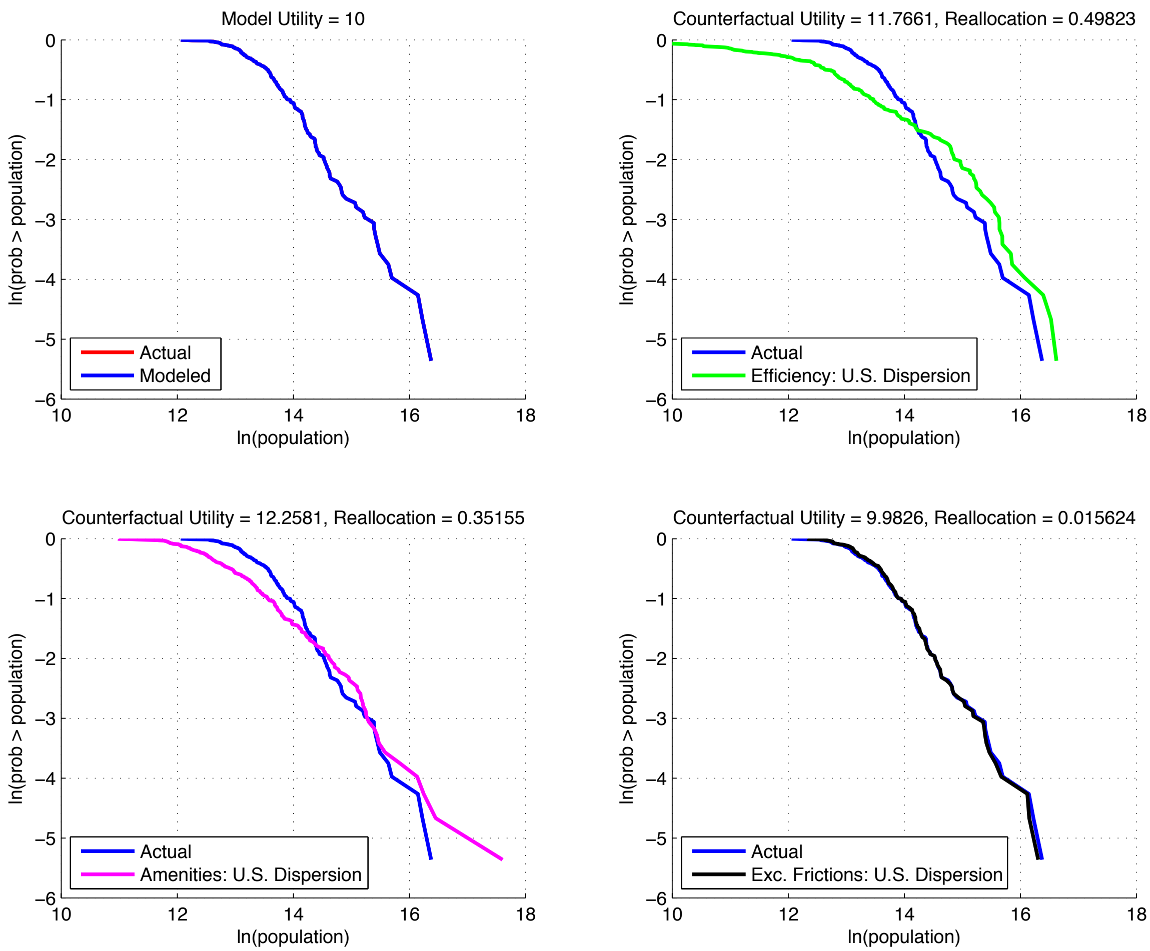

Figure 2: Counter-factuals Changing Dispersion in Characteristics to U.S. Levels (China)

Figure 2 shows the results. To be precise, in the top-right panel, we lower the dispersion in efficiency across Chinese cities to the level of the U.S., without changing the population-weighted average efficiency. In the bottom panels we follow an analogous procedure for amenities and excessive frictions. The welfare effects from reducing spatial differences are huge. If the dispersion in efficiency across Chinese cities were the same as that in the U.S., welfare would increase by 17.7%, and similarly, if the dispersion in amenities would drop to the level of the U.S., welfare would increase by 22.6%.

In terms of consumption equivalence, the respective numbers would be an order of magnitude larger. In contrast to the U.S., these figures suggest that reducing spatial differences could have enormously positive effects on China's well being. This is all the more striking given that in the case of the U.S. we *completely* eliminated spatial differences, whereas in the case of China we *only* reduced spatial differences to the level observed in the U.S. If we had completely smoothed out spatial differences in China, as we did in our counter-factual exercise for the U.S., the welfare effects would be even larger. In the case of efficiency, for example, welfare in China would increase by an astounding 55%, compared to a meager 1.2% in the U.S.

In terms of how reducing spatial differences in efficiency to the level of the U.S. affects China's largest cities, we find that Beijing would lose 87% of its population, and Shanghai a similar 88%. Most of its other large cities would also lose people, with the exception of Chongqing, whose population would grow 46%. Does this imply that the future will bring a demise of China's mega-cities as technology becomes more equally spread across the country? Not necessarily. In fact, Figure 2 seems to indicate the contrary: reducing differences in efficiency would make the city-size distribution more dispersed, implying that China's largest cities would become larger and China's smallest cities would become smaller. Hence, in contrast to the U.S., where smoothing efficiency made cities more equal in size, in China the opposite happens. This unexpected result illustrates the need to analyze policies in a model of a system of cities with multiple determinants of city sizes. Not doing so could lead to unintended consequences: a policy aimed at reducing dispersion across cities might actually increase it.

