UNIVERSITY OF WAIKATO

Hamilton
New Zealand

Pareto's Law and City Size in China:

Diverging Patterns in Land and People

Chao Li and John Gibson

Department of Economics

Working Paper in Economics 09/16

October 2016

Corresponding Author

John Gibson

Economics Department
University of Waikato
Private Bag 3105
Hamilton, 3240, New Zealand

Tel: +64 (7) 838 4289

Email: john.gibson@waikato.ac.nz

Chao Li

Economics Department
Auckland University of Technology
Private Bag 92006
Auckland, 1142, New Zealand

Email: chao.li@aut.ac.nz

Abstract

Using Pareto's Law as a benchmark, the very largest cities in China appear to have scope to absorb more migrants, contrary to the pro-small bias in urban policy. We use population census data from 2000 and 2010 and remote sensing data to study the evolution of the size distribution of Chinese cities in terms of land and people. Migrants without local *hukou* registration increasingly congregate in a few larger cities, so previous studies that rely on the count of local *hukou* holders wrongly make the city size distribution seem more even. Temporal comparisons show the city size distribution is diverging in terms of the urban resident population but converging in terms of land area. These divergent patterns suggest that growth in the resident population of large cities is not being assisted by fast enough area expansion, while area expansion of less populous cities is too fast for their slow growth in resident numbers.

Keywords
agglomeration
city size
hukou
migration
Pareto's law
China

JEL Codes
R12, O15

Acknowledgements

1. INTRODUCTION

China's urban population is forecast to be one billion by 2030; an increase of 350 million from 2010 (MGI, 2009). This urban expansion may occur by existing big cities joining Shanghai as a mega-city but also by China's central and local governments growing new cities in currently less urbanized areas. There may be quite different effects on wages and productivity, house prices, land and water use, food security, and environmental stress of taking one path versus the other. The scale of the required urban expansion, and China's need to import resources, will likely see these effects spilling into global markets. Consequently, there is debate about whether small and medium-sized cities in China should be favored over expansion of big cities.

Efforts to limit China's big cities have a long history. The 1990 'City Planning Law' (*Zhonghua Renmin Gongheguo Chengshi Guihua Fa*) mandated 'strictly controlling the size of large cities and developing medium-sized and small cities' (Xu 2009). New cities in this era were often just counties with new labels. This unsuccessful experiment of creating cities was ended in the late 1990s (Fan *et al*. 2012). More even-handed policy followed, with the Tenth Five-Year Plan (2001-2005) seeking balanced development of large, medium-sized, and small cities and the Eleventh Five-Year Plan (2006-2010) emphasizing development of metropolitan regions. In line with this balanced approach, the 'Urban and Rural Planning Law' (*Zhonghua Renmin Gongheguo Chengxiang Guihua Fa*) of 2008 dropped the key phrase 'strictly controlling the size of large cities' that had been part of the 1990 'City Planning Law' (Fan *et al.* 2012).

The policy pendulum is now moving against the biggest cities. In 2014 President Xi Jinping announced reforms to assist rural migrants into small towns but restrict access to bigger cities: '…the overall principle is to fully remove *hukou* restrictions in towns and small cities, gradually ease restrictions in medium-sized cities, set reasonable conditions for settling in big cities, and strictly control the population of megacities'.[1] Big city growth also may be limited by land use controls. Citing food security concerns, land on the outskirts of the biggest cities like Beijing and Shanghai is being classified as 'permanent basic farmland' to be used only for cultivation. In announcing these controls the Minister for Land and Resources claimed that good farmland has been 'eaten by steel and cement'.[2] Conversely, land use controls for small urban areas are less strictly enforced and fiscal decentralization creates incentives for local officials to convert more farmland to industrial or residential use than is actually needed (Lichtenberg and Ding 2009).

---

[1] A report on the speech is here:
http://english.peopledaily.com.cn/n/2014/0607/c90785-8738238.html

[2] Details are in *Xinhua* 2014-11-04:
http://news.xinhuanet.com/english/china/2014-11/03/c_133763130.htm

What is missing in this swing back to a pro-small bias is consideration of the evidence on China's evolving urban system. Using Pareto's Law as a benchmark, we find the very largest cities in China have scope to absorb more migrants, contrary to the pro-small bias in urban policy. One of the most robust empirical facts about the relative size of cities in market economies is that they follow either Pareto's or Zipf's Law (Gabaix 1999). [3] While economists and demographers tend to apply these rank-size laws to population, new research that builds a statistical representation of cities from the bottom up shows that Zipf's Law holds both for population and for area, to a good approximation, in Great Britain and the United States (Rozenfeld *et al.* 2011).

The Pareto Law negative relationship between logarithms of city size (in terms of people or land) and city rank is used here to identify the cities with too many migrants and those with scope to take more. We also use changes in these relationships over time to contrast trends in the distribution of city area and city population. A key feature of our analysis is use of census data from 2000 and 2010; previous studies may mislead because they use annual data that measure how many people have *hukou* household registration for each place, not who lives there. [4] The census counts show that non-*hukou* migrants increasingly congregate in a few larger cities; for example, there are urban districts in each of China's 287 prefectures but just 27 are home to 71 million of the 117 million non-*hukou* migrants residing in urban districts in the 2010 census. [5] Thus if cities are measured by the count of local *hukou* holders, as in prior studies, the size distribution wrongly seems more even because it ignores the funneling of migrants into a few big cities.

The claim by prior studies that, over time, China's cities became more evenly distributed in terms of population size, may not be reliable because those studies do not use data that count people where they actually live. In fact, contrary to existing claims, the census counts of residents in 2000 and 2010 show that the Pareto coefficient is falling, which implies a less even distribution (larger changes in city size are needed to change city rank), and is moving closer to the unitary value implied by Zipf's Law. [6] Yet when we study city area, whether measured more finely using Landsat or more coarsely using night-time lights, there

---

[3]   Zipf's Law is a special case of the Pareto distribution, with a Pareto exponent equal to one.

[4]   Examples include Anderson and Ge (2005), Liang (2010), Peng (2010), Chen *et al* (2013) and Li and Sui (2013).

[5]   Non-*hukou* migrants are people who move somewhere other than where their *hukou* registration is from without converting either their type of *hukou* (agricultural or non-agricultural) or their place of registration (*hukou suozaidi*). The problems for the interpretation of China's statistics due to these migrants are discussed in Li and Gibson (2013).

[6]   The same trend is apparent using the non-agricultural *hukou* registered population for those years, but the Pareto coefficients using the urban resident population are always closer to what Zipf's Law implies.

is a clear trend for the Pareto coefficients to be rising, implying that cities are becoming more equally sized in area. These divergent trends may reflect a mismatch between migrants funneling into a few large cities and governments trying to steer them into smaller cities; one aspect of smaller cities expanding in area is 'ghost towns' where empty new housing units sit on recently converted farmland.[7]

Another problem with prior studies is that county-level cities (*xianji shi*) are often included in samples along with urban districts (*shiqu*). But some county-level cities are just relabeled counties and do not differ from rural counties in economic performance (Li 2011 and Fan *et al.* 2012). More reason for doubt about county-level cities is that they lack urbanization externalities (as do counties), which are only found in urban districts (Li and Gibson, 2014b). Thus a focus of some studies on the apparent growth of small cities (for example, Anderson and Ge 2005) may be misplaced since county-level cities should be excluded (as is done here) when studying Pareto's Law.

