Urban Studies

http://usj.sagepub.com

The Mills Muth Model of Urban Spatial Structure: Surviving the Test of Time?

Christy Spivey

Urban Stud 2008; 45; 295

DOI: 10.1177/0042098007085964

Urban Studies
45(2) 295–312, February 2008

The Mills–Muth Model of Urban Spatial Structure: Surviving the Test of Time?

Christy Spivey

[Paper first received, September 2005; in final form, April 2007]

Abstract

This paper examines the viability of some basic predictions of the Mills–Muth model of city structure for modern cities using US data for the year 2000. The estimation strategy used to test the predictions is very similar to that of Jan Brueckner and David Fansler, who use 1970 data to find support for the model's basic comparative statics predictions—namely, that city area is increasing in population and income but decreasing in agricultural land value and commuting costs. This paper uses different measures for land values and commuting costs where possible and a measure of polycentricity to estimate a slightly modified empirical model. Despite the changing structure of cities, there is evidence that the Mills–Muth comparative statics predictions hold for modern US cities, that densely populated cities are more likely to have sub-centres and that market forces drive urban spatial structure.

> In fact, I believe that the remarkable fact is not that the chimp types so badly, but that it types at all; the broad predictions from the simple models remain more accurate than I would have expected, given the massive dispersion of employment in U.S. metropolitan areas and the pervasiveness in the U.S. of fragmented local government jurisdictions (Mills, 2000, p. 18).

I. Introduction

Does the chimp still really type? To what extent do the very broadest predictions of the original Mills–Muth model of urban spatial structure apply today? The statement above by Mills coupled with an empirical study by Brueckner and Fansler (1983), which finds that the most general implications do hold for a sample of 1970 urbanised areas, provoke these questions. Since urban spatial structure analysis grows extremely complicated with efforts to add increased realism to the models, and many researchers in urban economics and related areas still evoke implications of the Mills–Muth model in their work, it seems useful to know if the basic model still applies at the city level. Given the changing nature of cities over recent decades, especially the increased polycentricity of cities and the less

Christy Spivey is in the Department of Economics and Finance, Southern Illinois University at Edwardsville, Campus Box 1102, Edwardsville, Illinois 62026, USA. Fax: 618 650 3047.
E-mail: cspivey@siue.edu.

0042-0980 Print/1360-063X Online

DOI: 10.1177/0042098007085964

predictable commuting patterns, one might be sceptical that the model does in fact still hold substantial predictive power. Mills (2000) himself argues that the biggest failure of the model is the predicted location of all businesses in adjoining space in the CBD despite the fact that only about 10 per cent of metro-area employment in the 1990s was located in the CBD in some US cities.

McMillen (2006) provides a thorough review of the various ways that the Mills–Muth model has been empirically tested and an argument that, despite the changing nature of cities and a general consensus that the basic model is no longer accurate, the monocentric city model is still the dominant model of urban structure. Attempts to estimate directly the predictions of the Mills–Muth model fall into two categories: studies that look at one city at a time and try to determine whether the price of a housing unit, the capital–land ratio, land values and population density all decrease with increasing distance from the central city; and, studies that test the model's comparative statics predictions, that city area is increasing in population and income but decreasing in agricultural land value and commuting costs, with a cross-section of cities or by looking at one city over time. The comparative statics approach using a cross-section of cities is uncommon; in fact, the Brueckner and Fansler (1983) study is the only one that analyses more than a handful of cities. They compare 40 urbanised areas using 1970 data and, despite using measures of commuting cost that certainly are not perfect, they find strong support for the model's predictions.

What follows is in part an update to the Brueckner and Fansler (1983) study. I estimate a slightly modified version of their empirical model in order to test comparative statics results of the Mills–Muth model. However, while they test these implications with 1970 data for a relatively small sample of urbanised areas due to data limitations, I attempt to overcome these limitations and test the predictions for all US urbanised areas using 2000 data. I also use different measures of commuting costs, where available, and address the changing nature of cities with data on commuting patterns and the polycentricity of cities. The results suggest that the monocentric model of the city still has predictive power in the year 2000. The next section of the paper discusses the implications of the model tested by Brueckner and Fansler (1983) as well as their findings. Section 3 discusses the data used in the empirical analysis, section 4 presents the empirical model and the results, and section 5 contains a discussion and the concluding remarks.

2. Implications and a Test of the Mills–Muth Model

The simple Mills–Muth model, as outlined by Brueckner and Fansler (1983), assumes that consumers have the same income I at the CBD and have identical preferences over housing (residential lot size), q, and a composite numeraire good, z. Housing rents for price $p(x)$ per unit, where p depends on distance x from the CBD. Consumers also face a commuting cost t per round-trip mile and maximise utility subject to a budget constraint

$$\underset{z,q,x}{\text{Max}}\, U(z,q) \text{ s.t. } I = z + p(x)q + tx \qquad (1)$$

Because consumers are free to move around and p varies with x, an implication of the model is that, in equilibrium, all consumers reach the same utility level u. To keep a consumer indifferent between any two given locations, the price of housing must be lower at the location that is farthest from the CBD. Inputs to housing are assumed to be capital and land, with a constant returns to scale housing production function. Producers maximise profit per unit of land, $ph(K)-iK-r$, where r is land rent, i is the rental price of capital, K is capital per unit of land and h is amount of housing per unit of land.

