INTERNATIONAL REGIONAL SCIENCE REVIEW 26, 1: 86–116 (January 2003)

IS THE COST OF LIVING LESS IN RURAL AREAS?

JAMES A. KURRE
School of Business, Pennsylvania State University–Erie, Erie, PA, k12@psu.edu

There seems to be a general presumption that rural areas benefit from lower costs of living than urban areas. However, there is relatively little systematic data, other than anecdotal evidence, to support this. To test this presumption, this study develops spatial cost of living estimates for each of the sixty-seven counties of Pennsylvania. In addition to the overall cost of living, it generates indexes for each of six component subindexes: groceries, housing, utilities, transportation, health care, and miscellaneous goods and services. These estimates find that the average urban resident of Pennsylvania pays about 6 percent more than rural residents for a broad basket of goods and services. Moreover, urban residents pay more on average for all six major categories of goods, with the greatest difference (12.7 percent) occurring for housing costs. Interestingly, the article also finds that the efficiency of the local public sector can affect local cost of living. The article also identifies policy implications of these differences for economic development and other purposes.

There seems to be a general presumption that rural areas benefit from lower costs of living compared to urban areas. However, there is relatively little systematic data, other than anecdotal evidence, to support this.

For example, McHugh (1990, 32) said, "While it is known that cost of living differences between rural and urban areas exist and that they affect the pattern of economic development, there is no consistent and comprehensive measure of these cost differences currently available." More recently, Isserman (2001, 45) said that rural America offers "lower land costs, lower building costs, lower housing prices, lower labor costs, lower security costs, lower parking costs, and lower taxes." But he cited no data to support this, perhaps because it seems self-evident.

Lower costs in rural America seem intuitive, but might there be other costs that are higher in rural areas, more than offsetting the lower costs that Isserman (2001) cited? McHugh (1990) found that rural households spent more than urban households on some items, such as transportation, health care, and tobacco. And Marshall

This project was made possible by a grant from the Center for Rural Pennsylvania, a legislative agency of the Pennsylvania General Assembly. Responsibility for this work and its policy implications remain with the author. Special thanks to research assistants Christopher Collins and Matthew Dubowski, whose able assistance and attention to detail added greatly to this research.

DOI: 10.1177/0160017602238987

(2001) pointed out that we should expect rural telecommunications services to be more expensive due to greater distances, lower population densities, and the inability to take advantage of economies of scale. Kurre (2000) found that rural families commuted farther in 1990 than their urban counterparts, presumably incurring higher money travel costs (if not time costs, since they might be expected to face less congestion). Lack of scale in rural areas might also be expected to result in fewer local sellers of many goods and services and so less competition and less competitive prices.

It seems, then, that an a priori case could be made either for a lower or a higher cost of living (COL) in rural areas.[1] Given this, it is logical to turn to the data. The problem is that there is relatively little systematic data about rural costs of living. The work that has been done on spatial cost differences generally (not just urban-rural) leads to broad agreement that COL varies from place to place within the nation. Virtually every study that looks at spatial COL finds significant differences between places. For example, Walden (1997, 237) said, "It is now well-established that prices vary between states." McMahon (1991, 426) said, "Significant differences in the cost of living exist among different parts of the country, as well as among different rural and urban counties of the same state." And studies of interregional income variations and convergence have found that COL differences play a major role (for example, DuMond, Hirsch, and Macpherson 1999; Wojan and Maung 1998; Walden 1997; Deller, Shields, and Tomberlin 1996; Bishop, Formby, and Thistle 1994; and Eberts and Schweitzer 1994).