Given the huge decline in Beijing and Shanghai, it seems counter-intuitive that dispersion in city sizes increases. What happens is that some of the medium-sized cities with high amenities and low efficiency now become the new mega-cities, and they end up being larger than present-day Beijing and Shanghai. The country's three largest cities become Liuan in Anhui province (predicted population 16 million), Chongqing (predicted population 15 million) and Bazhong in Sichuan province (predicted population 13 million). With the exception of Chongqing, which today already qualifies as a mega-city, these are currently medium-large cities with populations between one and two million. But because they have above average amenities and below average productivities, improving their efficiency transforms them into huge cities.

Reducing the spatial differences in amenities to the level of the U.S. also leads to greater dispersion in China's city-size distribution. This is the result of many of the larger cities in China having poor amenities. One probable reason for this are formal and informal migratory restrictions in the past and present. Indeed, if large cities are kept artificially small through mobility restrictions, this will show up as low amenities in our model. Take the case of a highly efficient city with a predicted city size larger than observed in reality. In that case, some other force in our model must keep it from reaching that larger size. That counteracting force will be worse amenities. As a result, many of the highly efficient eastern coastal cities have low amenities. Giving them better amenities would make them grow tremendously. This is the case, for example, of several of the cities of Guangdong province, which constituted some of the first special economic zones under Deng Xiaoping's *Open Door Policy*. Shenzhen would grow to a population of 27 million, and Guangzhou would increase its population by 64%. This finding is in line with Au and Henderson (2006), who argue that China's mega-cities are too small. As is well known, migratory restrictions are not applied always and everywhere in China. The World Development Report 2009 gives the example of the cities of Chongqing and Chengdu, which are pursuing "an unabashedly urbanization-based growth strategy" (p.#221). If those cities are benefiting from government policies promoting rural-urban migration, then this should be reflected in our model as high amenities. Consistent with this, reducing the dispersion in amenities across cities would make Chongqing lose 83% of its population and Chengdu a more modest but still high 46%.

Although mobility restrictions often stem from government policy through the so-called *hukou* system, not all such restrictions are policy-based. Cities take time to grow, housing needs to be constructed, and other urban infrastructure needs to be built. This "time to build" implies that Chinese cities can only gradually converge to their steady-state population level. The city of Shenzhen is telling in this context: while the model predicts that it may be too small, it is unclear which part is due to policy restrictions and which part is due to the "time to build" constraint. Given that it has been China's fastest growing city since 1979 (WDR 2009), it might very well not have been able to grow much faster, even if people had been completely free to move. Other urban problems that afflict many mega-cities, but more so in China than in the U.S., include severe air pollution. Again, in our model such pollution will show up as a negative amenity, making cities such as Beijing less desirable places to reside. In that sense, the low amenities of the larger Chinese cities might also be due to environmental problems.

Overall we find that lowering the spatial dispersion of amenities or efficiency would lead to larger cities *and* huge welfare gains. Greater mobility would surely contribute to narrowing the dispersion in amenities, allowing the larger cities to grow further. Differences in efficiency are also bound to become smaller over time, as technology and efficient management practices diffuse spatially.

Therefore, as China further develops and matures, it is likely to see larger and more mega-cities.

Urban policies

While the previous section focused on the likely spatial evolution of the Chinese economy as it further matures, we now analyze the effects of specific urban policies. One common policy is to improve conditions in lagging locations. Another common policy is to favor growth in cities of certain sizes. In the former case cities get selected on their low productivity, their bad amenities, or their poor governance. An example would be the European Union's structural funds which subsidize local infrastructure and human capital formation in regions with an income per capita below a certain threshold. Another example are the attempts of China to spread development from the highly productive coastal cities to the less productive inland. In the latter case cities get selected on their size. For example, China's urbanization policy in the 1980s and 1990s was based on "controlling the big cities, moderating development of medium-sized cities, encouraging growth of small cities", although in more recent years this policy has gradually been phased out (Kamal-Chaoui et al., 2009).

Table 2. Urban Policies in China

Urban Policies	**Welfare Differences**		
	Efficiency	Amenities	Frictions
Improve in Worst Cities by 20%	4.9%	13.7%	0.2%
Improve in Smallest Cities by 20%	2.0%	4.5%	0.1%
Improve in Medium-Large Cities by 20%	5.2%	8.9%	0.1%
Improve In Largest Cities by 20%	12.1%	22.0%	0.4%

In what follows we analyze the effect of the two types of policies mentioned above, those that benefit lagging cities and those that benefit cities of different sizes. Table 2 reports the results. The first line estimates the effect of a policy that improves conditions in backward cities. If we increase efficiency by 20% in the bottom quartile of cities ranked by efficiency, the model predicts an increase in welfare of 4.9%. Corresponding policies that increase amenities and lower excessive frictions have a positive welfare effect of, respectively, 13.7% and 0.2%. The next lines report the effects of size-based policies that improve conditions in the smallest cities (bottom quartile), the medium-large cities (second quartile) and the the largest cities (top quartile). Two results stand out. First, targeting the worst cities is more effective than targeting the smallest cities. For example, improving amenities by 20% in the smallest cities increases welfare by 4.5%, compared to a much higher 13.7% if the policy focuses on the cities with worst amenities. Second, when comparing the different size-based policies, the welfare effects are greater, the larger the cities that are targeted. This is of course not surprising: improving efficiency (or any of the other characteristics) in the largest cities benefits many more people than the same policy applied to the smallest cities. For example, when increasing efficiency by 20% in the largest cities, with populations above 1.4 million, welfare goes up by 12.1%, compared to a much smaller 2.0% when applying the same policy to the smallest cities, with populations below 0.5 million.

From the point of view of expected benefits, this suggests that policy markers should concentrate their efforts on the country's largest cities. But there is of course also a cost of implementing these policies, and this cost may very well differ by city. It is therefore only if the cost is constant across cities that we can conclusively say that targeting the largest cities is better. For example, if we had two isolated cities, and we had to choose which one to connect by highway to the rest of the country, we would prefer to connect the largest one of the two. But many other policies are likely to have a cost that is increasing in the number of people. For example, building urban infrastructure or improving local schools will be more costly the greater a city's population. Enhancing efficiency in the largest cities will therefore require more resources than doing so in the smallest cities.

