BEBR

FACULTY WORKING
PAPER NO. 997

A Switching Regression Analysis of
Urban Population Densities

Jan K. Brueckner

College of Commerce and Business Administration
Bureau of Economic and Business Research
University of Illinois, Urbana-Champaign

BEBR

A Switching Regression Analysis of
Urban Population Densities

Jan K. Brueckner, Professor
Department of Economics

Abstract

New Theoretical work in urban economics suggests that urban population density contours are inherently discontinuous. This result calls into question the standard practice of estimating smooth density contours and points to the need for an estimation technique capable of capturing discontinuities. The paper applies such a technique (Quandt's switching regression method) to the estimation problem. Density contours with marked discontinuities emerge from the empirical work.

A Switching Regression Analysis of
Urban Population Densities

by

Jan K. Brueckner*

No empirical relationship in urban economics has been studied more
thoroughly than the relationship between population density and dis-
tance to the city center. While the pioneering paper of Clark (1951)
did not rely on an explicit theoretical framework, recent studies have
been designed to test the Muth (1969)-Mills (1972) model of urban spa-
tial structure, which predicts an inverse relationship between density
and distance. A representative group of recent studies includes Muth
(1969), Mills (1970), Kau and Lee (1976), McDonald and Bowman (1976),
Glickman and Oguri (1978), and Anderson (1982). A common practice is
to fit a negative exponential function to the data, although some
papers experiment with other specifications.

While past studies have generally estimated smooth population
density contours, a new body of theoretical work in urban economics
suggests that density contours are inherently discontinuous. The
new theoretical literature (see, for example, Anas (1978), Brueckner
(1980), Brueckner and von Rabenau (1981), and Wheaton (1982))[1] focuses
on an aspect of urban structure entirely omitted in the Muth-Mills
analysis: the durability of housing capital. Rather than invoking
the Muth-Mills assumption that housing capital is perfectly malleable,
which allows the city to be reconstructed each period as underlying
conditions change, the new literature recognizes that the replacement
of buildings is costly and infrequent. A key insight afforded by the

new analysis is that since population density (which depends on building heights and dwelling sizes) will reflect economic conditions (rents, land values, etc.) prevailing at a neighborhood's construction date, the spatial pattern of building ages in a city will be an important determinant of the pattern of densities. In particular, when the age pattern exhibits the discontinuities which result from spatial redevelopment waves (with old neighborhoods adjacent to brand new ones), the age-density link will generate a corresponding discontinuity in the population density contour. Given that discontinuous age patterns will be a typical outcome of the redevelopment process, the existence of density discontinuities becomes an inescapable implication of the theory.

As should be clear from the above discussion, the new urban models call into question the accepted practice of estimating smooth population density contours. While such a procedure will capture the overall spatial trend of densities in an urban area, it may mask great irregularity. To capture the discontinuous density patterns predicted by the new literature, an estimation technique capable of handling discontinuities is clearly required. The purpose of the present paper is to apply such a technique (Quandt's (1958) switching regression technique) to the estimation problem.[2] Density contours are estimated for 77 city samples using data compiled by Kau and Lee (1976) and Anderson (1982).

Application of the switching regression technique requires formulation of a criterion for choosing the number of "switches" (the number of potential discontinuities in the estimated contour). A novel feature of the present approach is that this choice is viewed as a problem in

model selection. The model selection criteria recently introduced by
Hannan and Quinn (1979), Akaike (1977), and Risannen (1978), which
modify the Information Criterion of Akaike (1973), are used to deter-
mine the number of switches in the estimated density contours.

The plan of the paper is as follows. The next section presents
a more detailed discussion of how discontinuous density contours are
generated in an urban model. The third section of the paper discusses
the switching regression technique and explains the model selection
criteria. The fourth section presents selected empirical results,
while the fifth section investigates the effect of joint estimation of
population density and building age contours. The last section offers
conclusions.

2. Building Age Variation and Density Discontinuities

To better grasp how discontinuous population density contours
emerge as a result of spatial variation in building ages, it is help-
ful to consider a concrete example based on the model of Brueckner
(1980). By assuming that producers are myopic with regard to future
housing prices (current prices are expected to last forever) and by
imposing Cobb-Douglas utility and production functions, the analysis
yields an especially simple theoretical result: the lifespan of
buildings (the length of the interval between construction and re-
placement) is the same regardless of location and initial construction
date. This result leads to a very simple evolutionary process for a
growing city, as seen in the following example.

First, suppose for purposes of illustration that the uniform life-
span of buildings is three years (time and distance are measured in

discrete units). Next, assume that the city grows spatially by one block each year, starting from a dimensionless point at time zero. Table 1 shows the city's age-distance profiles for t=3,4,8 under these assumptions (blocks are numbered moving outward from the city center). The cyclical nature of the age pattern emerging from the model is highlighted in the upper panel of Figure 1, which shows the pattern for t=8.

Table 1

Building Age Contours

(t=3)		(t=4)		(t=8)	
Block #	Age	Block #	Age	Block #	Age
1	2	1	0	1	1
2	1	2	2	2	0
3	0	3	1	3	2
		4	0	4	1
				5	0
				6	2
				7	1
				8	0

To deduce the implications of a cyclical building age pattern for the spatial behavior of population densities, the separate effects of age and distance on density must be considered. First, an implication of the model is that holding building age fixed, density declines with distance to the city center. Referring to Table 1, this result implies that density will be lower, for example, in block 6 than in block 3 at t=8. The model, however, is ambiguous about the partial effect of age on density. Whether density is an increasing or decreasing function of a block's construction date depends on the growth rates of the various exogenous variables in the model. For certain parameter values, density

is a decreasing function of the construction date for any given loca-
tion, while for other values the relationship is reversed.