Another set of studies that may give a misleading picture of China's evolving urban system are based on employment statistics. Some of these, such as Au and Henderson (2006a), come to similar conclusions as the current study; China's big cities are too small to reap all agglomeration economies, but the data they use are not reliable. For example, Au and Henderson (2006a) estimate an inverted U-shaped relationship between output per worker and city scale, in terms of 1997 employment, but the *Yearbook* data that they use excludes most private sector workers. Similarly, Au and Henderson (2006b) use employment data in urban yearbooks from 1991 to 1998 to estimate the effect of city scale (employment) on per worker productivity and to simulate the effect of doubling urban agglomerations, where this increased scale would follow from relaxing migration restrictions. It surely is true that the restrictions lower productivity, but the yearbook employment data that is used is only for subsets of total employment in 'directly reporting industrial enterprises' whose share of total employment varies from about two-thirds in some years to 40 percent in others. Holz (2013) shows that all-sector employment, which should be a more accurate measure of city scale, is only available in the decadal population census.

In the next section we describe our data and methods, paying attention to the restrictions on the sample (based on size thresholds and data availability). In Section 3 we report several results: (i) measuring cities by how many *hukou* registrations they have makes cities seem more evenly sized; (ii) the city size distribution is becoming less even over time in terms of population but more even over time in terms of land area; (iii) it is moves away from the middle of the population size distribution that are causing the divergence in city size; and, (iv) very large cities appear to have the most scope for absorbing more migrants. Our conclusions are in Section 4.

---

[7] For example, *China Daily* June 10, 2010:
http://www.chinadaily.com.cn/china/2010-06/10/content_9958431.htm

2. DATA AND MODEL SPECIFICATION

A review of data quality issues showed the most reliable information on city population is from the Population Census in 2010 and 2000. The census counts residents of an area as those living there at least six months, giving a more realistic measure of a city's size than the count of people whose *hukou* registration is from that place but who may live elsewhere. Existing studies rely on the *China City Statistical Yearbook* (NBS, 2011), which ignores non-*hukou* residents. When we need *hukou* counts, to contrast with results using counts of residents, we take them from the Ministry of Public Security (MPS, 2001, 2011). A further problem with *Yearbook* data is that they do not report on urban cores while the census reports on each individual district within a prefectural city, and contiguous districts are the best proxy for an urban core in China (Roberts *et al* 2012).[8]

We start with urban districts of all 287 prefectural cities in the 2010 census. Amongst these are 24 that were classified as Leagues, Regions or Autonomous Prefectures in 2000. To keep the same geographical coverage we treat these as if they were prefectural cities in 2000. The size thresholds set on the estimation samples (see below) exclude 20 of these 24, so our inclusive approach to treating them as cores of prefectural cities in 2000 should not matter.

Our interest is in using Ordinary Least Squares (OLS) to estimate:

$$\ln(Rank) = \alpha - \beta \ln(Size) + \varepsilon \qquad (1)$$

with $\hat{\beta}$ the Pareto exponent, ε a random error and *Size* and *Rank* may be in terms of people or land. The special case of $\hat{\beta} = 1$ is Zipf's law. Prior studies for China mainly use this specification (for example, Anderson and Ge 2005, Liang 2010 and Li and Sui 2013) or else a specification that shifts city ranks by 0.5 but gives similar results (Xu and Zhu, 2009; Chen et al, 2013). When considering population we estimate equation (1) with four measures of city size and city rank: the non-agricultural *hukou* population (NA) and the urban resident population (U) for 2000, and for 2010. Comparing results for NA and U helps to assess possible bias in prior studies that only use NA to measure city size. Comparing 2000 and 2010 shows if the city size distribution is converging. For city area we use remote sensing data from Landsat, which is more precise but available for limited years, and night-time lights

---

[8]   The urban core of a prefectural city is made up of adjacent districts (*shiqu*). The exceptions in our data are ten districts of Chongqing (Puling, Wansheng, Shuangqiao, Changshou, Jiangjin, Hechuan, Yongchuan, Nanchuan, Wanzhou and Qianjiang) excluded due to being largely non-urbanized and only recently upgraded from county-level city or county status, plus four districts of Wuhan (Caidian, Jianxia, Huangpi and Xinzhou) and one from Kunming (Dongchuan) that are similar to county-level cities or counties.

observed by the Defense Meteorological Satellite Program (DMSP) which give coarser resolution annual measures, and are used for China's cities by Gibson et al (2014).[9]

A typical pattern if predictions from equation (1) are compared with a scatter plot is for the lower tail to flatten out due to 'cities' too small to distinguish from rural areas (Brakman *et al.* 1999). Studies set lower thresholds to exclude these small cities (Giesen *et al.* 2010). For example, thresholds for China range from 80,000 (Xu and Zhu 2009) or 100,000 (Anderson and Ge 2005 and Liang 2010) to 200,000 and 500,000 (Chen *et al.* 2013), and also use relative values such as the smallest city in the top 70% of cities (Li and Sui 2013) and rolling sample approaches that constantly change the threshold (Peng 2010). Thresholds such as 80,000 once coincided with official city size definitions but are less relevant now due to growth in average city size.

Our approach is to set a threshold for the most reliable measure (the urban resident count) and hold constant the proportion of cities below that threshold in the other samples. A threshold of 0.3 million for U in 2010 excludes $n=36$ (one-eighth) of the total sample (only 2% of urban residents are in these small cities). In order to also drop the smallest one-eighth of cities for each of the other three samples, thresholds of 0.204 million for urban residents in 2000, and 0.147 and 0.197 million for the non-agricultural *hukou* in 2000 and 2010 are used. These four estimation samples, each of $n=251$, are used when we focus just on trends in city size in terms of people.

Fewer cities have their area separately distinguished, since adjacent cities that are separated by administrative boundaries may agglomerate into a single unit when viewed from space. This clumping together is especially apparent for cities in the Pearl River delta and Yangtze delta (Gibson et al, 2014). Any cities that join together at any time from 1992-2012, according to night lights, are treated as a single unit for all years, by merging the land area and population data for the separate cities covered by the agglomeration. After making these merges, and also dropping any cities with lit area less than one square kilometer (the spatial resolution of the DMSP data) a sample of 205 cities was available. This is the sample used when studying trends in the distribution of city area, and when comparing trends in the Pareto coefficients for land and people.

Comparing Pareto exponents shows the dynamics of the city size distribution and if size is converging but does not show where change occurs. Kernel density plots of relative city size can show this, and these are presented for all four population samples. The Markov

---

[9] The two sources of remote sensing data on city size are highly correlated, with a correlation coefficient of 0.86 for comparing Landsat and night light-derived measures of city area in 2000. The other source of information on city area is from *Yearbook* reports of built-up area, but comparisons with remote sensing data show this is unreliable (Gibson *et al* 2014, Liu *et al.* 2012) since local governments have an incentive to under-report land conversions.

transition matrix is another nonparametric approach that we use, dividing cities in the four population samples into six groups defined by cut-points at 0.4, 0.6, 1.2, 2, and 4 times the population of the average city in a particular sample.[10] The higher the Markov transition probability for moving into a new group between 2000 and 2010, the less stable are cities in the original size range.

We also consider deviations of actual city size from the size predicted by Pareto's Law, given a particular city rank:

$$\widehat{Size} = \exp\left(\frac{\ln(Rank) - \hat{\alpha}}{\hat{\beta}}\right) \tag{2}$$

where equation (2) is applied to each of our four population samples. These deviations show where, from the standpoint of Pareto's Law, cities are made 'too large' or 'too small' by either policy (showing up when using the NA *hukou* count to measure cities) or by the choices of migrants (showing up when using the urban resident count). We also see where these 'oversized' or 'undersized' cities move in the distribution over time and consider what may cause these moves.