Table 3. Urban Policies in China for Constant Population-Weighted Characteristics

Urban Policies	**Welfare Differences**		
	Efficiency	Amenities	Frictions
Improve in Worst Cities by 20%	2.7%	8.2%	0.0%
Improve in Smallest Cities by 20%	0.8%	2.2%	0.0%
Improve in Medium-Large Cities by 20%	1.9%	1.3%	0.0%
Improve In Largest Cities by 20%	0.0%	2.0%	0.1%

To remove the effect that certain policies benefit more people than others (and are therefore likely to be more costly to implement than others), Table 3 analyzes the same urban policies as before, but now keeps the ex-ante population-weighted characteristics constant. In other words, we leave ex-ante aggregate efficiency unchanged, but change the distribution of efficiency across cities, benefiting some cities at the cost of others.[3] We apply the same methodology for amenities and

[3] Of course, given that people will move as a result of the policy, the ex-post aggregate efficiency may be different.

excessive frictions. As can be seen, policies that target cities on their characteristics, rather than on their size, seem to be much more effective. Implementing such a policy requires a methodology to rank cities on the basis of their different characteristics. Our model provides a framework to do so. Note furthermore that within the class of size-based policies there is no clear picture of which size should be targeted. For example, in the case of amenities, focusing on the smallest cities is more effective, whereas in the case of efficiency, improving conditions in medium-sized cities has a greater impact. Which type of city benefits the most depends on the distribution of the different characteristics across cities of different sizes.

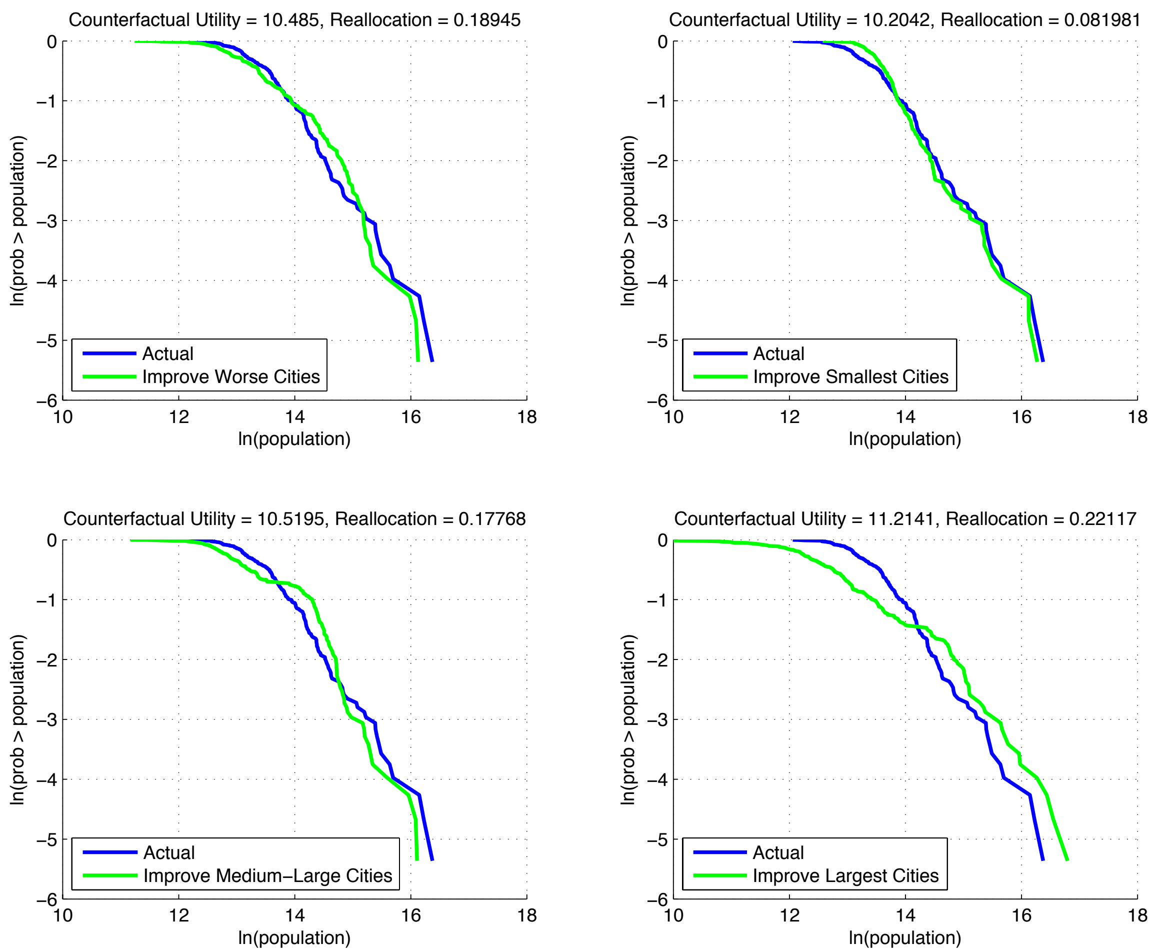

Figure 3: Counter-factuals Different Urban Policies (China)

Beyond their effect on welfare, these policies of course also affect the city size distributions. Figure

3 shows the counterfactual city size distributions when improving efficiency. These correspond to the urban policies reported in Table 2. All policies, except for the one that targets the biggest cities, tend to make the largest cities somewhat smaller. For example, when increasing efficiency by 20% in the bottom quartile of least efficient cities (top-right panel), the country's largest city continues to be Shanghai, but its population is now 21% lower. The other two largest cities, Beijing and Chongqing, would also lose, respectively, 22% and 6% of their population.

