Modeling the Drivers of Urban Land Use Change in the Pearl River Delta, China: Integrating Remote Sensing with Socioeconomic Data

Karen C. Seto and Robert K. Kaufmann

ABSTRACT. *This paper estimates econometric models of the socioeconomic drivers of urban land use change in the Pearl River Delta, China. The panel data used to estimate the models are generated by combining high-resolution remote sensing data with economic and demographic data from annual compendium. The relations between variables are estimated using a random coefficient model. Results indicate that urban expansion is associated with foreign direct investment and relative rates of productivity generated by land associated with agricultural and urban uses. This suggests that large-scale investments in industrial development, rather than local land users, play the major role in urban land conversion.* (JEL R14)

I. INTRODUCTION

Urbanization is one of the most widespread environmental changes of the twenty-first century. The wholesale transformation of agricultural and natural ecosystems to more intensive uses is among the biggest anthropogenic impacts on earth (Vitousek et al. 1997). Although urban areas cover less than 2% of the earth's total land surface (Grübler 1994), more than half of the world's population reside in urban regions. Urban growth generally leads to an increase in motorized transport, air, water and noise pollution, energy consumption, a loss of agricultural land, and a reduction in biological diversity. Given these impacts, understanding the mechanisms that drive urbanization is critical to an understanding of global environmental change.

Nowhere is urbanization more important than in China, where an unprecedented scale and rate of urban expansion has occurred over the last two decades. With 30% of its 1.3 billion inhabitants residing in urban areas and a UN projection that more than 50% of the population will be urban by 2030 (United Nations 2001), the urban Chinese landscape is likely to expand at a very rapid rate. Traditional agrarian communities and vast tracts of agricultural land are being metamorphosed by a changed economy and enveloped by sprawling urban formations. The extent and pace of urban transformation have led to concerns about the country's food security (Brown 1995), the effects of increased energy consumption (Sadownik and Jaccard 2001), and the sustainability of continued economic expansion. Prior to the economic and political reforms in the late 1970s, urban planning and urban growth were managed largely from Beijing. However, with the opening of the economy and the introduction of market-oriented policies, urban development has been increasingly unstructured and unregulated: more a mélange of individual opportunities than coordinated activities. Urban growth and market reforms have led to massive rural to urban migrations (Smith 2000), and generated a bi-level labor market,

The authors are, respectively, senior research scholar and fellow, Center for Environmental Science and Policy, Institute for International Studies, Stanford University; and associate professor, Department of Geography, Center for Energy and Environmental Science, Boston University.

with disparate wages between urban and rural workers (Wang, Maruyama, and Kikuchi 2000; Meng and Zhang 2001).

To formulate policy that promotes economic growth and urban development while minimizing environmental impacts, decision-makers must understand the factors that drive urban expansion. Without a clear sense of the conditions that *cause* urban encroachment on farmland, policies are likely to be ineffectual. In this paper, we combine time series land use data extracted from high-resolution satellite imagery, socio-economic data, and field observations to estimate econometric equations for the socioeconomic forces that drive urban growth between 1988 and 1996 in the Pearl River Delta, one of the fastest growing regions in China. This is the first paper to integrate socioeconomic data with land use data derived from a sequential time series of satellite imagery (nine consecutive years). Previous work with satellite imagery focuses on deforestation and decisions driven by micro-economic influences and individual households in two to four time periods (Chomitz and Gray 1996; Nelson and Hellerstein 1997; Cropper, Griffiths, and Mani 1999; Pfaff 1999; Geoghegan et al. 2001). Here, we focus on the macro-level socioeconomic environment and its effect on the rate at which agricultural land and natural ecosystems are converted to urban uses. Our goal is to develop a better understanding of the comparative significance of local and exogenous macro-socioeconomic factors that affect urbanization. The paper is organized as follows: Section 2 describes the study area and changes in policies that have led to economic and urban expansion. The remote sensing and socioeconomic data are described in Section 3, and the estimation techniques are described in Section 4. Results and discussion are provided in Section 5, and concluding comments are presented in Section 6.

II. STUDY AREA: PEARL RIVER DELTA

The study area is the Pearl River (*Zhujiang*) Delta, located in the southern Chinese province of Guangdong, between 21°N and 23°N, and crossed by the Tropic of Cancer

(Figure 1). The geographical area of the Delta included in this study is defined by one Landsat TM scene, and covers 26,000 km². The Delta is distinguished by temperate winters during a dry season between October through April, and long summers during a rainy season between May through September. The combination of fertile soil and a tropical/subtropical monsoon climate supports two to three crops per year, which makes the region well suited for agriculture. More than half of the region's agricultural area is planted with paddy rice. Other major crops include sweet potatoes, sugar cane, peanut, soybean, banana, oranges and lychee. The Delta also is one of the major mulberry and silkworms producing areas in China. These agricultural systems combine pond fish with mulberry trees to form a symbiotic relation where the trees feed the silkworm, the waste of which feeds fish, and the mud from the ponds are used as fertilizer for the mulberry trees. Since the mid-1980s, pond fish have become a major agricultural output, which reflect changes in government policies and consumer demand for luxury food items.

More than 21 million people reside in the Delta, which represents one-third of the total population of the province. They speak *guangdong hua*, or Cantonese, the dialect shared by 7 million residents in Hong Kong. Shared linguistic and cultural ties to overseas Chinese investors provide access to investment capital, technological advances, business acumen, and management experience. Between 1997 and 1999, more than 74% of total foreign direct investment in the province of Guangdong came from Hong Kong (Statistical Bureau of Guangdong various years). The influx of foreign capital has transformed the historically agrarian economy by developing and expanding textile, electronics, and food processing industries over the last two decades.