If population density is defined as $D(x, t, I, u) \equiv h(K)/q$, then the equilibrium for the city can be written as

$$r(\bar{x}, t, I, u) = r_a \tag{2}$$

$$\int_0^{\bar{x}} 2\pi x D(x, t, I, u) dx = N \tag{3}$$

where, $\bar{x}$ is the distance to the urban edge, r_a is the agricultural land rent and N is the urban population. Equation (2) is an arbitrage condition, which indicates that urban land rents must equal agricultural land rents at the urban edge. Equation (3) simply states that the urban population must be accommodated inside the city boundary. The following comparative statics results, first derived by Wheaton (1974) and requiring a utility function such that both goods are normal and have positive income effects, are the ones tested in the empirical estimation

$$\frac{\partial \bar{x}}{\partial N} > 0, \frac{\partial \bar{x}}{\partial I} > 0, \frac{\partial \bar{x}}{\partial r_a} < 0, \frac{\partial \bar{x}}{\partial t} < 0 \tag{4}$$

As population increases, so does the radius of the city. A higher level of income increases city size as demand for housing increases. A city becomes smaller with an increase in the value of agricultural land, which increases the opportunity cost of urban land. On the other hand, an increase in the commuting cost decreases city size because of the income effect and hence less housing demand.[1] It should be noted that the model presented here, in its simplest form, does not account for time costs of commuting. As McMillen (2006) points out, expanding the model to account for time costs leads to an ambiguous comparative statics prediction for income. While an increase in income leads residents to prefer living farther from the central city due to an increase in demand, it also increases the opportunity cost of time spent commuting, making housing closer to the central city more desirable. Thus, the net effect is ambiguous, although typically empirical studies find what the basic model predicts—that an increase in income leads to a larger city size.[2]

Brueckner and Fansler (1983) use a Box–Cox specification, with a single transformation parameter applied to both dependent and independent variables, and show that an urbanised area's total land area (and hence distance from the city centre to the urban–rural boundary) increases with population and income but decreases with agricultural land values. They use data from the 1970 US census and their dataset consists of only 40 urbanised areas with 1970 populations that range from 52 000 to 257 000. Because data on agricultural land values are available only by county, their sample includes only urbanised areas contained within a single county in an effort to measure accurately land values adjacent to the developed portion of the city. However, this clearly neglects a large number of urbanised areas. In addition, their proxies for commuting cost have no significant effect upon land area, although the coefficients are negative as expected. The two proxies are the percentage of commuters using public transport and the percentage of households owning at least one automobile. The intuition behind these proxies is that high levels of automobile usage and low levels of public transport usage indicate a low cost of commuting per mile. The former hopefully indicate low congestion levels, holding income constant, while the latter are associated with a high time cost per mile. Overall, their results, reproduced in Appendix 2, support the simplest predictions of the basic Mills–Muth model. They argue this is evidence that city size is determined by an organised, market-driven allocation of land use, not uncontrolled sprawl.

3. Data

The data employed are from the 2000 US Census of Population and Housing, the Texas Transport Institute and both the 1997 and 2002 US Censuses of Agriculture. I also use McMillen and Smith's (2003) estimated number of sub-centres in an urbanised area. The unit of observation is the United States urbanised area (UA) (see Appendix 3 for relevant Census definitions). The urbanised area is used as opposed to the metropolitan statistical area (MSA) because the boundaries are less artificial. While the physical shape of an MSA is defined by county boundaries, the shape of an urbanised area is driven to a larger extent by market forces and where people choose to work and live. Moreover, while the model assumes that agricultural land is adjacent to but outside the circular city's boundary, an MSA is much more likely to contain agricultural land than is an urbanised area. The sample used here consists of all 452 urbanised areas in the US as of 2000, which have populations ranging from just under 50 000 to over 17 million.

The measure used for agricultural land values is calculated using the 2000 Census and the 1997 and 2002 Censuses of Agriculture. Two measures of agricultural land values are available, the market value of agricultural products sold per acre and the estimated market value of agricultural land and buildings per acre. It is unclear which variable Brueckner and Fansler (1983) used, as they call their measure simply the agricultural land value per acre. Nevertheless, the measure that excludes buildings would seem to coincide better with the requirements of the theoretical model. Land values are available by county, but many urbanised areas have land area in more than one county. Thus, it is necessary to find a way to compute one land value per urbanised area in order to be consistent with an assumption of the model—namely, that agricultural land rent is constant beyond the urban-rural boundary. The Census of Population and Housing contains information on the percentage of an urbanised area's total land area that is comprised by any given county. The Census of Agriculture provides the value of agricultural land and a value for 2000 is imputed from the two different years available for the census based on the annual growth rate. Then, the land value for each county that makes up part of an urbanised area is weighted by the percentage of the urbanised area that falls in that county. The result is a weighted average land value for each urbanised area in dollars per acre. While it would be ideal to have the length of the urbanised area's boundary that falls in each county, such data are not readily available.