4.2. Mexico

Reducing spatial dispersion to the level of the U.S.

After analyzing the U.S. and China, we now look at a large Latin American economy, Mexico. We start by asking the same question as with China: how would Mexico's city-size distribution look like if it had the same dispersion in the different city characteristics as the United States. Figure 4 shows the results. Compared to China, the welfare effects are significantly smaller. If the dispersion in efficiency across Mexican cities were the same as in the U.S., welfare would increase by a mere 0.1%, compared to 17.7% in China. In the case of amenities, the effect would be a substantially larger 0.8%, but still much smaller than the 22.6% in the case of China.

The main reason for the smaller effect is that overall Mexico looks more like the U.S. than China, especially in the case of efficiency. The relative productivity of the city at the 80th percentile in Mexico is 43% higher than the city at the 20th percentile. The corresponding figures for China and the U.S. are, respectively, 71% and 32%. This suggests that the diffusion of technology has been greater in Mexico, reducing spatial differences. In the case of amenities, the smaller welfare effects suggest that Mexico suffers less from mobility restrictions than China. The finding that Mexico's spatial structure is closer to the U.S. than to China is further reinforced when we compare the effects of completely smoothing out spatial differences in the different city characteristics. In the case of efficiency, for example, this would lead to a welfare increase of 0.7% in Mexico. The corresponding figures in the U.S. and China are, respectively, 1.2% and 55%.

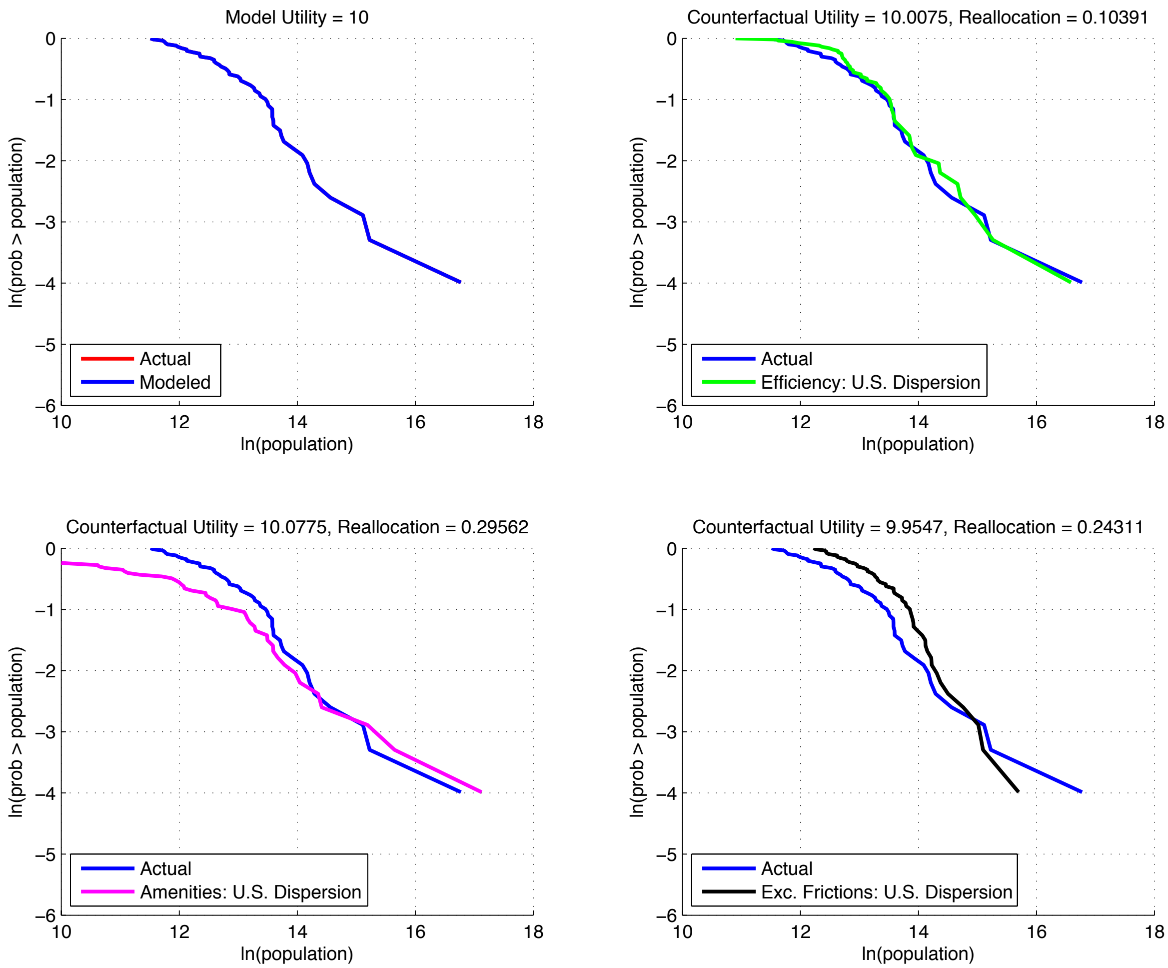

Figure 4: Counter-factuals Changing Dispersion in Characteristics to U.S. Levels (Mexico)

Another relevant question is how reducing spatial differences affects mega-cities relative to medium-sized and smaller cities. In the case of efficiency, we found that reducing differences to the level of the U.S. made the city-size distribution in China more disperse. In Mexico this is not the case: the country's mega-city, Mexico City, loses 17% of its population, dropping from 19.2 million to 16.0 million. This is similar to the U.S., where spatial smoothing also reduced the size of the country's largest city, New York. However, in contrast to the U.S., in Mexico no other city takes the place of Mexico City. Instead, some of the country's intermediate-sized cities now become substantially larger, without reaching the dimensions of Mexico City. This is the case of, for example, Leon, with a population that increases from 1.4 to 2.3 million. Other cases of medium-sized cities gaining population include Puebla (16.5%), Aguascalientes (34.2%) and Acapulco (115.6%). The latter example can be easily understood: Acapulco has one of the country's highest amenities, so that once it improves its efficiency, it grows tremendously.