Urbanization is strongly linked to economic growth, which accelerated in 1979 due to decentralization policies and market-oriented reforms. These include agricultural price reforms and the elimination of collective farm management. The latter was replaced with household and individual farming, also known as the household responsi-
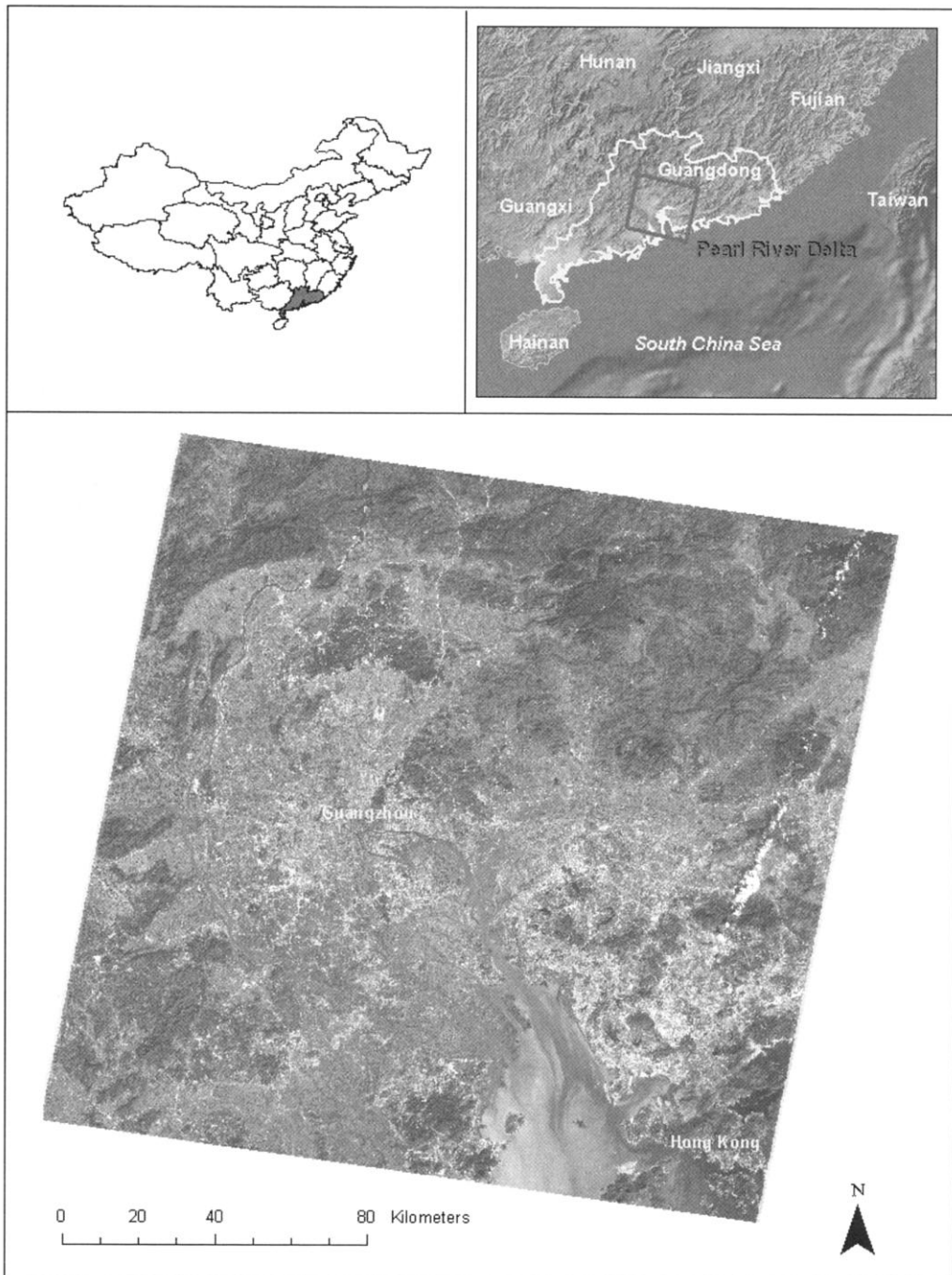
FIGURE 1

STUDY AREA: PEARL RIVER DELTA IN THE SOUTHERN PROVINCE OF GUANGDONG.
THE UPPER LEFT INSET SHOWS THE GUANGDONG PROVINCE IN RELATION TO THE REST OF THE
COUNTRY. THE UPPER RIGHT INSET HIGHLIGHTS THE PEARL RIVER DELTA. THE BOTTOM IMAGE IS
A COMPOSITE OF SAMPLE SATELLITE DATA OF THE LANDSAT THEMATIC MAPPER.

bility system. The household responsibility system allowed communes to distribute their land among farmers, who could sell a portion of their crops on the free market. Concurrently, price reforms led to price increases for grain and other agricultural outputs. In response, farmers began to choose their crops based on market conditions and dramatically increased agricultural production (Lin 1988).

Shortly after these reforms, the central government established special economic zones (SEZs) to draw foreign investment, promote exports and light industry, and test various free-market policies (Chu 1998). The most successful of the SEZs is Shenzhen, a region that receives the lion's share of foreign direct investment, located adjacent to Hong Kong. To complement this development, the Delta was given special tax and trade status and was designated as an Open Economic Region in 1985 (Lin 1997). In addition, the privatization of some government and communally held enterprises shifted economic control from the state to the competitive market. Amplifying these reforms, land use policies changed in 1988, when the Constitution of the People's Republic of China was amended to allow the transfer of land use rights (Sharkawy, Chen, and Pretorius 1995). Land ownership remains in the purview of the state, but land use rights are available by negotiation, bid, or auction. This arrangement allows farmers and collectives to rent their land to foreign and local ventures. These actions have converted large areas of farmland to urban uses.

The reforms also have generated competing demands on land use. Higher per capita income has increased the private ownership of vehicles, increased the demand for transportation networks, improved dietary standards, and raised the quality and quantity of living space. Prior to the reforms, China had a stringent *hukou*, or household registration system, which limited people's residential mobility; migration from rural to urban areas was nearly impossible. Additionally, the *danwei*, or work unit, provided basic goods and services such as housing, health care, food ration tickets, and education. Together, the *hukou* and *danwei* controlled labor mobility and curbed urbanization. Reforms have

relaxed the *hukou* and diminished the importance of the *danwei*. Additionally, the household responsibility system increased agricultural efficiency, which freed excess farm laborers and increased the number of "floating" workers. Given the economic opportunities in the coastal regions, there have been massive migrations from rural to urban areas and unparalleled rates of urbanization. Urban areas in the Delta increased more than 300% between 1988 and 1996 (Seto et al. 2002), most of which were converted from agricultural land (Seto, Kaufmann, and Woodcock 2000).

III. DATA SOURCES

To quantify the factors that drive urbanization, we estimate econometric equations for annual rates of land use change using land use change data assembled from satellite imagery and socioeconomic data collected from statistical yearbooks. These data are compiled at the county level using a geographic information systems (GIS). County-level administrative boundaries (1:1,000,000) are obtained from the Institute for Remote Sensing Application (IRSA) in Beijing and the Center for International Earth Science Information Network (CIESIN) at Columbia University. The remote sensing and socioeconomic data are combined to form a data panel that includes eight observations (the dependent variable, land use change, consumes one observation) from 1988 to 1996 for eleven counties.[1] Treating the data as a panel, rather than solely a cross-section or a time series, has several advantages, including the ability to control for heterogeneity across counties, an increase in the degrees of freedom and therefore a reduction in collinearity among explanatory variables, an increase in statistical efficiency, and a reduction in omitted variable bias (Hsiao 1986).

