AD-A095 455 RAND CORP SANTA MONICA CA F/6 5/3
DETERMINANTS OF SPATIAL VARIATIONS IN THE COST OF LIVING.(U)
OCT 79 K NEELS
UNCLASSIFIED RAND/P-6444 NL

1 OF 1
AD A
095465

END
DATE
FILMED
3-81
DTIC

DETERMINANTS OF SPATIAL VARIATIONS IN THE COST OF LIVING

Kevin Neels

October 1979

P-6444

DETERMINANTS OF SPATIAL VARIATIONS IN THE COST OF LIVING

14 RAND/P-

Kevin Neels

October 1979

ACKNOWLEDGMENTS

The author wishes to thank Matthew Neels for his assistance in constructing and auditing the data used in the analysis. Thanks are also due to Peter Rydell and John Mulford for their comments and suggestions, and to Karen Stewart for her help in preparing the draft and final versions of the paper.

SUMMARY

The steady rise in the rate of inflation in recent years has focused a great deal of attention on the dynamics of price change and the importance in all types of analysis of carefully controlling for changes in the cost of living. We have seen a great deal of effort expended in careful attempts to control for inflation and allow for the measurement of real incomes and prices. It has long been recognized that there are large and persistent spatial differences in the cost of living. In spite of this, however, it has remained common for researchers and analysts to assume that prices vary longitudinally but not cross-sectionally, and to ignore spatial variation in the cost of living.

In order to know when such an assumption is justifiable and when it is likely to do significant harm to the analysis, we must understand the determinants of spatial variations in the cost of living. This paper discusses the factors which are likely to cause such variations and reports the results of an empirical investigation of their relative strengths.

The empirical analysis is based upon the Bureau of Labor Statistics' data on Urban Family Budgets. The model is estimated on a pooled time series/cross-sectional sample constructed from data on thirty-eight metropolitan areas over the period from 1969 to 1977.

Much of the cross-sectional variation in the cost of living is explained by differences in climate. Metropolitan areas with cold winters have a higher cost of living than comparable areas located in warmer climates. Metropolitan population size is also related to the cost of living in several ways. Large metropolitan areas have a higher cost of living than small areas. In addition, small areas which are located close to large areas have a higher cost of living than those located far away.

I. INTRODUCTION

The steady rise in the rate of inflation in recent years has focused a great deal of attention on the dynamics of price change and the importance in all types of economic analysis of carefully controlling for changes in the cost of living. We have seen a great deal of effort expended in careful attempts to control for inflation and allow for the measurement of real incomes and prices.

In its commentary on changes in consumer prices, the Bureau of Labor Statistics has long noted that rates of inflation differ by region of the country [1,5]. It has also long been recognized that there are large and persistent spatial differences in the cost of living. In spite of this, it has remained common for researchers and analysts to assume that prices vary longitudinally but not cross-sectionally, and to ignore spatial variation in the cost of living.

In order to know when such an assumption is justifiable and when it is likely to do significant harm to the analysis, we must understand the determinants of spatial variation in the cost of living. In this paper I will report the results of some empirical investigations of the factors which help to determine the cost of living in a metropolitan area.

The following section discusses the components which make up the cost of living and reviews the research which has attempted to uncover their determinants. Here I develop a simple model of the cost of living. Section 3 describes the data sources and the statistical procedures used in the study. Section 4 presents my empirical findings. The last section presents conclusions.

II. DETERMINANTS OF THE COST OF LIVING

By the definition of the Bureau of Labor Statistics, the cost of living in an area is the sum of four components: the cost of commodities purchased by consumers; the cost of personal services; the cost of housing; and the tax burden associated with living in the area. An increase in the size of any component leads directly to an increase in the cost of living. This definitional relationship leads to a simple equation:

$$CL = CC + CS + CH + TB \tag{1}$$

where CL is the cost of living, CC is the cost of commodities, CS is the cost of services, CH is the cost of housing, and TB is the tax burden.

The costs of commodities, services, and housing can be further decomposed into price times quantity:

$$CC = (PC)\ (QC) \tag{2a}$$

$$CS = (PS)\ (QS) \tag{2b}$$

$$CH = (PH)\ (QH) \tag{2c}$$

PC, PS, and PH are the prices of commodities, services, and housing; QC, QS, and QH are the quantities associated with a particular standard of living. Because of the difficulty of measuring either the price or quantity of public services, the tax burden is best regarded as a lump sum transfer from households to the government.