For example, if small cities are made too large by policy biases, it should show up when using the NA *hukou* count, and the Kernel density and Markov transition matrix would show movement of small cities into the middle-sized groups. This would be a case of planning causing a more even city size distribution, as suggested by existing studies that show Pareto exponents rising over time. This is attributed either to the growth of small cities (for example, Anderson and Ge 2005 and Xu and Zhu 2009) or to restrictions on large cities (for example, Chen *et al.* 2013 and Li and Sui 2013). But if this pattern is not apparent when using data on the urban resident population, it suggests a statistical artefact in prior studies due to them measuring city size in terms of *hukou* registrations rather than by the count of how many people actually live in each city.

The comparison between planned city sizes (based on the non-agricultural *hukou*) and the distribution of city sizes resulting from decisions of migrants can also be examined more directly. The number of non-*hukou* migrants (M) in each city in 2010 can be calculated as:

$$M\,2010 = U\,2010 - NA\,2010 \tag{3}$$

---

[10] For the urban resident population (*U*) in the 2010 census, these relative cut-points correspond to 0.56 million (m), 0.84m, 1.69m, 2.80m, and 5.63m. These ranges are similar to those announced by the *State Council* in 2014 for adjusting standards for categorizing city sizes; 0.5 million, 1m, 3m, 5m, and 10 million except that the very small cities are divided into two groups with another threshold at 0.2 million in the *State Council* guidelines.

and we can compare the actual stock of incoming and outgoing migrants in 2010 with the movement of people needed in order for the city size distribution to exactly follow Pareto's Law. That is, we can use equation (2) to calculate the predicted size for each city in terms of both urban residents and non-agricultural *hukou* holders and then examine hypothetical flows based on the deviation of the two actual population series from these two predictions:[11]

$$\widehat{M\,2010} = \widehat{Size} - Size \tag{4}$$

The comparison of the actual stock of migrants with the hypothetical number that would hold under Pareto's Law can identify whether it is large or small cities that have already taken in enough migrants (for their rank) and what type of cities can potentially take in more.

3. RESULTS

The results of estimating equation (1) on the four population samples are reported in Table 1, with the raw data and fitted trends shown in Figure 1. Measuring city size by the number of *hukou* registrations (NA) makes Pareto exponents significantly larger than if cities are measured by their number of residents.[12] Larger Pareto exponents imply a more even distribution of city sizes (since small changes in city size are associated with larger changes in rank). Intuitively, ignoring non-*hukou* residents, as existing studies have done, leads one to miss the fact that many migrants congregate into a few large cities, and being blind to this pattern wrongly implies a more even city size distribution. But even with the smaller Pareto coefficients for city size measured in terms of residents, the null hypothesis of $\beta = 1$ is rejected so Zipf's Law does not hold exactly. Likewise, prior studies for China find evidence against the parallel growth of cities (Chen *et al.* 2013).

The second finding in Table 1 is that Pareto exponents fell significantly from 2000 to 2010, regardless of whether city population is measured by counting residents or by counting the number of non-agricultural *hukou* registrations. Thus, the population size distribution of China's cities became less even over the decade and this divergence is not something noted in the prior literature. Indeed, some studies claim the reverse, of a convergence in the city size distribution (for example, Anderson and Ge 2005, Xu and Zhu 2009).

While Pareto coefficients are getting smaller over time for city size in terms of population, they are getting larger for city size in terms of land area (Table 2). The Pareto coefficient for city area measured by night lights rose from 0.71 in 1995 to 0.94 in 2010.[13]

---

[11] This comparison is restricted to the *n*=244 cities common to the NA 2010 and U 2010 samples.

[12] Equality of the Pareto exponents using NA versus U is rejected at *p*=0.000 for both 2000 and 2010.

[13] The Landsat data are only available for 1995, 2000 and 2008 (we are grateful to Dr Xiangzheng Deng for these data) so we also use the DMSP data for the same years, and for 2010 to compare with the population estimates.

The distribution of city area according to Landsat always more closely follows Zipf's Law (results in the top right-hand panel of Table 2 show no rejection of the hypothesis that $\beta = 1$), but the Pareto coefficients also rise over time. Yet when the same sample of 205 cities is used, based on aggregating the population of nearby cities whose lights merge into one agglomeration, the Pareto coefficients on city size according to population decline over time (bottom panel of Table 2) in line with what Table 1 shows with a larger sample. Thus there are contrasting changes in the distribution of city size according to land and people, with cities becoming more equal in area and less equal in population.

Table 1: Rank-Size Regressions
for Urban Resident, and Non-Agricultural Registered Population 2000 and 2010

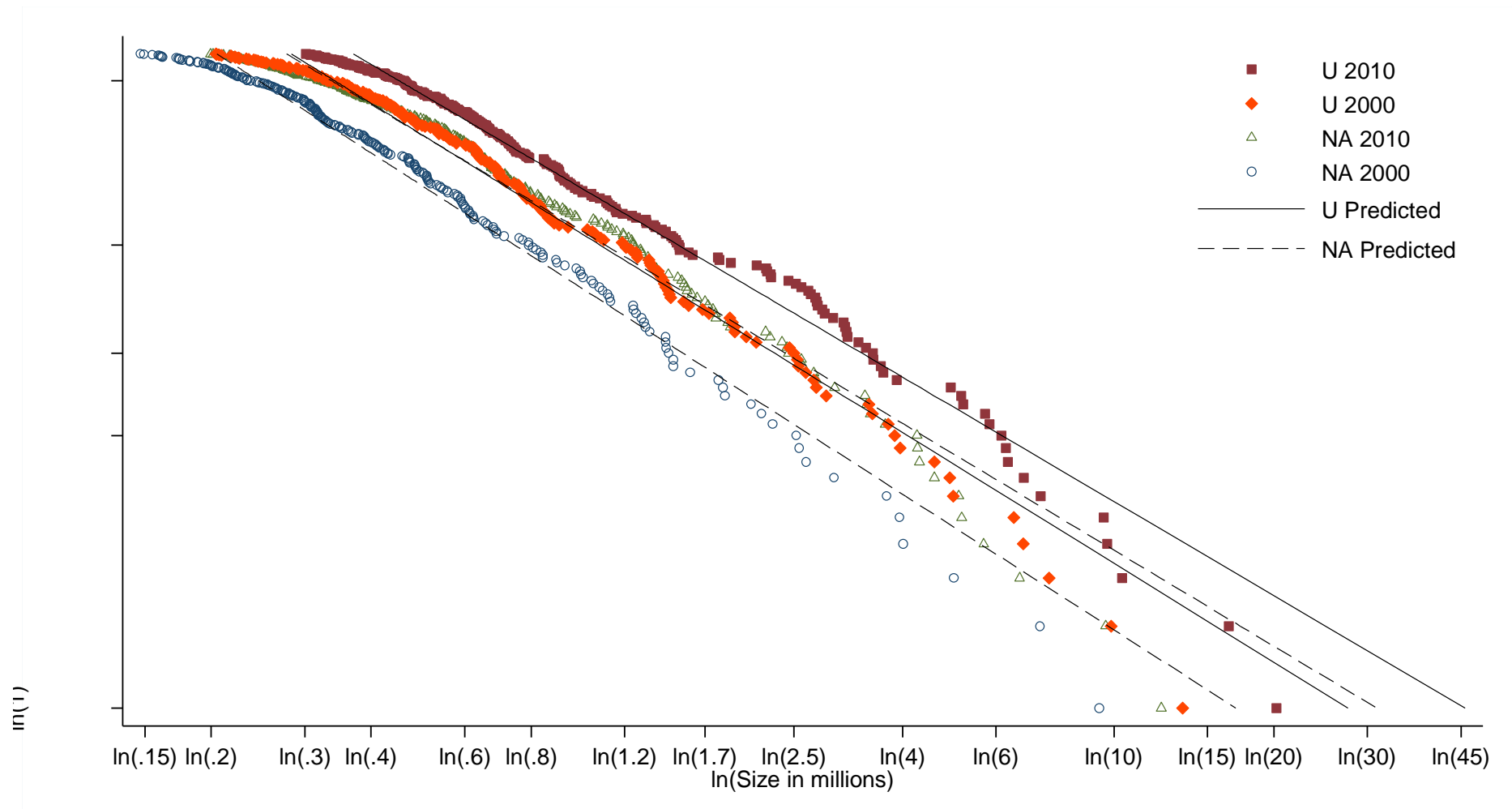|	(1) OLS	(2) OLS	(3) OLS	(4) OLS
	Ln (Rank NA 2000)	Ln (Rank U 2000)	Ln (Rank NA 2010)	Ln (Rank U 2010)
Ln(NA 2000)	1.251 (0.011)***			
Ln(U 2000)		1.207 (0.011)***		
Ln(NA 2010)			1.169 (0.013)***	
Ln(U 2010)				1.147 (0.009)***
Constant	3.542 (0.012)***	4.002 (0.010)***	4.026 (0.012)***	4.386 (0.008)***
Observations	251	251	251	251
Adjusted R^2	0.98	0.98	0.97	0.98

