

Infill versus Outspill: the microstructure of urban expansion

Stephen Sheppard
Williams College

Abstract

The continuing pace of urbanization around the world has generated an ongoing process of expanding urban land use. This is verified both in everyday experience and in empirical studies such as Burchfeld, Overman, Puga and Turner [2006] or Sheppard, Civco and Angel [2006]. Several factors contribute to the growth in urban land use: for example increasing population and income, falling transportation costs, and changing values of agricultural land surrounding the city. In this paper we investigate how the local structure of urban expansion is affected by these factors. Do these forces encourage increasing urban land use at the periphery, or encourage infill of vacant areas within the city, or both? It is difficult to empirically test our understanding of these processes using data from a single city or even a single country because the data do not offer sufficient independent variation in the relevant factors to reliably distinguish their separate impacts. This paper makes use of newly collected data from a global sample of cities with land use data classified from satellite imagery in all cities. The analysis distinguishes increasing urban land use at the periphery from increasing urban land use through infill of built up areas, and examines the extent to which this process differs around the world.

The financial support of the World Bank and the National Science Foundation (SES-0433278) helped to make this research possible. Special thanks are due to Alison Kraley, whose efforts have been central to this work. Thanks are also due to Solly Angel and Daniel Civco who collaborated on earlier stages of this research.

1. Introduction

The model of urban land use and urban form derived from von Thunen by Alonso (1965), Muth (1968) and Mills (1972) is remarkably simple and widely used. Although essentially partial equilibrium and static in nature (at least in its original form), in some ways the theory must be regarded as a tremendous success in serving as a foundation for policy-oriented as well as commercial analysis. Many of the central predictions of the theory are generally supported: population density and land values tend to increase as we move towards the center of the city; general trends towards lower transport costs appear to have resulted in “flatter” density gradients; etc.

Despite these successes, there remain a variety of puzzles, anomalies or simply empirical regularities observed in real world cities that are either in direct contradiction with the standard model or appear difficult to explain with it. The increasing importance of suburban employment locations, the non-annular arrangement of land uses, and dynamic processes such as gentrification are examples.

Additional examples relate to the timing and spatial location of land development activity. Sometimes identified by planners and others concerned with urban policy as indicative of either irrationality or clever and exploitative speculation, development patterns that leave large patches of vacant land within the built-up area have been the target of policies designed to encourage “infill” and “densification.”

In this paper we provide empirical estimates of the determinants of these different types of development within cities. We base these estimates on data measured from a global sample of cities selected to be representative of the global urban population in places having population in excess of 100,000 persons. For each city, we measure total urban land use at two time periods (using techniques described below), and use these to measure the change in total urban land. For each time period we measure the amount of “internal” open space that is contained within the built up area (land with structures or other impervious urban surfaces). Increase in urban land cover (development) that occurs within this internal open space we characterize as “infill.” The remaining increase in urban land cover takes place on undeveloped land that is external to the built area. Some might characterize this as “sprawl” but to avoid any implied comment on the economic efficiency of the development we refer to it as “outspill” which at least seems a fitting companion to “infill.”

Our empirical estimates provide the first analysis (of which we are aware) of infill development in a global context. They permit an initial evaluation of the determinants of the different types of urban expansion, and may allow some preliminary tests of alternative theoretical perspectives on urban expansion.

2. Microstructure of urban expansion

In the simple version of the Alonso-Muth-Mills model, urban expansion would occur naturally as the (exogenously specified) population or incomes grew. The expansion that would result would be completely outspill, as the bid-rents increased to generate conversion of more agricultural land to urban use. In actual cities, we frequently observe development that leaves some internal open space within the built up area and generates new development at or even at some distance removed from the urban periphery. Why do we observe this “leapfrog” development?

There would seem to be at least four general types of answer to this question.

1. **Some vacant land is preserved as a public good or for public purposes**

Most cities have some internal open space that is maintained as a common area for recreational or ceremonial purposes. This land may be purchased in the market and collectively owned via the public sector (such as parkland or palace gardens) or it might be preserved as open space through land use regulations that prevent the owner from developing it. In this case the cost to society comes in the form of reduced income/wealth to the land-owner plus the distortion in land values and housing costs.

Such land can serve either as a public good in providing open space (whether or not public access is available) or as a public good in providing species habitat (assuming there is a public benefit to providing for the continued survival of the particular species). Land in this form can also serve a public purpose by providing for greater separation between incompatible land uses (for example a buffer zone that separates residential land use from the noise of a roadway). Depending on the particular urban context and source of externality this can offer the community an effective abatement strategy.

2. Land is not homogeneous

Some land is more costly to develop due to slope, soil or subsoil conditions, need for drainage, etc. While this might be obvious to the landowner (and hence reflected in the price paid for the land) and to all residents of the city, it is not always obvious to an external observer (particularly when viewed from the perspective of a satellite in orbit). Land which is partially under water or marshy can sometimes be detected and when we identify such areas we exclude them from the measured area designated as “internal open space.”

While some areas can be more expensive to develop and therefore skipped in favor of more easily developed areas further from the urban center, an increase in land values and the value of built space can induce development of such land.

3. Land owners and developers may have idiosyncratic payoff functions

It need not reflect a failure to optimize to observe land-owners owning and maintaining large vacant parcels within the city. There may be landowners who want to keep land in use in agriculture or open space for reasons of culture, tradition or entertainment and the value of these benefits to the owner may exceed the profits that can be obtained through development that in other circumstances would appear to be optimal.

Alternatively a fraction of landowners may simply be misinformed about market opportunities or for other reasons fail to optimize as theory would predict. If land prices are changing unevenly throughout the urban area then a process of adaptive updating of the expected payoff from development might generate “unemployed land” in the same way as such models generate unemployed labor.

4. Developers may hold some vacant land as part of a portfolio

As observed by Titman (1985), vacant land can be viewed as an option to develop structures in the future at a density and of a type that is optimal at the time the option is exercised. Since land development is difficult or impossible to reverse, land owners who face variability in the demand for built structures (or other forms of variability such as the regulatory variability studied by Mayo and Sheppard (2001)) will

generally hold vacant land in their portfolio in order to reduce the variability of their income. The greater is the volatility in the value of built structures, the greater will be value of this option and the greater will be incentive to hold some land vacant.

Given the uneven nature of provision of transport infrastructure and other desired locations within the urban area, this motive can create localized pockets of vacant land in the city. A recently published study by Cunningham (2006) undertakes a detailed, parcel-by-parcel analysis of properties in Seattle, Washington. The study tests for impacts of uncertainty about future house prices, and finds strong support for an options model of land development such as we have developed here. He finds that a one standard deviation increase in price uncertainty increases vacant land prices modestly (1.6 percent) and reduces the likelihood of development of the parcel (that is increases the likelihood of the parcel remaining vacant) by 11 percent.