When reducing the dispersion in amenities across cities, the effect is quite different. In that case, many of the larger cities lose population, such as Guadalajara (-56%), but not Mexico City, which grows from 19.2 million to 27.4 million, and Tijuana, which increases its residents from 1.5 million to 6.3 million. This reflects those latter cities having bad amenities. Once again, it is important to understand that what we call amenities should be interpreted in a broad sense. For example, any feature that holds back city growth but that does not distort the labor supply, shows up as a bad amenity. For example, if pollution and a complex geography puts the brakes on the growth of Mexico City, this will be reflected as a negative amenity.

Urban policies

As in the case of China, we now analyze the effect of two types of policies — those that benefit lagging cities and those that benefit cities of different sizes. Table 4 shows the results for Mexico. If we improve efficiency by 20% in the bottom quartile of least efficient cities, the model predicts welfare in Mexico will increase by 0.4%. Analogous policies that improve amenities would have a positive welfare effect of 0.7%, whereas policies that lower excessive frictions would have no noticeable effect. As for the scale-dependent policies, improving conditions in the smallest cities by 20% would have very small effects, in no case improving welfare by more than 0.1%. Doing the same in larger cities has, as expected, a greater effect. One important difference with China is that the welfare effects of any of these policies are much smaller.

Once again, the finding that targeting the largest cities has a greater impact is not surprising. As in the case of China, if we were to control for the fact that improving conditions in larger cities benefits more people, then we would find that the best policy is the one that targets the most lagging cities. Table 5 reports those results. As can be seen, improving efficiency in the most backward cities improves welfare by 0.2%, compared to 0.1% or less when targeting cities of different sizes.

Mexico's changing city-size distribution

Until now we have compared urban systems across countries, but obviously our methodology can easily be applied to compare urban systems across time. When reducing spatial efficiency differences to those observed in the U.S., welfare gains in Mexico rise from 0.2% in 1989 to 0.6% in 1994, and

Table 4. Urban Policies in Mexico

Urban Policies	**Welfare Differences**		
	Efficiency	Amenities	Frictions
Improve in Worst Cities by 20%	0.4%	0.7%	0.0%
Improve in Smallest Cities by 20%	0.1%	0.1%	0.0%
Improve in Medium-Large Cities by 20%	0.5%	0.3%	0.1%
Improve in Largest Cities by 20%	2.1%	1.1%	0.5%

Table 5. Urban Policies in Mexico for Constant Population-Weighted Characteristics

Urban Policies	**Welfare Differences**		
	Efficiency	Amenities	Frictions
Improve in Worst Cities by 20%	0.2%	0.1%	0.0%
Improve in Smallest Cities by 20%	0.0%	0.0%	0.0%
Improve in Medium-Large Cities by 20%	0.1%	0.1%	0.0%
Improve in Largest Cities by 20%	0.1%	0.1%	0.2%

then drop again, to 0.3% in 2000 and to 0.0% in 2005.[4]

We now interpret this finding in light of the severe economic crisis that hit Mexico in 1994, which had profound effects on efficiency (Meza and Quintin, 2007). As argued before, equalizing efficiency across locations tends to help intermediate-sized cities at the expense of larger cities. We would therefore expect the positive effect on welfare to be larger, the greater the dispersion in efficiency. If so, the temporal pattern suggests that the dispersion in efficiency must have gone up during the economic crisis. This is equivalent to saying that the larger cities are likely to have been less hard hit than the smaller cities. These intuitions are borne out by the data. The standard deviation of the log of efficiency increased from 0.27 in 1989 to 0.32 in 1994 and then started to decline to 0.23 in 2000 and 0.16 in 2005. Furthermore, when comparing the correlation between changes in efficiency over time and city size in Mexico, they are always negative, with the exception of 1989-1994 when the correlation was positive, indicating that larger cities suffered less than smaller cities. This is

[4]The numbers for 2005 are slightly different than the ones in Figure 3. Because we want to use a constant sample of cities across all years, the number of cities analyzed here is a bit lower.

reflected by the counter-factual predictions for Mexico City. We already know that reducing spatial differences in efficiency to the level of the U.S. will have a negative effect on the country's capital, but the effect is much larger in 1994. While the predicted change in Mexico City's population in 1989 and 2005 is, respectively, -11% and 4%, in the crisis year 1994 the predicted change is a much larger -53%.

In addition to analyzing the spatial impact of the crisis, our results also shed light on the long-run spatial development of Mexico's urban system. The drop in the dispersion of efficiency suggests a tendency toward the spatial convergence of efficiency over time. In fact, by 2005 the spatial dispersion of efficiency in Mexico had converged to that of the U.S. Another long-term trend we observe is the worsening in Mexico City's amenities. If the country's capital had the country's average level of amenities, it would have lost 41.7% of its population in 1989. By 2005 the situation was completely reversed: with average amenities, the city would gain 40.0% in population.

5. SOME POLICY CONCLUSIONS

We have applied a simple model of a system of cities to do policy analysis. Our focus has been on the United States and two large emerging economies, China and Mexico. Our findings allow us to state a number of relevant policy conclusions.