Notes
Standard errors are in parentheses. *** Significantly different from 0 (constant) or 1 (for the Pareto exponent) at p=0.01 confidence level. NA is the non-agricultural *hukou* registered population and U is the urban resident population. See Appendix Table 1 for details.

Regressions cannot show where the change over time in the distribution of city size by population occurs but the kernel densities in Figure 2 give some clues, especially if using the urban resident population (we contend this is the most correct measure of city size). In particular, it seems that between 2000 and 2010 there was a movement away from the middle of the distribution, with a fall in the proportion of cities that are from 0.4 to two times the mean size, and rises in the proportions that are either smaller than 0.4 of the mean or larger than twice the mean.[14]

---

[14] This contrasts with Anderson and Ge (2005) whose Kernel density plot shows a rise in the number of cities in the center of the distribution. But their sample mixes together county-level cities, which often are not very urbanized, with genuine urban cores (districts).

Figure 1: Rank-Size Plot

Legend:
- ■ U 2010
- ◆ U 2000
- △ NA 2010
- ○ NA 2000
- —— U Predicted
- – – – NA Predicted

X-axis: ln(Size in millions) — ln(.15) ln(.2) ln(.3) ln(.4) ln(.6) ln(.8) ln(1.2) ln(1.7) ln(2.5) ln(4) ln(6) ln(10) ln(15) ln(20) ln(30) ln(45)

Y-axis: ln(1)

Notes: The number of observations is 251 for all four samples (NA 2000, U 2000, NA 2010 and U 2010) whose details are in Appendix Table 1.

11

Figure 2: Relative City-Size Distribution
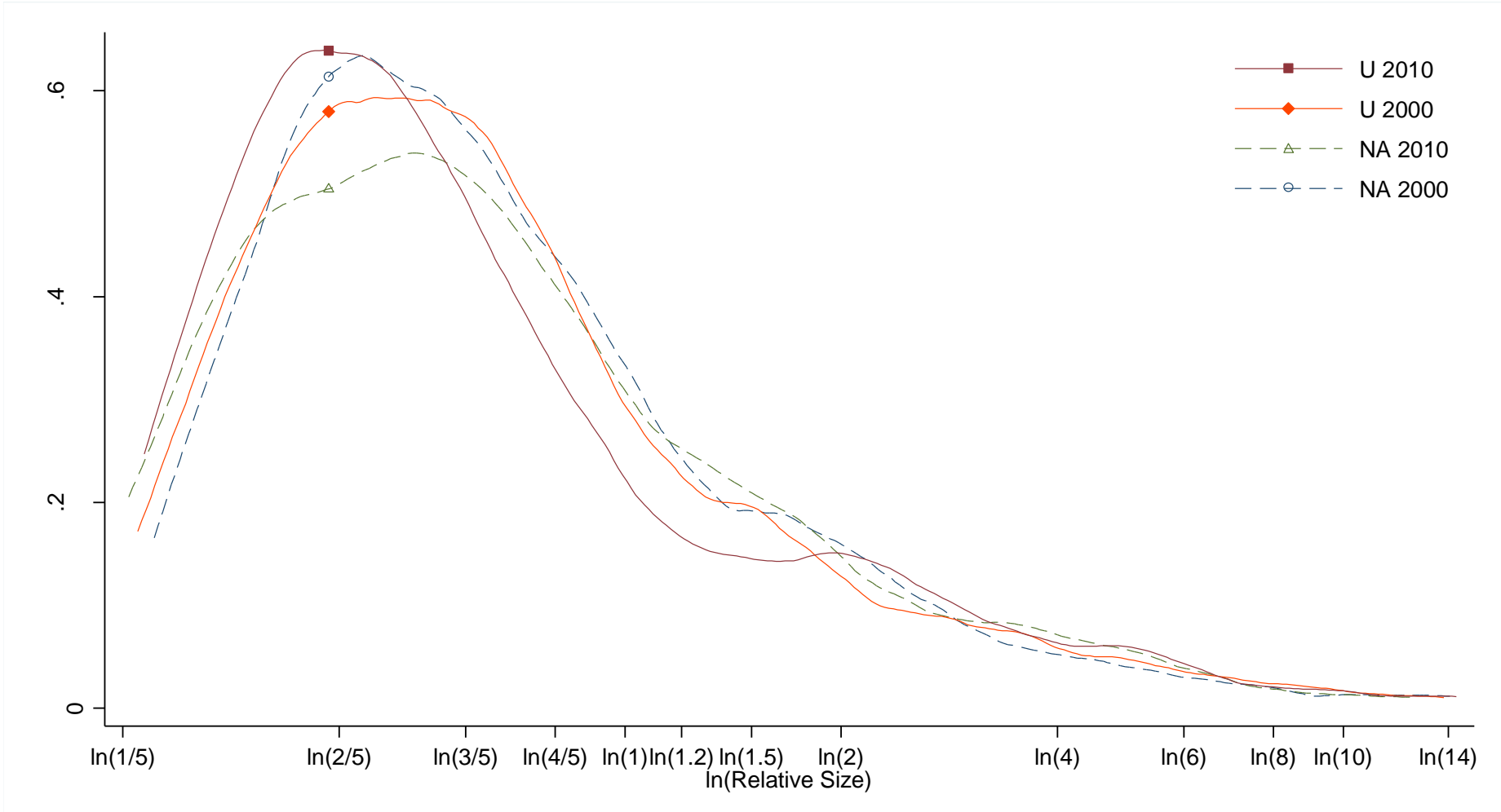
Note
City size is relative to the mean of each distribution, as described in Appendix Table 1.