In the case of explanation (2) and (4), any factor that increased the demand for developed space and that would in the ordinary course of events give rise to an expansion of urban land use (such as an increase in population or income, or a decrease in transportation costs or agricultural land rents) would increase the probability of development for any interior parcel left vacant. Such parcels only remain vacant if the value of the vacant land (the option to build in the future) exceeds the value of the land developed for the current market conditions. Factors which increase the value of developed land will tend to push some of the vacant parcels into development.

A similar process could be at work in explanations (1) and (3) – depending on the particular circumstances. For example, an increase in the value of developed land certainly increases the opportunity cost to the public sector of holding land vacant. On the other hand it is very rare for public sector lands to be given over to urban development in response to an increase in the market value of the land¹.

It would be quite possible to develop a formal mathematical model to illustrate each of these possible explanations, for the now we proceed to consider the data that have been obtained and the estimates based upon them.

¹ One recent exception was the practice of British schools in the south of England selling off their playing fields to developers in order to raise cash for their operations.

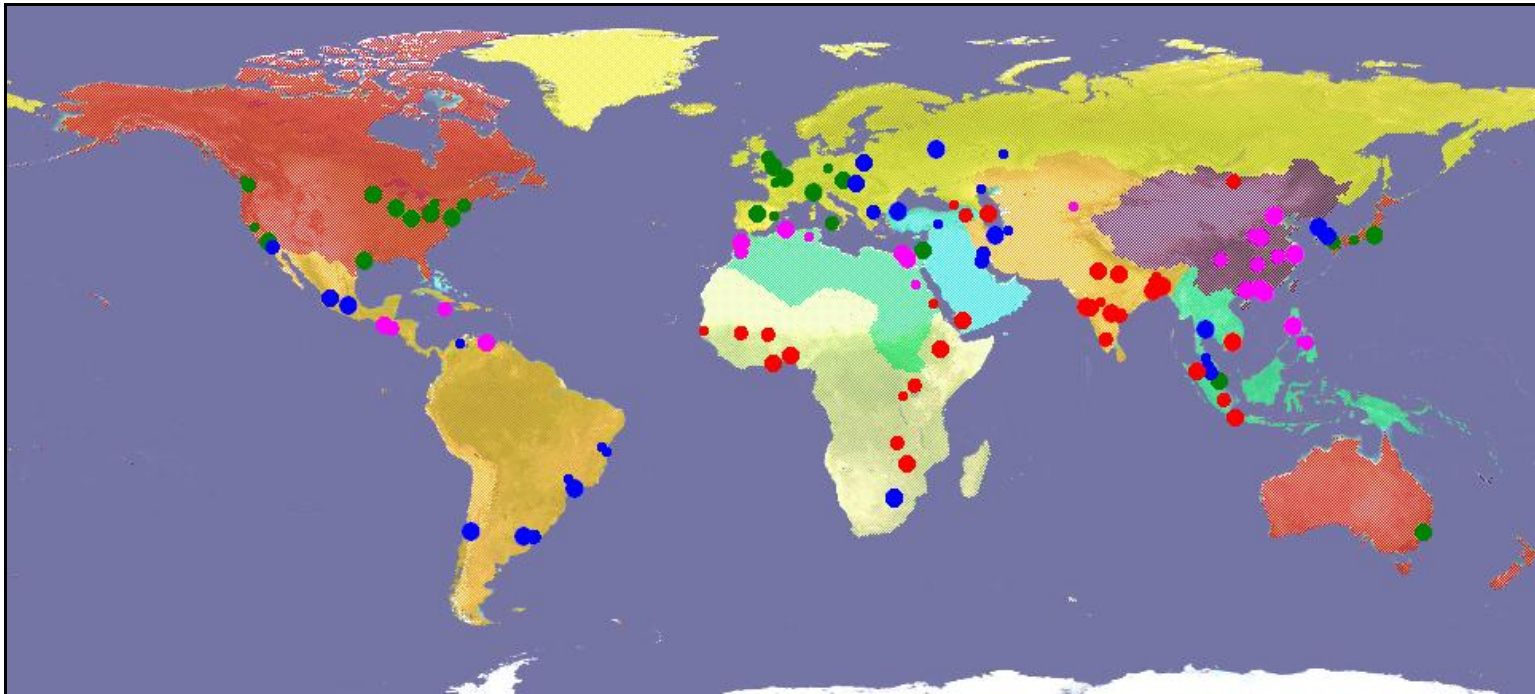
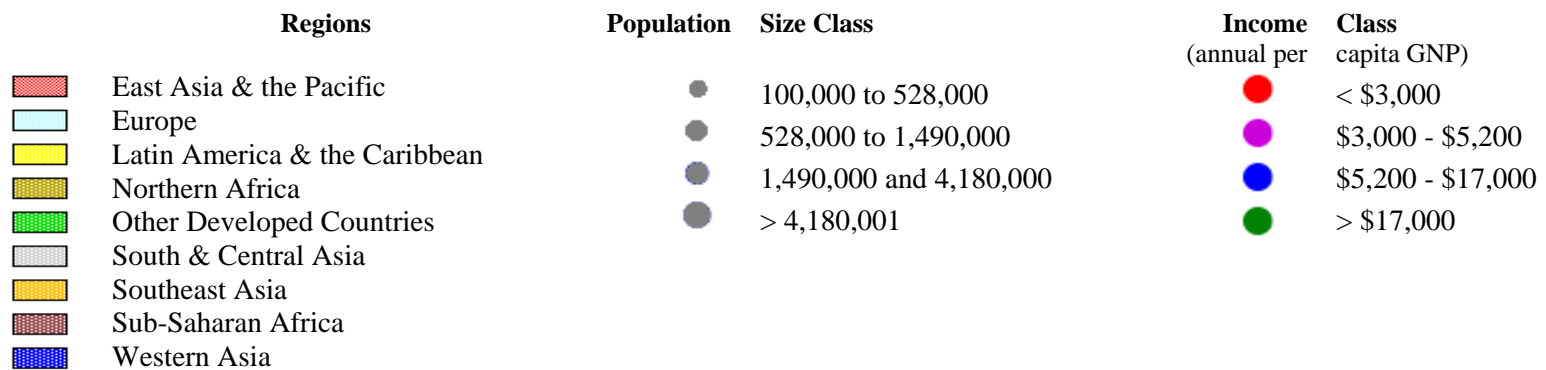


Figure 1: The Global Sample of 120 Cities, by Regions, Population Size, and Income Classes



3. Measuring infill and outspill

The data used for the empirical analysis combine information on urban land cover in a global sample of cities, population for jurisdictional areas within each city, measures of income and relevant economic variables for the country in which each city is located. Our data are collected for a sample of 120 cities randomly selected from a larger random sample constructed by UN Habitat for data collection in connection with their Urban Observatory program. The larger sample has been constructed to be representative of the global urban population in cities having population over 100,000 persons. The map in Figure 1 above shows the global distribution of the sample. The entire sample of cities is listed in a table in the Appendix.