Unfortunately, data on COL variation are relatively scarce. There is currently no official government program in the United States to provide information on this important topic. Johnston, McKinney, and Stark (1996, 568) said, "There is without doubt a need for data on regional variations in prices or costs of living." Deller, Shields, and Tomberlin (1996, 110) said, "Our findings . . . are . . . limited by the availability of good regional price data . . . regional scientists need to develop a research program to address the shortcomings of our data." And Koo, Phillips, and Sigalla (2000, 135) said, "There is a great need for information about regional COL's." But a panel of price research experts convened in 1999 by the National Bureau of Economic Research to "present their individual views on what research agendas they would propose as meriting highest priority over the next 20 years" (Abraham et al. 2000, 31) did not even mention the topic of spatial price comparisons. This is very disappointing to regional scientists.

The private sector has responded to some extent, though, to this lack of data. Runzheimer International makes a substantial part of its revenues from estimating living and travel costs in different locations worldwide. More than two thousand clients pay for this kind of spatial COL data from them, and according to Runzheimer (1998), more than half of U.S. companies pay salary differentials based on geographic COL differences. The federal government also adjusts salaries for some of its employees based on COL differences, despite having no official

COL statistics available to the public (U.S. Office of Personnel Management, 1997, 1999).

Probably the most widely available source of data on COL differences comes from the American Chamber of Commerce Researchers Association (ACCRA), which has been publishing data on COL differences in American urban areas since 1968. (The ACCRA data are explained in detail in the next section.) More recently, other firms have begun to provide COL data on the web. HomeFair (http://www2.homefair.com) offers comparison of costs in pairs of cities, including some foreign cities, in their "Salary Calculator." HomeFair bases its COL comparisons on data collected by its own staff and that of the Center for Mobility Resources. Its five major categories for U.S. data are housing costs (33 percent), utilities (8 percent), consumables (16 percent), transportation (10 percent), and other services (33 percent). Taxes are not included.

DataMasters (http://www.datamasters.com) offers similar bilateral COL comparisons, using the ACCRA data as the basis for their own data. But DataMasters expands on the ACCRA data by incorporating information on tax differences between locations, which are conspicuously absent from the ACCRA database. DataMasters claims that this results in as many as two-thirds of the index values decreasing, some by double-digit amounts. And ReloSmart (http://www.VirtualRelocation.com) provides a similar service in its Cost of Living Tool, but it requires detailed information on the user's income, expenses, mortgage, and even expected interest rates to provide a customized COL value for each user.

All of these Web sites typically provide comparisons only for pairs of cities rather than data for all places at one time. And unlike ACCRA, they are not always willing to provide detailed information about their data sources and methodology. Most important for our purposes, most of these COL data are for urban areas. For example, the requirements of the ACCRA program concerning the types of products and number and types of pricing locations make it very difficult for a rural area to participate. Given the paucity of rural COL data, those who wish to study rural-urban COL differentials must typically create their own estimates. The next section reviews some of those approaches.

Techniques for Estimating COL Differences

Primary Price Data Collection

The most obvious approach to constructing a COL index is to actually collect price data from all sites to be included in the study, for a representative basket of goods and services that a typical household would consume, as the Bureau of Labor Statistics does in calculating the Consumer Price Index.

Collection of such data by a private group (the Rewards Group) in Britain allowed Borooah et al. (1996) and Johnston, McKinney, and Stark (1996) to

construct spatial COL indexes for twelve regions of the United Kingdom. An American example is the Florida Price Level Index, generated annually by the Florida Department of Education (Denslow 1999; Florida Department of Education 1996-2001). The index uses 117 items that are priced regularly in all counties of the state. Data collection for this project is quite expensive, running hundreds of thousands of dollars per year. As mentioned above, the resulting COL index has been used since 1974 to adjust state aid to Florida schools (McMahon 1991, 430).