Table 2: Rank-size Regressions (OLS) using City Areas in 1995, 2000, 2008 and 2010, and City Population in 2000 and 2010

Ln(DMSP)	Ln(Rank DMSP)				Ln(Landsat)	Ln(Rank Landsat)			
	1995	2000	2008	2010		1995	2000	2008	2010
1995	0.705				1995	0.987			
	(0.020)***					(0.028)***			
2000		0.821			2000		1.007		
		(0.015)***					(0.028)***		
2008			0.905		2008			1.017	
			(0.011)***					(0.026)***	
2010				0.943	2010				
				(0.010)***					
Constant	7.149	7.905	8.911	9.574	Constant	8.947	9.083	9.303	
	(0.083)***	(0.069)***	(0.059)***	(0.057)***		(0.134)***	(0.134)***	(0.128)***	
Adjusted R^2	0.860	0.933	0.967	0.978	Adjusted R^2	0.857	0.864	0.883	
Test p-value	0.000	0.000	0.000	0.000	Test p-value	0.654	0.792	0.504	

Ln(NA)	Ln(Rank NA)		Ln(U)	Ln(Rank U)	
	2000	2010		2000	2010
2000	1.056		2000	1.051	
	(0.014)***			(0.012)***	
2010		1.011	2010		1.031
		(0.014)***			(0.010)***
Constant	3.496	3.845	Constant	3.844	4.140
	(0.017)***	(0.014)***		(0.012)***	(0.010)***
Adjusted R^2	0.963	0.963	Adjusted R^2	0.973	0.980
Test p-value	0.000	0.436	Test p-value	0.000	0.003

Notes: Standard errors in parentheses. *** Significantly different from 0 at 1% confidence level. The test *p*-value is for testing the null hypothesis that the Pareto exponent equals 1 (Zipf's Law). DMSP is city area using night lights from satellites F12 (1995), F15 (2000), F16 (2008) and F18 (2010) with a 50% luminosity threshold. NA is the non-agricultural *hukou* population and U is the urban resident population. Number of observations is 205 excluding areas that are missing data or are less than 1 km^2 according to DMSP in all four years 1995, 2000, 2008 and 2010, and includes adjacent cities (according to administrative boundaries) that merged into single agglomerations according to DMSP (and the same merges are applied to data from Landsat, NA and U to maintain spatial consistency).

Further evidence on dispersion away from the medium size range comes from the Markov transition matrices for city population groups reported in Table 3 according to residents (left panel) or non-agricultural *hukou* holders (right panel). Cut-points of 0.4, 0.6, 1.2, 2, and 4 times the mean city size define groups we label as 'small', 'small-medium', 'medium', 'large-medium', 'large', and 'very large' in the table. The small-medium, medium, and large-medium cities have the lowest odds of staying in the same resident size range from 2000 to 2010, with probabilities from 47% to 65%. In contrast, 79% of small cities, and (100%) 86% of (very) large cities stay in the same size range between the two censuses. For cities starting in the middle size ranges and moving into a different group, usually it is a move downwards, indicating a fall in size relative to the mean.

While upward moves are less common they do include the only instances of moves to non-adjacent groups, which are both from Guangdong province; Foshan went from medium size in 2000 to very large in 2010 and Huizhou went from small-medium size to large-medium. The example of Foshan shows the error in using the count of non-agricultural *hukou* holders as a measure of city size; in 2010 there were 3.7 million people with non-agricultural *hukou* registration from Foshan but the census shows almost twice as many (6.8 million) residents.

The obscuring of city size dynamics by using the count of local *hukou* holders to measure city size is seen by comparing the two panels in Table 3. When the *hukou* count is used, there is no clear pattern of lower transition probabilities along the main diagonal, below the very large size group. In contrast, resident population size classes from small-medium to large-medium show a substantial chance of dispersion. In other words, cities in the middle of the distribution had lower odds of staying in the same size groups between 2000 and 2010 than did cities who started in the tails of the distribution.

The different patterns for the two transition matrices likely reflects the fact that the number of non-agricultural *hukou* holders registered in a particular city evolves only slowly over time since *hukou* status is inherited and is hard to convert, and so this indicator misses the rapid changes in size that can come from migrants voting with their feet. For example, an average of 35% of the cities that were large-medium, medium, or small-medium in 2000 fell into the next lower size category by 2010. In contrast, when the size groupings are according to the *hukou* count, just 26% of cities in these size ranges, on average, had fallen into a lower size range by 2010 due to the slower evolution of the spatial pattern of *hukou* registrations.

Table 3: Markov Transition Matrices for City Size Groups

| | Using the Urban Resident Population | | | | | | | Using the Non-agricultural *hukou* Population | | | | | |
| | U 2000 | | | | | | | NA 2000 | | | | | |
U 2010	S	SM	M	LM	L	VL	NA 2010	S	SM	M	LM	L	VL
S	**79.2%**	41.9%					S	**78.5%**	37.3%				
(89)	(57)[a]	(26)					(80)	(51)[a]	(25)[a]				
SM	12.5%	**46.8%**	31.8%				SM	12.3%	**43.3%**	17.1%			
(59)	(9)	(29)	(21)				(49)	(8)	(29)	(12)			
M		9.7%	**65.2%**	32.1%			M	3.1%	13.4%	**71.4%**	24%		
(58)		(6)	(43)	(9)			(67)	(2)	(9)	(50)	(6)		
LM		1.6%	1.5%	**50%**			LM	1.5%	4.5%	10%	**60%**	29.4%	
(16)		(1)	(1)	(14)			(31)	(1)	(3)	(7)	(15)	(5)	
L				17.9%	**85.7%**		L			1.4%	12%	**58.8%**	
(17)				(5)	(12)		(14)			(1)	(3)	(10)	
VL			1.5%		14.3%	**100%**	VL				4%	11.8%	**100%**
(12)			(1)		(2)	(9)	(10)				(1)	(2)	(7)
Total	91.7%	100%	100%	100%	100%	100%	Total	95.4%	98.5%	100%	100%	100%	100%
(251)	(72)	(62)	(66)	(28)	(14)	(9)	(251)	(65)	(67)	(70)	(25)	(17)	(7)

Notes

Transition probability (in %) calculated by dividing the number of cities that move to a size range in 2010 by the total of cities in the range they left in 2000. The number of cities in each cell is in ( ), with 251 cities in total. The abbreviations S, SM, M, LM, L, and VL are for small, small medium, medium, large medium, large and very large, and are based on cut-points of 0.4, 0.6, 1.2, 2, and 4 times the mean city size (as measured by either U or NA, in either 2000 or 2010).