For each of the cities in our sample, we obtained Landsat thematic mapper satellite images for dates that are near the national census dates and for which cloud-free images are available. Images were obtained for two time periods: approximately 1990 and approximately 2000. The images provide data on reflected light intensity in 7 spectral bands (3 visual and 4 infrared). These data are used to classify each point as urban (covered with impervious built structures or surfaces), water, or non-urban (everything else) in each time period.

Commercial data sources are available that provide this information, but the usual practice is to “fill in” small interior open spaces and classify them as urban. This essentially destroys the ability to detect interior open spaces and as a result eliminates the ability to identify infill development. Our approach has been to regard such spaces as non-urban so that we can distinguish between new capital investment and building at the urban periphery and “infill” development building inside the built up area. The contrast is illustrated in the figure below, which shows urban land classification for Addis Ababa, Ethiopia. The image on the left is from a commercial data provider, and the one on the right is the data used for our analysis. Further details concerning the remote sensing data and classification procedure are provided in Angel, Sheppard and Civco (2005).

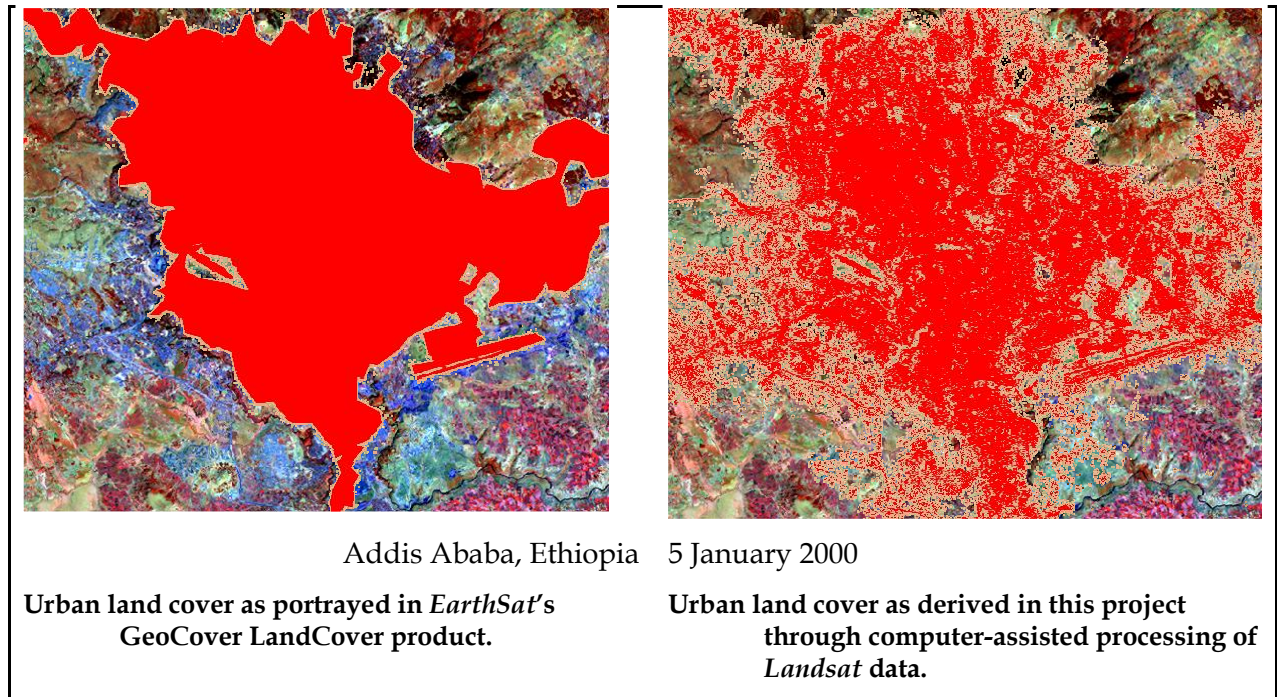


Figure 2: Alternative measures of urban land use and detection of interior open space

The remote sensing data provide us with measures of total urban land use at two points in time. These measures are then matched with population data for jurisdictional boundaries in each area, obtained from the Center for International Earth Science Information Network's Global Rural-Urban Mapping Project. Using growth rates observed for each jurisdiction during 1980 through 2000 we interpolate to obtain population estimates for the dates of each image. Similarly, we interpolate national per capita GDP to provide an estimate of income levels in each city at the date of the image. Data on the value of agricultural land (approximated by agricultural output per hectare), the cost of motor fuel, and other variables were obtained from World Development Indicators or from sources described more fully in Angel, Sheppard and Civco (2005).

We classify land use in the satellite images for the entire jurisdiction containing the central portion of the city included in the sample, plus additional jurisdictions required to cover an area larger than the entire contiguously built up urban area, plus additional jurisdictions required to cover nearby "satellite" cities. Since we do not have commuting or labor market data for most of the 59 countries containing the cities in our sample, we cannot precisely replicate the construction of "metro areas" used in the US or Europe. We nevertheless endeavor to approximate this construction, building up "metro areas" by taking the central

city jurisdiction and adding adjacent jurisdictions required to cover the built-up area plus nearby clusters of development from which commuting to the central contiguous area might be possible.

The image below illustrates the basis for our calculations to obtain the data for our analysis. Each pixel in the image identifies an area that is 28.5 meters by 28.5 meters from the Landsat Thematic Mapper image. Each pixel has been classified as urban, water, or other. Those pixels that were under urban cover in the later image (T_2) are colored in the image, and include those that were urban in the T_1 image.