Consider the implications of the first case by referring to the
upper panel of Figure 1. Within each of the three segments of blocks,
1-2, 3-5, and 6-8, age declines with distance, so that more distant
blocks have more recent construction dates. When density is a decreas-
ing function of the construction date, it follows that the negative
pure distance effect on density felt within each block segment is rein-
forced by the negative effect of lower ages at greater distances, and
density declines unambiguously within the segment. Between block seg-
ments, however, age increases discontinuously from zero to 2, and the
negative distance effect is swamped by a much larger positive effect
due to a higher age. The result is a discontinuous increase in popula-
tion density between blocks 2 and 3 and blocks 5 and 6. The implied
spatial pattern of densities is shown in the second panel of Figure 1.[3]

When density is an increasing function of a block's construction
date, the outcome is quite different. Within block segments, the nega-
tive pure distance effect is offset by the positive effect of lower
ages at greater distances, and the net change in population density is
ambiguous (density may even increase with distance within segments).
Between segments, however, the discontinuous increase in age leads to
a discontinuous decrease in density, with the negative pure distance
effect reinforced by a large negative age effect. The implied spatial
pattern of densities is shown in the bottom panel of Figure 1 (density
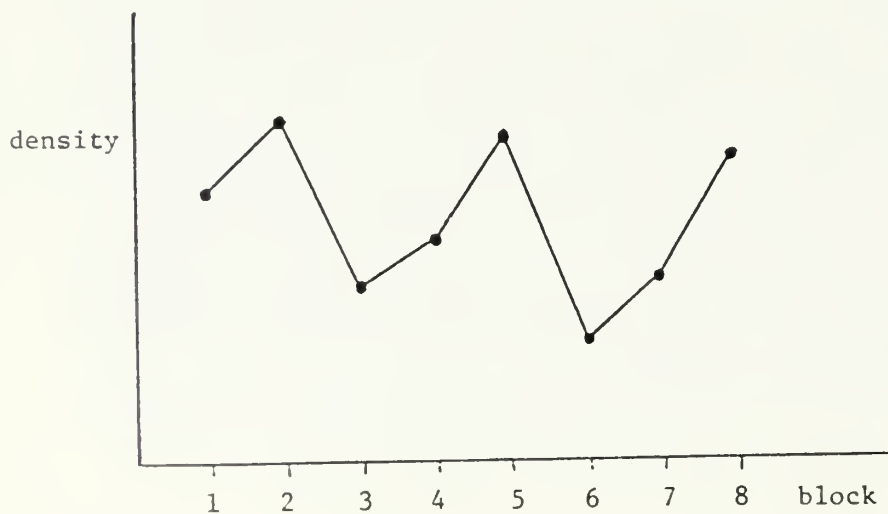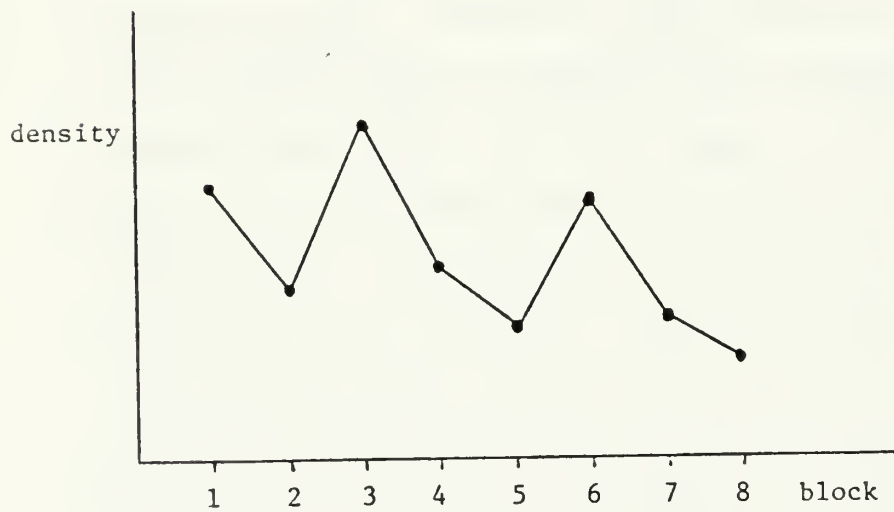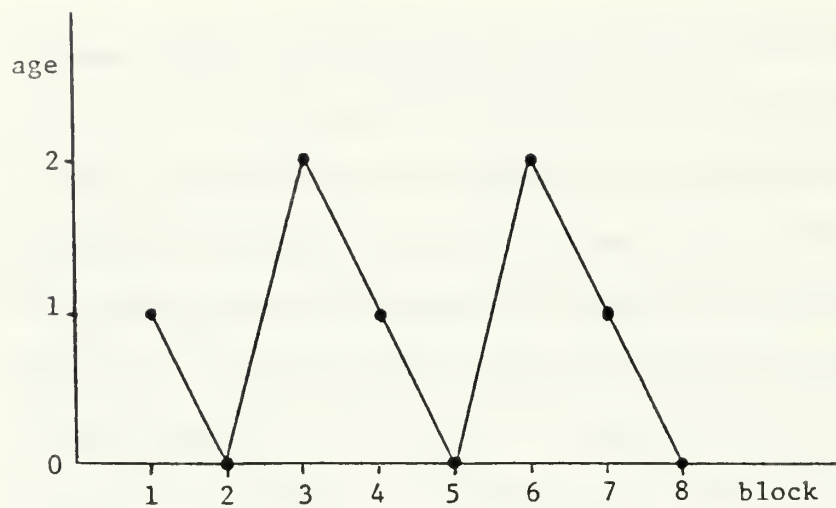is assumed to increase within segments).[4]

Figure 1.

A previous paper (Brueckner (1981)) applied Quandt's (1958) switching regression technique to the estimation of discontinuous age contours like the one shown in Figure 1 for several U.S. cities. Although the estimated contours did not exhibit the orderly cyclical pattern shown in the Figure, marked discontinuities emerged nevertheless. The remainder of the present paper shifts attention to the lower panels of Figure 1 and the estimation of density contours. The next section describes the estimation technique.