One popular approach has been to build on or update the last spatial COL indexes that were estimated by the U.S. government through its Bureau of Labor Statistics' Family Budget program in 1982 (U.S. Department of Labor 1982). The Family Budgets are the most recent effort by the federal government to provide an official spatial COL index. Indexes were provided for twenty-four large metro areas, as well as an aggregate index for all metro areas and one for all nonmetro areas. Some researchers have updated the Family Budget data using price data from the Consumer Price Index for selected metro areas. This approach was used by McMahon (1991) and Eberts and Schweitzer (1994). However, caution is warranted in basing a current COL index on the Family Budget data for 1981. In fact, the last release of the Family Budget explains the termination of the program in this way:

> The expenditure data on which the budgets are based are now 20 years old and continuation of the program would require a revision of concepts, more current expenditure data, and extensive collection of price data, for which funding was not available. (U.S. Department of Labor 1982, 44)

Clearly, the Bureau of Labor Statistics had some misgivings about the accuracy of the program at the time that it was discontinued (see also Watts 1980).

A related approach is to calculate the cost of self-sufficiency for families of various types in various places. The key question in these studies is, "How much would a household have to earn to attain self-sufficiency?" Examples include Pearce and Brooks (1997) and Zimmerman and Garkovich (1997). They vary from the studies mentioned above in that they define their market basket to focus on lower-income households and use published data rather than actually collecting the data themselves. This results in assuming that some costs, such as food, are invariant over space. This approach frequently is tied to the "living wage" issue, a type of alternative minimum wage applied to government contractors, which some localities are adopting. (For information, see the *Living Wage Issue Guide* of the Economic Policy Institute at http://epinet.org/.)

As mentioned earlier, the most widely used spatial price index in the United States is the *ACCRA Cost of Living Index* (see ACCRA 1997), which is published quarterly by ACCRA (formerly the American Chamber of Commerce Researchers Association). The *Index* measures the cost of buying a specific basket of goods and services in a large number of urban areas around North America each quarter.

Volunteers from each area, often from chambers of commerce, agree to price the fifty-nine or so items that the ACCRA team has identified as representing the standard of living of a midmanagement executive household. Table 1 presents a list of items and their relative weights in the market basket. The midmanagement executive is a salaried employee, typically in the top quintile of income, earning perhaps double the average household income in the area. The total basket of goods and services to be priced is divided into six subcategories: groceries, housing, utilities, health care, transportation, and miscellaneous. The ACCRA team chooses items to represent each category based on the Bureau of Labor Statistics' 1992 Consumer Expenditure Survey, which is also the basis for the Consumer Price Index. These items are priced by volunteers in each area, and the ACCRA home office checks the data and calculates the indexes.[2] The index for each quarter is unique, since a different group of areas typically participates in each edition. The index value of 100.0 represents the average cost of the specified basket for that quarter's particular group of participants.

The ACCRA *Index* has some limitations, however. ACCRA does not attempt to include the government sector in its cost of living calculations since no acceptable way has been found of integrating the myriad forms of taxes imposed by multiple levels of government in different jurisdictions as well as somehow evaluating the service levels provided. While the ACCRA basket has been carefully designed to cover all major categories of consumer expenditures, it does not adequately represent some of the subcategories such as vehicle or furniture purchases, shelter repairs, and educational expenses. And since participation in the index is voluntary, the rather large group of areas that typically participate cannot be assumed to be a random sample of American communities. It comes as no surprise that some of the most expensive areas have not participated on a regular basis. Of course, this has the effect of making other areas seem more expensive, in relative terms, than they would if these high-cost areas were represented. On the other hand, the smallest and most rural areas are often unable to participate, and this may have the opposite effect. Finally, although ACCRA provides a very detailed guide for the volunteers on how to collect price data, and the ACCRA head office then looks for anomalous data before publication, it is possible that those actually collecting the data make some errors that remain undiscovered, as might be expected in any data collection program of this magnitude. A cynic might also suggest that a chamber of commerce may have a vested interest in seeing low COL numbers for its local area, which would make the place seem more attractive. Despite these drawbacks, the *ACCRA Cost of Living Index* is the most widely used of all the spatial price index databases.