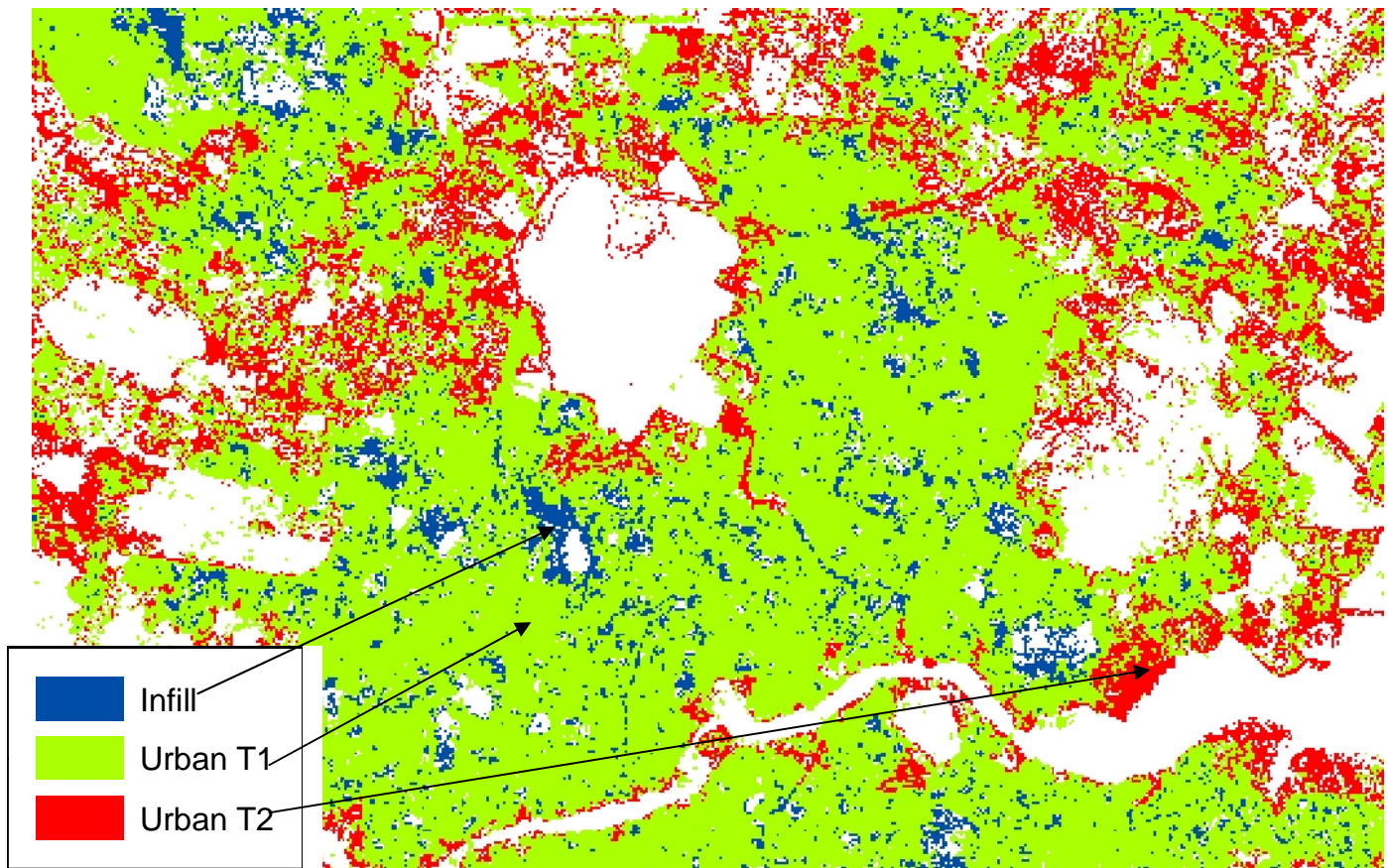


Figure 3: Hyderabad, India

Figure 3 shows both the total urban land use at T_1 in light green and the total urban land use at time T_2 (combined red and blue and light green). The red and blue areas illustrate the increase in urban land use. This change is decomposed into development that takes place in an area that was undeveloped but surrounded by developed land in T_1 and development taking place on previously undeveloped land that was not interior open space. Note that “surrounded” in this context means that it is not possible to construct a path from the undeveloped land to the edge of the picture without crossing at least 28.5 meters of developed urban land.

4. Sample data

The procedure discussed and illustrated in the preceding section provides data on 120 cities. The analysis reported in this paper includes 119 cities (extremely small jurisdictional boundaries in Pretoria require correction and checking before the data can be included in the full sample). The data are summarized in Table 1 below.

Table 1: Descriptive statistics for variables

Variable	Mean	Std. Dev	Min	Max	Obs
T ₁ Urban	398.858	647.247	10.706	3847.187	119
T ₂ Urban	542.256	776.907	20.258	4391.454	119
T ₁ Open	65.558	141.661	0.012	822.869	119
T ₂ Open	94.215	193.135	0.237	1109.965	119
Infill	21.881	41.119	0.007	261.672	119
Outspill	121.517	147.691	6.072	1003.179	119
Leapfrog	16.745	82.920	0	639.679	119
Δ Open Space	28.658	70.390	-81.889	443.709	119
Δ Population (millions)	9.293	94.854	-0.899	1035.279	119
Δ Real GDP per cap	1811.031	2377.095	-4125.445	7593.987	119
T ₁ Agrent	3138.589	14121.420	84.900	150542.900	118
T ₂ Agrent	2993.892	10440.210	68.837	109631.400	118
Δ Agri Rent	-144.698	3926.792	-40911.500	3364.859	118
Δ Fuel Cost	0.283	0.186	-0.170	0.880	120
EastAsia	0.133	0.341	0	1	120
LAC	0.133	0.341	0	1	120
NorthAfrica	0.067	0.250	0	1	120
SCentralAsia	0.133	0.341	0	1	120
SEAsia	0.1	0.301	0	1	120
SubSahAfrica	0.1	0.301	0	1	120
WestAsia	0.067	0.250	0	1	120

The table reports means for the seven regional indicator variables, showing that 73.3 percent of our sample cities are in these regions. The remaining 26.7 percent are in developed countries (Europe and other developed) which are taken as datum in the analysis. All land areas are reported in square kilometers. The sample provides tremendous variation in the relevant variables, with cities ranging in size from about 10 km² (Saidpur, Bangladesh) to over 3800 km² (Chicago, IL). City sizes range from barely 100,000 (Zugdidi, Georgia) to over 27 million (Tokyo, Japan). When comparing these measurements to other data sources, it is important to remember that these are aggregates for the set of jurisdictional boundaries that we have constructed as comprising the metropolitan area. Our primary objective was to capture virtually all of the area where urban expansion related to the economic

conditions in the city would occur. As a result there may be differences between our measures of population and other variables and statistical sources whose area of interest is defined differently.

It also bears repeating that our measure of “total urban area” is **not** the land area of the jurisdictions included in our metropolitan regions. It is our estimate of the total square kilometers of land that is covered by urban use.

Table 2 provides information on the amount of infill and outspill (and hence on total urban expansion), how this total urban expansion is divided between the two types of expansion, and how this structure varies across regions of the world.