3. Estimation Technique

The basic elements of the estimation technique are simply described. The data, which consist of observations on census tract population densities and tract distances from the city center, are arranged in ascending order of distance. To estimate a density contour with, say, one switch, a pair of semi-log regressions[5] is computed for each possible location of the switchpoint (one regression for each segment of the contour). The switchpoint location and associated segment parameters which yield the highest value of the likelihood function are chosen. Assuming that errors are i.i.d. normal and that segments one and two have n_1 and $n_2 = T-n_1$ observations respectively (T is the sample size), the maximized likelihood function equals

$$-(n_1 \log \hat{\sigma}_1 + n_2 \log \hat{\sigma}_2) + c, \qquad (1)$$

where c is a constant and $\hat{\sigma}_i^2 = SSR_i/n_i$, with SSR_i being the sum of squared residuals for segment i. The switchpoint location (the value of n_1) is chosen to maximize (1). For higher numbers of switches, the

procedure is similar except that the number of possible joint switch-
point locations rises dramatically (the expression analogous to (1)
is $-\sum_{i=1}^{s+1} n_i \log\hat{\sigma}_i + c$, where s is the number of switches).[6]

Although estimation of the density contour for a given number of
switches s is straightforward, the problem posed by the fact that the
correct value for s is unknown is a source of difficulty. In esti-
mating age contours, Brueckner (1981) solved this problem by applying
an ad hoc method based on a suggestion of Quandt (1958). The method
exploits Quandt's observation that the distribution of $-2\log L_s/L_{s+1}$
will be approximately χ^2 with four degrees of freedom, where L_s and
L_{s+1} give the likelihood function values with s and s+1 switches re-
spectively.[7] Under the method, the null hypothesis of zero switches
is first tested against the alternative of one switch using the χ^2
statistic with s=0. If the null hypothesis is rejected, a new null
hypothesis of one switch is tested against an alternative of two. The
process continues until the null hypothesis of s* switches cannot be
rejected in favor of the alternative of s*+1 (s* is then identified as
the optimal number of switches).

While this method has the appealing feature that an extra switch
is justifiable only if the resulting increase in the log likelihood
is large enough, its drawbacks are that the distribution of the test
statistic may not be close to χ^2 because of the discrete nature of the
switchpoint location and that the individual χ^2 tests are not indepen-
dent.[8] In view of these difficulties, the present paper views the
choice of s as a problem in model selection. Under Akaike's (1973)
Information Criterion for model selection (denoted AIC), the quantity

$$\frac{-2\log L_r}{T} + \frac{2K_r}{T} \qquad (2)$$

is computed for each model specification r, where L_r is the maximized value of the likelihood function for specification r, K_r is the number of unknown parameters in specification r, and T is the number of observations. The model specification is chosen to minimize (2). Note that as the number of parameters increases, the second term in (2) rises while the increase in the likelihood value reduces the first term. The model size which minimizes (2) thus strikes a balance between complexity and goodness of fit.

Recognizing that the criterion (2) is inconsistent, tending to favor excessively large models, Hannan and Quinn (1979), Akaike (1977), and Risannen (1978) have proposed consistent modifications of it. Hannan and Quinn (hereafter HQ) propose replacing the last term in (2) by

$$\frac{2K_r \log\log T}{T} \qquad (3)$$

while Akaike and Risannen propose replacing the last term by

$$\frac{K_r \log T}{T} \qquad (4)$$

(the resulting criterion has been denoted BIC).

While the HQ and BIC criteria have been applied mainly in the context of times series models, they offer an attractive method for choosing the number of switches in the present problem. In computing

the number of parameters in a contour with s switches, the location of each switchpoint is counted as a parameter, with each regression adding three additional parameters. A contour with s switches thus contains $3(s+1) + s = 4s + 3$ parameters. Substituting for L_r and K_r in the modified versions of (2), the model selection criteria can be written

$$\frac{2(\sum_{i=1}^{s+1} n_i \log\hat{\sigma}_i - c)}{T} + \begin{cases} \dfrac{(8s+6)\log\log T}{T} & \text{(HQ)} \\ \\ \dfrac{(4s+3)\log T}{T} & \text{(BIC)} \end{cases} \tag{5}$$

In the next section, these criteria are applied to the problem at hand.

4. Basic Results

Since presentation of all the estimation results would consume a great deal of space, attention in what follows is restricted to illustrative highlights. Before viewing summary results for the two data sets, it will be helpful to consider the case of a selected city, Birmingham, Alabama, using results from the Kau-Lee data set. Table 2 shows the values of the HQ and BIC criteria from (5), as well as the value of the log likelihood function, for zero up through five switches.[9]

Table 2

Criterion Values for Birmingham
(Kau-Lee Data)

s	log likelihood	HQ	BIC
0	-49.28	2.53	2.61
1	-38.21	2.26	2.44
2	-30.36	2.14	2.42
3	-22.61	2.02	2.41
4	-18.18	2.06	2.56
5	-15.73	2.19	2.80

The Table shows that the HQ and BIC criteria are both minimized at s=3, indicating the optimality of three switches in each case. Under the likelihood ratio approach, adding a switch is justified as long as the log likelihood increases by at least one half the χ^2 critical point, a value equal to 6.64 at the 99 percent confidence level. Referring to Table 2, it is clear that the likelihood ratio approach also indicates the optimality of three switches, so that the three methods yield the same answer in this instance.

A graph of the estimated density contour for Birmingham is shown in Figure 2 along with the underlying point scatter (to save space, parameter estimates are not presented). LogD, with density measured in people per acre, is on the vertical axis, while distance x, measured in miles, is on the horizontal axis. The dramatic discontinuities in the estimated contour, as well as the underlying density variations generating them, are clear from the Figure. Note that the pattern of discontinuities for Birmingham does not bear close resemblance to either of the patterns shown in Figure 1. Given that the orderly patterns of Figure 1 follow from the uniform-lifespan property of the model, which is unlikely to hold exactly in the real world, the divergence between Figures 1 and 2 should not be viewed as disconfirmation of the hypothesis that building age variation affects the density contour. Erratic building age patterns, consistent perhaps, with a richer model, could generate a contour like the one shown in Figure 2 (joint estimation of age and density contours is carried out in Section 5).