As detailed in Kurre (1992, 2-3), there are typically immense logistical problems and high costs to conducting a program of primary price collection, especially in rural areas. Unless such a data collection program already exists for some reason, the direct pricing technique is not a practical approach for most areas.

TABLE 1. The *ACCRA Cost of Living Index* Market Basket

Weight	*Item*
16.0000 percent: *GROCERY ITEMS*	
0.5840	T-bone steak, price per pound
0.5840	Ground beef or hamburger, price per pound, lowest price
0.7600	Sausage, price per pound, Jimmy Dean brand, 100 percent pork
0.7584	Frying chicken, price per pound, whole fryer
0.5648	Tuna, 6.0-6.125 oz. can, chunk light, Starkist or Chicken of the Sea
0.6176	Whole milk, half-gallon carton
0.1280	Eggs, one dozen grade A large
0.4720	Margarine, one pound cubes, Blue Bonnet or Parkay
0.4720	Parmesan cheese, 8 oz., grated, Kraft
0.3472	Potatoes, 10 lb. sack, white or red
0.7056	Bananas, price per pound
0.3472	Iceberg lettuce, approx. 1.25 pound head
1.4608	White bread, 24 oz. loaf or 24-oz. equivalent, lowest price
1.0768	Cigarettes, carton, king size (85 mm.), Winston
1.1024	Coffee, 13 oz. can, vacuum packed, Maxwell House, Hills Brothers or Folgers
0.5328	Sugar, 4 lb., cane or beet, lowest price
0.7328	Corn flakes, 18 oz., Kellogg's or Post Toasties
0.0928	Sweet peas, 15-17 oz. can, Del Monte or Green Giant
0.0928	Tomatoes, 14.5 oz. can, Hunt's or Del Monte
0.2752	Peaches, 29 oz. can, halves or slices, Hunt's, Del Monte or Libby's
0.6048	Facial tissues, 175-count box, Kleenex brand
0.6656	Dishwashing powder, 50 oz., Cascade
0.3456	Shortening, 3 lb. can, all vegetable, Crisco
0.8304	Frozen orange juice, 12 oz. can, Minute Maid
0.1840	Frozen corn, 16 oz., whole kernel, lowest price
0.0689	Baby food, 4.0-4.5 oz. jar, strained vegetables, lowest price
0.5600	Soft drink, 2 liter, Coca Cola, excluding any deposit
28.0000 percent: *HOUSING*	
5.3816	Apartment, monthly rent; two bedroom, unfurnished, excluding all utilities except water, 1½ or 2 baths, approx. 950 sq. ft.
22.6184	Home purchase, consisting of monthly principal and interest payment on a 30-year first mortgage, based on 75 percent loan with current conventional fixed rate mortgage, on 1,800 sq. ft. living area new house, with approximately 8,000 sq. ft. lot, in appropriate urban area with all utilities
8.0000 percent: *UTILITIES*	
6.9280	Total energy costs at current rates for average monthly consumption of all types of energy during the previous 12 months for the type of home specified above
1.0720	Telephone, private residential line, customer owns instruments. Price includes: basic monthly rate; additional local user charges, if any, for a family of four; touch tone fee; all other mandatory monthly charges, such as long distance access fee and 911 fee; and all taxes on foregoing

(continued)

TABLE 1 (continued)