Measuring commuting cost is most problematic. There are perhaps no measures better than the ones used by Brueckner and Fansler (1983) available for all urbanised areas. The only other possibility that the census offers is the average commute time or the percentage of workers whose commute lasts more than a certain amount of time. Although it is not a monetary cost, longer commutes will be positively correlated with monetary costs and opportunity costs. In addition, the Texas Transport Institute provides several possible measures of commuting cost, but only for 85 large urbanised areas. One is a travel time index, which is a measure of congestion during peak periods. More specifically, it is the ratio of the travel time during the peak period to the time required to make the same trip at free-flow speeds. Another possibility is the thousands of miles travelled per day by vehicles per mile of freeway lane. This captures commuting cost at all times of day, whereas the travel time index does so for peak periods of congestion. Finally, the institute also calculates a monetary cost of congestion, measured as the value of travel delay and extra fuel consumed in traffic congestion. Delay is the extra travel time compared with some standard, in this case 65 miles per hour on freeways and 30 miles per hour on city streets.

The income measure, available from the 2000 Census, is simply the median family income in the urbanised area. Other useful measures are available to address the increasing polycentricity of urbanised areas and investigate how this might affect the empirical results. The census does have information on the percentage of workers living in an MSA who work in the central city of the same MSA. In addition, McMillen and Smith (2003) have estimated the number of sub-centres in over 60 areas using commuting costs from the Texas Transport Institute. They identify sub-centres as local peaks in the predictions from non-parametric regressions of employment density on distance from the city centre. Table 1 presents some basic descriptive statistics of the key variables.

4. Empirical Model and Results

Since data requirements are not sufficient to allow non-parametric estimation, a Box–Cox equation is estimated, where the area of the urbanised area in square miles is related to population, average agricultural land value, median family income and a measure of commuting costs. Allowing for some form of non-linearity seems reasonable. For example,

Table 1. Basic descriptive statistics

	Mean	*Minimum*	*Maximum*	*Standard deviation*
All urbanised areas ($N = 452$)				
Area (square miles)	159	12	3,353	302
Population	425 527	49 776	17 800 000	1 245 656
Income	49 359	25 967	91 741	10 234
Land value per acre	650	2	7,685	836
Fraction of HH owning at least 1 car	0.91	0.68	0.97	0.03
Fraction using public transport	0.02	0	0.30	0.03
Average travel time	22.89	14.70	42.49	4.89
Number of sub-centres	0.36	0	33	2.29
On coast, Great Lake, Mexican border	0.18	0	1.00	0.38
Percentage working in central city	0.48	0	0.97	0.26
Largest urbanized areas (with data from Texas Transportation Institute) ($N = 85$)				
Area (square miles)	536	33	3,353	547
Population	1 686 604	112 331	17 800 000	2 511 807
Income	52 597	25 967	81 226	9 235
Land value per acre	999	12	7 685	1 310
Fraction of HH owning at least 1 car	0.90	0.68	0.95	0.04
Fraction using public transport	0.04	0.01	0.30	0.04
Average travel time	25.27	19.15	37.05	3.58
Congestion Cost ($millions)	618	7	10 358	1 354
Travel time index	1.21	1.04	1.76	0.13
Vehicle miles of travel per freeway mile	13 780	5 533	22 999	3 110
Number of sub-centres	1.92	0	33	5.01
On Coast, Great Lake, Mexican border	0.32	0	1.00	0.47
Percentage working in central city	0.53	0.23	0.97	0.20

the land area of an urbanised area might increase with population at a decreasing rate. Similar suppositions can be made for the other covariates. Thus, given a vector x of positive covariates, the model to be estimated, which incorporates transformations on the dependent as well as the independent variables,[3] takes the following form

$$y^{(\theta)} = \alpha + \sum_{k=1}^{4} \beta_k x_k^{(\lambda)} + \varepsilon \tag{5}$$

where,

$$y^{(\theta)} = \begin{cases} \dfrac{y^{\theta} - 1}{\theta} & \text{when } \theta \neq 0 \\ \log(\mathrm{y}) & \text{when } \theta = 0 \end{cases}$$

and

$$x^{(\lambda)} = \begin{cases} \dfrac{x^{\lambda} - 1}{\lambda} & \text{when } \lambda \neq 0 \\ \log(\mathrm{x}) & \text{when } \lambda = 0 \end{cases}$$