[a] A city that was below the threshold size in 2000 occurs in this size range in 2010 or vice versa. These cities are counted in the column total, but not in the particular cell, to restrict attention to cities that were in the size categories considered here (above the threshold for sample inclusion) in both years.
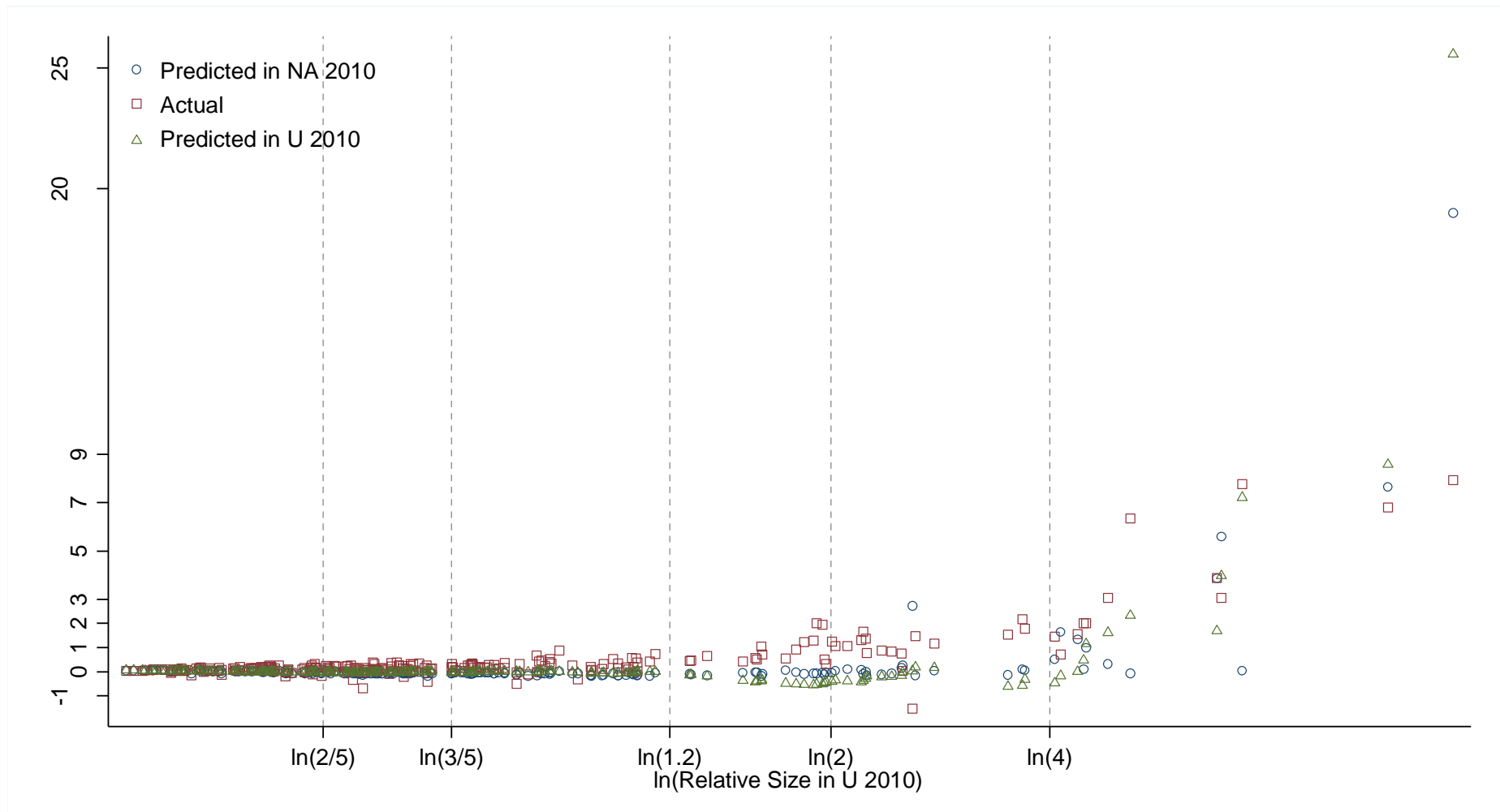
Migrants voting with their feet may change the city size distribution, but by at least one criteria they are still too few, since the largest cities remain undersized in terms of Pareto's Law. Consider the top seven cities by urban resident population in 2010; all are well below the Figure 1 trend line showing the size needed to fit an exact Pareto distribution. The extra people needed to shift the data points for those seven cities on to the trend line is one indicator of how many more migrants it would take to have cities of the right size for their rank. Why do these cities have too few migrants? Some migrants may be deterred by high housing costs, which restricted urban land expansion in the largest cities exacerbates. Li and Gibson (2014) use a hedonic model to compare apartment prices across Chinese cities; for a city like Beijing (ranked 2nd by residents) prices are 230% higher than for a city like Changsha (ranked 27th). It may be that those wanting to move into very large Chinese cities are deterred by the high housing costs and instead go into large and large-medium cities like Changsha. In fact, most cities in Figure 1 of between two million and five million residents (including Changsha) seem larger than what the Pareto distribution would predict (so are above the trend line). The deviation above the trend line for this group of cities was less apparent in 2000, when house prices were more equal.

The destination choices of non-*hukou* migrants, and the scope for cities to absorb more, are shown by applying equations (3) and (4) and comparing the results to the size of each city. In Figure 3 this is done for all 244 cities that are common to the NA and U samples for 2010 (so each city is in the largest 87.5% of cities according to both counts). The stock of non-*hukou* migrants (from equation (3)) ranges from over seven million for Shanghai and Shenzhen (ranked 1st and 3rd by residents) to -1.5 million for Shantou (ranked 18th by residents, but 5th in terms of *hukou* registrations), and these stocks are shown by the red squares in the figure. The hypothetical number of extra non-*hukou* migrants to give an exact Pareto distribution in terms of the number of residents is shown by the green triangle markers, and to give an exact Pareto distribution in terms of non-agricultural *hukou* registrations is shown by the blue circles. These hypothetical values come from equation (4) and they can be negative, which corresponds to cities that were larger in 2010 than what would be predicted from the rank of that city.

This exercise suggests that the scope to absorb more non-*hukou* migrants is mainly limited to the very large cities, defined as those with four or more times the mean number of residents. For example, to achieve the exact Pareto distribution, Shanghai would need to absorb sufficient migrants to get the total resident population to just over 40 million people, making it slightly larger than Tokyo. Similarly, Beijing would need to get to a resident population of almost 30 million, making it slightly larger than Delhi but a little less populous than Jakarta.

Figure 3: Stock of Non-*hukou* Migrants 2010 and Hypothetical Stock Needed for Exact Pareto Distributions versus City Population

When moving from these specific examples, we consider the six size groups (from 'very large' to 'small') that are shown by the vertical lines on Figure 3, with results for these size classes in Table 4. Consider the 12 'very large' cities; in 2010 these were home to 47 million non-*hukou* migrants. For each of those cities their number of urban residents exceeded their non-agricultural *hukou* (so there are no net out-migrants). In the next class of cities, the migrants total 19 million and are funneled into 16 cities. One city in this size range (Shantou) has 1.5 million out-migrants (that is, there are fewer urban residents than people registered with non-agricultural *hukou* from here). The large-medium and medium size classes each hold 14-15 million migrants while the small-medium and small size classes each hold 8-9 million migrants, and these are spread over many cities. The three smallest size classes also include 22 cities that are the source of 3.5 million out-migrants (that is, the number of people with non-agricultural *hukou* from these cities exceeded their number of urban residents in the 2010 census).

Table 4: Stocks of Non-*hukou* Migrants and Hypothetical Number Needed for Exact Pareto Distributions in 2010

| City Size Groups (based on 2010 resident count) | Actual | | Hypothetical Extra Migration to Give an Exact Pareto Distribution | | | |
| | | | Urban Residents 2010 | | Non-agricultural *hukou* 2010 | |
	In	Out	In	Out	In	Out
Very Large	46.58		52.63	-0.64	41.02	-0.07
	(12)		(10)	(2)	(11)	(1)
Large	19.21	-1.52	0.42	-4.41	3.59	-0.71
	(16)	(1)	(3)	(14)	(9)	(8)
Large-Medium	13.63			-6.26	0.05	-1.12
	(16)			(16)	(1)	(15)
Medium	15.12	-0.94	0.38	-0.98	0.02	-4.51
	(53)	(5)	(25)	(33)	(1)	(57)
Small-Medium	8.77	-1.74	0.03	-1.13	0.09	-2.75
	(52)	(7)	(8)	(51)	(10)	(49)
Small	8.28	-0.83	1.31	-0.38	1.92	-0.89
	(72)	(10)	(44)	(38)	(53)	(29)
Total	111.59	-5.04	54.77	-13.80	46.69	-10.05
	(221)	(23)	(90)	(154)	(85)	(159)

Notes

The number of actual and hypothetical migrants is in millions, with the number of cities in ( ). *N*= 244 based on cities common to the samples for NA 2010 and U 2010.