Land Use Data

Data on agricultural land use can be obtained from county-level statistical year-

---

[1] The eleven counties include: Jiangmen, Sanshui, Shunde, Nanhai, Foshan, Dongguan, Shenzhen, Panyu, Zengcheng, Huadu, and Guangzhou.

books. However, we chose not use them for three reasons. First, the quality of Chinese statistics, in particular, land use data, have long been questioned (Smil 1993; Chow 1994). We suspect that political and economic pressures, combined with inconsistencies in reporting, have resulted in over- and underestimates of the quantity of agricultural land. A recent study using high-resolution satellite imagery confirms this bias—in general, government statistics understate agricultural land and the rate at which it is converted to non-agricultural uses (Seto, Kaufmann, and Woodcock 2000). Second, there are no area estimates of urban land. Third, although county-level statistics on cultivated land and forested areas refer to a particular geographic area, we cannot use this information to calculate rates of urban land use change. Increases in urban area could have come from the conversion of natural vegetation or water bodies, which are not documented in the statistical yearbooks. Given this, it is not possible to estimate types of urban land use change solely from statistical yearbooks.

Data on land use and urbanization are extracted from Landsat Thematic Mapper (TM) images for each of the nine years, 1988 to 1996. The images vary in the month in which they were acquired, but they were all taken during the dry season, between October and March. The images are georeferenced to a Universal Transverse Mercator projection, with a registration error of less than 0.3. The images are resampled to 30-meter pixel resolution, and radiometrically corrected using a relative calibration technique (Song et al. 2001).

We use a two-step process to extract annual rates of urban land use change. In the first step, we classify the 1988 and 1996 images using a Bayesian maximum likelihood classifier to generate a two-date map of land use change.[2] The map has nine classes: *natural vegetation, agriculture, water, urban, agriculture → water, natural vegetation → water, water → agriculture, natural → urban,* and *agriculture → urban.* The *natural vegetation* class includes both shrubs and forest. The *urban* class includes two points along the urban development continuum: 1) areas that have been cleared for construc-

tion; and 2) new urban areas, roads, and old villages and towns. The *agriculture* class includes fish ponds and crops. Crop types include orchards, rice fields, and field crops. The two-date map with these nine land use and land use change classes has an overall area-weighted accuracy of 93.5% as indicated by a field-based accuracy assessment. The satellite data and remote sensing change detection methodology are described in greater detail in Seto et al. (2002).

Of the nine classes, we focus on the two urbanization classes. The *natural → urban* class includes new urban areas that were previously forest, shrub, or water, and represents the urbanization of land with no prior economic return. The *agriculture → urban* class includes new urban areas that were used previously for fish aquaculture, rice, fruit orchards, or field crops, and describes the urbanization of land that previously generated an economic return. We distinguish between the conversion of agricultural and nonagricultural lands because we assume that the opportunity cost of converting agricultural land is greater than the opportunity cost of converting shrub, water, or forest.

After identifying pixels that urbanized between 1988 and 1996, we identify the year in which the urban land use change occurred. This information is used to generate annual data on land use change. Two methods are tested to extract estimates of annual rates of land use change. The first technique is the aforementioned Bayesian maximum likelihood methodology used to develop the two-date 1988–1996 map (Seto et al. 2002). The second method uses econometric techniques to identify the year in which land use changes occur in a time series of images

---

[2] The Bayesian maximum likelihood classifier categorizes each pixel into a land use class based on the probability density function of each class. A discriminant function is generated from the mean and covariance statistics of the classes. A pixel is assigned to a class based on the probability of that pixel value occurring within the probability density function of that class. For a detailed description of the technique, please see Richards and Jia (1999).

(Kaufmann and Seto 2001).[3] The results indicate that both methods generate accurate estimates of the date of land use change. Although the econometric technique provided slightly less statistically biased results than the maximum likelihood method, the latter was used in this study because it simpler to employ.

Other Data

County-level data for the socioeconomic variables that we hypothesize to correlate with urbanization are obtained from Guangdong Statistical Yearbooks (Statistical Bureau of Guangdong various years). These variables are selected based on field observations and interviews with local land users. They include: demographic data, investments in capital construction, agricultural and urban output, and wage rates in different sectors. Foreign direct investments (FDI) from overseas Chinese investors affect urbanization through three channels. First, foreign investments fund much of the construction of residential and commercial complexes. Some county officials (e.g., Shenzhen and Shunde) use land subsidies and tax incentives to direct a share of FDI to light industries. Other counties (e.g., Zhuhai, Panyu, and Dongguan) encourage the construction of large-scale housing projects, often aimed at the overseas Chinese retirement market. Regardless of the sectoral allocation of investment funds (construction of apartments or factories), or the scale of the project, overseas investments are an important source of the funds required to convert land to urban uses. Although data on FDI from overseas Chinese in Hong Kong, Macau, and Taiwan are available for some of the counties in the study, they are not available for all counties for all years. To proxy FDI, we use total completed investments in capital construction (CICC). CICC is not disaggregated into foreign and domestic components, but that most of it is from overseas investors is well-documented (Xu and Li 1990; Eng 1997).

Rates of land use change also are affected by rents associated with various uses. The land market in the Pearl River Delta is poorly developed, and consequently, there is relatively little information about rents. We proxy agricultural rents by dividing a coun-

ty's value of gross agricultural output, which is available from (Statistical Bureau of Guangdong various years) by agricultural land, which is obtained from the analysis of the satellite imagery. Similarly, we proxy industrial rents by dividing a county's value of gross industrial output (Statistical Bureau of Guangdong various years), by urban land, which is obtained from the analysis of satellite imagery. Finally, we use urban and agricultural population and wages rates in the urban and agricultural sectors as measures of the social and economic structure.

IV. STATISTICAL TECHNIQUES TO MODEL THE DRIVERS OF URBAN LAND USE CHANGE

Conceptual Framework

Studies that model land use change can be divided between two broad categories: spatially explicit and aspatial. Spatially explicit models come from the disciplines of geography and landscape ecology, and usually attempt to explain the location of changes in land use as a function of the behavior of individual land users (Mertens and Lambin 1997), or of various growth "rules," as in the case of simulation models (Clarke, Hoppen, and Gaydos 1997). The unit of observation in these studies often is the individual household, pixel, or smallest landscape grain. Aspatial models have been developed from the von Thünen tradition, in which land use patterns are explained by bid-rent models that specify accessibility and distance to the markets as the major determinants of land use (Alonso 1964; von Thünen 1966). Both types of models assume that land users are rational decision-makers who choose land use from a suite of options.