Table 2: Regional variation in infill and outspill

Region	Infill	% Infill	Outspill	% Outspill	Freq.
East Asia	12.608	6.159	234.849	93.841	16
	<i>14.714</i>	<i>4.962</i>	<i>282.265</i>	<i>4.962</i>	
Europe	22.174	14.231	116.851	85.769	16
	<i>31.165</i>	<i>9.166</i>	<i>111.874</i>	<i>9.166</i>	
LAC	19.489	17.339	68.138	82.661	16
	<i>26.967</i>	<i>10.955</i>	<i>68.371</i>	<i>10.955</i>	
North Africa	7.248	9.263	77.807	90.737	8
	<i>8.553</i>	<i>7.578</i>	<i>69.979</i>	<i>7.578</i>	
Other Developed	73.309	22.830	184.889	77.170	16
	<i>85.016</i>	<i>20.012</i>	<i>157.840</i>	<i>20.012</i>	
South Central Asia	7.815	7.554	74.883	92.446	16
	<i>9.628</i>	<i>5.155</i>	<i>55.208</i>	<i>5.155</i>	
Southeast Asia	16.461	12.646	113.593	87.354	12
	<i>22.892</i>	<i>7.007</i>	<i>124.924</i>	<i>7.007</i>	
Sub-saharan Africa	9.590	9.216	75.170	90.784	11
	<i>13.376</i>	<i>7.131</i>	<i>54.653</i>	<i>7.131</i>	
West Asia	9.562	9.179	96.786	90.821	8
	<i>12.056</i>	<i>4.686</i>	<i>100.448</i>	<i>4.686</i>	
Total	21.881	12.525	121.517	87.475	119
	<i>41.119</i>	<i>11.235</i>	<i>147.691</i>	<i>11.235</i>	

The data provided in Table 2 reveal some interesting detail about the global variation in the micro-structure of urban expansion. We see a simple recipe for urban land development: urban expansion consists, on average, of **1 part infill to 7 parts outspill**. There is, however, considerable regional variation in this urban expansion cuisine. Ranging from **1 part infill to 15 parts outspill in East Asia** to **1 part infill combined with about 3 parts outspill in other Developed countries**.

Of course, Table 2 also indicates that there is considerable variation among cities within each of these regions as well, driven by the differences in city size (which vary tremendously within in each region) as well as income and other factors. In comparing the levels of inspill versus outspill across regions and the relative proportions, it is important to keep in mind that these factors and more might influence the patterns of development. Naturally the nature and enforcement of land use regulation and development controls could play an important role. For this study we do not explicitly consider these policies, but efforts to collect these data for each of the sample cities are nearing completion.

Developing a deeper understanding of the determinants of these types of urban expansion, and whether there are significant regional effects in the types of land development requires a more sophisticated analysis of the data, and this is presented in the following section.

5. Estimation

Before presenting estimates of the determinants of infill and outspill (components of the **change** in urban land use) it is of interest to examine estimates of the determinants of the **level** of urban land use. These estimates use the cross-sectional variation in the data to estimate the parameter coefficients. Urban land use should be a function of urban population, household incomes, the value of agricultural land around the city, and the costs of transportation within the city. Letting $L_{i,t}$ denote total urban land use in city i at time t , and using logarithms for all variables, estimate the following relationship:

$$L_{i,t} = \beta_0 + \beta_1 Pop_{i,t} + \beta_2 Inc_{i,t} + \beta_3 AgriVal_{i,t} + \beta_4 Fuel_{i,t} + \varepsilon_{i,t}$$

Naturally, there may be cultural, policy, and economic factors that are important in determining urban land use and that are constant within broad global regions. To capture these we include regional indicators.

A further potential difficulty with estimating this relationship

Table 3: Determinants of total urban land use

<i>Variable</i>	<i>Coefficient</i>
Population	0.723 9.76**
Income	0.709 7.37**
Agri Rent	-0.185 3.86**
Fuel Cost	-0.099 1.68
East Asia	0.296 1.09
West Asia	0.235 0.85
LAC	0.007 0.04
North Africa	0.158 0.45
South Central Asia	-0.182 0.87
Southeast Asia	-0.165 0.7
Sub-saharan Africa	0.65 2.48*
Constant	3.915 3.19**
Observations	232
R ²	0.76

is the potential endogeneity between right hand side variables and the error term $\varepsilon_{i,t}$ of the model. A city that finds itself with a large shock in urban land development may generate feedbacks on population and income that destroy the assumed independence between these variables and the model error. Tests of variable endogeneity cannot reject this hypothesis for our sample data, so we adopt an instrumental variable estimation strategy that uses the city location and the “biome” (soil and biological covering type) in the region of the city as instruments. We also account for the “clustering” of observations (each city is observed twice) and present the resulting estimates in Table 3.

Tests of endogeneity and relevance of our instruments supports maintenance of the hypothesis of independence from model error, and indicates that they are sufficiently strong to provide appropriate estimates. Table 3 reveals that the parameter estimate for each of our variables is estimated with the appropriate sign and all but the impact of fuel costs are statistically significant at the 5% level or better. Even the impact of fuel costs is significant at the 10% level.

The logarithmic form of the estimated equation in Table 3 permits easy interpretation. Both population and income are seen to be important determinants of the level of urban land use. Doubling of either population or income (which happens with some frequency over the roughly 10 year span between our satellite images in this sample) is associated with an increase of total urban land of about 70%. Doubling the value of agricultural land (or real agricultural value added per hectare) is associated with an 18.5% decrease in urban land use. Doubling the real cost of motor fuel is associated with about a 10% decrease in urban land use. These are all consistent with what would be predicted by a simple urban model.

Our central interest for this paper is to understand the determinants of the changes in urban land use, and in particular the determinants of infill versus outspill. For this analysis we estimate a differenced version of the model, along with a couple of other changes. The basic model to be estimated is:

$$\Delta L_{i,k} = \beta_0 + \beta_1 \cdot \Delta Pop_i + \beta_2 \cdot \Delta Inc_i + \beta_3 \cdot \Delta AgriVal_i + \beta_4 \cdot \Delta Fuel_i + \varepsilon_i$$

The dependent variable includes an index k that identifies whether infill (increase in urban land use through conversion of interior open space) or outspill is being evaluated. Here time indices have been dropped since it is the difference between the two observations that generates the differenced variables. As indicated in Table 1 above, some cities experience declines in each of these variables during the interval between our satellite observations. We estimate equations using the actual changes

in each variable. Although few of the regional indicator variable estimates indicated presence of a statistically significant regional fixed effect in the **levels** of urban land use, there is the potential for unmeasured regional factors to be present influencing the **expansion** of urban land use. The large differences between infill plus outspill across regions revealed in Table 2 above are suggestive of this possibility. We therefore include regional indicator variables in each equation. Table 4 presents estimates for these models.

