An Empirical Analysis of Determinants of Interstate Living-Cost Differentials, 2005

Cebula, Richard and Toma, Michael

Jacksonville University, Armstrong Atlantic State University

10 January 2007

JRAP 38(3): 222-228.

The Journal of
Regional
Analysis
& Policy

An Empirical Analysis of Determinants of Interstate Living-Cost Differentials, 2005

Richard J. Cebula and Michael Toma
Armstrong Atlantic State University

Abstract. This empirical study investigates determinants of interstate living-cost differentials for the year 2005. It seeks to supplement existing related studies, most of which have investigated such differentials but at the metropolitan area or county levels and for earlier time periods. OLS and two-stage least squares (2SLS) results imply that the overall cost of living in a state is positively a function of per capita income and the relative amount of shoreline on major bodies of water, and negatively a function of the presence of right-to-work laws, heating degree days, and toxic chemical releases. Interestingly, the 2SLS estimate adopts two additional amenity/dis-amenity variables (population density and crime rate) as instrumental variables.

1. Introduction

A number of studies have investigated determinants of geographic living-cost differentials (L-CDs) in the U.S., including Cebula (1980) and Cebula and Todd (2004), Cobas (1978), Ostrosky (1983), McMahon (1991), Nord (2000), and Kurre (2003). The study of geographic living cost differentials and their determinants is relevant for a variety of reasons, perhaps the most important of which is that such differentials have consistently found to be statistically significant in explaining geographic mobility in the U.S. (Renas, 1978, 1980, 1983; Cebula, 1978, 1993; Cebula and Alexander, 2006). To the extent that the findings in the present study shed light on living-cost differentials and, by extension, migration, they provide insight to policy makers about the role of factors over which they have no control such as climate or topography and those factors over which they might be able to exert some influence such as right-to-work legislation or pollution abatement.

Most of the L-CD related published research to date has tended to be national in scope and has stressed L-CDs for those metropolitan areas for which geographically comparable data were available. There are exceptions to this pattern, however. For example, Kurre (2003) and Cebula and Todd (2004) address L-CDs at the county level in single individual states, Pennsylvania and Florida, respectively. Furthermore, certain studies, the best known of which may be that by McMahon (1991), investigate the topic at the state level. This empirical note seeks to supplement the existing literature on geographic L-CDs by investigating determinants of interstate L-CDs for the year 2005. Unlike previous related studies, certain quality-of-life/environmental factors are expressly integrated into the present study.

2. The framework

Following previous research, including the recent study by Kurre (2003), the perspective underlying this analysis is that factors tending to elevate demand in a geographic area tend to elevate the overall level of prices in that area, whereas factors tending to elevate supply or reduce production costs in that area tend to lead to a lower overall level of prices in that area. In addition to factors that may directly alter general demand and supply conditions in the economy, this analysis also considers certain quality of life/environmental factors that can be expected to be capitalized into housing prices. Given that housing prices are the single largest component of overall cost-of-living indicators, amenities such as coastal location or dis-amenities such as cold weather, crime, and air pollution that may influence migration (Riew 1973) and housing demand should also have an effect (either direct or indirect) on overall living-cost differentials.

As indicated above, the dependent variable is the state level cost of living index for 2005. Data is unavailable for two states, Maine and New Hampshire, whereas the data set is incomplete for New Jersey, leaving a total of 47 states available for empirical investigation.

Interstate differentials in the index are modeled initially with ordinary least squares analysis and subsequently with two-stage least squares (2SLS), in the latter case to account for possible endogeneity in the model. Ultimately the conclusions from the 2SLS and OLS results are very similar, yet the 2SLS model to some degree addresses, albeit indirectly, the issue of amenity-induced compensating differentials in wages across regional economies (Roback 1982, 1988; Blomquist, Berger and Hoehn 1988; and Clark, et al. 2006). The findings in this research can be interpreted as potentially consistent with the notion that utility-maximizing individuals may be willing to accept lower wage rates in exchange for living in high-amenity areas, and, conversely, may demand higher wages in low-amenity areas. To some extent, our methodology could be interpreted as addressing this issue by including certain amenities and dis-amenities as explanatory variables in both the OLS and 2SLS estimations. The issue is also addressed, in part, in the 2SLS estimation by instrumenting the per capita income variable (which is used to reflect compensating wage rate differentials) with population density, a variable reflecting a portfolio of desirable and undesirable amenities not directly addressed. Furthermore, it is should be noted that another dis-amenity variable, the crime rate, is adopted in the 2SLS estimation as the instrument for the unemployment rate. Thus, amenities and dis-amenities are considered directly and indirectly in the model, partly on the assumption that such factors may be capitalized into housing prices that affect overall geographic living-cost differentials.