In examining temporal variations in the cost of living, the standard practice has been to hold quantities constant and focus attention on the changes over time in prices. This practice is based upon the fact that over short periods of time, tastes are relatively fixed. Hence, the assumption that the same physical collection of goods and services produces the same level of utility and hence that consumption does not change is an accurate one. This leads to a simple temporal model of the cost of living:

$$CL = f(PC, PS, PH, TB) \quad (3)$$

In considering spatial variations in the cost of living, it is necessary to pay more attention to the quantities consumed. The typical market basket varies dramatically over space, and it is dangerous to assume that a given collection of goods and services provides the same level of utility everywhere. Heating fuel, for example, will not be as important to a Florida resident as it will to a resident of Minnesota. Thus, a spatial model of the cost of living must consider both prices and quantities:

$$CL = f(PC, PS, PH, QC, QS, QH, TB) \quad (4)$$

A full examination of spatial variations in consumption patterns is well beyond the scope of this paper. As a result, I make the simplifying assumption that it is not the composition of the market basket that matters, but rather only its size. The quantity of goods and services needed to provide a given level of utility varies from place to place. This formulation makes sense if one considers, for example, the relationship between climate and interior warmth. To pursue the earlier example, the maintenance of a given interior temperature will require more heating fuel in Minnesota than in Florida. Minnesota will have a correspondingly higher cost of living.

If I label this "quantity requirement" QR, I get the following simplified model of the cost of living:

$$CL = f(PC, PS, PH, TB, QR) \quad (5)$$

My model of quantity requirements is also relatively simple. Much of the variation in this variable will be due to climate-related differences in energy requirements. To capture this effect, I include a variable for the mean January temperature in the locality. Any remaining differences are likely to be due to a host of diverse factors. To account in some way for them, I include a set of dummy variables identifying localities in the northeast, north central, and southern regions of the country. The excluded category includes localities in the western region. Putting all these together leads to the following equation for quantity requirements:

$$QR = f(MJT, DNER, DNCR, DSR) \qquad (6)$$

The prices of commodities such as food, apparel, and other manufactured goods in a locality are influenced by many different forces. First, one must consider the costs of production. These will depend upon the price and availability of factors of production. Resource endowments will vary in a non-uniform way from region to region. Capital is traded in national markets and can be expected to have a similar price in all locations. Labor, in contrast, is a local good. Its price will vary from place to place. Second, one must account for the transportation costs from the point of production to the point of consumption. Metropolitan areas located far from major production centers will have to pay more to import commodities and will have correspondingly higher costs of living. Lastly, one must consider the effects of sales and excise taxes which increase the prices paid by consumers. Combining all these factors leads to the following model of the price of commodities:

$$PC = f(DNER, DNCR, DSR, W, POP, DLM, LST, SST, SGT) \qquad (7)$$

DNER, *DNCR*, and *DSR* are the region dummies which also appeared in equation (6). They are intended to capture variations in production costs due to differences in resource endowments and miscellaneous other factors. *W* is the local wage rate. *POP* is the population of the metropolitan area. I expect its coefficient to be negative, both because of agglomeration economies associated with metropolitan area size and because large metropolitan areas are usually diversified manufacturing centers which can produce many of their own goods. *DLM* is the over-the-road distance to the nearest large metropolitan area. Its coefficient should be positive, reflecting the effects of transportation costs. Finally, *LST*, *SST*, and *SGT* are the local sales tax, state sales tax, and state gasoline tax rates. Their coefficients should, of course, be positive.

Personal services are a highly labor intensive set of goods which are almost entirely produced locally. Their prices should depend largely upon local wage rates. This leads to a relatively simple equation:

$$PS = f(W) \tag{8}$$

The price of housing shows major variations from one place to another. Prices will depend upon the costs of production, which in turn depend upon the prices of land and materials. Residential land prices have been the subject of much theoretical and empirical analysis. Muth has developed a model of the conversion of land from rural to urban uses in which the average price of land in a metropolitan area depends upon the total housing demand in the area and the rent which land can earn in its most profitable rural use [8]. This model implies that land prices will be higher in larger and richer metropolitan areas than in smaller and poorer ones. Prices will also be higher in metropolitan areas where the agricultural land is rich and background rural prices are higher. Witte has examined residential land prices empirically and has generally confirmed Muth's theoretical results [10].

Housing is a prime example of a good which is traded only at the local level. Using data from the Housing Assistance Supply Experiment, Rydell has looked at the effects of local market conditions on housing prices [9]. He finds that in a loose housing market rents drop slightly while capital value and purchase prices drop substantially.