First, reducing spatial differences across cities does not necessarily imply larger cities becoming smaller. If efficiency were more equally spread across space, mega-cities would lose in Mexico, but gain in China. The finding that mega-cities may not always decline has to do with the multiple determinants of the city-size distribution and how those relate tho each other. In China many of the intermediate-sized cities have great amenities but low efficiency. Improving their efficiency gives them a huge boost, allowing them to become larger than China's current mega-cities. In other words, Shanghai and Beijing would lose population, but other cities would take their place and outpace them. In contrast, Mexico City would lose population, but remain the country's largest city.

Second, in a mature economy like the U.S., the welfare effects of completely smoothing out spatial differences are small. Since policy is unlikely to actually be able to completely eliminate differences across cities, these already small effects should be interpreted as upper bounds. If, in addition, one were to take into account the reallocation cost of people implied by policies that reduce spatial differences, these small positive welfare effects are bound to become negative. This suggests there is little room for policies that focus on lowering differences across U.S. cities.

Third, emerging economies are typically characterized by larger spatial differences than more developed economies. As China and Mexico continue to grow and develop, it is therefore likely that their spatial differences will converge to those observed in more mature economies such as the United States. Reducing spatial differences in efficiency to the level of the U.S. would improve welfare by 17.7% in China and by a much smaller 0.1% in Mexico. The huge welfare effects in China can be explained by the large spatial disparity in efficiency across its cities, a situation which is largely absent in Mexico. A similar policy that lowers spatial differences in amenities would increase welfare by 22.6% in China and by 0.8% in Mexico. The much larger effect in China is likely to be related to the existing mobility restrictions under the *hukou* system. Taken together, our findings suggest that China stands much to gain from the spatial reallocation of people and economic activity, as technologies diffuse across space and mobility restrictions are lifted. The effects in Mexico are substantially lower, probably reflecting the country having a longer history as a developing market economy and no formal mobility restrictions.

Fourth, urban policies often focus either on improving conditions in lagging locations or on favoring cities of particular sizes. Our results suggest that targeting lagging cities tends to be more effective, except when the cost of implementing such policies does not increase with city size. In that case, targeting the largest cities is better, because the benefits are shared by many more people. The often observed policy in developing countries of favoring smaller cities is in general unwarranted. In order to implement policies based on city characteristics other than size, there is a need for a methodology to rank cities on their different characteristics. Our model provides a framework to do so.

Fifth, country-wide shocks are likely to have important spatial effects. We analyze the impact of Mexico's 1994 crisis, and find that it increased the dispersion of efficiency across cities. This implies that the country's largest cities, such as Mexico City, were spared relative to the country's smaller cities. It suggests that large cities are better able to absorb country-wide shocks than small cities.

Sixth, from a methodological point of view, we believe there is much to be learned from a quantitative analysis of urban systems. Analyzing policy interventions for just one city, without taking into account its effects on other cities, is likely to be misleading. Given that the data requirements to implement this framework are limited, our hope is that this framework will provide a useful tool for policy makers in different countries.

REFERENCES

[1] Au, C.-C. and Henderson, J.V., 2006. "Are Chinese Cities Too Small?," *Review of Economic Studies*, 73, 549-576.

[2] Bai, C.-E., Hsieh, C.-T. and Qian, Y., 2006. "The Return to Capital in China," *Brookings Papers on Economic Activity*, 37, 61-102.

[3] Davis, S.J. and Haltiwanger, J., 1992. "Gross Job Creation, Gross Job Destruction, and Employment Reallocation," *Quarterly Journal of Economics*, 107, 819-63.

[4] Desmet, K. and Rossi-Hansberg, E., 2012. "Urban Accounting and Welfare," *American Economic Review*, forthcoming.

[5] Glaeser, E.L. and Gottlieb, J.D., 2008. "The Economics of Place-Making Policies," *Brookings Papers on Economic Activity*, 39(Spring), 155-253.

[6] Kamal-Chaoui, L., Leman, E. and Rufei, Z., 2009. "Urban Trends and Policy in China", OECD Regional Development Working Papers, 2009/1.

[7] Kehoe, T. and Ruhl, J.R., 2010. "Why Have Economic Reforms in Mexico Not Generated Growth?," *Journal of Economic Literature*, 48, 1005-27.

[8] McGrattan, E. and Prescott, E., 2009. "Unmeasured Investment and the Puzzling U.S. Boom in the 1990s," Research Department Staff Report 369, Federal Reserve Bank of Minneapolis.

[9] Meza, F. and Quintin, E., 2007. "Factor Utilization and the Real Impact of Financial Crises," *B.E. Journal of Macroeconomics*, 7 (Advances), Article 33.

[10] UN-Habitat, 2010. *The State of African Cities 2010: Governance, Inequality and Urban Land Markets.*

[11] Restuccia, D. and Rogerson, R., 2008. "Policy Distortions and Aggregate Productivity with Heterogeneous Plants," *Review of Economic Dynamics*, 11, 707-720.

[12] *World Development Report 2009: Reshaping Economic Geography*, World Bank, Washington D.C.

A. TECHNICAL APPENDIX

A.1. Step-by-Step Guide

The purpose of this section is to explain step-by-step how to implement the model of Desmet and Rossi-Hansberg (2012) with the aim of evaluating urban systems. We will describe how to do this for a given time period t, but it can be repeated for several years, as we do below for Mexico, to compare urban systems over time. In what follows, given that the time period is fixed, we eliminate the subscript t from the notation. In order to simplify the exposition and make this as close to a practical guide as possible, we will not discuss or justify the technical details of the model we use or its general features. The interested reader can consult the original paper, where we discuss this at length.

We divide the methodology into 5 practical steps to be followed sequentially by anyone with a basic knowledge of economics. We use a standard urban model with elastic labor supply so that labor taxes and other frictions create distortions. Cities have heterogeneous productivities and amenities to be determined by the data. Agents live in monocentric cities that require commuting infrastructures that city governments provide by levying labor taxes. City governments can be more or less efficient in the provision of the public infrastructure. We refer to this variation as a city's "excessive frictions."