---

[3] The econometric method uses the two-date technique to first identify pixels that changed between the two end dates of the study period. Following the identification of these change pixels, the methodology is a three-step procedure that uses time series and panel econometric techniques. In the first step, regression equations are estimated for each of the land use classes. In the second step, the regression equations for each class are used to calculate values for land use change classes for each of the possible dates of land conversion. In the third step, the date of land use change is identified by using tests for predictive accuracy.

Here, we estimate aspatial models of urbanization of land use change in the Pearl River Delta. We choose an aspatial model for two reasons. First, we are interested in the quantity of land use change. This focus is chosen to address concerns about the effect of land use changes on carbon stocks (Dye, Hinchliffe, and Woodcock in review), the loss of agricultural land associated with economic development, and the effect of agricultural land loss on domestic and international agricultural markets. Equally important, a spatially explicit model is unlikely to account for the market and policy environments that drive land use changes in the Pearl River Delta. While economic reforms in China have instituted usufruct rights, many of the factors that affect land use choices are exogenous to the farm, household, or village, and are beyond the control of individual land users. For example, a substantial portion of the funds for large-scale residential and commercial development projects in the Delta comes from overseas Chinese investors, many of whom invest a portion of their investments in the construction of schools, museums, and other civic buildings in their home villages. *Guanxi,* or personal connections, also play an important role in the location of projects. These factors cannot be explained in bid-rent models, or decisions by individual farmers or households.

Panel Methods to Analyze Land Use Change

A model for the relation between land use change and its socioeconomic determinants can be estimated most simply by:

$$Y_{it} = \alpha + \beta X_{it} + u_{it}, \qquad [1]$$

in which Y_{it} is the percentage of total agricultural land in a county that is converted to urban uses (*agriculture → urban*), or the percentage of total natural areas converted to urban uses (*natural → urban*) in year t, X is a matrix of variables thought to influence the rate of land conversion, μ is an error term, and α and β are regression coefficients that are estimated from the panel data.

Equation 1 can be specified using a variety of assumptions about the spatial heterogeneity of the relation between X and Y. In the simplest specification (but most restrictive), equation 1 hypothesizes that the regression coefficients (α, β) are the same for the entire period and for each county. Under these conditions, the effect of a unit increase in an explanatory variable such as per capita GDP on the conversion of agricultural land to urban uses is the same for each county for each time period.

Spatial heterogeneity and unobservable variables may cause the relation between socioeconomic variables and urban growth to vary among counties. If only intercepts vary (α_i) among counties, fixed/random effects estimators can be used to estimate equation [2]:

$$Y_{it} = \alpha_i + \beta_i X_{it} + u_{it}. \qquad [2]$$

Alternatively, both the intercepts (a_i) and the coefficients (β_i) can vary among counties. In this case, a random coefficient model can be used to estimate equation [3]:

$$Y_{it} = \alpha_i + \beta_i X_{it} + u_{it}. \qquad [3]$$

There is no way to choose among equations [1–3] *a priori*. Instead, we use standard statistical procedures to choose the specification that is consistent with the data, the proper method to estimate the equation, and the way to evaluate results (Hsiao 1986). To choose among equations 1, 2, and 3, we use F-tests that compare their residual sum of squares (RSS). If the restrictions implicit in equation [1] relative to equation [2] or equation [3] cause the residual sum of squares to increase in a statistically significant fashion, the less restrictive specification is chosen.

The results of these F-tests indicate that we can reject the null hypothesis that the intercepts and slopes are the same among counties (equation [3] vs. [1]), which means that we cannot use equation [1] to estimate the determinants of land use change (Table 1). Similarly, the F-tests indicate that we can reject the null hypothesis that the slopes are the same among counties (equation [2] vs. [3]), which means that we cannot use equa-

TABLE 1
RESULTS OF F-TESTS FOR MODEL SPECIFICATION

F-test	*agriculture → urban*	*natural → urban*
Slopes and intercepts equal	$F(50,33)$ 3.19 (0.000)	$F(40,44)$ 3.07 (0.000)
Intercepts vary, slopes equal	$F(40,33)$ 2.45 (0.002)	$F(30,44)$ 3.17 (0.003)

tion [2] to estimate the determinants of land use change (Table 1).

Based on the F-tests, we estimate equation [3] using the random coefficient model, which assumes that the value of β_i for any county is equal to a mean value for all counties (β) plus or minus some random error (λ_i):

$$\beta_i = \beta + \lambda_i. \qquad [4]$$

We use a technique developed by Swamy (1970) to estimate a mean value for β and its variance.

The regression results are evaluated using a variety of diagnostic statistics. The degree to which the estimation results are robust is evaluated using the Hendry forecast test (Curthbertson, Hall, and Taylor 1992). We use the Hendry forecast test to evaluate the null hypothesis that the in-sample and out-of-sample error variances are equal ($\sigma_1^2 = \sigma_2^2$). We compare the in-sample and out-of-sample variance for ten-county sub-samples. The test statistic is distributed as a χ^2 with degrees of freedom equal to the number of out-of-sample observations. If the parameter estimates are stable, the in- and out-of-sample variances are equal, indicating a high quality model.

We test for spatial autocorrelation by calculating Moran's I, an indicator of similarity among neighboring map cells (Griffith 1987), or in our case, bordering counties. If contiguous counties are similar to each other, Moran's I will approach 1. Alternatively, Moran's I will approach −1 if contiguous land uses are dissimilar. We calculate Moran's I values and Z statistics, which test the null hypothesis that the values of Moran's I are statistically different from 0.

Finally, we examine the residual for cointegration. Testing for cointegration is impor-

tant because spurious regressions are possible in panel data, although Phillips and Moon (1999) show that panel techniques may generate efficient estimates for the long-run relation among variables even when time series estimates for individuals do not cointegrate. To explore the properties of the individual time series and residual we use tests developed Kao (1999). These tests are designed to evaluate the results generated by a fixed effects estimator, but currently there is no test of cointegration for residuals from a random effects estimator. Consequently, conclusions about cointegration must be interpreted carefully.

Causal Order

The results of equation [3] identify correlations among variables. Successful policy depends on identifying a stronger linkage among variables, in particular, how changes in one variable can affect changes in another variable. This linkage cannot be identified statistically, but we can test the variables for relations that extend beyond simple correlations using the concept of Granger causality. We test for Granger causality in the two models of urban land use change using a procedure for panel data developed by Granger and Huang (Granger and Huang 1997). This procedure follows three steps. In the first step, we estimate restricted and unrestricted models that specify a dependent variable as a function of its own lagged values and lagged values of possible causal variables. In the second step, these restricted and unrestricted models are used to forecast out-of-sample values for Y. In the third step, we use a test of predictive accuracy to determine whether the restricted model forecasts Y as accurately than the unrestricted model.