Two variables are used characterize the overall demand for goods and services in an area: per capita income and the unemployment rate. Higher per capita income in an area (*PCIj*) is expected to raise the overall demand for goods and services in the area and hence the overall level of prices in the area, *ceteris paribus*. As observed by Kurre 2003, however, a greater effective demand *may* permit firms in some industries to benefit from scale economies that reduce production costs and potentially offset rising prices induced by greater overall demand. Nevertheless, based on Cebula 1980 and 1989, Cebula and Todd 2004, Cobas 1978, Ostrosky 1983, and Kurre 2003, it is expected that *COLj* is directly a function of *PCIj*. On the other hand, higher unemployment rates will reflect a reduced demand for goods and services and create downward pressure on prices, thus reducing the overall cost of living. Thus, *COLj* is expected to be negatively related to *UNj*, *cete-risparibus*.

While *PCIj* and *UNj* may jointly help to explain L-CDs, other amenity-like factors very likely play roles as well. For example, coastal location is deemed by many to be highly desirable. Indeed, it is hypothesized here that for many there is a value in closer proximity to large bodies of water, i.e., in this case, any of the Gulf of Mexico, the Pacific Ocean, the Atlantic Ocean, or the Great Lakes. Accordingly, following Cebula and Todd 2004, it is expected that coastal location (*COASTj*) exercises a positive impact on *COLj* because many people may be willing to pay a premium for living in a coastal area. Alternatively stated, the value of coastal location is capitalized into housing prices and thereby acts to elevate the overall cost of living. Secondly, the role of climate has been shown to be a determinant of migration and thus may be capitalized in housing prices (Cebula 1978) and thereby affect geographic living-cost differentials. In this analysis, climate is proxied by heating degree days (*HDDj).* This variable is lower in warmer climates, reflecting the desirable feature of warmer temperatures that presumably may be capitalized into housing prices. Accordingly, *COLj* is expected to be negatively related to *HDDj*, *ceteris paribus*, since increases in *HDDj* reflect colder climate. Next, the effect of toxic chemical releases (*TOXICj*), measured on a per square mile basis, on *COLj* is considered. This variable reflects interstate pollution conditions, and as an undesirable feature, is expected to influence *COLj* through its impact on housing prices. In particular, higher levels of toxic chemical releases in a state are expected to be capitalized into housing prices and are thusly expected to reduce the overall cost of living in that state.

On the other hand, the greater the overall land area (*AREA)* in a state, the greater may be the distances that commuters of all walks of life may need to traverse. Similarly, the greater the land area in a state, the greater may be the distances over which goods and services providers may need to traverse both their inputs and outputs. In both of these cases, greater land area implies greater pecuniary transportation costs and hence higher overall living-cost levels, *ceteris paribus.*

Finally, Section 14(b) of the Taft-Hartley Act provides that each state shall have the right to enact "right-to-work" laws, laws which provide workers/employees the legal right to refuse to join unions in their place of employment. By nature, states with right-to-work laws tend to be states with weaker labor union influence. As argued in Cebula 1980 and Ostrosky 1983, unit labor costs are likely to be lower in states with right-to-work laws. Accordingly, it is argued that because of lower unit labor costs in states with right-to-work legislation (*RTWj*), the overall price level of goods (and services) created in such states will be lower, *ceteris paribus*.