When attention shifts from actual migration to the hypothetical pattern needed to produce exact Pareto distributions for city size and rank, the relevant values are shown in the last four columns of Table 4. Consider the results for urban residents in 2010; an extra 52.6 million migrants could go into ten of the very large cities while 0.6 million could leave the other two very large cities and the resulting size distribution would sit exactly on the trend line shown in Figure 1. The other main change to get an exact Pareto distribution is for the large and large-medium cities to have about ten million fewer migrants (by moving them into the very large size class). This result is just another way of noting the pattern from Figure 1; cities between 1.2 and 4 times the mean city size (in terms of residents these are cities of from two million to five million people in 2010) seem larger than what would be predicted from their rank under an exact Pareto distribution while the very largest cities are smaller than what is predicted. Finally, for the three smallest size groups in Table 4, the extra inward or outward migration needed to get to an exact Pareto distribution never amounts to more than one million people per group.

Less migration is needed to get an exact Pareto distribution for the non-agricultural *hukou* population of each city, requiring 46.7 million coming in and ten million going out. Measuring cities by how many *hukou* registrations they have leads to a seemingly more even distribution than is truly the case since the funneling of most non-*hukou* migrants into just a few host cities is ignored. Furthermore, since the Pareto distribution for the non-*hukou* count is more evenly spread than is the one for the resident count (Table 1 and Figure 1) it takes less movement to get to this target. The second difference between using the *hukou* count and the resident count, in terms of hypothetical migration to get cities the right size for their rank, is a lack of apparent 'queuing' for the very large cities. When cities are measured in terms of residents, the large and large-medium size groups seem to have 'too many' migrants; removing ten million of them and transferring them into the very large cities would give sizes more consistent with the Pareto distribution. But this pattern is not apparent when city size is measured by the *hukou* count, and it is the small-medium and medium sized cities that appear to be the most over-sized (for their rank).

Despite differences between the two sets of results for hypothetical migration, a key point from Table 4 is their similarity in showing that it is only the very large cities with scope to accept many more non-*hukou* migrants, in terms of having city sizes that more closely follow a Pareto distribution. This finding contrasts with the views of leaders such as Xi Jinping and potentially informs about China's evolving urban system. The limited capacity of small and medium sized cities to absorb migrants, either in terms of the actual stock in 2010 or the hypothetical number of extra migrants to get to an exact Pareto distribution, shows that an urbanization process of nearby rural-urban migrants going to live in small, local cities is unlikely to succeed in transforming China into a fully urbanized country. Instead, it is the

agglomeration processes that not only absorb nearby rural-urban migrants but also take in inter-regional rural-urban and urban-urban migrants that are a key to China's urban transition.

Moreover, it is the very large cities, and not the small towns, that provide agglomeration-related productivity advantages. These advantages appear to operate only in the tertiary sector and not in the secondary sector activities like construction and manufacturing that increasingly left the urban districts and moved into smaller towns and counties between 2000 and 2010 (Li and Gibson 2014b). Thus, a focus on directing migrants into small cities will not put them into the places where they are likely to be the most productive. This misallocation will be especially costly as China rebalances the economy by developing the under-sized services sector and reducing reliance on the over-sized manufacturing sector (Ghani 2012), since it is the services sector that benefits the most from locating in larger agglomerations.

A related issue concerns population density, which may contribute to the agglomeration effects discussed above. Some existing studies already note that the density of China's cities is falling relative to comparator cities elsewhere (Du *et al.* 2014) and that urban area expansion can shift from being land saving to land using as patterns of urban development become less dense (Deng *et al.* 2015). The finding of diverging patterns in the Pareto coefficients for land and people that are described here imply that the trends in population density will vary along China's city size distribution.

4. CONCLUSIONS

There is ongoing debate over China's evolving urban system and especially on the policy question of whether small and medium-sized cities should be favored over expansion of big cities. In this chapter we uncover three facts that are missed by prior studies. First, if the population of cities is measured by their number of non-agricultural *hukou* registrations – as in previous studies – Pareto exponents seem larger than they actually are. These prior studies miss the funneling of non-*hukou* migrants into just a few large cities and once this fact is missed, statistical inquiries into Pareto and other distributions tend to find a more even city size distribution than truly exists. This bias is exacerbated by studies that include county-level cities in their samples, despite such 'cities' lacking an urban core (Roberts *et al.* 2012). Second, the population size distribution of China's cities has become less even over time, with movement out of the middle of the distribution between 2000 and 2010. In contrast, previous studies have highlighted an apparent move towards a more even population distribution for cities in China. The final new fact that is revealed here is that while the city size distribution is becoming less even over time in terms of population, it is becoming more even over time in terms of land area.

These three facts are found mainly by regressing city rank on city size, in terms of land and people, where we use both the *de jure* population from the *hukou* registration system and the *de facto* population from resident-based census counts. The Pareto exponents from such regressions provide one benchmark for evaluating an urban system, with the special case of a Pareto coefficient of unity (Zipf's Law) seeming to hold in market economies (Gabaix 1999).[15] Moreover, these relationships are not just empirical regularities; theoretical models of city growth can generate Zipf's Law for both land and people by assuming Cobb-Douglas preferences for goods and housing, random growth with small frictions, and small urbanization externalities (Rozenfeld *et al.* 2011). These relationships appear to hold more for geography-based definitions of cities rather than for legal-based ones, which is consistent with the evidence in Table 2 that Pareto coefficients are closer to the Zipf's Law value of unity when cities were defined according to the measurements from space rather than according to administrative boundaries.

In addition to uncovering these facts we also use the predictions from the estimated Pareto distribution as a normative standard to judge China's city size distribution. We are relying on this as a proxy for what city sizes look like in societies with a less distorted urban past than China. The logic of this exercise is that the parts of China's urban system deviating from a Pareto distribution may be a legacy of the command economy era and restrictions on labour movement and land supply that are less apparent in market economies. When viewed from this normative standard, even though China's largest cities look gigantic relative to other cities (for example, Shanghai with 20 million residents in 2010 is almost double the size of the 3rd largest city and triple the size of the 7th largest) they actually are too small according to the size they ought to be as the top ranked cities in the urban hierarchy. Thus, it is the very large cities that appear to have too few people and that have the most scope for absorbing more migrants in order to become the right size for their rank. In contrast, there is almost no scope for small to medium cities to absorb more migrants when using this benchmark, notwithstanding the policy biases and statements from political leaders in favor of directing migrants towards these small cities.

These normative results suggest that the focus of the Eleventh Five-Year Plan (2006-2010) on the development of metropolitan regions was broadly appropriate, while the current swing towards a pro-small policy bias is not. For example, rather than the recent policy to set land outside of the largest cities aside as 'permanent basic farmland' a better focus for rules on urban land expansion might be to more strictly regulate farmland conversion by small and medium-sized Chinese cities. The finding that the trend in the Pareto exponent for the distribution of city size in terms of land area is moving in the opposite direction to that for

---

[15] Some studies suggest Pareto's Law only fits the upper tail (for example, Eeckhout 2004 and 2009) and propose the lognormal as a better representation of the city size distribution, but this may be because they use legal definitions of cities rather than geography-based definitions.

city size in terms of population supports this suggestion. These divergent patterns suggest that growth in the resident population of large cities is not being assisted by fast enough area expansion, while area expansion of less populous cities is too fast for their slow growth in resident numbers. Also, reforms to governance and public finance that see less tax revenue transferred from local to central government or fewer responsibilities left for local government to fund from their own budgets may reduce the reliance of these smaller cities on revenue from land auctions. With less need for land conversion in small cities, the diverging patterns of cities becoming more equally sized in area and less equally sized in population may be reversed.

REFERENCES

Anderson, G., & Ge, Y. (2005). The size distribution of Chinese cities. *Regional Science and Urban Economics*, *35*(6), 756-776.