In the first step, we analyze the time series

in both of the models (*agriculture → urban* and *natural urban*) for all possible directions of causal order. We do so by specifying the following unrestricted model:

$$Y_t = \alpha + \beta Y_{t-1} + \gamma X_{t-1} + \lambda \mathbf{Z}_{t-1} + \mu_t, \qquad [5]$$

in which X is the variable thought to "Granger cause" variable Y, $\mathbf{Z}$ is a matrix that contains all of the other variables from the equation for land use change, and μ is a normally distributed random error term. The regression coefficients α, β, γ, and λ are estimated using the technique (OLS, fixed effect estimator, random effects estimator, or the random coefficient model) as determined by the F-tests described above on the unrestricted model.

To test whether variable X "causes" variable Y in a Granger sense (Granger 1969), we estimate a restricted version of equation [5] where we impose the restriction $\gamma = 0$:

$$Y_t = \alpha + \beta Y_{t-1} + \mu \mathbf{Z}_{t-1} + \mu_t. \qquad [6]$$

This restriction tests the null hypothesis: Do lagged values of X have information about current values of Y beyond the information that is contained in the lagged value of Y and the other variables in $\mathbf{Z}$? If the restriction $\gamma = 0$ is not rejected, lagged values of X have no explanatory information about current values of Y beyond that in lagged values of Y and $\mathbf{Z}$. In this case, there is no evidence that X "Granger causes" Y. Alternatively, rejecting the restriction $\gamma = 0$ indicates that the lagged value of X has explanatory power about the current value of Y beyond the lagged values of Y and $\mathbf{Z}$. In this case, variable X is said to "Granger cause" variable Y.

Conclusions about causal order depend on the variables included in $\mathbf{Z}$. That is, X may appear to cause Y if $\mathbf{Z}$ does not include a variable that ultimately drives Y. To include all available information, $\mathbf{Z}$ includes all of the variables in the *agriculture → urban* or the *natural → urban* models other than X and Y. Including these variables limits the degrees of freedom such that we can only include one lag in equations [5] and [6].

To evaluate whether the restriction $\gamma = 0$

reduces the explanatory power of equation [6] relative to equation [5], we estimate each equation with data for 8 of the 11 counties. Using 8 of the 11 counties in-sample generates 165 combinations that can be used to estimate equations [5] and [6]. Using these 165 estimates, we generate 165 pairs of out-of-sample forecasts for the 3 out-of-sample counties.

To assess whether the unrestricted model (equation [5]) generates a more accurate prediction for Y than the restricted model (equation [6]), we use a test for predictive accuracy developed by Diebold and Mariano (Diebold and Mariano 1995). This test compares the accuracy of the two models by comparing the absolute value of the forecast errors, which is given by:

$$d_{it} = |Y_{it} - \hat{Y}_{itu}| - |Y_{it} - \hat{Y}_{itr}|, \qquad [7]$$

in which Y_{it} is the observed value of Y for county i at time t, $\hat{Y}_{itu}$ is the out-of-sample forecast of Y for county i at time t generated by the unrestricted model and $\hat{Y}_{itr}$ is the out-of-sample forecast of Y for county i at time t generated by the restricted model. The values of d_t are weighted and summed as follows to generate the S_{3a} statistic:

$$S_{3a} = \frac{\sum_{t=1}^{N} I_+ d_t) rank(|d_t|) - \frac{N(N+1)}{4}}{\sqrt{\frac{N(N+1)(2N+1)}{24}}}$$

$$I_+(d_t) + 1 \quad if \ d_t > 0$$
$$= 0 \ otherwise \qquad [8]$$

The value of the S_{3a} statistic is evaluated against a *t*-distribution. Values that exceed the critical threshold ($t = 2.06, p < .05$) indicate that the forecast error associated with unrestricted model can be distinguished from the forecast error associated with the restricted model in a statistically meaningful fashion. The model with the smaller forecast error is indicated by the sign on the test statistic. The sign is determined by the order in which the error terms are subtracted from each other. The test statistic will be negative

if the absolute forecast error associated with unrestricted model is smaller than the absolute forecast error generated by restricted model. Under these conditions, the unrestricted model is said to forecast Y more accurately than the restricted model. This would indicate that eliminating the lagged value of X from the unrestricted model reduces its explanatory power in a statistically significant fashion. In this case, we could say that X "Granger causes" Y.

The interpretation of a single out-of-sample comparison is tempered by the fact that we make 165 comparisons. To account for the repeated testing, we count the number of combinations in which the unrestricted model generates out-of-sample forecasts for Y that are more accurate than the out-of-sample forecasts for Y that are generated by the restricted model. The percentage of such occurrences is an indicator of the strength of the causal order.

V. RESULTS

We use the random coefficient model to estimate the models for the rate of *natural/urban* land use change (equation [9]) and the rate of *agriculture → urban* land use change (equation [10]):

Each of these variables is correlated with the corresponding rates of land use change in a statistically significant manner, as indicated by the *t*-statistics associated with b (Table 2). The diagnostic statistics indicate that the results are reliable. None of the values for Moran's I are statistically significant at the p < 0.05 level. At the p < .10, the values for Moran's I indicate the existence of spatial autocorrelation for three years in both models (Table 3). These results are not unexpected given the aggregate level of the countries. The reliability of the random coefficient estimates is reinforced by the results of the Hendry Forecast Test. In only one case for each model does the test statistic exceed the critical threshold. The in-sample and out-of-sample variances differ statistically (p < 0.05) when Shenzhen is excluded from the *natural → urban* model and when Dongguan is excluded from the *agriculture → urban* model (Table 4). Of the eleven counties, Shenzhen and Dongguan show the most dramatic urban growth, therefore it is not surprising that the variance of the regression equation changes when these counties are excluded from the estimation sample. Finally, the analysis of cointegration is uncertain. Univariate tests on the variables in equations [9] and [10] give mixed results regarding the

$$natural \rightarrow urban = \alpha + \beta_1 \left(\frac{\text{Completed investments in capital construction}}{\text{Population}} \right)$$

$$+ \beta_2 \log(\text{Wages in non-state, non-collective units})$$

$$\beta_3 \left(\frac{\text{Value of gross agricultural output/Agricultural land}}{\text{Value of gross industrial output/Urban land}} \right) \quad [9]$$

$$agriculture \rightarrow urban = \alpha + \beta_1 \left(\frac{\text{Value of gross agricultural output}}{\text{Agricultural population}} \right)$$

$$+ \beta_2 \left(\frac{\text{Completed investments in capital construction}}{\text{Population}} \right)$$

$$+ \beta_3 \log(\text{Average total wages})$$

$$+ \left(\frac{\text{Value of gross agricultural output/Agricultural land}}{\text{Value of gross industrial output/Urban land}} \right) \quad [10]$$