Table 4: Determinants of infill and outspill

	Infill KM ²			Outspill KM ²		
	1	2	3	1	2	3
Δ Population	0.054 4.46**	0.054 4.46**	0.048 4.93**	0.282 7.60**	0.282 7.60**	0.233 5.90**
Δ GDP per cap	0.008 2.76**	0.008 2.76**	0.006 2.91**	0.027 3.69**	0.027 3.69**	0.013 1.84
T ₁ Agrent		-0.002 1.83			-0.015 3.01**	
T ₂ Agrent	-0.002 1.83			-0.015 3.01**		
Δ Agri Rent	-0.004 1.93	-0.005 1.91	0 0.71	-0.033 2.74**	-0.048 2.82**	0.002 0.73
Δ Fuel Cost	-33.14 1.47	-33.14 1.47	-41.599 1.93	-30.777 0.54	-30.777 0.54	-103.847 1.45
EastAsia	-21.303 1.79	-21.303 1.79	-31.679 2.66**	184.317 1.95	184.317 1.95	94.692 1.24
LAC	-8.301 0.69	-8.301 0.69	-12.278 1.07	-12.232 0.43	-12.232 0.43	-46.592 1.54
NorthAfrica	-17.851 1.42	-17.851 1.42	-25.397 2.27*	24.355 0.58	24.355 0.58	-40.828 1.05
SCentralAsia	-17.127 1.67	-17.127 1.67	-21.145 2.22*	0.467 0.02	0.467 0.02	-34.238 1.11
SEAsia	-17.699 1.06	-17.699 1.06	-24.69 1.6	38.702 0.78	38.702 0.78	-21.682 0.41
SubSahAfrica	-9.438 0.95	-9.438 0.95	-12.189 1.31	6.163 0.2	6.163 0.2	-17.599 0.55
WestAsia	-0.727 0.06	-0.727 0.06	-4.765 0.36	55.542 0.98	55.542 0.98	20.661 0.31
Constant	31.988 2.20*	31.988 2.20*	36.991 2.76**	87.433 2.54*	87.433 2.54*	130.647 3.03**
Observations	116	116	116	116	116	116
R ²	0.26	0.26	0.25	0.27	0.27	0.18
Wu-Hausman		0.21848			0.22205	
P-value		0.80412			0.80126	
Endog χ^2		0.4997			0.50783	
P-value		0.77892			0.77576	

Just as there was a possibility for endogeneity between right hand side variables and the model error in the estimation of the determinants of levels of urban land use, there is a possibility of endogeneity in modeling the changes in urban land use. Table 4 reports calculations of the Wu-Hausman and Chi-square statistics as tests of endogeneity of right hand side variables. In this case, we cannot support a hypothesis of endogeneity so OLS estimates appear to be appropriate. These are reported in the table.

The table presents estimates for 3 models for infill and outspill. Model 3 is the simple differenced model with regional indicators presented above. Models 1 and 2 add the level of agricultural land value (at the beginning or end of the interval) to the models. This addition is worthy of investigation because the **level** of agricultural land values provides a measure of what might be characterized as a “resistive force” against urban land development quite apart from any **change** in this level. The same observation would be true if we had available to us data on the levels of land use regulation and enforcement. The level of regulation would be expected to be as important (perhaps more important) in explaining the change in urban land use because it limits the responsiveness of urban land development to changes in income, population, and transport costs.

The estimates indicate that changes in population and income always have the expected sign and are statistically significant. Changes in agricultural land values typically have the expected sign and are correctly signed and are statistically significant (or close) when the level of agricultural land values are included. Changes in fuel costs present a more mixed picture. While the direction of the impact is of the expected sign, the parameter estimates are of low precision so that they are not statistically significant (or at best only marginally so).

Most of the regional indicators are not statistically significant, although those for East Asia are close and are of large magnitude supporting the observations made above: cities in this region are experiencing large changes in urban land use, have much lower changes coming through infill and much larger changes coming through outspill.

Our final series of estimates focuses attention on the role these variables play in determining the emerging levels of open space in the cities in our sample. Infill development can only take place if there are interior open spaces to be filled in. The role of interior open space in potentially providing public good benefits to residents also gives rise to interest in the factors that influence the amount of

this space and the creation of new interior open spaces (through “leapfrog” development). Table 5 below presents some preliminary estimates of these processes.

Table 5: Determinants of changes in open space

	Change in Open Space				Leapfrog Open Space			
	1	2	3	IV	1	2	3	IV
Δ Population	0.105 <i>5.25**</i>	0.105 <i>5.25**</i>	0.092 <i>5.27**</i>	0.185 <i>1.68</i>	0.006 <i>0.3</i>	0.006 <i>0.3</i>	-0.002 <i>0.14</i>	0.026 <i>0.23</i>
Δ GDP per cap	0.013 <i>3.07**</i>	0.013 <i>3.07**</i>	0.01 <i>2.96**</i>	0.019 <i>2.25*</i>	0.013 <i>1.88</i>	0.013 <i>1.88</i>	0.01 <i>1.96</i>	0.007 <i>0.82</i>
T ₁ Agrent		-0.004 <i>2.34*</i>		-0.004 <i>2.08*</i>		-0.002 <i>1.53</i>		-0.001 <i>0.63</i>
T ₂ Agrent	-0.004 <i>2.34*</i>				-0.002 <i>1.53</i>			
Δ Agri Rent	-0.006 <i>1.78</i>	-0.01 <i>1.96</i>	0.002 <i>2.75**</i>	-0.013 <i>1.9</i>	-0.005 <i>1.44</i>	-0.007 <i>1.47</i>	0.001 <i>1.59</i>	-0.004 <i>0.62</i>
Δ Fuel Cost	-18.363 <i>0.68</i>	-18.363 <i>0.68</i>	-36.418 <i>1.22</i>	6.117 <i>0.27</i>	-61.816 <i>2.17*</i>	-61.816 <i>2.17*</i>	-74.273 <i>2.22*</i>	-28.694 <i>1.24</i>
EastAsia	-2.711 <i>0.1</i>	-2.711 <i>0.1</i>	-24.857 <i>0.98</i>	37.567 <i>1.42</i>	-40.946 <i>1.98</i>	-40.946 <i>1.98</i>	-56.226 <i>2.14*</i>	-9.201 <i>0.52</i>
LAC	-19.533 <i>1.11</i>	-19.533 <i>1.11</i>	-28.023 <i>1.57</i>	24.973 <i>1.01</i>	-29.773 <i>1.61</i>	-29.773 <i>1.61</i>	-35.631 <i>1.82</i>	-2.738 <i>0.12</i>
NorthAfrica	-2.3 <i>0.11</i>	-2.3 <i>0.11</i>	-18.406 <i>0.93</i>	52.925 <i>1.75</i>	-27.498 <i>1.48</i>	-27.498 <i>1.48</i>	-38.61 <i>1.89</i>	-0.887 <i>0.03</i>
SCentralAsia	-11.62 <i>0.66</i>	-11.62 <i>0.66</i>	-20.196 <i>1.17</i>	39.169 <i>1.5</i>	-25.843 <i>1.49</i>	-25.843 <i>1.49</i>	-31.76 <i>1.77</i>	-0.115 <i>0</i>
SEAsia	18.305 <i>0.58</i>	18.305 <i>0.58</i>	3.385 <i>0.11</i>	64.201 <i>2.04*</i>	-40.402 <i>1.92</i>	-40.402 <i>1.92</i>	-50.697 <i>2.14*</i>	-8.104 <i>0.36</i>
SubSahAfrica	-0.837 <i>0.05</i>	-0.837 <i>0.05</i>	-6.709 <i>0.39</i>	50.377 <i>1.63</i>	-12.084 <i>0.74</i>	-12.084 <i>0.74</i>	-16.135 <i>1.02</i>	8.634 <i>0.32</i>
WestAsia	16.994 <i>0.8</i>	16.994 <i>0.8</i>	8.375 <i>0.4</i>	64.299 <i>1.71</i>	-3.664 <i>0.19</i>	-3.664 <i>0.19</i>	-9.611 <i>0.5</i>	12.564 <i>0.34</i>
Constant	21.505 <i>0.95</i>	21.505 <i>0.95</i>	32.183 <i>1.39</i>	-37.98 <i>1.28</i>	39.541 <i>1.84</i>	39.541 <i>1.84</i>	46.908 <i>2.07*</i>	6.132 <i>0.25</i>
Observations	116	116	116	116	116	116	116	116
R ²	0.18	0.18	0.16	0.1017	0.17	0.17	0.16	0.0891
Wu-Hausman		8.83279	Anderson	53.224**		4.6828	Anderson	53.224**
P-value		0.00029	Hansen J	5.63		0.01135	Hansen J	4.171
Endog χ^2		17.26876				9.84374		
P-value		0.00018				0.00729		