3. Empirical model and results

Based on the factors described above, the reduced-form equation to be estimated initially is given by:

$$COLj = \beta_0 + \beta_1 PCIj + \beta_2 UNj + \beta_3 AREAj + \beta_4 RTWj + \beta_5 COASTj + \beta_6 HDDj + \beta_7 TOXICj + u \quad (1)$$

where *COLj* is the average cost of living for a four-person family living in state *j* in 2005, expressed as an index, with the average value of *COLj* = 100.00; β_0 is the constant; *UNj* is the percentage unemployment rate in state *j*, 2005; *PCIj* is the per capita income in state *j*, 2005; AREAj is *the geographic area of state j, divided by 1,000,000, expressed in square miles; RTWj* is a binary variable indicating whether state *j* is a right-to-work state: *RTWj* = 1 for those states where right-to-work laws are in effect, and *RTWj* = 0 otherwise; *COASTj* is a relative measure of the amount of land along coastal areas of states with major bodies of water including the Gulf of Mexico, the Pacific Ocean, the Atlantic Ocean, and the Great Lakes; defined as the number of miles of general coastline on a major body of water divided by the land area of the state; *HDDj* is the average annual heating degree days in state j; *TOXICj* is toxic chemical releases into the air, water, or land of state j in thousands of pounds, divided by *AREAj*; and *u* is the stochastic error term.

Note that the first four variables in Equation (1) above are indicators of general demand and supply conditions that may influence interstate living-cost differentials. The last three variables characterize desirable or undesirable amenities that are expected to be capitalized into home prices and thereby affect the cost of living across the states. Based on Section II above, it is expected that:

$$\beta_1 > 0, \beta_2 < 0, \beta_3 > 0, \beta_4 < 0, \beta_5 > 0, \beta_6 < 0, \beta_7 < 0 \quad (2)$$

Descriptive statistics for the data are provided in Table 1, while the sources for all data used include the Missouri Economic Research and Information Center cost of living data series (*COL*)[1]; the Information Please database by Pearson Education, Inc. (*COAST*), the primary sources are the National Oceanic and Atmospheric Administration of the U.S. Department of Commerce[2] and web-sites of the Department of Natural Resources of various Great Lakes states; U.S. Census Bureau, Statistical Abstract of the United States, 2005, Table 640; and 2006, Table 295 (*RTW, CRIME*); North Carolina state data center web-site (*PCI, AREA, TOXIC, HDD, UN, DENSITY*).[3]

Correlation coefficients among the explanatory variables shown in equation (1) are reported in Table 2. For the purposes of the specification given in Equation (1) above, arguably only the correlation (-0.45) between *PCI* and *RTW* is at all even marginally noteworthy; however, as shown in the OLS estimation summarized in equation (3) below, since the *t*-values on both of the estimated coefficients for these two variables are robustly statistically significant, the matter hardly requires much attention.

Estimating equation (1) by OLS, using the White 1980 correction for heteroskedasticity, yields equation (3):

$$\begin{aligned} COLj = {} & 65.76 + 0.002\, PCIj - 1.65\, UNj + 28.2\, AREAj - \\ & (+6.36)\ (+9.39)\ \ (-1.31)\ \ (+1.55) \\ & 10.19\, RTWj + 280.0\, COASTj - 0.002\, HDDj - \\ & (-3.38)\ \ (+3.94)\ \ (-2.93) \\ & 4.11\, TOXICj, \\ & (-3.20) \end{aligned}$$

$$R^2 = 0.79,\ adjR^2 = 0.76,\ F = 21.8 \quad (3)$$

where terms in parentheses beneath coefficients are *t*-values.

In equation (3), five of the seven estimated coefficients are statistically significant at beyond the one percent level with the expected signs. The F-statistic is significant at the one percent level, attesting to the overall strength of the model. The coefficient of determination in equation (3) is 0.79. Thus, the variables in the model jointly explain nearly four-fifths of the variation in the cost of living by state.

The estimated coefficient on variable *PCIj* is positive and statistically significant at well beyond the one percent level, implying that the higher the per capita income in a state, the higher the demand for goods and services and hence the higher the overall level of prices of goods and services in the state. The coefficient on variable *UNj* is negative and but not significant at generally acceptable levels (the five percent level or, at the very least, the ten percent level). Theoretically, higher unemployment rates reduce overall

[1]Available on-line at www.missourieconomy.org/indicators/cost_of_living/index.stm.