Estimation via maximum likelihood is fairly straightforward, partly since the log-likelihood function incorporates a Jacobian term that prevents θ from becoming too small. Assuming that $\varepsilon \sim \mathrm{N}(0, \sigma^2)$, maximisation of the following log-likelihood function via the full-information method with respect to β, θ and λ yields consistent estimates and is asymptotically efficient

$$\begin{aligned} \ln L = & -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) \\ & + (\theta - 1)\sum_{i=1}^{n} \ln y_i \\ & - \frac{1}{2\sigma^2}\sum_{i=1}^{n} (y_i^{(\theta)} - x_i^{(\lambda)}\beta)^2 \end{aligned} \tag{6}$$

Consistent standard errors are calculated via the Berndt–Hall–Hall–Hausman (BHHH) method, which guarantees a non-negative definite Hessian as long as the number of observations is greater than the total number of parameters to be estimated.

Table 2 presents the maximum likelihood estimates using the complete sample of 452 urbanised areas and one of the proxies for commuting costs used by Brueckner and Fansler (1983), the percentage of households owning at least one vehicle. A dummy variable, not transformed, equal to one if the urbanised area borders a Great Lake, Mexico or any coast, is included, since such geographical constraints can limit growth that might occur in their absence.

The results of the estimation of Specification (a) are consistent with the comparative statics predictions of the Mills–Muth model. However, only the coefficients on population and income are statistically significant at above the 10 per cent level. The coefficient on income is negative, however, which indicates that, on average, the effect of the increasing opportunity cost of time as income rises may outweigh that of increased demand for more affordable housing farther from the central city. The coefficient on the percentage of households owning at least one car, while positive, is not quite significant at the 10 per cent level. Moreover, the percentage of households owning at least one car is certainly not an optimal measure of commuting costs, although it may have been a better measure in the 1970s than it is today. The results may simply be reflecting that, in larger cities, people are more likely to need a car to get to their destination. Bordering a large body of water or Mexico does not have a significant effect on land area, although the coefficient is positive, so that these cities are larger than average.

Specification (b) includes as a covariate the number of sub-centres in the urbanised area, as estimated by McMillen and Smith (2003). The rationale for including the number of sub-centres is to account for the increasing polycentricity of urbanised areas, since the predictions of the Mills–Muth model are

Table 2. Maximum likelihood estimates (dependent variable: area in square miles)

	Specification (a)		*Specification (b)*		*Specification (c) double-log model*	
	Estimate	*X^2 statistic*	*Estimate*	*X^2 statistic*	*Estimate*	*T-statistic*
Constant	–3.36		–1.56		–2.09	–2.02**
Population	1.06	925.76***	0.65	900.20***	0.91	54.46***
Income	–0.43	15.71***	–0.27	15.26***	–0.39	-3.89***
Land value per acre	–0.04	3.64*	–0.03	4.33**	–0.03	–1.63*
Fraction of HH owning at least 1 car	0.91	1.72	0.82	1.21	0.63	1.22
Number of sub-centres			–0.03	3.56*	–0.02	–2.79***
On Coast, Great Lake, Mexican border	0.05	0.70	0.06	1.10	0.04	0.97
Transformation parameters		*Z-statistic*		*Z-statistic*		
θ	0.07	2.17**	0.08	2.49***		
λ	0.01	0.44	0.06	1.50		
Number of observations	452		452		452	
Log-likelihood/R^2	–2104.89		–2103.10		0.90	

***indicates significance at the 1 per cent level; **at the 5 per cent level; and *at the 10 per cent level.

based on a monocentric model. However, once the number of sub-centres is included, there are few differences in the results. The signs of the coefficients remain the same, though the magnitudes, in terms of absolute value, for most of them are reduced. Part of the reason that few differences between Specifications (a) and (b) are seen may be because the number of sub-centres could only be estimated for just over 60 urbanised areas. Although these 60-plus urbanised areas are for the most part the most populous urbanised areas and thus perhaps the most likely to have sub-centres,[4] the assumption was made that the number of sub-centres for the other urbanised areas was zero. Interestingly, however, the coefficient on the number of sub-centres is negative and statistically significant at the 10 per cent level. Conditional on population, having more sub-centres is associated with a smaller land area, suggesting that sub-centres are more likely to develop in densely populated areas. Alternatively, perhaps the vertical expansion that comes with many sub-centres mitigates horizontal expansion of cities.

Because the transformation parameters on the covariates are not significantly different from zero and the transformation parameter, theta, on the independent variable is marginally significant in both Specifications (a) and (b), it seems reasonable to test the Box–Cox specification against a double-log specification. Using a likelihood ratio test, the hypothesis that both transformation parameters are equal to zero cannot be rejected at the 1 per cent level of significance for Specification (b). Even though the null hypothesis is rejected for Specification (a) and is very close to being rejected for Specification (b), having a transformation parameter on the independent variable that is always insignificantly different from zero and a transformation parameter on the dependent variable that is positive and fairly significant is consistent with the correct specification being log-linear with some heteroscedasticity of the errors. Thus, a double-log estimation with heteroscedasticity-consistent standard error is presented as Specification (c). The results are similar, although the estimates are less precise than with the Box–Cox specification. The estimation reveals that the elasticity of land area with respect to population is 0.91, with respect to income is –0.39 and with respect to land values is –0.03.