TABLE 2

STATISTICAL RESULTS FROM PANEL ANALYSIS

Variable	Agriculture → Urban[a]	Natural → Urban[b]
Constant	−0.109	0.028
	(−4.95)	(3.13)
$\dfrac{\text{Value of gross agricultural output/Agricultural lan}_a}{\text{Value of gross industrial output/Urban land}}$	1.496	6.53
	(1.97)	(3.55)
$\dfrac{\text{Value of gross agricultural output/Agricultural population}_b}{\text{Value of gross industrial output/Nonagricultural population}}$		−1.39
		(−3.37)
$\dfrac{\text{Value of gross agricultural output}_c}{\text{Agricultural population}}$	−3.965	
	(−4.58)	
$\dfrac{\text{Completed investments in capital construction}_d}{\text{Population}}$	1.24	2.85
	(2.74)	(3.27)
Log(Average total wage)[e]	0.032	
	(5.98)	
DF_ρ^*	−3.018[†]	−1.79[a]
DF_t^*	−2.68[†]	−1.46[#]

[a] Variable name: RATIO2
[b] Variable name: RATIO1
[c] Variable name: GVACAP
[d] Variable name: CCCAP
[e] Variable name: LOGWAGE

[†] Exceeds .01 threshold
[*] Exceeds .05 threshold
[#] Exceeds .10 threshold

TABLE 3

RESULTS OF TESTS FOR SPATIAL AUTOCORRELATION

Years	*natural → urban*		*agriculture → urban*	
	Moran's I	Z	Moran's I	Z
1988–1989	0.154	1.313	−0.162	−0.325
1989–1990	−0.005	0.490	**−0.465**	**−1.784**
1990–1991	−0.069	0.155	**−0.460**	**−1.792**
1991–1992	0.015	0.763	−0.052	0.276
1992–1993	0.141	1.239	−0.144	−0.208
1993–1994	**0.224**	**1.640**	−0.322	−1.053
1994–1995	0.037	0.828	−0.210	−0.547
1995–1996	0.116	1.291	−0.101	−0.006

Note: Values in bold are significant at $p < 0.10$

presence of a unit root while the test of cointegration reject the null of no cointegration at $p < .05$. But this result is uncertain because the tests are designed to analyze residuals from fixed effects estimators, not random coefficient models.

VI. DISCUSSION

As indicated by the results for equation [10], the annual rate at which agricultural land is converted to urban uses is correlated with investment in capital construction, the

TABLE 4
RESULTS OF HENDRY FORECAST TESTS

County excluded	*agriculture → urban* Test Statistic $\chi^2(1)$	*natural → urban* Test Statistic $\chi^2(1)$
Jiangmen	2.634	0.202
Sanshui	1.547	0.026
Shunde	0.772	0.029
Nanhai	0.928	0.001
Foshan	1.498	0.006
Dongguan	**5.228**	2.187
Shenzhen	1.595	**5.218**
Panyu	0.163	0.807
Zengcheng	2.849	0.560
Huadu	1.097	0.344
Guangzhou	1.914	2.215

Note: Values in bold are significant at $p < 0.05$

productivity of land productivity in agriculture versus urban uses, agricultural labor productivity, and off-farm wage rates. As expected, the sign on the coefficient associated with investment in capital construction is positive. This is consistent with observations that much of the capital construction in the Pearl River Delta is directed at residential/industrial structures and new roads. The sign on the coefficient associated with returns to agricultural land relative to industrial uses also is positive. This result indicates that the most productive agricultural lands are being converted to urban uses. This pattern is driven in part, by migration. Migrants tend to move to the areas with high farmland productivity. Consequently, urban centers emerge from areas where people first moved—land with highest productivity. The increased population puts positive pressure on urbanization. The conversion of agricul-

tural land is negatively related to agricultural labor productivity. This variable is an indicator of agricultural wages. As on-farm income increases, the opportunity cost of converting farmland increases, and farmers are less inclined to convert agricultural land to nonfarm uses. Finally, average off-farm wages have a positive effect on agricultural land conversion.

The analysis of causal order among the variables in the *agriculture → urban* model (equation [10]) is not conclusive (Table 5). The independent variables in equation [10] "Granger cause" the conversion of agricultural land to urban uses in relatively few of the 165 sub-samples (Table 5). At the $p < .05$ threshold, random chance should generate about eight instances (8/165 = .05) in which the test statistic exceeds the critical. There is some indication that investment in capital construction 'Granger causes' the

TABLE 5
ANALYSIS OF CAUSAL ORDER FOR THE *Agriculture → Urban* Equation

	Dependent Variable				
Causal Variable	AG → URBAN	CCCAP	RATIO2	LOGWAGE	GVACAP
AG > URBAN	—	12	14	3	6
CCCAP	19	—	10	18	8
RATIO2	0	33	—	2	17
LOGWAGE	0	18	11	—	24
GVACAP	5	26	14	16	—

Note: Values refer to the number of combinations in which the out-of-sample forecast for the change from agriculture to urban land generated by the restricted model is less accurate than the out-of-sample forecast generated by the unrestricted model.

conversion of agricultural land to urban uses—the out-of-sample predictive accuracy of the restricted version of equation [6] is inferior in 19 of 165 instances. On the other hand, it does not appear as though we have reversed the direction of causality. There are relatively few cases in which the *agriculture → urban* variable "Granger causes" any of the independent variables in equation [10]. Together, these results imply that the conversion of agricultural land to urban uses ultimately may be driven by some exogenous variable(s) that is not included in the model due to the lack of information.

As indicated by the results for equation [9], the annual rate at which forests, shrubs, and water bodies are converted to urban uses is correlated with investment in capital construction, the return to agricultural land relative to industrial uses, and labor productivity in agriculture relative to industry. Again, investment in capital construction tends to accelerate the rate at which natural ecosystems are converted to urban uses. Despite the similar signs, the interpretation of the relative land productivity variable is different from the *agriculture → urban* model (equation [10]). For this model, we interpret relative land productivity as a proxy for the opportunity cost of land conversions. For counties where the productivity of agricultural land is high relative to urban uses, the opportunity cost associated with converting agricultural land is high relative to converting forest and water. This higher opportunity cost tends to favor the conversion of natural ecosystems, and hence, a positive relationship with natural land conversion. Relative labor productivity is defined as the ratio of agricultural labor

productivity to industrial labor productivity. If agricultural labor productivity is high relative to industrial labor productivity, there is little incentive to convert natural ecosystems to industrial uses. Rather, higher relative agricultural labor productivity suggests that agricultural land produces greater returns than industrial land. Therefore, there is less incentive to urbanize natural areas. On the contrary, there may be more pressure to convert natural areas to agricultural land to take advantage of the relative high labor productivity.