[2]Available on-line at www.infoplease.com/ipa/A0001801.html.

[3]Available on-line at http://sdc.state.nc.us.

Table 1. Descriptive statistics

Variable	Mean	Standard Deviation	Maximum	Minimum
COL	103.0	15.54	157.8	86.9
AREA	74,258.9	87,513.2	571,949.0	1,045.0
COAST	0.008	0.020	0.117	0.00
CRIME	414.7	171.9	794.0	78.0
DENS	0.171	0.221	1.034	0.001
HDD	4,896.1	2,214.5	8,812.0	0.00
PCI	29,301.4	4,659.2	45,898.0	22,263.0
RTW	0.468	0.504	1.00	0.00
TOXIC	0.666	0.724	2.97	0.001
UN	4.87	0.969	7.10	2.70

N = 47 for all variables.

Table 2. Correlation coefficients

	AREA	*COASTj*	*CRIMEj*	*DENSITYj*	*HDDj*	*PCINCj*	*RTWj*	*TOXICj*	*UNj*
AREAj	1.00								
COASTj	-0.13	1.00							
CRIMEj	0.21	0.01	1.00						
DENSITYj	-0.34	0.30	0.17	1.00					
HDDj	0.15	-0.21	-0.50	-0.05	1.00				
PCINCj	-0.02	0.14	0.01	0.60	0.33	1.00			
RTWj	0.03	-0.26	0.12	-0.36	-0.37	-0.45	1.00		
TOXICj	-0.29	-0.01	0.34	0.30	-0.22	-0.06	-0.09	1.00	
UNj	0.23	-0.15	0.39	0.01	-0.12	-0.18	-0.15	0.33	1.00

demand for goods and services and create downward pressure on the cost of living; however there is no credible evidence of that phenomenon in this model. Next, the coefficient on variable *AREAj* is positive, but not significant at an acceptable level. Thus, although larger land area may theoretically imply higher transportation costs for firms and consumers, the effect, while beginning to approach a marginally acceptable level of statistical significance, is nevertheless not a statistically significant factor in the determination of overall living-cost differentials. By contrast, the estimated coefficient on variable *RTWj* is negative and significant at the one percent level. This result implies that the overall level of prices of goods and services is lower in states having right-to-work laws, presumably because such legislation leads to weaker labor unions and hence to lower unit labor costs in the state (Cebula 1980).

Turning to the three variables in Equation (1) that represent amenities that may be capitalized into housing prices, note that there is strong statistical evidence that all of these variables significantly influence interstate living-cost differentials. The coefficient for variable *COASTj* is positive and significant at the one percent level, implying that the desirability of coastal location along or relatively near the Gulf of Mexico, the Pacific Ocean, the Atlantic Ocean, or the Great Lakes results in higher overall price levels. Plausibly, this may very well be largely reflected in the higher price of housing in coastal areas, as people capitalize a premium for the value of coastal location into housing prices. Next, the estimated coefficient for variable *HHDj* is negative and statistically significant at the one percent level. *HDD* values are lower in warmer climates and higher in colder climates. Thus, it appears that colder temperatures/colder climates act to negatively impact the overall cost of living, presumably through the capitalization of this (dis-)amenity into housing prices. Another (dis-) amenity captured in Equation (1) is pollution as modeled by toxic chemical releases. The estimated coefficient for variable *TOXICj* in Equation (3) is negative and statistically significant at beyond the one percent level. States featuring higher levels of toxic chemical pollution are presumably deemed as less desirable locations, and this is plausibly capitalized into housing prices, thereby creating downward pressure on the overall cost of living.

The results provided in the above estimation may require additional consideration. In particular, the economic variable *COLj* is contemporaneous with the economic variables *PCIj* and *UNj*. Potentially, this situation could result in simultaneity bias in the estimation. Accordingly, the equation is re-estimated using two-stage least squares (2SLS), with the instruments for *PCIj* and *UNj* being population density in 2004 (*DENSITYj*) and the overall crime rate in 2003 (*CRIMEj*), respectively. The choice of instruments was based on the empirical findings that there is a relatively high degree of correlation between *PCIj* and *DENSITYj*, as well as between *UNj* and *CRIMEj*, whereas these lagged instruments are not correlated with the error terms in the system. The variables used as instruments are defined as follows (data sources provided above):

DENSITYj = the number of persons per square mile (in thousands) in state j, 2004;
CRIMEj = the number of violent crimes per 100,000 population in state j, 2003.