Table A1 in Appendix 1 presents the same estimations using the other proxy for commuting costs used by Brueckner and Fansler (1983), the percentage of workers using public transport. Again, the signs of the coefficients are all consistent with the theory, but now the coefficients on income and land value are not statistically significant. This proxy is just as problematic as the previous one. A large fraction of workers using public transport may not reflect a high commuting cost, especially in the year 2000. Instead, it may simply reflect that public transport is better and more heavily used in densely populated cities. Unfortunately, a very convincing measure of commuting cost is not available for all urbanised areas. The only other option is using travel time to work. One might expect this to be positively correlated with a monetary measure of commuting cost, but it will also be positively correlated with the physical size of the city. Since the theory would predict that the coefficient on commuting cost would be negative, the positive correlation may cause us to observe the opposite. A better measure might be travel time per mile, but the distance to work is not readily available. Nevertheless, Table A2 presents the estimations with the average travel time to work as a measure of commuting costs. The coefficient on average travel time is indeed positive for the whole sample of urbanised areas, but when the sample is restricted to cities in which over 85 per cent of the population work in the central city of their MSA of residence, the coefficient becomes negative. The non-transformed variables for the number of sub-centres and for being on a large body of water are left out because there is hardly any variation in them when the sample is restricted to cities that have a large fraction of workers commuting to the central city. This result provides evidence that the predictions of the Mills–Muth model may in fact hold up better for monocentric cities. Overall, the predictions of the model hold up fairly well despite which problematic measure of commuting cost is used.

In order to use the better measures of commuting cost that are available, it is necessary to restrict the sample to 85 urbanised areas, the cities for which the Texas Transport Institute calculates these measures. Table 3 presents estimates with the institute's measure of annual congestion cost, in millions of dollars. It is the value of travel delay and extra fuel consumed in traffic congestion annually, where delay is the extra travel time compared with some standard, in this case 65 miles per hour on freeways and 30 miles per hour on city streets. One benefit of this measure is that it is an actual monetary cost of congestion. All covariates, including the commuting cost, have significant coefficients with the expected sign. Bordering a large body of water or the coast now is negatively correlated with the physical size of the city, although it is still not statistically significant. This is a reasonable outcome for larger cities that might be constrained by geography, whereas across all cities, such geographical locations afford cities opportunity for economic growth to a certain point. Although both transformation parameters are statistically different from zero when the number of sub-centres is included, I have also presented the double-log specification. Similar results are found when the congestion cost is converted to a per person or per peak traveller measure. Moreover, similar results are found when using the other two measures provided by the institute, the travel time index and the daily vehicle miles of travel per mile of freeway lanes. These results are presented in Tables A3 and A4.

In all of the estimations using the Texas Transport Institute measures of commuting cost, the coefficient on income is now positive and significant, where it was negative and usually significant when the estimation was performed on the whole sample of urbanised areas. The cities for which the Texas Transport Institute provides data are for the most part the most populous cities in the country, so this indicates that in larger urbanised areas the trade-off between housing demand and commuting cost elasticities is different from that in smaller cities. In more populous ones, the price effect of increased demand for housing when income rises outweighs the effect of the increase in aversion to time spent commuting. The trade-off between these two elasticities would be an interesting topic for further study.

Compared with the Brueckner and Fansler (1983) results, presented in Appendix 2, these results are quite similar, which is fairly impressive given the amount of time that passed and the changes in cities that occurred in the interim. They obtain estimates with the expected sign for all covariates. Moreover, all are significant at the 5 per cent level with the exception of the proxies for commuting cost. Most of the variables I use are measured in a comparable way, save perhaps for the value of agricultural land and the commuting cost variables for estimation on the smaller sample of urbanised areas. My estimations using the alternative land value measure, the estimated value of land and buildings per acre, do not yield significant results. Brueckner and Fansler (1983) find the coefficient on land value to be significant, so if this is the measure they use, then perhaps this is evidence that

Table 3. Maximum Likelihood Estimates (dependent variable: area in square miles)

	Specification (a)		*Specification (b)*		*Specification (c) double-log model*	
	Estimate	*X^2 statistic*	*Estimate*	*X^2 statistic*	*Estimate*	*T-statistic*
Constant	–33.76		–19.18		–13.04	–4.28***
Population	1.20	69.46***	0.35	72.35***	1.13	8.69***
Income	0.69	5.88**	0.24	5.40**	0.40	1.40
Land value per acre	–0.12	2.57*	–0.08	3.17*	–0.04	–1.29*
Congestion cost ($millions)	–0.54	6.90***	–0.36	8.72***	–0.14	–1.76
Number of sub-centres			–0.12	4.21**	–0.02	–3.67***
On Coast, Great Lake, Mexican border	–0.22	0.41	–0.13	0.11	–0.02	–0.26
Transformation parameters		*Z-statistic*		*Z-statistic*		
θ	0.26	2.8***	0.29	3.5***		
λ	0.11	1.06	0.21	2.37**		
Number of observations	85		85			85
Log-likelihood	–512.20		–510.09			0.90