Weight	*Item*
10.0000 percent: *TRANSPORTATION*	
1.0700	Commuting fare, typical one way, up to 10 miles
3.7000	Auto maintenance, average price for computer or spin balance of one front wheel
5.2300	Gasoline, one gallon, unleaded regular, national brand, cash price at self-service pump, including all taxes
5.0000 percent: *HEALTH CARE*	
0.8750	Hospital room, semi-private, average cost per day
1.7545	Doctor, office visit, general practitioner's routine exam of established patient, average charge
1.7545	Dentist, office visit, adult teeth cleaning and periodic oral exam
0.6160	Antibiotic ointment, 1/2 oz. tube, Polysporin
33.0000 percent: *MISCELLANEOUS GOODS AND SERVICES*	
3.0822	Hamburger sandwich, quarter pound patty with cheese, McDonald's
3.0822	Pizza, 11-12" thin crust cheese pizza, Pizza Hut or Pizza Inn
3.0822	Fried chicken, thigh and drumstick, Kentucky Fried Chicken or Church's
0.6171	Man's barbershop haircut, no styling
0.6171	Woman's beauty salon visit, including shampoo, trim and blow-dry
0.6171	Toothpaste, 6-7 oz. tube, Crest or Colgate
0.6171	Shampoo, 15 oz. bottle, Alberto VO5
0.6171	Dry cleaning, man's two-piece suit
4.3131	Man's dress shirt, 100 percent cotton pinpoint Oxford, long sleeves
1.3629	Boy's underwear, three briefs, size 10-14, cotton, lowest price
4.3131	Man's denim jeans, Levi's brand, 501 or 505, rinsed or washed or bleached, size 28/30-34/36
1.5873	Major appliance repair, home service call, washing machine, minimum labor charge excluding parts
0.9438	Newspaper subscription, daily and Sunday home delivery of large city paper, monthly rate
1.3365	Movie, first run, indoor, evening, no discount
1.3365	Bowling, average price per game, evening rate
2.2638	Tennis balls, can of three extra-duty, yellow, Wilson or Penn brand
1.5213	Child's game, "Monopoly", No. 9 edition
0.5643	Liquor, J&B Scotch, 750 ml. bottle
0.5610	Beer, 6-pack of 12 oz. containers, Miller Lite or Budweiser, excluding deposit
0.5643	Wine, 1.5 liter bottle Chablis blanc, Gallo
100.0000 percent TOTAL	

Note: This table lists the items that made up the American Chamber of Commerce Researchers Association (ACCRA) market basket for a midmanagement executive household in the third quarter of 1997. These are not the precise item definitions; exact definitions are given in the pricing manual (ACCRA 1997) to ensure uniformity in pricing methods across areas.

Estimation of a Complete Set of Demand Equations for All Commodities in All Places

This approach is based solidly in microeconomic theory and starts with an extensive set of demand equations, one for each commodity consumed. Theoretically, the equations would have interaction variables to allow for substitution and complementarity among all commodities consumed. However, this approach gets very complicated very quickly and typically is not operational due to data requirements (Crawford 1996). Ravallion and van de Walle (1991), using a partial version of this approach, found urban costs to be about 10 percent higher than rural costs in Java, Indonesia.

Estimation of a Regression Model of the Factors That Explain (Predict) COL in an Area

This technique involves identifying the factors that cause (or at least are correlated with) the COL differences between places. Starting with a COL database for a set of areas, this approach fits a structural equation to those data. Then data for areas that did not participate in the initial study can be applied to this equation to estimate their COL values. Many of those who currently work in the field of spatial COL use a variation of this approach. This includes Walden (1997, 1998); Cebula (1993); Cebula, Alexander, and Koch (1992); Kurre (1992); and McMahon (1991).

This approach relies on the preexistence of a COL database generated by one of the other techniques. Since there are few of these, relatively little work has been done in this field. However, there are three databases that have been used: the ACCRA database, the Florida Price Level Index, and an updated version of the Bureau of Labor Statistics' Family Budgets. Among these three, the ACCRA database seems clearly superior. It is obviously more current than the Family Budget studies (which ended in 1982) and has broader coverage than the Florida data (which apply, of course, only to counties in Florida). While the ACCRA database is certainly not without its problems (see Koo, Phillips, and Sigalla 2000 for discussion), it seems to be the best currently available database on which to build a COL study.