These considerations lead to the following model of local housing prices:

$$PH = f(DNER, DNCR, DSR, POP, GRT) \tag{9}$$

DNER, *DNCR*, and *DSR* capture the effects of background rural land prices. *POP*, metropolitan area population, will also be related to residential land prices. The expected sign of its coefficient will be positive. *GRT* is the growth rate of metropolitan area population. It stands

$$CL = f(DNER, DNCR, DSR, MJT, LST, SST, SGT, DLM, POP, GRT, W) \quad (12)$$

Substituting equation (10) into this and solving for *CL* yields the following reduced form:

$$CL = f(DNER, DNCR, DSR, MJT, LST, SST, SGT, SLM, POP, GRT, UR) \quad (13)$$

The coefficient of mean January temperature should be negative. The coefficients of the tax variables--*LST*, *SST*, and *SGT*--should all be positive. The distance to the nearest large metropolitan area can take a positive or a negative sign, depending upon whether the transport cost or wage gradient effect is stronger. Metropolitan area population will take a positive or negative coefficient depending upon whether the agglomeration economy or residential land price effect is stronger. The coefficient of *GRT* can also be positive or negative. The unemployment rate should be negatively related to the cost of living.

III. DATA SOURCES AND ANALYTICAL PROCEDURES

The model as it was finally estimated is shown below:

$$\begin{aligned} ln\ (UFB) = {} & a_0 + a_1\ CPI + a_2\ DNER + a_3\ DNCR + a_4\ DSR \\ & + a_5\ MJT + a_6\ LST + a_7\ SST + a_8\ SGT + a_9\ DLM \\ & + a_{10}\ ln(POP) + a_{11}\ GRT + a_{12}\ DUR \end{aligned} \tag{14}$$

The dependent variable is the log of the intermediate budget for an urban family of four. This statistic is published annually by the Bureau of Labor Statistics for thirty-eight metropolitan areas. I created a pooled time series/cross-sectional sample by combining data for the thirty-eight areas from the years 1969 through 1978.

The first independent variable is the national consumer price index for the appropriate year. My main interest in this study was in the cross-sectional rather than the temporal variation in the cost of living. I chose to include the *CPI* in the equation rather than to adopt the alternative of separately accounting for all the factors which have caused the overall price level to increase in recent years.

The mean January temperature was entered in degrees Fahrenheit. The local sales tax variable is a zero-one dummy indicating whether the state in which the metropolitan area is located authorizes localities to impose their own supplemental sales taxes. The state sales tax variable is the sales tax rate for the appropriate state. The gasoline tax is entered in cents per gallon. All four variables were taken from various issues of the U.S. Statistical Abstract.

DLM is the over-the-road distance from the metropolitan area to the nearest metropolitan area with a population of four million or over. Values for this variable were obtained from the Interstate Commerce Commission's tables of intercity distances.

Metropolitan area population was entered in logarithmic form. The population data came from the U.S. Statistical Abstract and from the Census Bureau's P-25 and P-26 Current Population Reports. *GRT* is the percent change in metropolitan area population over the preceding five years. This variable was derived from the same sources.

DUR is the difference between the metropolitan area unemployment rate and the U.S. unemployment rate for the same year. These data came from the BLS reports, entitled *Geographic Profile of Employment and Unemployment*. These documents provide unemployment rates for the thirty largest metropolitan areas. For the smaller metropolitan areas in the sample, I used estimates derived from the state unemployment rates contained in the U.S. Statistical Abstract. I found the relationship between state and metropolitan area unemployment rates to be very strong, and so these estimates should be quite accurate.

The use of pooled time series and cross-sectional data increased the number of observations which I had to work with, but also introduced special statistical problems. Because cross-sectional differences in the cost of living tend to remain stable over time, the residuals for successive observations on a particular metropolitan area will be correlated. This has the well-known effect of decreasing the efficiency of the coefficient estimates. More importantly, it biases the standard errors of those coefficient estimates. The presence of correlation means that you add more real information by expanding the cross-section than by adding another observation to the time series. Failure to take this into account will often tend to make your results look better than they really are.

In order to deal with this problem, I estimated the equation using a variance component procedure. This is a version of the Aitken Generalized Least Squares estimator which assumes that the error covariance matrix is block diagonal. Each block corresponds to a cluster of observations on a particular metropolitan area. These blocks themselves have the form:

$$\begin{bmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \ddots & \vdots \\ \vdots & & \ddots & \vdots \\ \rho & \cdots & & 1 \end{bmatrix}$$

where the off-diagonal element is the correlation between the residuals for pairs of observations on a metropolitan area. I estimated this correlation from the data using a maximum likelihood procedure.

IV. RESULTS

The results of the analysis are shown in Table 1. Not surprisingly, the strongest variable in the equation is the consumer price index. This says merely that there is more longitudinal than cross-sectional variation in the cost of living. This was not an unforeseen result.

After controlling for other factors, the northeast region has a higher cost of living than the western region. In the north central and southern regions, the cost of living is lower. These effects are fairly weak, however. Only the southern region has a cost of living which begins to differ significantly from that in the western region.

As expected, the variable for mean January temperature took a negative coefficient. Localities with mild winters have a significantly lower cost of living than those with severe winters.