Tables 3 and 4 show the optimal numbers of switches under the three different criteria for the cities in the Kau-Lee and Anderson data sets
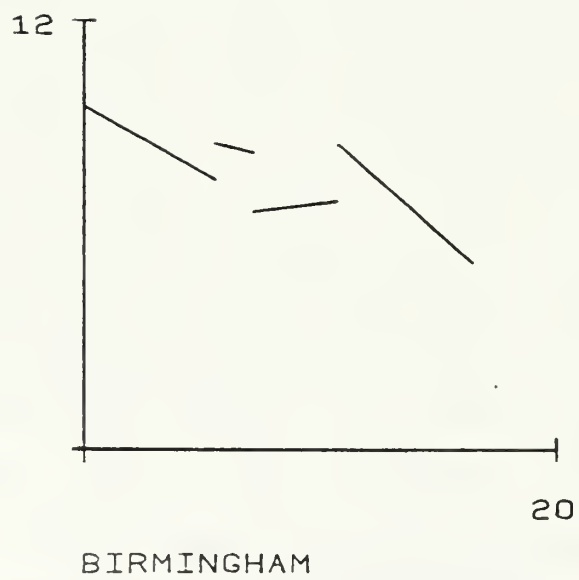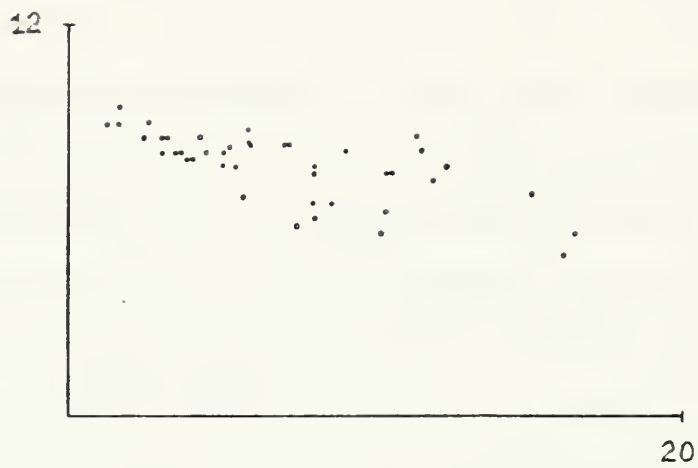
Figure 2.

Table 3

Estimated Number of Switches for
Cities in Kau-Lee Data Set*

	s_{LR}	s_{HQ}	s_{BIC}		s_{LR}	s_{HQ}	s_{BIC}
Akron	1	1	1	New Orleans	0	2	0
Baltimore	1	1	1	Oklahoma City	2	3	1
Birmingham	3	3	3	Omaha	1	3	0
Boston	1	1	0	Philadelphia	0	0	0
Buffalo	0	2	0	Phoenix	2	2	2
Chicago	0	0	0	Pittsburgh	0	2	0
Cincinnati	0	2	0	Portland	0	4	0
Cleveland	1	1	0	Providence	0	0	0
Dallas	1	2	1	Richmond	1	1	1
Dayton	1	1	1	Rochester	1	1	0
Denver	0	0	0	Sacramento	0	0	0
Detroit	3	4	2	Salt Lake City	0	3	0
Flint	0	2	0	San Antonio	0	2	0
Ft. Worth	1	2	1	San Diego	0	0	0
Houston	2	2	2	San Jose	4	4	3
Jacksonville	0	2	2	Seattle	1	1	0
Kansas City	0	0	0	Spokane	1	3	1
Los Angeles	0	0	0	St. Louis	0	0	0
Louisville	0	0	0	Syracuse	2	2	2
Memphis	3	3	2	Tacoma	1	2	1
Miami	0	0	0	Toledo	0	0	0
Milwaukee	0	0	0	Utica	2	2	2
Nashville	0	4	0	Wichita	1	1	1
New Haven	2	3	0	Washington	1	1	0

*Most cities have 43 observations; the remainder have either
42 or 41.

Table 4

Estimated Number of Switches for
Cities in Anderson Data Set

	s_{LR}	s_{HQ}	s_{BIC}	obs
Akron	0	2	0	65
Birmingham	1	1	1	50
Buffalo	2	2	1	56
Cincinnati	2	2	2	58
Cleveland	2	2	1	51
Columbus (OH)	2	2	1	62
Denver	1	1	0	58
Des Moines	0	1	0	49
Evansville	1	3	0	38
Flint	0	0	0	47
Fort Wayne	2	3	2	45
Grand Rapids	1	1	0	72
Indianapolis	1	1	1	60
Kalamazoo	1	1	1	24
Kansas City	0	0	0	54
Lansing	0	2	0	54
Memphis	1	1	0	54
Milwaukee	3	3	2	124
New Haven	0	2	0	35
New Orleans	1	2	1	69
Richmond	2	5	2	55
Rochester	2	5	2	57
Rockford	2	2	1	46
Spokane	1	5	3	49
St. Louis	5	5	1	65
Syracuse*	2	7	2	59
Toledo	0	1	0	50
Topeka	0	0	0	30
Wichita	3	4	3	55

*The HQ criterion value for Syracuse is essentially constant from four through seven switches, with seven giving the lowest value. Results for higher numbers of switches were not computed.

(note that many cities are common to both data sets). The results of the likelihood ratio approach are contained in the first column of each Table; s subscripts are self explanatory. The number of census tract observations for each Anderson city is also listed (Kau-Lee cities all have around 40 observations; see the note to Table 3). A number of features of the results deserve note. First, inspection of the Tables shows that agreement among the three criteria is the exception rather than the rule. However, it is interesting to note that a strong pattern characterizes the results in that the inequalities $s_{HQ} \geq s_{LR} \geq s_{BIC}$ hold in almost every city (the only exceptions are Jacksonville in Table 3 and Spokane in Table 4). Thus, in the vast majority of cases, the HQ criterion yields at least as many switches as the likelihood ratio approach, which in turn yields at least as many switches as the BIC criterion. Another feature of the Tables is that there is little correspondence between the results for those cities appearing in both data sets. This tendency, which at first appears quite disturbing, seems to be largely a consequence of the fact that the spatial coverage of the data sets is different, with the Kau-Lee data typically extending much farther out from the center in a given city than the Anderson data. In addition, the number of observations for each common city differs between the data sets.