Estimation of COL Data from Expenditure Data

This approach is based on the premise that variations in expenditures (i.e., prices times quantities) can be used to approximate prices alone. Early work was done by Ghelfi (1988) and McHugh (1990) and more recently by Voicu and Lahr (1999). The problem is that quantities will vary as prices do—the law of demand. As a result, changes in expenditures tend to reflect both price and quantity changes, rendering use of expenditure data suspect. Voicu and Lahr attempted to address this

issue cleverly through use of commodity weights based on elasticities, but it is not clear that their technique measures price variation alone across regions. It may be measuring, at least partially, differences in the quantities of goods purchased and, therefore, differences in the standard of living between areas. Ghelfi found that urban expenditures were about 10 percent higher than rural in Wisconsin using this approach.

Conclusions

A few other studies have used different approaches and/or different data. For example, Nord (2000) used data on food security to estimate metro-nonmetro COL differentials, finding costs to be about 16 percent lower in nonmetro areas. But a review of the literature suggests that the regression approach is most useful in estimating spatial COL differences, and that is the approach used in this study.

Estimating the COL

The basic approach of this project is to specify a model, using existing data, that will allow us to identify key determinants—or at least correlates—of COL differences across space in the United States. Once the model is specified and estimated, it can be applied to data for the sixty-seven counties of Pennsylvania to generate a COL estimate for each county. This will permit comparison of costs for the state's rural counties versus its urban counties. The ACCRA data are for urban areas (some in predominantly urban counties and some in predominantly rural counties), so that is the basic geographical unit for fitting the model. Since the focus of the project is to estimate COL for counties, most of the independent variables use county-level data.

ACCRA data for urban areas from the third quarter of 1997 were used as the dependent variable to calibrate the model. Choice of this time frame was due primarily to issues of data availability for the independent variables. Although 321 urban areas participated in the *ACCRA Cost of Living Index* for that quarter, 18 were eliminated due to inappropriateness for the study (e.g., areas in Alaska and Saskatchewan) or data problems (such as no county-equivalent area). The resulting database, used for fitting the model, included 303 urban areas from all over the country.

Basic economic theory can be used to identify the determining factors in this model. The fundamental idea is relatively simple: factors that increase demand in an area will tend to cause local prices to be higher; those that tend to increase supply will cause prices to be lower. Of course, economies of scale and agglomeration economies play a role in this process as well. Based on this theory as well as previous work in the field, the key variables to be examined are population, income, density, growth, utilities, government, and unemployment rate.[3]

POPULATION

Population is included as a measure of demand; if there are more people competing to purchase the supply of a product, the price will be driven higher. On the other hand, a larger demand may mean that firms can take advantage of economies of scale, leading to lower costs—at least for some production processes. Given the offsetting nature of these two factors, it is not immediately clear which would predominate. Previous research (Knapp, Gorski, and Cox 1988; Cebula 1989; Deller and McConnon 1990; and Raper 1990) has typically found that larger population has tended to be associated with higher COL, and that will be our working hypothesis. Data on population were taken from the Regional Economic Information System (REIS) of the U.S. Bureau of Economic Analysis (1999).

INCOME

Income is expected to affect COL in much the same way as population. If two cities have the same population but one has a higher income per capita, the richer city would experience greater demand for most goods with concomitant upward pressure on prices. Of course, the economies of scale effect could have the same impact here as well. It is not immediately clear that all prices would be affected equally, however. Higher incomes will result in greater demand for goods with income elastic demands while the demand for inferior goods would fall. In other words, the price of shoes and canned peas might not be affected as much as the price of champagne and facials. Data on income were also taken from REIS (U.S. Bureau of Economic Analysis 1999).

The impact of these two variables, population and income per capita, might be better measured as an aggregate—the total income of residents of the area. This study will consider the effect both ways: population and income per capita separately and in the form of aggregate income (population times income per capita). One or the other of these approaches may be more fruitful, especially if there are multicollinearity problems between the population variable and one of the income measures. It should be mentioned that independent variables that are highly correlated (such as population and aggregate income) are not introduced into the estimating equation simultaneously, to avoid problems of multicollinearity.