Au, C. & Henderson, V. (2006a). Are Chinese cities too small? *Review of Economic Studies*, *73*(3), 549-576.

Au, C. & Henderson, V. (2006b). How migration restrictions limit agglomeration and productivity in China. *Journal of Development Economics*, *80*(2), 350-388.

Brakman, S., Garretsen, H., Van Marrewijk, C., & Van Den Berg, M. (1999). The return of Zipf: Towards a further understanding of the rank-size distribution. *Journal of Regional Science*, *39*(1), 183-213.

Chen, Z., Fu, S., & Zhang, D. (2013). Searching for the parallel growth of cities in China. *Urban Studies*, *50*(10), 2118-2135.

Deng, X., Huang, J., Rozelle, S., Zhang, J., & Li, Z. (2015). Impact of urbanization on cultivated land changes in China. *Land Use Policy*, *45*(1), 1-7.

Du, S., Shi, P., & Van Rompaey, A. (2014). The relationship between urban sprawl and farmland displacement in the Pearl River Delta, China. *Land*, *3*(1), 34-51.

Eeckhout, J. (2004). Gibrat's law for (all) cities. *American Economic Review*, *94*(5), 1429-1451.

Eeckhout, J. (2009). Gibrat's law for (all) cities: reply. *American Economic Review*, *99*(4), 1676-1683.

Fan, S., Li, L., & Zhang, X. (2012). Challenges of creating cities in China: Lessons from a short-lived county-to-city upgrading policy. *Journal of Comparative Economics*, *40*(3), 476-491.

Gabaix, X. (1999). Zipf's law for cities: an explanation. *The Quarterly Journal of Economics*, *114*(3), 739-767.

Ghani, E. (2012). Are China and India converging? Available on-line at: *http://www.voxeu.org/article/are-china-and-india-converging*

Gibson, J., Li, C., & Boe-Gibson, G. (2014). Economic Growth and Expansion of China's Urban Land Area: Evidence from Administrative Data and Night Lights, 1993–2012. *Sustainability*, *6*(11), 7850-7865.

Giesen, K., Zimmermann, A., & Suedekum, J. (2010). The size distribution across all cities–double Pareto lognormal strikes. *Journal of Urban Economics*, *68*(2), 129-137.

Holz, C. (2013). Chinese statistics: classification systems and data sources. *Eurasian Geography and Economics*, *54*(5-6), 532-571.

Li, C. and Gibson, J. (2013). Rising regional inequality in China: fact or artifact? *World Development* 47(1): 16-29.

Li, C. and Gibson, J. (2014). Spatial price differences and inequality in the People's Republic of China: housing market evidence. *Asian Development Review*, *31*(1), 92-120.

Li, C. & Gibson, J. (2014b). Urbanization Economies in China: Nature, Location, and Effects. *Working Paper* No. 14/02, Department of Economics, University of Waikato.

Li, L. (2011). The incentive role of creating 'cities' in China. *China Economic Review*, *22*(1), 172-181.

Li, S., & Sui, D. (2013). Pareto's law and sample size: a case study of China's urban system 1984-2008. *GeoJournal*, *78*, 615-626.

Liang, S. (2010). Research on China's urban development strategy based on arable land preservation. *Journal of Urban Planning and Development*, *137*(3), 329-336.

Lichtenberg, E., & Ding, C. (2009). Local officials as land developers: Urban spatial expansion in China. *Journal of Urban Economics*, *66*(1), 57-64.

Liu, Z., C. He, Q. Zhang, Q. Huang, & Y. Yang (2012). Extracting the dynamics of urban expansion in China using DMSP-OLS nighttime light data from 1992 to 2008. *Landscape and Urban Planning*, *106*(1), 62–72.

McKinsey Global Institute (MGI). (2009). *Preparing for China's Urban Billion*. http://www.mckinsey.com/insights/urbanization/preparing_for_urban_billion_in_china

Ministry of Public Security (MPS). (2001). Population Statistics of the People's Republic of China by County 2000 (*Zhonghua Renmin Gongheguo Quanguo Fenxianshi Renkou Tongji Ziliao 2000 Nian*). Beijing: Qunzhong Press.

Ministry of Public Security (MPS). (2011). Population Statistics of the People's Republic of China by County 2010 (*Zhonghua Renmin Gongheguo Quanguo Fenxianshi Renkou Tongji Ziliao 2010 Nian*). Beijing: Qunzhong Press.

National Bureau of Statistics (NBS). (2003). Tabulation on the 2000 Population Census of the People's Republic of China by County (*Zhongguo 2000 Nian Renkou Pucha Fenxian Ziliao*). Beijing: China Statistics Press.

National Bureau of Statistics (NBS). (2011). China City Statistical Yearbook 2011 (*Zhongguo Chengshi Tongji Nianjian 2011*). Beijing: China Statistics Press.

National Bureau of Statistics (NBS). (2012). Tabulation on the 2010 Population Census of the People's Republic of China by County (*Zhongguo 2010 Nian Renkou Pucha Fenxian Ziliao*). Beijing: China Statistics Press.

Peng, G. (2010). Zipf's law for Chinese cities: Rolling sample regressions. *Physica A: Statistical Mechanics and its Applications*, *389*(18), 3804-3813.

Roberts, M., Deichmann, U., Fingleton, B., & Shi, T. (2012). Evaluating China's road to prosperity: A new economic geography approach. *Regional Science and Urban Economics*, *42*(4), 580-594.

Rozenfeld, H., Rybski, D., Gabaix, X., & Makse, H. (2011). The area and population of cities: New insights from a different perspective on cities. *American Economic Review*, *101*(5), 2205-2225.

Song, S., & Zhang, K. (2002). Urbanization and city size distribution in China. *Urban Studies*, *39*(12), 2317-2327.

Xu, Z. (2009). Productivity and agglomeration economies in Chinese cities. *Comparative Economic Studies*, *51*(3), 284-301.

Xu, Z., & Zhu, N. (2009). City size distribution in China: Are large cities dominant? *Urban Studies*, *46*(10), 2159-2185.

APPENDIX

Table 1: Details on the Four Population Sub-Samples Used in the Estimation

	Sub-Sample Name			
	NA 2000	U 2000	NA 2010	U 2010
Share of population from districts in all prefectures, that are in sub-sample	97.8%	97.9%	98.0%	97.8%
Actual City Size				
Minimum (threshold for inclusion in the sub-sample)	0.147	0.204	0.199	0.301
Median	0.386	0.552	0.571	0.698
Mean (used to divide cities into the six size classes in Tables 2 and 3)	0.666	0.975	0.979	1.407
Maximum	9.382	13.460	12.286	20.218
Predicted City Size (equation (2))				
Minimum	0.205	0.283	0.277	0.370
Median	0.356	0.501	0.500	0.676
Mean	0.722	1.074	1.126	1.571
Maximum	16.953	27.558	31.274	45.794
Stock of non-hukou migrants in 2010 (equation 3)				
Minimum			-1.523	
Median			0.166	
Mean			0.437	
Maximum			7.931	
Hypothetical extra migrants for exact Pareto distribution to hold according to NA or U				
Minimum			-0.176	-0.613
Median			-0.037	-0.008
Mean			0.150	0.168
Maximum			18.988	25.576

Notes

NA is the non-agricultural *hukou* registered population and U is the urban resident population. Each sub-sample has 251 observations, and represents the largest 87.5% of cities for each indicator and each year, except migration calculations that are based on the union of NA and U (*n*=244). Numbers are in millions.

Sources: MPS (2001 and 2011); NBS (2003, 2011and 2012).