To give a more complete idea of the nature of the results, Figure 3 presents graphs of the HQ density contours for Dallas, Detroit, Fort Worth, New Haven, Utica, and Milwaukee and of the BIC contours for Rochester and Spokane (the first five are Kau-Lee cities; the last three are from the Anderson data set). These contours, which are perhaps more dramatic than average, indicate the striking nature of the
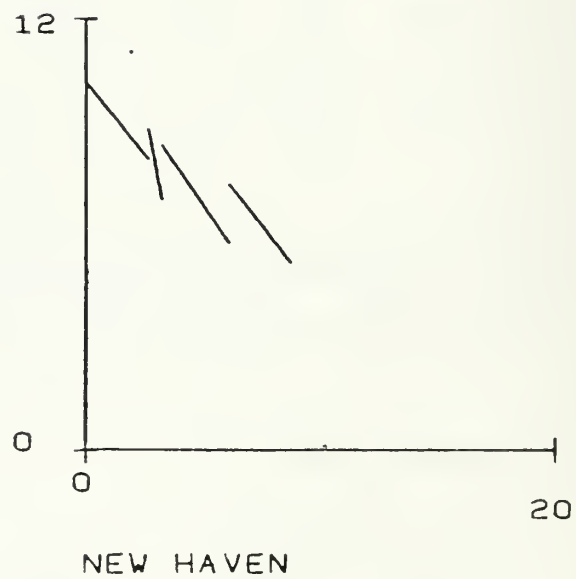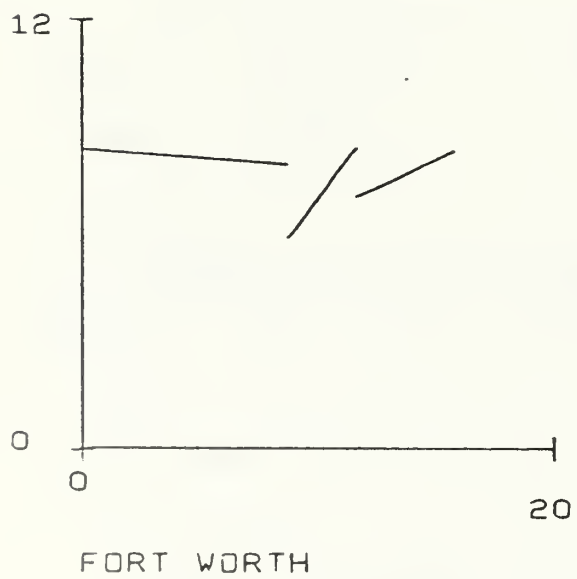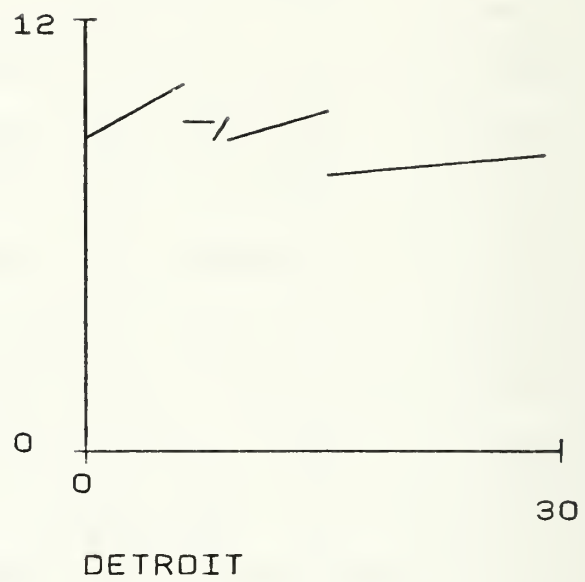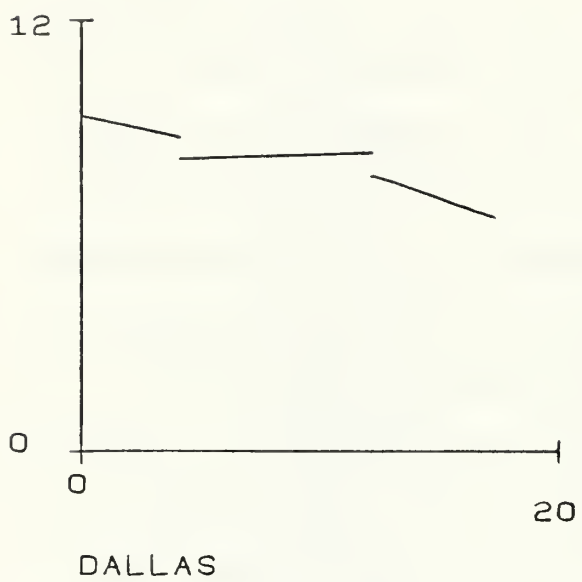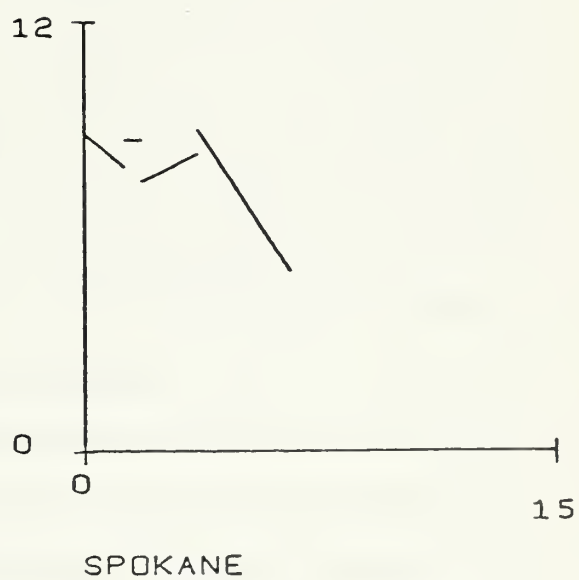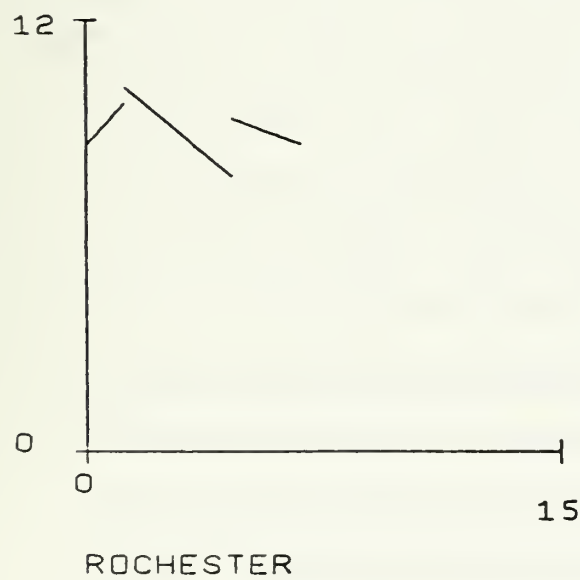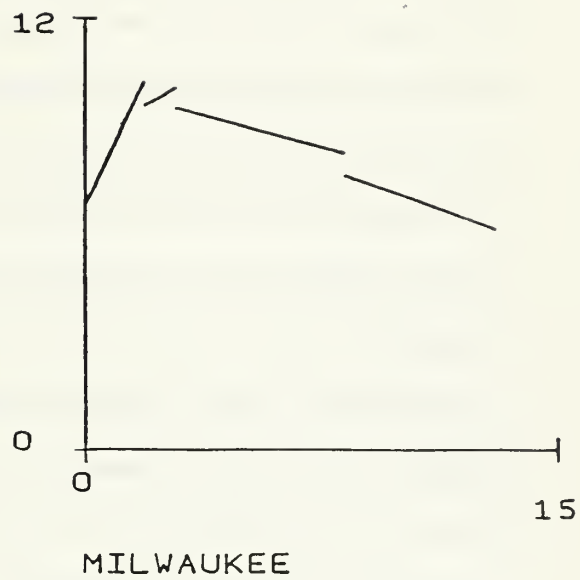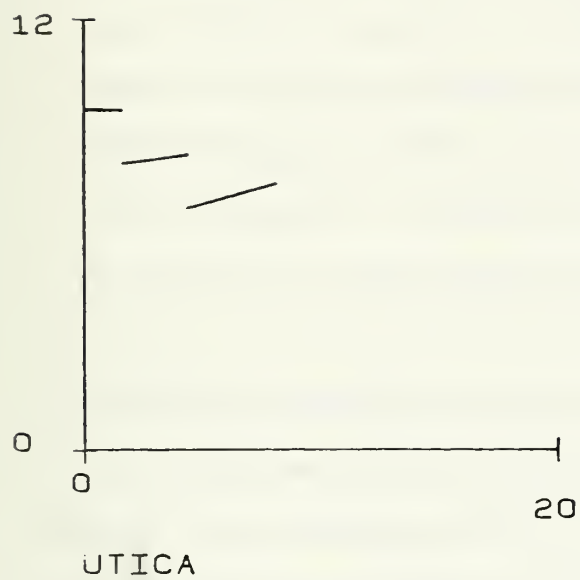
Figure 3.