DENSITY

Density of population—as distinct from sheer numbers of people—may also have an effect on COL. If two cities each have a million inhabitants, but one has them concentrated into a land area that is only a fourth of the other's, we may expect that city to have greater congestion and resulting transportation problems, higher land costs, and worse problems with environmental issues. As a result, COL may be expected to be higher in that place. Data on population density were calculated from

the population data discussed above and land area data taken from *ArcView GIS* from Environmental Systems Research Institute, Inc. (1996).

GROWTH

Aside from sheer size—either in terms of people or the money they have to spend—the rate of change in that size may have an important effect on the area's COL. The logic is that increasing demand resulting from growth may run ahead of supply response, causing higher prices in growing areas until supply can catch up. Thus, faster growth recently may mean higher prices. This approach recognizes that short- and long-run elasticities of supply may vary across types of products, with housing responding more slowly to increases in demand than more easily transportable products such as groceries. Data on growth rates were calculated from the population and income data discussed above.

UTILITIES

The price of utilities in a local area can certainly have an impact on local COL; electric and gas prices play their part in determining a homeowner's budget. However, these prices are a little different from most local prices in that they have traditionally been heavily regulated in this country. Unlike typical market prices, which are determined by supply and demand, electricity and gas prices reflect political influences in an important way. Political appointments to state public utility commissions can affect the cost of heating or cooling the homestead. Utility costs can also affect local COL in an indirect way since they will have an impact on the cost of production of goods and services produced locally. Given their special status, it makes sense to include the price of local utilities as separate determinants, unlike other prices. We include the actual prices (per kilowatt hour or per thousand cubic feet of natural gas) rather than a measure like "average gas bill" since the latter would also include the quantity of the good purchased, which is not our intent. (A true COL index, rather than a spatial price index, might wish to consider this, however.)

Since data on utility rates are not available by county, state-level data were used as a proxy; the state utility rate was used for each of that state's counties. While this is clearly not the preferred solution, it was decided that inclusion of state-level data would introduce less error into the estimates than complete exclusion of the energy cost variables. State-level data on cost per kilowatt hour of electricity and dollars per thousand cubic feet of natural gas delivered to residential consumers both came from the Energy Information Administration of the U.S. Department of Energy (1998a, 1998b).

GOVERNMENT

Amount, type, and quality of services provided by local governments vary dramatically in this country. Like utility costs, these have both direct and indirect effects on the COL in a locality. A government that provides excellent education and efficient garbage collection saves residents the costs of providing similar services out of their own budgets. Similarly, effective police protection and local street maintenance helps keep costs low for local producers, resulting in lower prices for locally produced goods and services. Of course, there is no such thing as a free lunch, and those services have to be paid for somehow. Measurement of local government efficiency must involve two components: the amount and quality of services provided and the cost to local residents in terms of taxes and other charges by local government. Actually measuring these, especially the services provided, is a bit tricky, but the idea is simple conceptually: governments that provide better service for a lower cost (or more and better services for the same cost) are more efficient and contribute to a lower COL.

For this analysis, the government sector variable attempts to measure the efficiency of local governments in a county by comparing revenues raised from citizens with the employment by local government (as a proxy for services provided). Note that although the geographical level is the county, the data include information for all local governments in the county, including counties, municipalities, townships, special districts, and school districts.

The revenue portion of this variable counts all revenue raised from "own" sources by governments in a county. This includes all general tax revenues such as income, property, and sales taxes as well as revenues from liquor stores, insurance and utility revenues, and charges and miscellaneous revenues. It attempts to measure dollars paid to local governments in the county for all purposes. It does not include intergovernmental revenue, since we wish to measure the amount that local taxpayers have to pay for locally provided government goods and services.