When new development at the urban periphery “captures” undeveloped open space by surrounding it with urban land use, new interior open space is created. We refer to the additional open space created in this way as “leapfrog open space”. The total change in open space results from the combination of this leapfrog open space minus the infill development that occurs. The table presents models for both

processes. As before, there is concern about endogeneity and in this case the concerns appear to be justified. The reported Wu-Hausman and Chi-square tests suggest acceptance of the hypothesis of variable endogeneity, which would mandate an IV approach in order to obtain consistent estimates. Table 5 includes IV estimates for both types of land use changes making use of the same instruments (city location and surrounding natural environment) used in modeling the levels of urban land use.

6. Concluding remarks

This paper has presented a first evaluation of the microstructure of urban expansion in a unique global data set being assembled for a representative sample of cities. We have measured and modeled total urban land use and changes in urban land use. We have divided the change in urban land use into “infill” – land development taking place within previously undeveloped spaces inside the city and “outspill” that takes places on undeveloped sites at the urban periphery.

Our results are consistent with what would be expected from a standard urban model, and in particular are consistent with at least two alternative views of why such spaces exist within cities. Increases in population and income increase both infill and outspill, and well as the total amount of internal open space. Income appears to be particularly important in this process, further establishing the recognized link between the process of economic growth and development and urbanization.

There are many directions in which this research can and should be extended. We mention two that are high priority in the present research project. First, testing the possible explanation that interior open space exists because of the desire for land developers to hold undeveloped land in their portfolio as a hedge against future variability in the demand for urban development requires that we bring some indication of the volatility of development pressures into the model. Second, policy interest is naturally focused on the ability of land use regulation and development controls to influence the nature of urban land development. These data are being collected at present and will be studied more closely in a future paper.

Finally, we note that all of our analysis is on development of previously undeveloped land. We are not able to model the process of urban **re**-development (involving the demolition or radical modification of existing structures) because remote sensing data and analysis techniques do not permit us to

distinguish between old and new construction. Nevertheless, these processes are an interesting part of the process of urban expansion. Their more careful study in a global context will have to wait for improved remote sensing techniques or other data sources.

7. References

- Angel, S., Sheppard, S. and Civco, D. (2005), *The Dynamics of Global Urban Expansion*, Washington: The World Bank. Available online at www.williams.edu/Economics/UrbanGrowth/WorkingPapers.htm or <http://www.citiesalliance.org/publications/homepage-features/feb-06/urban-expansion.html>
- Brueckner, J., (1987), The Structure of Urban Equilibria, Chapter 20 in *Handbook of Regional and Urban Economics*, E. Mills, ed., New York: Elsevier.
- Burchfeld, N., Overman, H., Puga, D. and Turner, M. (2006), Sprawl: A View from Space, *Quarterly Journal of Economics*, **121**, 587-633.
- Cunningham, C. (2006), House price uncertainty, timing of development, and vacant land prices: Evidence for real options in Seattle, *Journal of Urban Economics*, **59**, 1-31.
- Irwin, E. and Bockstael, N., (2006) Measuring and Modeling Urban Sprawl: Data, Scale and Spatial Dependencies, unpublished working paper.
- Sheppard, Stephen, (2006), Infill versus Outspill: the microstructure of urban expansion, Williams College Working Papers in Economics.
- Titman, Sheridan, (1985) Urban Land Prices Under Uncertainty, *American Economic Review*, **75**, 505-514.
- Wu, Weiping and Yusuf, Shahid, (2004), Shanghai: remaking China's future global city, in Gugler, J. (ed) *World Cities Beyond the West*, Cambridge: Cambridge University Press.

Appendix: The Global Sample of 120 Cities

No.	City				Country		
	Name	Population in 2000	Size Class	Rank	Name	GNP/cap. in PPP (\$)	Income Class
Eastern Asia							
1	Shanghai	12,900,000	4	1	China	3,547	2
2	Beijing	10,800,000	4	2	China	3,547	2
3	Seoul	9,887,779	4	1	Republic of Korea	13,958	3
4	Hong Kong	6,927,000	4	4	China	3,547	2
5	Guangzhou	3,893,000	3	9	China	3,547	2
6	Pusan	3,830,000	3	2	Republic of Korea	13,958	3
7	Zhengzhou	2,070,000	3	23	China	3,547	2
8	Yulin	1,558,000	3	46	China	3,547	2
9	Yiyang	1,343,000	2	67	China	3,547	2
10	Leshan	1,137,000	2	88	China	3,547	2
11	Ulan Bator	738,000	2	1	Mongolia	1,491	1
12	Changzhi	593,500	2	185	China	3,547	2
13	Anqing	566,100	2	196	China	3,547	2
14	Ansan	549,900	2	15	Republic of Korea	13,958	3
15	Chinju	287,100	1	24	China	13,958	3
16	Chonan	114,600	1	47	Republic of Korea	13,958	3
Europe							
1	Paris	9,624,000	4	1	France	23,225	4
2	Moscow	9,321,000	4	1	Russian Fed.	6,644	3
3	London	8,219,226	4	1	United Kingdom	22,652	4
4	Milan	4,251,000	4	1	Italy	22,875	4
5	Madrid	4,072,000	3	1	Spain	18,314	4
6	Warsaw	2,269,000	3	1	Poland	9,114	3
7	Vienna	2,070,000	3	1	Austria	25,694	4
8	Budapest	1,825,000	3	1	Hungary	11,301	3
9	Thessaloniki	789,000	2	2	Greece	15,280	3
10	Palermo	684,300	2	7	Italy	22,875	4
11	Sheffield	640,048	2	9	United Kingdom	22,652	4
12	Astrakhan	486,100	1	36	Russian Fed.	6,644	3
13	Leipzig	446,491	1	19	Germany	23,913	4
14	Le Mans	194,825	1	34	France	23,225	4
15	Castellon	144,500	1	40	Spain	18,314	4
16	Oktyabrsky	111,500	1	147	Russian Fed.	6,644	3
Latin America and the Caribbean							
1	Mexico City	18,100,000	4	1	Mexico	8,182	3
2	Sao Paulo	17,800,000	4	1	Brazil	6,781	3
3	Buenos Aires	12,600,000	4	1	Argentina	11,131	3
4	Santiago	5,538,000	4	1	Chile	8,412	3
5	Guadalajara	3,908,000	3	2	Mexico	8,182	3