The analysis of causal order among the variables in the *natural → urban* (equation [9]) is somewhat conclusive (Table 6). Two of the independent variables in equation [9], investment in capital construction and relative rates of labor productivity, appear to "Granger cause" the conversion of natural ecosystems to urban uses in 21 of the 165 sub-samples. This rate of change appears to feedback on investment in capital construction. The *natural → urban* variable "Granger causes" investment in capital construction in 32 of the 165 sub-samples. This result would indicate that high rates of conversion attract additional investments. On the other hand, there is little evidence for a causal relation between the *natural → urban* variable and the other two variables on the right hand side of equation [9].

The results from the models verify the difference between the dynamics of agricultural land conversion and conversion of natural vegetation and water. The relative labor productivity ratio serves as a good proxy for how wage differentials in the agricultural and industrial sectors affect land conversion. If

TABLE 6
ANALYSIS OF CAUSAL ORDER FOR THE *Natural → Urban* Equation

Causal Variable	Dependent Variable			
	NATURAL > URBAN	CCCAP	RATIO2	RATIO1
NATURAL > URBAN	—	32	9	10
CCCAP	21	—	1	6
RATIO2	8	14	—	23
RATIO1	21	17	14	—

Note: Values refer to the number of combinations in which the out-of-sample forecast for the change from natural vegetation to urban land generated by the restricted model is less accurate than the out-of-sample forecast generated by the unrestricted model.

agricultural labor productivity is high relative to industrial labor productivity, farmers and villages will have less incentive to convert their land to urban uses directly or indirectly through leases. This is not surprising. If farmers and villages receive relatively high returns on agriculture relative to industry, there is little incentive to convert natural vegetation for industrial use. Rather, the difference in the return on labor would encourage the conversion of natural vegetation to agricultural areas. Only once industrial labor productivity exceeds that of agricultural land productivity would land users convert natural areas to industrial uses.

The difference in interpretation of the relative land productivity variable for the two models can be viewed as the effects of opportunity costs versus proximity. In the model of agricultural land conversion, the variable for relative land productivity should be interpreted as a measure of distance to urban activity. Generally, urban development occurs first in regions with the highest land productivity. This is a direct result of population movements. Usually, people inhabit areas with the highest land productivity and therefore land conversion is positively correlated with land productivity. Based on this effect, the ratio of land productivity is an indirect measure of proximity to urban areas.

The econometric results for equations [9] and [10] and the causal relations among these variables suggest several options for slowing urban encroachment of agricultural land, should policymakers make this a priority. First, policymakers can subsidize agricultural production. This would raise the returns to agricultural land, which could slow conversion. Currently farmers and village leaders lease or convert their land to urban uses because this option generates greater revenue streams. One cause for the relatively poor returns to agriculture may be that the household responsibility system divided land into plots based on family size. This fragmentation discourages the production of some high value crops (e.g., bananas and lychees) because they require larger tracts of land compared to lower value crops (e.g., sweet potatoes and rice). This effect may prompt families with smaller plots to forego agri-

cultural production altogether, and lease their land to commercial interests. Land use planners in the Guangdong Province have cracked down on the leasing of agricultural land for these uses. However, they remain ineffective in controlling all transactions of this nature.

Second, the land market should be reformed. Currently, land exchanges occur in two markets. The first is a land use rights transfer that grants usufruct rights by auction. Applicants have usufruct rights for a fixed period after a bid is made on the open market. These lump-sum payments can reflect the market values of the potential land uses. In the second land transfer mechanism, land leases are obtained through private negotiations. These negotiations can allow the government to target industries and projects through land subsidies. Many of these subsidies in the early 1990s inflated the return on investment for projects that had inadequate funding, business expertise, or strategic planning. This in turn led to a number of half-completed development, largely residential, projects throughout the Delta. Although there are no official estimates of the fraction of development projects that are abandoned, field observations and interviews confirm that the numbers are likely to be high. When developers first offered to finance large projects, village leaders (largely composed of farmers) were quick to negotiate contracts to lease out their land. These lease agreements usually involved the payment of one lump sum, rather than a steady income stream of rents. Therefore, while farmers were able to compare current agricultural prices to lease rents, they did not consider future revenues from agricultural income stream.

Third, the feedback loop that includes income growth, consumer demand, and urban expansion indicates that higher incomes have fuelled a shift from agricultural to non-agricultural livelihoods. Similarly, township and village governments have been agents of urbanization within their levels of jurisdiction by facilitating development of town village enterprises in the region and by endorsing local efforts cluster privately owned firms through the development of industrial

estates and technology corridors. The results indicate that large-scale urbanization in the Pearl River Delta has been caused largely by exogenous factors, such as international capital movements. The models suggest and field interviews verify that local land users do not have much influence over the siting of large development projects.

VII. CONCLUSION

In this paper, remotely sensed land use data are integrated with socioeconomic data to develop statistically meaningful models of urbanization. Important economic factors of agricultural land, natural vegetation, and water conversion are identified and quantified. The ratio of agricultural land productivity and industrial land productivity is one of the essential variables that cause the conversion of both natural ecosystems and agricultural land. Investments in capital construction per capita, a proxy for foreign direct investment, is also identified as a driver of both types of urban land use change. We also find that investments in capital construction are driven also by changes in agricultural land use. The ratio of agricultural to industrial labor productivity are correlated with conversion of natural areas. However, only agricultural labor productivity, not the ratio of agricultural to industrial labor productivity, is correlated with agricultural land conversion. Wages and income are not correlated with conversion of natural areas, but are correlated with conversion of agricultural areas. The results also indicate that the causality of these relations is neither clearly evident nor unidirectional.

References

Alonso, W. 1964. *Location and Land Use.* Cambridge: Harvard University Press.

Brown, L. R. 1995. *Who Will Feed China? Wake-up Call for a Small Planet.* New York: W. W. Norton and Co.

Chomitz, K. M., and D. A. Gray. 1996. "Roads, Land Use, and Deforestation: A Spatial Model Applied to Belize." *The World Bank Economic Review* 10 (3): 487–512.

Chow, G. C. 1994. *Understanding China's Economy.* Singapore: World Scientific.

Chu, D. K. Y. 1998. "Synthesis of Economic Reforms and Open Policy." In *Guangdong: Survey of a Province Undergoing Rapid Change,* ed. Y. M. Yeung and D. K. Y. Chu, 485–504. Hong Kong, Chinese University Press.