Step 1: Urban productivity

In order to use the model we need measures of efficiency, amenities, and "excessive frictions" for all the cities included in the analysis. None of these are directly observable, so we need to use the structure of a model to calculate them. Let's start with efficiency. Suppose value-added in a city i is of the form

$$Y_i = A_i K_i^{\theta} H_i^{1-\theta}$$

where A_i denotes city productivity, K_i denotes total capital and H_i denotes total hours worked in the city. The share $1-\theta$ denotes the share of labor in production, which we are assuming is constant across cities and, therefore, also in the aggregate. Hence, it can be obtained from the labor share in the national accounts. In the U.S., and several other countries, $1-\theta$ is approximately $2/3$.

If we have data on Y_i, K_i and H_i we can invert the equation above and obtain $A_i = Y_i / K_i^{\theta} H_i^{1-\theta}$. Data on Y_i is, in general, relatively easy to find. Finding data on K_i and H_i is more problematic.

For H_i one can use household surveys and then aggregate up to the city level, using appropriate weights. Data on K_i are trickier to obtain. If capital markets work well, one can assume a national interest rate r and obtain K_i from the first-order condition of the firm problem, namely, $K_i = \theta Y_t / r$. If national capital markets are presumed to work badly, one would need data on local interest rates r_i to estimate K_i. Finally, one could use data on capital stocks by industry at a more aggregate level and assign capital to cities according to the share of each industry in the city.

- The result of this step is a value of θ and a productivity measure A_i for each city i.

Step 2: Excessive frictions

The next step is to obtain a measure of the excessive frictions in a city. These excessive frictions capture the excessive distortions that result from local governments taxing or distorting markets in order to provide urban infrastructure and services. By "excessive" we mean the frictions over and above what the city size would predict.

Our identification is based on the concept of a "labor wedge," which measures the extent to which the marginal utility of leisure differs from the local wage (which, in turn, is related to the marginal product of labor). Assume log-utility in consumption and leisure of the form

$$\sum_{t=0}^{\infty} \beta^t \left[\log c_i + \psi \log\left(1 - h_i\right) + \gamma_i\right],$$

where c_i denotes an agent's consumption, h_i hours of work as a fraction of available time, and γ_i the utility derived from the amenities in city i. Then, the "labor wedge," τ_i, is implicitly defined by

$$\psi \frac{c_i}{1 - h_i} = \left(1 - \tau_i\right) w_i = \left(1 - \tau_i\right)\left(1 - \theta\right) \frac{Y_i}{H_i}$$

where the second equality comes from the first-order condition of the firm's problem with respect to the labor input.

We can use this equation to calculate τ with the data discussed in Step 1, plus data on consumption per capita at the city level, c_i, and a value for the parameter ψ, which determines the relative value of leisure. A good value for ψ lies between 1 and 1.5, as estimated by McGrattan and Prescott (2009) using aggregate data. Instead of taking the value of ψ from the literature, one could use aggregate data from a particular country, set the aggregate τ to zero, and calculate a country-specific ψ using the above equation. This way of calculating ψ implies that we are measuring the labor wedge relative to the aggregate. Once we have a value for ψ, we need to get data on consumption at the level of cities. One good way to get reasonable estimates of city-level consumption is to use data on retail sales at the city level and multiply them by the ratio of consumption to retail sales in the aggregate. Alternatively, we could use household-level surveys and aggregate up to the city level.

We assume the labor wedge acts like a tax on labor that is used to finance local infrastructure. The need to hire workers to provide infrastructure is assumed to be proportional to the amount of total commuting in the city, which, for the monocentric city model, is given by $TC_i = \frac{2}{3}\pi^{-\frac{1}{2}}N_{it}^{\frac{3}{2}}$. Hence, the government budget constraint requires that

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

where N_i denotes city i's population, κ denotes commuting costs per unit of distance and g_i is our measure of excessive frictions. A higher g_i implies that the city requires more expenditures, and therefore more frictions, to provide infrastructure conditional on city size. Essentially, it is a measure of city governance. We have everything we need to compute g_i for each city except for the parameter κ. We compute this parameter in Step 3 below.

- The result of this step is a value of ψ and a labor wedge τ_i for each city i.

Step 3: Commuting costs

To calibrate the parameter κ we use equation (1) in logs to obtain

$$\ln \tau_i - \frac{1}{2}\ln N_i = \alpha + \ln g_i$$

where $\alpha = \ln\left(\frac{2}{3}\right) + \ln\kappa - \frac{1}{2}\ln\pi$. We can calculate the left-hand side of this equation for each city since we obtained τ_i for each city in Step 2. To identify g_i we impose the condition that $E(\ln g_i) = 0$. In this sense the frictions we calculate are "excessive." They represent the frictions over and above commuting costs and the congestion created by city size. Hence the mean of the left-hand side, $E\left(\ln\tau_i - \frac{1}{2}\ln N_i\right) = \alpha$. We can calculate κ from this equation. After subtracting the mean α from the left-hand side, the resulting residuals are the g_i's.

- The result of this step is a value of κ and a measure of excessive frictions g_i for each city i.

Step 4: Amenities

We calculate amenities as the utility that agents get from living in a given city on top of their consumption and leisure choices. It is given by the term γ_i in the specification of utility discussed in Step 2. The system of cities is in equilibrium if, given the values of γ_i, the number of residents in a city is such that they get the same value of utility as they would get in any alternative city. This value of utility is arbitrary, so let's normalize it to $\bar{u} = 10$. So we are looking for values of γ_i such that the observed population in each city, N_i, gives its residents utility $\bar{u}$. Using some approximations discussed in Desmet and Rossi-Hansberg (2012), this implies that

$$\gamma_i = C_1(\bar{u}) - \log(C_2(A_i, r)) + \kappa \left(\frac{N_i}{\pi}\right)^{\frac{1}{2}} \left(\frac{(1+\psi)}{C_2(A_i, r)} + \frac{2}{3} g_i\right) \tag{2}$$

where $C_1(\bar{u}) = \bar{u} + (1+\psi)\log(1+\psi) - \psi \log \psi$ and $C_2(A_i, r) = (1-\theta) A_{it}^{\frac{1}{1-\theta}} / (r/\theta)^{\frac{\theta}{1-\theta}}$. We already have all the data and parameters used in the equation above from the previous steps, so we can directly use the equation to obtain γ_i.