Figure 3.
(continued)

discontinuities which can result from application of the switching

regression method. Note that in the cases of Dallas, Detroit, Fort

Worth, Utica and Milwaukee, each discontinuity involves a decrease in

density (as in the bottom panel of Figure 1), while in the cases of

New Haven and Rochester, discontinuities involve an increase in den-

sity (as in the middle panel of Figure 1). Spokane, like Birmingham

in Figure 2, has a mixture of these two patterns. A complete set of

figures showing the estimated contours for all cities is available on

request.

Returning to the results of Tables 3 and 4, a natural question is

whether the estimated numbers of switches for the various cities are

related in a systematic way to underlying city characteristics. First,

it seems plausible that for cities of a given population, the switching

regression method would isolate greater detail (and hence show more

discontinuities) the greater the number of census tract observations.

For the same reason, the number of switches should fall with popula-

tion, holding the number of observations fixed. This suggests that

the estimated number of switches should be positively influenced by

a variable such as observations per capita. In addition, since the

number of different generations of buildings coexisting in a city will

rise with the city's overall age, and since under the simple example

of Section 2, density discontinuities occur at generational boundaries

(see Figure 1), the number of switches should increase with a city's

age. Finally, the absolute magnitude of a city's population might be

expected to affect the number of switches in a positive direction.[10]

To test these hypotheses, a cross section regression was computed

relating the number of segments (s+1) in the estimated contours to
OBSPOP (observations per ten thousand 1970 urbanized area population),
POP (1970 urbanized area population in thousands), and YEARS (the
number of years between 1970 and the decade the city first achieved 10
percent of its 1970 population). From the above discussion, positive
signs were expected for all the coefficients. Since the nearly iden-
tical sample sizes for the Kau-Lee cities (see the note to Table 3)
result in OBSPOP and POP being almost exactly inversely proportional,
the regression was computed for the Anderson data set alone (Syracuse
was deleted; see Table 4). The results using the number of HQ segments
as the dependent variable are as follows (variables are in log form; t-
ratios are in parentheses):

Constant	OBSPOP	POP	YEARS	R^2
-8.031	0.883	0.636	1.116	.2747
(-2.66)	(2.16)	(1.97)	(2.43)	

Expectations are nicely confirmed by the regression; all coefficients
are significant with the correct sign at nearly the 5 percent level.
Results using the BIC dependent variable are poorer, however, with posi-
tive but insignificant coefficients. This is not surprising given that
the BIC criterion yields less dispersion in the dependent variable.[11]

5. Joint Estimation of Age and Density Contours

While the empirical results summarized in Section 4 appear to con-
firm the hypothesis that real-world density contours exhibit discon-
tinuities, the present section offers a more stringent evaluation of
the underlying theory by testing the hypothesis that density discon-
tinuities occur in the same locations as discontinuities in the contour

of building ages. Recall from Section 2 that abrupt changes in building
ages were responsible for density discontinuities. The test involves
simultaneously estimating density and age contours with and without
the constraint that the switchpoints occur in the same locations. The
building age measure is average age of structures based on six age
classes, a variable used in Brueckner (1981). The cities chosen for
study were Milwaukee and Baltimore.[12]

Before proceeding to the results of the joint estimation, it is
possible to use the data to validate a main premise of the underlying
theory: that building age and distance have important partial effects
on density. Table 5 presents the results of OLS regressions of log
density on distance and age for the two samples.