Clarke, K. C., S. Hoppen, and L. Gaydos. 1997. "A Self-Modifying Cellular Automation Model of Historical Urbanization in the San Francisco Bay Area." *Environment and Planning, B* 24:247–61.

Cropper, M., C. Griffiths, and M. Mani. 1999. "Roads, Population Pressures, and Deforestation in Thailand, 1976–1989." *Land Economics* 75 (Feb.): 58–73.

Cuthbertson, K., S. G. Hall, and M. P. Taylor. 1992. *Applied Econometric Techniques.* Ann Arbor: University of Michigan Press.

Diebold, F. X., and R. S. Mariano. 1995. "Comparing Predictive Accuracy." *Journal of Business and Economics Statistics* 13:253–63.

Dye, D. G., T. C. Hinchliffe, and C. E. Woodcock. In review. "Modeling the Effects of Recent Land Use Change on the Carbon Cycle in the Zhu Jiang Delta Region of Southern China." *Global Biogeochemical Cycles.*

Eng, I. 1997. "The Rise of Manufacturing Towns: Externally Driven Industrialization and Urban Development in the Pearl River Delta of China." *International Journal of Urban and Regional Research* 21:554–68.

Geoghegan, J., S. C. Villar, P. Kelpeis, P. M. Mendoza, Y. Ogneva-Himmelberger, R. R. Chowdhury, I. B. L. Turner, and C. Vance. 2001. "Modeling Tropical Deforestation in the Southern Yucatán Peninsular Region: Comparing Survey and Satellite Data." *Agriculture, Ecosystems and Environment* 85 (1–3): 25–46.

Granger, C. W. J. 1969. "Investigating Causal Relations by Econometric Models and Cross-Spectral Models." *Econometrica* 37:424–38.

Granger, C. W. J., and L. Huang. 1997. "Evaluation of Panel Data Models: Some Suggestions from Time Series." University of California, San Diego, Department of Economics Working Paper Series.

Griffith, D. A. 1987. *Spatial Autocorrelation: A Primer.* Washington, D.C.: The Association of American Geographers.

Grübler, A. 1994. "Technology." In *Changes in Land Use and Land Cover: A Global Perspective,* ed. W. B. Meyer and B. L. Turner II. New York: Cambridge University Press.

Hsiao, C. 1986. *Analysis of Panel Data.* New York: Cambridge University Press.

Kao, C. 1999. "Spurious Regression and Residual-Based Tests for Cointegration in Panel Data." *Journal of Econometrics* 90 (1): 1–44.

Kaufmann, R. K., and K. C. Seto. 2001. "Change Detection, Accuracy, and Bias in a Sequential Analysis of Landsat Imagery in the Pearl River Delta, China: Econometric Techniques." *Agriculture, Ecosystems and Environment* 85 (1–3): 95–105.

Lin, G. C. S. 1997. *Red Capitalism in South China: Growth and Development of the Pearl River Delta.* Vancouver: UBC Press.

Lin, J. Y. 1988. "The Household Responsibility System in China's Agricultural Reform: A Theoretical and Empirical Study." *Economic Development and Cultural Change* 36:S199–S224.

Meng, X., and J. S. Zhang. 2001. "The Two-Tier Labor Market in Urban China: Occupational Segregation and Wage Differentials Between Urban Residents and Rural Migrants in Shanghai." *Journal of Comparative Economics* 29 (3): 485–504.

Mertens, B., and E. F. Lambin. 1997. "Spatial Modeling of Deforestation in Southern Cameroon." *Applied Geography* 17:143–62.

Nelson, G. C., and D. Hellerstein. 1997. "Do Roads Cause Deforestation? Using Satellite Images in Econometric Analysis of Land Use." *American Journal of Agricultural Economics* 79 (1): 80–88.

Pfaff, A. S. P. 1999. "What Drives Deforestation in the Brazilian Amazon? Evidence from Satellite and Socioeconomic Data." *Journal of Environmental Economics and Management* 37:26–43.

Phillips, P. C. B., and H. R. Moon. 1999. "Linear Regression Limit Theory for Nonstationary Panel Data." *Econometrica* 67 (5): 1057–1111.

Richards, J. A., and X. Jia. 1999. *Remote Sensing Digital Image Analysis: An Introduction.* New York: Springer Verlag.

Sadownik, B., and M. Jaccard. 2001. "Sustainable Energy and Urban Form in China: The Relevance of Community Energy Management." *Energy Policy* 29 (1): 55–65.

Seto, K. C., R. K. Kaufmann, and C. E. Woodcock. 2000. "Landsat Reveals China's Farmland Reserves, but They're Vanishing Fast." *Nature* 406:121.

Seto, K. C., C. E. Woodcock, C. Song, X. Huang, J. Lu, and R. K. Kaufmann. 2002. "Monitoring Land-Use Change in the Pearl River Delta Using Landsat TM." *International Journal of Remote Sensing* 23 (10): 1985–2004.

Sharkawy, M. A., X. Chen, and F. Pretorius. 1995. "Spatial Trends of Urban Development in China." *Journal of Real Estate Literature* 3 (1): 47–59.

Smil, V. 1993. *China's Environmental Crisis: An Inquiry into the Limits of National Development.* Armonk, N.Y.: M. E. Sharpe.

Smith, C. J. 2002. "The Transformative Impact of Capital and Labor Mobility on the Chinese City." *Urban Geography* 21 (8): 670–770.

Song, C., C. E. Woodcock, K. C. Seto, M. Pax Lenney, and S. A. Macomber. 2001. "Classification and Change Detection Using Landsat TM Data: When and How to Correct Atmospheric Effects?" *Remote Sensing of Environment* 75:230–44.

Statistical Bureau of Guangdong. various years. *Statistical Yearbook of Guangdong.* Beijing: China Statistical Publishing House.

Swamy, P. A. V. B. 1970. "Efficient Inference in a Random Coefficient Model." *Econometrica* 38:311–23.

United Nations. 2001. *World Population Prospects: The 2000 Revision.* New York.

Vitousek, P. M., H. A. Mooney, J. Lubchenco, and J. M. Melillo. 1997. "Human Domination of Earth's Ecosystems." *Science* 277 (5325): 494–99.

von Thünen, J. 1966. *Isolated State.* New York: Pergamon Press.

Wang, T. H., A. Maruyama, and M. Kikuchi. 2000. "Rural-Urban Migration and Labor Markets in China: A Case Study in a Northeastern Province." *Developing Economies* 38 (1): 80–104.

Xu, X.-Q., and S.-M. Li. 1990. "China's Open Door Policy and Urbanization in the Pearl River Delta Region." *International Journal of Urban and Regional Research* 14:49–69.