***indicates significance at the 1 per cent level; **at the 5 per cent level; and *at the 10 per cent level.

it is more difficult to find the expected signs and significance using modern data due to the increasing complexity of cities. It is not that the Mills–Muth model does not hold to a large extent for modern cities, but perhaps that the method of testing its comparative statics predictions for a cross-section of cities requires better measurement of the variables of interest to get the results predicted by the theory. Their estimation involves only one transformation parameter on both dependent and independent variables, and it converges to a much higher value of 0.53. Since the value of 1 lies at the edge of the confidence interval for the transformation parameter, they also present a linear specification and the corresponding elasticities evaluated at the sample means. The elasticities indicate that a 1 per cent increase in population results in an increase in land area of approximately 1.1 per cent, a 1 per cent increase in land values results in a decrease in land area of approximately 0.25 per cent, and a 1 per cent increase in income increases area by about 1.5 per cent. The elasticity with respect to population is quite similar in the current study, although the elasticity with respect to land value is smaller and the elasticity with respect to income is of the opposite sign for the whole sample of urbanised areas. Brueckner and Fansler (1983) only include urbanised areas contained within a single county in their study. This, perhaps along with other data limitations, reduces the sample to only 40 urbanised areas, the populations of which range from 52 000 to 257 000, out of a total of 248 urbanised areas in 1970. In this study, when the sample is restricted to urbanised areas associated with one county, 234 urbanised areas remain, with populations ranging from 50 000 to 2.7 million. Estimation of the model when only one-county urbanised areas are included does not reveal stronger support for the model using 2000 data. However, urbanised areas contained in one county do not have a much larger percentage of workers commuting to the central city.

5. Discussion and Conclusion

The Mills–Muth model and its assumptions are no doubt highly stylised. While many studies have attempted to add needed realism to the simple monocentric model, they often simultaneously highlight the simple model's artfulness and success in capturing some essential features of cities. For example, assuming that all urban residents earn the same income is of course unrealistic. This issue has been addressed by a number of studies, including Wheaton (1976) and Hartwick *et al.* (1976), who analyse comparative statics results of an equilibrium in which the city has several income classes. The results show that most of the crucial predictions of the model still hold under these circumstances. Moreover, the model treats housing as a single commodity, floor space. Clearly, houses are characterised by a vector of amenities and attributes, and the literature on hedonic pricing makes it clear that these different attributes matter when it comes to the value of a house. However, several studies have included a vector of housing attributes in an analysis of urban spatial structure, and it turns out that once again many of the important predictions of the model remain intact (see, for example, Büttler, 1981; and Brueckner, 1983).

In addition, making the assumption that an urbanised area is monocentric is a large one, one that seems especially sensitive to the passage of time. The development of secondary employment centres in many urban areas has become quite a widespread phenomenon and the literature indicates that the degree of polycentricity has been increasing over time. Certainly, casual observation suggests that, at least in some cities, commuters pass one another in opposite directions on their way to work. A comprehensive model should then allow for heterogeneity of preferences as well as sub-centres, otherwise people would move to reduce commuting costs. However, by treating each sub-centre as a miniature urbanised area, Muth (1969) shows that patterns of land use around the sub-centres follow the predictions of the model. The biggest 'failure' of the basic Mills–Muth model, that it does not allow for employment sub-centres, is not a failure at all. Instead, its simplicity has proved it to be very versatile. This study supports Brueckner and Fansler's (1983) conclusion that city structure, while it has grown increasingly complex, is still governed by market forces and not uncontrolled sprawl. It also suggests, with the regard to the Mills–Muth model, that the chimp still types, if at an increasingly slower rate.

Notes

1. These comparative statics predictions do not rule out multiple urban forms. For example, Anas *et al.* (1998, p. 1435) point out that most of the predictions "follow from the weaker assumption that employment is dispersed in a circularly symmetric manner, so long as it is less dispersed than residences". Nevertheless, the comparative statics predictions are a cogent test of the monocentric model. A favourable test indicates, at a minimum, that market forces shape urban spatial structure, not indiscriminate 'urban sprawl'.
2. See Mankin (1972), who finds that when leisure and commuting distance are complements, it is possible that a rise in wage income will reduce commuting distance.
3. Davidson and MacKinnon (1993) point out that using more than one transformation parameter can help to deal with the presence of heteroscedasticity. In a simpler Box–Cox model, the single transformation parameter is forced to play two roles; more specifically, it affects both the properties of the residuals and the functional form of the regression function. When the transformation parameter on the dependent variable is allowed to differ from those on the independent variables, then the former primarily affects the properties of the error terms, while the latter primarily affect the functional form.
4. In fact, McMillen and Smith (2003) show that population is a strong predictor of the number of sub-centres.