Interestingly, these two instruments also reflect facets of the quality of life not already explicitly addressed in the model. One may consider *DENSITYj* as an indicator of a portfolio of desirable amenities often associated with more urban areas, such as access to cultural offerings in the arts, or greater access to health care, or professional sporting venues, or the vibrancy of large urban areas not modeled in Equation (3). Conversely, high-density areas may also be characterized by poor public schools, traffic congestion, pedestrian congestion, poverty, or even high property taxation. In either case, the use of *DENSITYj* as an instrument creates a channel through which such desirable or undesirable features may indirectly influence interstate living-cost differentials while more directly influencing the income variable (*PCIj*). Similarly, the use of *CRIMEj* as an instrument for *UNj* also creates an indirect channel through which other undesirable social ills associated with crime may influence the overall cost-of-living.

The 2SLS estimation of equation (1), using *DENSITYj* and *CRIMEj* as the instruments for *PCIj* and *UNj*, respectively, while adopting the White (1980) correction for heteroskedasticity, yields equation (4):

$$\begin{aligned} COLj = {} & \underset{(+2.07)}{43.86} + \underset{(+4.80)}{0.003}\,PCIj - \underset{(-0.84)}{1.34}\,UNj + \underset{(+1.67)}{29.8}\,AREAj - \\ & \underset{(-2.04)}{7.80}\,RTWj + \underset{(+4.02)}{277.2}\,COASTj - \underset{(-3.06)}{0.002}\,HDDj - \\ & \underset{(-2.48)}{3.86}\,TOXICj, \end{aligned}$$

$$F = 15.9 \qquad (4)$$

where terms in parentheses beneath coefficients are *t*-values.

The conclusions to be inferred from the 2SLS estimation results shown in Equation (4) differ little from those shown in the OLS estimate in Equation (3), suggesting that little simultaneity bias is introduced into the system from any endogeneity involving *COLj*, *PCIj*, and *UNj*. Of the five variables found in the OLS estimation to be significant at the one percent level, in Equation (4), three (*PCIj*, *COASTj*, and *HDDj*) remain similarly significant, while two variables (*RTWj* and *TOXICj*) are now statistically significant at beyond the five percent level. Furthermore, in this 2SLS estimate, the variable *AREAj* is now nearly significant at the ten percent level, while *UNj* remains insignificant at an even marginally acceptable level. Of the 2SLS results, the most interesting finding would seem to be that the t-statistic on the variable *PCIj* is reduced by nearly one-half to 4.80 from 9.39. This outcome suggests at least some evidence of simultaneity in the system. Overall, however, the OLS and 2SLS results are very similar and yield the same basic conclusions.

4. Conclusion

This empirical study finds that, for the year 2005, the overall cost of living in the j^{th} state was an increasing function of the state's per capita income and location of the state on the Gulf of Mexico, the Pacific Ocean, the Atlantic Ocean, or the Great Lakes. In addition, the overall cost of living was found to be lower in those states having right-to-work laws, colder temperatures, and higher levels of toxic chemical releases. While the OLS estimate presented may possess some simultaneity bias, the 2SLS estimation suggests that this problem is limited in nature. Additionally, to some degree, the 2SLS methodology indirectly addresses the potential issue of amenity-related compensating differentials in wage rates raised by Roback 1982 and 1988 and considered further by Clark, et al.

2006. Our methodology controls for this issue to at least some extent by including amenities in the above models estimated. Amenities and dis-amenities are addressed directly and indirectly in the model on the presumption that these factors are likely to be to at least some degree capitalized into housing prices that in turn affect overall geographic living-cost differentials.