The Global Sample of 120 Cities (continued)

No.	City				Country		
	Name	Population in 2000	Size Class	Rank	Name	GNP/cap. in PPP (\$)	Income Class
Latin America and the Caribbean (continued)							
6	Guatemala City	3,242,000	3	1	Guatemala	3,633	2
7	Caracas	3,153,000	3	1	Venezuela	5,174	2
8	San Salvador	1,408,000	2	1	El Salvador	4,307	2
9	Montevideo	1,236,000	2	1	Uruguay	8,130	3
10	Tijuana	1,167,000	2	7	Mexico	8,182	3
11	Kingston	912,500	2	1	Jamaica	3,370	2
12	Ribeirão Preto	502,333	2	23	Brazil	6,781	3
13	Valledupar	274,300	1	16	Colombia	5,618	3
14	Guarujá	269,104	1	70	Brazil	6,781	3
15	Ilhéus	161,898	1	85	Brazil	6,781	3
16	Jequié	130,207	1	102	Brazil	6,781	3
Northern Africa							
1	Cairo	10,600,000	4	1	Egypt	3,253	2
2	Alexandria	4,113,000	3	2	Egypt	3,253	2
3	Casablanca	3,541,000	3	1	Morocco	3,195	2
4	Algiers	2,760,740	3	1	Algeria	4,979	2
5	Marrakech	736,500	2	4	Morocco	3,195	2
6	Port Sudan	384,100	4	2	Sudan	1,512	1
7	Aswan	219,017	4	15	Egypt	3,253	2
8	Tébessa	163,279	4	13	Algeria	4,979	2
Other Developed Countries							
1	Tokyo	26,400,000	4	1	Japan	23,828	4
2	Los Angeles	16,373,645	4	2	United States	31,338	4
3	Chicago	9,157,540	4	3	United States	31,338	4
4	Philadelphia	6,188,463	4	6	United States	31,338	4
5	Houston	4,669,571	4	10	United States	31,338	4
6	Sydney	3,664,000	3	1	Australia	24,013	4
7	Minneapolis	2,968,806	3	14	United States	31,338	4
8	Pittsburgh	2,358,695	3	21	United States	31,338	4
9	Cincinnati	1,979,202	3	23	United States	31,338	4
10	Fukuoka	1,341,470	2	10	Japan	23,828	4
11	Tacoma	700,820	2	62	United States	31,338	4
12	Springfield	591,932	2	72	United States	31,338	4
13	Modesto	446,997	1	94	United States	31,338	4
14	St. Catharine's	389,600	1	14	Canada	25,456	4
15	Victoria	317,506	1	16	Canada	25,456	4
16	Akashi	293,117	1	60	Japan	23,828	4

The Global Sample of 120 Cities (continued)

No.	City				Country		
	Name	Population in 2000	Size Class	Rank	Name	GNP/cap. in PPP (\$)	Income Class
South and Central Asia							
1	Mumbai	18,100,000	4	1	India	2,220	1
2	Kolkata	12,900,000	4	2	India	2,220	1
3	Dhaka	12,300,000	4	1	Bangladesh	1,427	1
4	Teheran	7,225,000	4	1	Iran	5,460	3
5	Hyderabad	6,842,000	4	4	India	2,220	1
6	Pune	3,489,000	3	9	India	2,220	1
7	Kanpur	2,450,000	3	11	India	2,220	1
8	Jaipur	2,145,000	3	13	India	2,220	1
9	Coimbatore	1,292,000	2	23	India	2,220	1
10	Vijayawada	1,237,000	2	28	India	2,220	1
11	Rajshahi	1,016,000	2	4	Bangladesh	1,427	1
12	Ahvaz	997,000	2	7	Iran	5,460	3
13	Shimkent	360,100	1	4	Kazakhstan	4,215	2
14	Jalna	244,523	1	158	India	2,220	1
15	Gorgan	188,710	1	33	Iran	5,460	3
16	Saidpur	114,000	1	25	Bangladesh	1,427	1
Southeast Asia							
1	Metro Manila	10,900,000	4	1	Philippines	3,668	2
2	Bangkok	7,281,000	4	1	Thailand	5,846	3
3	Ho Chi Minh City	4,615,000	4	1	Vietnam	1,854	1
4	Singapore	3,567,000	3	1	Singapore	21,832	4
5	Bandung	3,409,000	3	2	Indonesia	2,807	1
6	Medan	1,879,000	3	4	Indonesia	2,807	1
7	Palembang	1,422,000	2	5	Indonesia	2,807	1
8	Kuala Lumpur	1,378,000	2	5	Malaysia	8,217	3
9	Cebu	718,821	2	10	Philippines	3,668	2
10	Ipoh	566,211	2	2	Malaysia	8,217	3
11	Bacolod	429,076	1	7	Philippines	3,668	2
12	Songkhla	342,475	1	2	Thailand	5,846	3
Sub-Saharan Africa							
1	Addis Ababa	2,639,000	3	1	Ethiopia	648	1
2	Johannesburg	2,335,000	3	2	South Africa	8,667	3
3	Accra	1,976,000	3	1	Ghana	1,804	1
4	Harare	1,752,000	3	1	Zimbabwe	2,372	1
5	Ibadan	1,731,000	3	3	Nigeria	808	1
6	Pretoria	1,508,000	3	4	South Africa	8,667	3
7	Kampala	1,212,000	2	1	Uganda	1,164	1
8	Bamako	1,131,000	2	1	Mali	683	1
9	Ouagadougou	1,130,000	2	1	Burkina Faso	931	1

The Global Sample of 120 Cities (continued)

No.	City				Country		
	Name	Population in 2000	Size Class	Rank	Name	GNP/cap. in PPP (\$)	Income Class
Sub-Saharan Africa (continued)							
10	Ndola	568,600	2	3	Zambia	715	1
11	Banjul	399,386	1	1	Gambia	1,542	1
12	Kigali	351,400	1	1	Rwanda	1,019	1
Western Asia							
1	Istanbul	9,451,000	4	1	Turkey	5,731	3
2	Tel Aviv-Jaffa	2,181,000	3	1	Israel	18,895	4
3	Baku	1,936,000	3	1	Azerbaijan	2,358	1
4	Sana'a	1,653,300	3	1	Yemen	760	1
5	Yerevan	1,406,765	2	1	Armenia	2,222	1
6	Kuwait City	1,190,000	2	1	Kuwait	14,471	3
7	Malatya	437,000	1	14	Turkey	5,731	3
8	Zugdidi	104,947	1	6	Georgia	1,722	1
	Total	415,605,624					