References

Anas, A., Arnott, R. and Small, K. (1998) Urban spatial structure, *Journal of Economic Literature*, 36, pp. 1426–1464.

Brueckner, J. K. (1983) The economics of urban yard space: an 'implicit-market' model for housing attributes, *Journal of Urban Economics*, 13, pp. 216–234.

Brueckner, J. K. and Fansler, D. A. (1983) The economics of urban sprawl: theory and evidence on the spatial sizes of cities, *Review of Economics and Statistics*, 65(3), pp. 479–482.

Büttler, H. (1981) Equilibrium of a residential city, attributes of housing, and land-use zoning, *Urban Studies*, 18, pp. 23–39.

Davidson, R. and MacKinnon, J. G. (1993) *Estimation and Inference in Econometrics*. New York: Oxford University Press.

Hartwick, J., Schweizer, U. and Varaiya, P. (1976) Comparative static analysis of a residential economy with several classes, *Journal of Economic Theory*, 13, pp. 396–413.

Mankin, W. (1972) A new look at the Muth model, *American Economic Review*, 62(5), pp. 980–981.

McMillen, D. P. (2006) Testing for monocentricity, in: R. Arnott and D. McMillen (Eds) *A Companion to Urban Economics*, pp. 128–140. Malden, MA: Blackwell.

McMillen, D. P. and Smith, S. C. (2003) The number of subcenters in large urban areas, *Journal of Urban Economics*, 53, pp. 321–338.

Mills, E. S. (2000) A thematic history of urban economic analysis, *Brookings-Wharton Papers on Urban Affairs*, 1, pp. 1–52.

Muth, R. F. (1969) *Cities and Housing*. Chicago, IL: University of Chicago Press.

Wheaton, W. C. (1974) A comparative static analysis of urban spatial structure, *Journal of Economic Theory*, 9, pp. 223–237.

Wheaton, W. C. (1976) On the optimal distribution of income among cities, *Journal of Urban Economics*, 3, pp. 31–44.

Appendix 1: Additional Results

Table A1. Maximum likelihood estimates (dependent variable: area in square miles)

	Specification (a)		Specification (b)		Specification (c) double-log model	
	Estimate	*X^2 statistic*	*Estimate*	*X^2 statistic*	*Estimate*	*T-statistic*
Constant	–8.20		–5.96		–5.82	–5.25***
Population	1.18	960.73***	0.76	935.06***	0.94	56.04***
Income	–0.16	2.34	–0.11	2.34	–0.15	–1.49
Land value per acre	–0.02	0.95	–0.02	1.29	–0.01	–0.83
Fraction workers using public transport	–0.16	44.30***	–0.20	43.38***	–0.13	–6.72***
Number of sub-centres			–0.02	3.155*	–0.02	–2.85***
On Coast, Great Lake, Mexican border	0.05	1.12	0.07	1.63	0.04	1.21
Transformation parameters		*Z-statistic*		*Z-statistic*		
θ	0.05	1.67*	0.06	1.95**		
λ	0.002	0.06	0.04	1.10		
Number of observations	452		452			452
Log-likelihood/R^2	–2083.60		–2082.02			0.91

***indicates significance at the 1 per cent level; **at the 5 per cent level; and *at the 10 per cent level.

Table A2. Maximum likelihood estimates (dependent variable: area in square miles)

	Specification (a)		*Specification (b)*		*Specification (c)* *UAs with more than 85% working in central city*	
	Estimate	*X^2 statistic*	*Estimate*	*X^2 statistic*	*Estimate*	*X^2 statistic*
Constant	–4.39		–2.17		–111.20	
Population	1.16	880.30***	0.67	846.02***	111.73	50.99***
Income	–0.48	17.97***	–0.30	18.51***	–8.57	0.68
Land value per acre	–0.05	6.37***	–0.04	7.31***	0.00	1.14
Average travel time to work	0.25	4.10**	0.23	4.38**	–0.02	2.670***
Number of sub-centres			–0.03	4.35**		
On Coast, Great Lake, Mexican border	0.04	0.52	0.06	0.95		
Transformation parameters		*Z-statistic*		*Z-statistic*		*Z-statistic*
θ	0.07	2.26**	0.09	2.59***	–1.50	–3.61***
λ	0.01	0.21	0.06	1.44	–0.92	–3.46***
Number of observations	452		452		27	
Log-likelihood	–2103.70		–2101.52		–92.67	

***indicates significance at the 1 per cent level; **at the 5 per cent level; and *at the 10 per cent level.

Table A3. Maximum likelihood estimates (dependent variable: area in square miles)