Table 5

Regressions of log D on Distance and Age

	Constant	Distance	Age	R^2	Obs
Baltimore	9.702	−0.142	0.019	.3643	152
	(34.30)	(−4.43)	(3.33)		
Milwaukee	9.944	−0.268	0.010	.6601	124
	(28.74)	(−9.18)	(1.39)		

*t-ratios in parentheses

As can be seen, distance has the expected negative effect on density
while age has a positive effect, indicating higher densities in older
neighborhoods at a given distance (the Milwaukee age coefficient,
however, is not significantly different from zero). Recalling the
idealized model of Section 2, these results suggest that the middle
panel of Figure 1, which exhibits upward density discontinuities, is

relevant (as noted earlier, it would be surprising if the data con-

formed exactly to this pattern).

Turning now to the joint estimation problem, the procedure is as

follows. First, it is assumed that the covariance between the age and

density error terms for the same census tract equals zero, which means

that the joint log likelihood is just the sum of the individual age

and density log likelihoods.[13] Next, it is assumed that the age and

density contours have the same number of switches and that the switch-

points occur in the same locations for both contours. Under these re-

quirements, the joint log likelihood is maximized for each value of s,

and the optimal value s* is chosen using either the HQ or BIC criterion

(with s switches, the number of parameters is $6(s+1) + s = 7s + 6$). To

test for the effect of the identical switchpoint constraint, switch-

point locations are found which separately maximize the individual age

and density log likelihoods, holding the number of switches for each

contour at s*. The sum of the individual log likelihoods is then com-

pared to the joint log likelihood. If the difference exceeds one-half

the appropriate χ^2 critical point, the null hypothesis of equal switch-

point locations is rejected. Note that this procedure involves an

asymptotic likelihood ratio test conditional on the number of switches

for the individual contours equalling the optimal number for the joint

problem. Since the number of elements in the parameter space rises by

s* when the identical-switchpoint constraint is relaxed, s* is the

appropriate number of degrees of freedom for the χ^2 test.

Table 6 shows the results of applying the above procedure to the

Baltimore and Milwaukee data.[14] With joint estimation, the optimal

Table 6

Results of Joint Estimation Procedure

	log likelihood	switchpoints
Baltimore		
Joint:	−521.47	1.47, 2.38
Age:	−444.97	1.47, 2.38
Density:	−70.57	1.42, 1.92
Total:	−515.54	--
Difference:	5.93	--
Milwaukee		
Joint:	−491.86	2.98, 6.72
Age:	−395.96	2.65, 6.72
Density:	−90.81	2.98, 8.24
Total:	−486.77	--
Difference:	5.09	--

number of switches equals two for both cities under both the HQ and
BIC criteria. The second column of the Table shows the switchpoint
locations (in miles from the city center) for the jointly and indivi-
dually estimated contours (note the changes which occur when the
identical-switchpoint constraint is relaxed). The first column shows
the jointly maximized log likelihood, the sum of the individually maxi-
mized log likelihoods, and the difference between these two quantities.
Unfortunately, this difference exceeds 4.605 (one half the χ^2 critical
value with two degrees of freedom) for both cities, indicating rejec-
tion of the null hypothesis of identical switchpoints in both cases.
The discrepancy, however, is not great; the null hypothesis can in fact
be accepted at the (admittedly stringent) 99.5 percent level in the
case of Milwaukee.[15] The upshot is that while the test is unfavorable
to a strict interpretation of the theory, the verdict is not a resound-
ing one.[16]

6. Conclusion

The main contribution of this paper has been to show that popula-
tion density contours in real-world cities are highly irregular.
Tables 3 and 4 indicate that contours with zero switches are the excep-
tion rather than the rule, and the graphs in Figures 2 and 3 establish
that when switches occur, they often involve striking discontinuities.
These conclusions confirm the predictions of the new durable housing
literature, which implies that density contours are inherently discon-
tinuous. A principal lesson of the results is that while estimation of
a smooth Muth-Mills density contour may be acceptable as a technique

for isolating an overall spatial trend, the procedure will mask con-
siderable irregularity in the spatial structures of many cities.

The paper also makes an important methodological contribution in
that the number of switches is chosen by model selection criteria. Use
of the HQ and BIC criteria circumvents difficulties associated with the
ad hoc likelihood ratio method, yielding a simple and intuitively
appealing solution to the choice problem. Application of this approach
in other econometric problems involving structural change could no
doubt prove fruitful.

Footnotes

*I wish to thank the National Science Foundation (Grant
#SES-8308119) and the Office of Real Estate Research at the University
of Illinois for financial support. Other acknowledgments are also
due. First, I am indebted to James Kau and Cheng-few Lee and to John
Anderson for providing me with their data sets. Also, I wish to thank
Paul Newbold for suggesting the model selection approach used in the
paper. Finally, I wish to thank my able and diligent research
assistant, Chuan Lin, for his substantial contribution to this
research. Any errors, however, are my own.

[1] See the references in the third or fourth papers above for a more
complete list of studies in this area.

[2] Kau, et al. (1983) estimate discontinuous density contours using
a restrictive method which requires a discontinuity to occur between
each pair of adjacent observations.

[3] Note that when attention is restricted in Figure 1 to locations
with the same age buildings, density falls with distance.

[4] Although the model of Brueckner and von Rabenau (1981) is dif-
ferent from the one underlying the above example, it nevertheless
gives rise to similar density discontinuities. In the models of Anas
(1978) and Wheaton (1982), however, the assumption that buildings last
forever prevents the emergence of discontinuous building age patterns
and the striking density discontinuities which accompany them. The
models do generate modest density discontinuities nevertheless.

[5] Use of the semi-log form seemed natural given the popularity of
the negative exponential density function. The convexity of the den-
sity contours in Figure 1 anticipates the use of semi-log regressions.

[6] In the estimation process, care must be taken to avoid locating a
switchpoint between two observations with the same distance value, and
enough observations must be included in each segment to make a
regression feasible (the required number will exceed three when
distance values are repeated).