Lastly, these results may be of interest to state-level or regional policy makers because they present evidence of the role of different factors on the overall cost of living that has been previously shown to affect geographic mobility (Renas 1978, 1980, 1983; Cebula 1978, 1993; Cebula and Alexander 2006] and hence, population growth and regional economic growth and development. In general, given the inability of policy makers to affect many of the causal variables impacting the cost of living, such as climate, topography, or land area, or their very limited ability to meaningfully change per capita income and unemployment, the results presented here may serve to remind policy makers to further consider the role of right-to-work legislation or pollution control/abatement.

While legislators may influence the latter two factors (variables), these factors may present interesting challenges to policy makers. For example, to the extent that employers are attracted to a state by relatively lower wages in right-to-work states, the influx of new establishments/firms might then act to increase per capita income and/or to reduce unemployment, both of which would tend to create upward pressure on the cost of living. The latter effects would, of course, slow economic growth over time. Furthermore, efforts in any given state at a reduction in pollution through abatement statutes and regulations, while providing potentially beneficial health outcomes, would then also potentially increase the cost of production (as well as the cost of living), *ceteris paribus*, perhaps discouraging firms from relocating to that particular state.

Perhaps the most useful practical implication of the results in this study for policy makers is simply to highlight the role that certain factors play with respect to the cost of living. This information would then perhaps provide some insights into the ultimate impact of these factors on migration and the resulting issues associated with increasing or decreasing population growth. Potentially, policy makers could then direct their attention to the implications of increasing (or decreasing) growth in terms of infrastructure and human resource needs.

References

Blomquist, G., M. Berger, and J. Hoehn. 1988. New estimates on the quality of life in urban areas. *American Economic Review* 78: 89-107.

Cebula, R.J. 1978. An empirical note on the Tiebout-Tullock Hypothesis. *Quarterly Journal of Economics* 92(4): 705-11.

______. 1980. An empirical note on determinants of geographic living cost differentials. *Land Economics* 56: 477-81.

______. 1989. The analysis of geographic living cost differentials: A brief empirical note. *Land Economics* 65: 64-7.

______. 1993. The impact of living costs on geographic mobility. *Quarterly Review of Economics and Finance* 33(1): 101-5.

______ and G. Alexander. 2006. Determinants of net interstate migration, 2000-2004. *Journal of Regional Analysis and Policy* 36: 115-127.

______ and S. Todd. 2004. An empirical note on determinants of geographic living cost differentials for counties in the state of Florida, 2003. *Review of Regional Studies* 34: 112-9.

Clark, D., W. Herrin, T. Knapp, and N. White. 2003. Incomplete compensation and migration behavior: Has anything changed between 1990 and 2000? *Journal of Regional Analysis and Policy* 36(2): 140-152.

Cobas, J.A. 1978. A method to estimate BLS family budgets for all standard metropolitan statistical areas. *Social Science Quarterly* 59: 546-52.

Kurre, J.A. 2003. Is the cost of living less in rural areas? *International Regional Science Review* 26: 86-116.

McMahon, W.W. 1991. Geographical cost of living differences: An update. *AREUEA Journal*, 19: 426-50.

Nord, M. 2000. Does it cost less to live in rural areas? Evidence from new data on food security and hunger. *Rural Sociology* 65: 104-25.

Ostrosky, A.L. 1983. Determinants of geographic living cost differentials: Comment. *Land Economics* 59: 212-3.

Renas, S. 1978. The cost of living, labor market opportunities, and the migration decision. *Annals of Regional Science* 15(1): 95-104.

______. 1980. An empirical note on the Tiebout-Tullock Hypothesis: Comment. *Quarterly Journal of Economics* 94(3): 619-23.

______. 1983. The cost of living, labor market opportunities, and the migration decision: More on problems of misspecification and aggregation bias. *Annals of Regional Science* 17(1): 98-110.

Riew, J. 1973. Migration and public policy. *Journal of Regional Science* 13(2): 65-76.

Roback, J. 1982. Wages, rents, and the quality of life. *Journal of Political Economy* 90: 1257-1278.

_______. 1988. Wages, rents, and amenities: Differences among workers and regions. *Economic Inquiry* 26: 23-41.

U.S. Census Bureau. 2005. *Statistical Abstract of the United States, 2005*. Washington, D.C.: U.S. Government Printing Office.

White, H. 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity *Econometrica* 48: 817-38.