	Specification (a)		*Specification (b)*		*Specification (b)* *Double-Log Model*	
	Estimate	*X^2 statistic*	*Estimate*	*X^2 statistic*	*Estimate*	*T-statistic*
Constant	–30.01		–23.18		–13.10	–5.64***
Population	1.52	155.06***	0.82	146.12***	1.08	16.63***
Income	0.78	6.60***	0.48	6.30***	0.44	1.81*
Land value per acre	–0.08	1.41	–0.07	1.63	–0.03	–1.00
Travel time index	–7.48	23.64***	–8.55	22.14***	–2.35	–4.47***
Number of sub-centres			–0.04	0.89	–0.01	–1.59
On Coast, Great Lake, Mexican border	–0.13	0.40	–0.12	0.24	–0.04	–0.46
Transformation parameters		*Z-statistic*		*Z-statistic*		
θ	0.19	2.35**	0.22	2.55***		
λ	0.06	0.79	0.11	1.26		
Number of observations	85		85		85	
Log-likelihood/R^2	–503.83		–503.38		0.92	

*** indicates significance at the 1 per cent level; ** at the 5 per cent level; and * at the 10 per cent level.

Table A4. Maximum likelihood estimates (dependent variable: area in square miles)

	Specification (a)		*Specification (b)*		*Specification (b) Double-Log Model*	
	Estimate	*X^2 statistic*	*Estimate*	*X^2 statistic*	*Estimate*	*T-statistic*
Constant	–24.18		–9.40		–7.22	–1.93*
Population	2.80	157.03***	0.70	140.47***	1.00	16.49***
Income	1.21	4.61**	0.41	4.450**	0.38	1.39
Land value per acre	–0.07	0.50	–0.06	1.04	–0.02	–0.72
Vehicle miles of travel per freeway mile	–1.63	8.20***	–0.67	8.21***	–0.49	–1.99**
Number of sub-centres			–0.08	2.41	–0.02	–3.77***
On Coast, Great Lake, Mexican border	–0.10	0.25	–0.06	0.04	–0.02	–0.29
Transformation parameters		*Z-statistic*		*Z-statistic*		
θ	0.17	1.99**	0.23	2.51***		
λ	0.00	–0.06	0.12	1.12		
Number of observations	85		85		85	
Log-likelihood/R^2	–511.55		–510.35		0.90	

*** indicates significance at the 1 per cent level; ** at the 5 per cent level; and * at the 10 per cent level.

Appendix 2: Brueckner and Fansler Results

Table A5. Maximum likelihood estimates

	Specification (a)		*Specification (b)*	
	Coefficient	*T-statistic*	*Coefficient*	*T-statistic*
Constant	–16.71	–3.05	–18.72	–1.31
N	0.0155	9.04	0.0154	9.16
r_a	–0.0715	–2.86	–0.0705	–2.74
y	0.0791	3.23	0.0791	3.23
PUBLIC	–0.0467	–0.20		
CARS			0.1117	0.16
λ	0.53		0.53	

Table A6. Linear estimates

	Specification (a)		*Specification (b)*	
	Coefficient	*T-statistic*	*Coefficient*	*T-statistic*
Constant	–41.07	–2.28	–63.47	–1.24
N	0.0004	10.03	0.0004	9.88
$\mathrm{r_a}$	–0.0303	–3.09	–0.0289	–2.89
y	0.0062	3.03	0.0062	3.05
PUBLIC	–0.2444	–0.41		
CARS			0.2475	0.46
R^2	0.7982		0.7985	

Table A7. Elasticities from Linear Equations

	Specification (a)	*Specification (b)*
N	1.097	1.086
r_a	–0.234	–0.231
y	1.497	1.496

Appendix 3: Census Definitions

Central City

This is the largest city of a Metropolitan area (MA). Central cities are a basis for establishment of an MA. Additional cities that meet specific criteria also are identified as central cities. In a number of instances, only part of a city qualifies as central, because another part of the city extends beyond the MA boundary.

Metropolitan Statistical Area (MSA)

This is a geographical entity defined by the US federal Office of Management and Budget for use by federal statistical agencies, based on the concept of a core area with a large population nucleus, plus adjacent communities having a high degree of economic and social integration with that core. Qualification for an MSA requires the presence of a city with 50 000 or more inhabitants, or the presence of an Urbanised Area (UA) and a total population of at least 100 000 (75 000 in New England). The county or counties containing the largest city and surrounding densely settled territory are central counties of the MSA. Additional outlying counties qualify to be included in the MSA by meeting certain other criteria of metropolitan character, such as a specified minimum population density or percentage of the population that is urban. MSAs in New England are defined in terms of minor civil divisions, following rules concerning commuting and population density.

Urbanised Area (UA)

This is an area consisting of a central place(s) and adjacent territory with a general population density of at least 1000 people per square mile of land area that together have a minimum residential population of at least 50 000 people. The US Bureau of the Census uses published criteria to determine the qualification and boundaries of UAs.