The sales tax variables took the wrong sign and were completely insignificant in the regression. In contrast, a strong positive relationship existed between the gasoline tax rate and the cost of living. In fact, the relationship is suspiciously strong. A one cent rise in the gas tax would increase the cost of living by roughly $90. This far exceeds what the average household would pay directly as a result of the tax rate change. Part of the explanation may be that high gasoline taxes mean high transportation costs and hence also higher prices for all goods and services. A more likely explanation may be the existence of a reverse causation. That is, where the overall level of prices is low, state governments may be able to subsist on lower revenues and be less likely to levy high gasoline taxes.

The variable for the distance to the nearest major metropolitan area bore a significant relationship to the cost of living. In this case, the wage gradient effect won out over the transport cost effect. The relationship between metropolitan area population and the cost of living was positive and strongly significant. Large metropolitan areas have a higher cost of living than small ones. Among the smaller metropolitan areas, those which are close to the large areas have a higher cost of living than those which are far away.

Table 1

REGRESSION RESULTS

Variable	Coefficient	T-Statistic
Consumer price index	.00718	70.39
Northeast region dummy	.0208	.61
North central region dummy	-.0254	-.75
Southern region dummy	-.0250	-1.22
Mean January temperature	-.00243	-1.92
Local sales tax dummy	-.00145	-.18
State sales tax rate	-.00406	-.70
State gasoline tax	.00879	2.17
Distance to large SMSA	-.0000483	-1.74
Log of SMSA population	.0217	2.92
Five year growth rate	.000233	.77
Local unemployment rate minus U.S. unemployment rate	-.000296	-.14
Constant	8.21	65.18

Sample Size = 154 $R^2 = .97$

From these relationships, we can begin to discern the outlines of a geography of prices. The peaks of the price surface occur in the largest areas. These peaks cast a price shadow over their nearby neighbors. We can observe a price gradient extending out from these peaks to a considerable distance.

The coefficient for the five year population growth rate took a positive sign, but was not significantly different from zero. Finally, we note that the unemployment variable, though it had the expected sign, was insignificant. There is apparently no relationship between local labor market condition and the cost of living.

V. CONCLUSION

This study has confirmed and quantified many of the hypotheses that have been advanced about spatial variations in the cost of living. It is higher in the northeast and lower in the south. It is also substantially higher in the "frostbelt" than in the sunbelt. If Los Angeles were somehow to acquire the climate of Minneapolis, the cost of living in Los Angeles would increase by 11 percent. Finally, it is higher in large metropolitan areas than in small areas. An increase in metropolitan area population from a half to two million will increase the cost of living by three percent.

It is worth noting some of the study's negative findings. Sales taxes appear to have little power to explain differences in the cost of living. There also appears to be little or no relationship between the cost of living and local labor market condition. The latter result must be qualified, however, by the statement that there may exist a complex dynamic relationship which the relatively simple static specification adopted here was not able to detect.

The most noteworthy result to emerge from the study was the finding that the high cost of living in a large metropolitan area also influences the cost of living in nearby smaller areas. A metropolitan area located 500 miles from a large SMSA will have a cost of living 2 percent lower than a comparable metropolitan area located 100 miles away. The identification of the price shadow cast by the largest metropolitan areas adds a new dimension to our understanding of metropolitan dominance.

REFERENCES

1. Richard C. Bahr and Frank Ptacek, "Southern Cities Again Paced Consumer Price Rise in 1974," Monthly Labor Review, 98(5), 1975.

2. Robert F. Engle, "An Exploratory Policy-Oriented Econometric Model of a Metropolitan Area: Boston," Discussion Paper 77-23, Department of Economics, University of California, San Diego, 1977.

3. Werner Z. Hirsch, The Economics of State and Local Government, McGraw-Hill, New York, 1970.

4. Oded Izraeli and Mitchell Kellman, "Changes in Money Wage Rates and Unemployment in Local Labor Markets: The Latest Evidence," Journal of Regional Science, 19(3), 1979.

5. Paul Monson, "West, Small Cities Led Increases in Consumer Prices in 1975," Monthly Labor Review, 99(7), 1975.

6. Leon Moses, "Toward a Theory of Intraurban Wage Differentials and their Influence on Travel Patterns," Papers of the Regional Science Association, Vol. 9, 1962.

7. Thomas Muller, Growing and Declining Areas: A Fiscal Comparison, The Urban Institute, URI13400, 1975.

8. Richard Muth, "The Derived Demand for Urban Residential Land," Urban Studies, 8(3), 1971.

9. C. Peter Rydell, Effects of Market Condition on Prices and Profits of Rental Housing, The Rand Corporation, P-6008, 1977.

10. Ann Dryden Witte, "The Determination of Interurban Residential Price Differentials: A Derived Demand Model with Empirical Testing," Journal of Regional Science, 15(3), 1975.

D

FI

3