[7] With one extra switch, four extra parameters (3 regression para-
meters plus the switchpoint location) are introduced.

[8] Another peculiarity of this method is that the log likelihood
increase in going from, say, zero switches to one may be too small to
justify the extra switch even though the null hypothesis of one switch
might be rejectable in favor of the alternative of two. In such a
case, zero switches are chosen.

[9] In computing the Birmingham results and all subsequent results
where the city has fewer than 100 observations, the minimum number of

observations per regression was set at five (this was the smallest possible value given repeated distance observations). For cities with over 100 observations, the minimum was set arbitrarily at 10 observations per regression.

[10] This prediction does not follow, however, from the model of Section 2. For a given founding date, faster population growth in one city than another (resulting in a higher final population) would merely lengthen the segments in the Figure 1 population density contour without affecting their number. However, total population seems to be a plausible determinant of the number of switches (results were poorer when it was deleted).

[11] At this point, the presence of a potentially serious measurement problem should be noted. While the predictions of the theory sketched in Section 2 relate to population density on residential land ("net" density), the empirical work uses observations on gross density (population per unit total land area). As a result of this measurement error, a new element affecting the spatial behavior of population density is introduced: spatial variation in the fraction of land devoted to residential uses. Although the effect of smooth variation in this fraction would appear to be innocuous, a danger lies in the possibility that discontinuous changes due to the presence of parks, railroad yards, etc., could introduce spurious discontinuities into the estimated gross density contour when the underlying net density contour is smooth. Unfortunately, it is impossible to ascertain the extent to which this phenomenon influences the results of the estimation.

[12] The Anderson Milwaukee data was used. To appraise the effect of the measurement problem noted in footnote 11, Baltimore census tract areas were remeasured using detailed maps, with only the built-up sections (those with streets) contributing to tract area. Since the resulting sample includes only tracts within the Baltimore city limits, it has more limited spatial coverage than the Kau-Lee sample. The fact that the density estimation results for this new sample showed no striking divergence from the results for other cities suggests that the consequences of the measurement problem may not be terribly serious.

[13] Results of the joint estimation procedure changed little when a non-zero covariance was allowed. However, it is easy to see that the individual maximization procedure described below cannot incorporate this assumption.

[14] To get an idea of the computational efficiency achieved in the switching regression program, each city's joint estimation results for zero up through five switches, which involved an enormous number of calculations, required only about five seconds of central processor time on a CDC Cyber 175 computer.

[15]Note that in contrast to the standard situation, <u>acceptance</u> of a null hypothesis implies confirmation of the theory in the present case.

[16]To save space, graphs of the estimated density and age contours are not presented. A short verbal description of the contours is useful, however. In both cities, the innermost density and age segments are upward sloping for both the jointly and individually estimated cases. For both density and age, the two outer contours are either flat or downward sloping for both cities in both estimation cases. The sympathetic movements of age and density evidently reflect the positive association between those variables shown in Table 5. Another salient feature of the graphs is that little change occurs when the identical switchpoint constraint is relaxed. This, of course, reflects the relatively small likelihood differences between joint and individual estimation. Finally, it should be pointed out that the estimated age contours, which show a dramatic increase in age over the inner third of both cities, with discontinuous decreases in evidence out to the city borders, do not mimic the regular contours shown in Figure 1.

References

Akaike, H., "Information Theory and the Extension of the Maximum Likelihood Principle," in B. N. Petrov and F. Csaki, eds., 2nd International Symposium on Information Theory (Budapest: Akailseoniai-Kiudo, 1973).

Akaike, H., "On Entropy Maximization Principle," in P. R. Krishnaiah, ed., Applications of Statistics (Amsterdam: North-Holland, 1977).

Anas, Alex, "Dynamics of Urban Residential Growth," Journal of Urban Economics 5 (1978), 66-87.

Anderson, John E., "Cubic Spline Urban Density Functions," Journal of Urban Economics 12 (1982), 155-167.

Brueckner, Jan K., "A Vintage Model of Urban Growth," Journal of Urban Economics 8 (1980), 389-402.

Brueckner, Jan K., "Testing a Vintage Model of Urban Growth," Journal of Regional Science 21 (1981), 23-35.

Brueckner, Jan K. and Burkhard von Rabenau, "Dynamics of Land-Use for a Closed City," Regional Science and Urban Economics 11 (1981), 1-17.

Clark, Colin, "Urban Population Densities," Journal of the Royal Statistical Society (Series A) 114 (1951), 490-496.

Glickman, Norman J. and Yukio Oguri, "Modeling the Urban Land Market: The Case of Japan," Journal of Urban Economics 6 (1978), 505-525.

Hannan, E. J. and B. J. Quinn, "The Determination of the Order of an Autoregression," Journal of the Royal Statistical Society (Series B) 41 (1979), 190-195.

Kau, James B. and C. F. Lee, "Functional Form, Density Gradient, and the Price Elasticity of Demand for Housing," Urban Studies 13 (1976), 193-200.

Kau, James B., C. F. Lee, and Rong C. Chen, "Structural Shifts in Urban Population Density Gradients: An Empirical Investigation," Journal of Urban Economics 13 (1983), 364-376.

McDonald, John F. and H. Woods Bowman, "Some Tests of Alternative Population Density Functions," Journal of Urban Economics 3 (1976), 242-252.

Mills, Edwin S., "Urban Density Functions," Urban Studies 7 (1970), 5-20.

Mills, Edwin S., Urban Economics (Glenview, Ill.: Scott Foresman, 1972).

Muth, Richard F., Cities and Housing (Chicago: University of Chicago Press, 1969).

Quandt, Richard, "The Estimation of the Parameters of a Linear Regression System Obeying Two Regimes," Journal of the American Statistical Association 53 (1958), 873-880.

Rissanen, J., "Modeling by Shortest Data Description," Automatica, 14 (1978), 465-471.

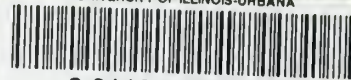Wheaton, William C., "Urban Residential Growth Under Perfect Foresight," Journal of Urban Economics 12 (1982), 1-21.