- The result of this step is a measure of amenities γ_i for each city i.

Step 5: Counter-factual exercises

In Steps 1 to 4 we have computed the relevant parameters of the model together with three city characteristics for each city, namely, (A_i, g_i, γ_i). We can now do counter-factual exercises by changing any of these characteristics and recomputing the equilibrium and the utility $\bar{u}$ associated with the new equilibrium. The difference between 10 and the new counter-factual utility, $\bar{u}^c$, is the utility gained or lost if cities had the counter-factual city characteristics. For example, we can ask what would happen if we replace A_i in each city for the mean of A_i, $E(A_i)$. Such a counter-factual exercise allows us to evaluate a scale-dependent policy in which we improve the productivity of the less productive cities and worsen the productivity of the most efficient cities. This could be implemented by, for example, taxing firms in large cities and subsidizing them in small ones in a way that is budget neutral. The variety of policies that one can evaluate with this methodology is obviously enormous and depends on the interest of policy makers. We illustrate some examples in the next section.

To do the counter-factuals we first determine our counter-factual city characteristics. Denote them by $(A_i^c, g_i^c, \gamma_i^c)$. Thus, for the example above, we would have that $A_i^c = E(A_i)$ for all cities i, while $g_i^c = g_i$, and $\gamma_i^c = \gamma_i$. With the counter-factual characteristics in hand we can solve equation (2) evaluated at the counter-factual city characteristics to obtain counter-factual values of city sizes. These city sizes depend on the value of $\bar{u}$. The counter-factual value of $\bar{u}$ is determined by the value that solves the labor market clearing condition $\bar{N} = \sum_i N_i$, where $\bar{N}$ is the total urban population in the cities we are analyzing. That is, we solve for the value of $\bar{u}^c$ that solves the implicit equation

$$\sum_i \frac{\pi}{\kappa^2}\left(\frac{\log\left(C_2\left(A_i^c,r\right)\right)-C_1\left(\bar{u}^c\right)+\gamma_i^c}{\frac{(1+\psi)}{C_2\left(A_i^c,r\right)}+\frac{2}{3}g_i^c}\right)^2=\bar{N}.$$

Unfortunately, there is no simple algebraic solution to this equation. However, solving for $\bar{u}^c$ can easily be done using a non-linear equation solver or by trying out different values of $\bar{u}^c$ until the equation above is satisfied. With $\bar{u}^c$ in hand we can use the term inside the sum to calculate the counter-factual size of any city i, N_i^c.

So far we have emphasized policies that directly change the distribution of city characteristics. Of course, a similar methodology can be used to do counter-factual exercises where we change the values of particular parameters, like commuting costs, interest rates, or the value of leisure. The choice of counter-factual city characteristics can also be influenced by our prior on the importance of externalities in cities. For example, we can take the view that productivity in a city is in part the result of a city's size. Hence, when we do counter-factual exercises we might want to equalize the exogenous, but not the endogenous, part of productivity. The two are different in the presence of externalities. For example, suppose $A_i = A_i^E N_i^\omega$, where A_i^E is the exogenous productivity of the city. Then we might want to equalize A_i^E and let the resulting productivities across cities be determined by this equation given a value of ω (which has been estimated in the urban literature repeatedly to be around 0.02). Note that by adding externalities we are changing the nature of the counter-factual exercise, but Steps 1 to 4 remain unchanged.

- The result of this step is a triplet of counter-factual city characteristics $(A_i^c, g_i^c, \gamma_i^c)$ for each city i, a counter-factual utility level $\bar{u}^c$, and a counter-factual city size N_i^c for each city i.

A.2. Parameter Values

To implement Steps 1 through 4 in the previous section, we need to get parameter values for the elasticity of substitution between consumption and leisure (ψ), the income share of capital (θ) and the real interest rate (r). For the U.S. our parameter values come from McGrattan and Prescott (2009). For the other countries, we assume that the preference parameter (ψ) is equal to that of the U.S., whereas for the income share of capital and the real interest rates we rely either on local data or on other country-specific studies. Table 6 gives further details.

Table 6. Parameter Values

Parameter	Value	Comments
1. United States		
Elasticity of substitution consumption-leisure ψ	1.4841	McGrattan and Prescott (2009)
Income share of capital θ	0.3358	McGrattan and Prescott (2009)
Real interest rate r	0.02	Standard number in the literature
2. China		
Elasticity of substitution consumption-leisure ψ	1.4841	Same choice as U.S.
Income share of capital θ	0.5221	Bai et al. (2006)
Real interest rate r	0.2008	Bai et al. (2006)
3. Mexico		
Elasticity of substitution consumption-leisure ψ	1.4841	Same choice as U.S.
Income share of capital θ	0.3	Kehoe and Ruhl (2010)
Real interest rate r	0.02	Consistent with K/Y and θ
		in Kehoe and Ruhl (2010)

A.3. Identification of City Characteristics

As explained in Steps 1 through 4, to compute efficiency, amenities and excessive fractions we need data on population, income, consumption, hours worked and possibly (but not necessarily) capital at the level of cities. Table 1 in the main text gives a brief overview of how we collected data on these different variables for the U.S., China, and Mexico.