What do taxpayers get in return for their money? Ideally, we would like to include a measure of the quantity and quality of public services. Unfortunately, there are no generally accepted measures of this complicated bundle of goods and services. For example, Walden (1998) used Scholastic Aptitude Test scores as a proxy for quality of government services, with mixed results. To cast our net more broadly, we use government full-time equivalent (FTE) employment as a proxy for the quantity and quality of government services. The underlying assumption is that areas that have more workers provide more and better services. Of course, this is not a perfect measure since it implicitly assumes that government workers are equally productive across all areas.

The resulting variable used in the analysis is the ratio of local revenue to local employment, or government cost per (FTE) employee, with a lower level expected to

correlate with lower COL. Unfortunately, the most recent data available for this purpose are from the 1992 Census of Government, which covers data for fiscal 1991-92. Data came from the Compendium of Government Finances and the Compendium of Public Employment (U.S. Census Bureau, 1997a, 1997b).

UNEMPLOYMENT RATE

Like the population and income variables, an area's unemployment rate can provide a measure of demand in the local economy. Higher unemployment in a locality can be expected to mean lower demand from area residents for most goods and services and concomitant lower prices than in an area with an economy that is humming along near full employment. On the supply side, higher unemployment would mean less upward pressure on wages, keeping costs of production lower. To the extent that residents buy locally produced goods and services, this will contribute to a lower COL. Bartik (1991) explains the effects on COL in the short run and the long run. Data came from the Bureau of Labor Statistics.

REGIONAL DUMMY VARIABLES

Aside from all the factors discussed above, there are certainly others that affect the COL in various places. It would be surprising indeed if a relative handful of variables would explain everything there is to know about COL differences from place to place. Therefore, regional dummy variables are included to account for some of these effects in an attempt to improve the accuracy of the estimates. The Census Bureau's regional definitions were used.

COL MODEL

The resulting model takes the form

$$\mathrm{COL}_i = f(\overset{+}{\mathrm{POP}}_i,\ \overset{+}{\mathrm{DENSITY}}_i,\ \overset{+}{\mathrm{INCOME\ PER\ CAP}}_i,\ \overset{+}{\mathrm{GROWTH\ RATE}}_i,\ \overset{+}{\mathrm{UTILITIES}}_i,\ \overset{+}{\mathrm{GCOST}}_i,\ \overset{-}{\mathrm{UNEMPT}}_i,\ \mathrm{REGION}_i)$$

where

COL_i:	overall COL in area *i*,
POP_i:	population of area *i*,
$\mathrm{DENSITY}_i$:	people per square mile in area *i*,
$\mathrm{INCOME\ PER\ CAP}_i$:	income per capita of residents in area *i*,
$\mathrm{GROWTH\ RATE}_i$:	rate of growth of population or aggregate income in area *i*,
$\mathrm{UTILITIES}_i$:	utility rates (prices) in area *i*,
GCOST_i:	government cost per unit of service in area *i*,

$UNEMPT_i$: unemployment rate of area *i*, and
$REGION_i$: dummy variable for the census region in which area *i* is located.

The sign above each variable indicates the type of effect it is expected to have on COL. For example, the positive sign above DENSITY means that a greater density is expected to cause the COL to be higher.

The final estimates for the overall COL were created by testing different models that used various combinations of the hypothesized variables listed above. Many different models were tested, using data for the 303 areas around the country for which *ACCRA Cost of Living Index* data and complete data for the independent variables were available. The calibration process started with "all" variables in the model, then winnowed to get to the best model. Linear and quadratic versions of the independent variables were tested when appropriate. In each case, the decision to keep or eliminate a variable was based on the sign of the coefficient compared to its expected sign, statistical significance or lack thereof, and intercorrelation of the variables. This process was deemed appropriate since the goal of this project is not to test hypotheses about the causal determinants of COL differences through space but rather to create a model that will identify patterns of COL variability and let us estimate COL for counties that do not participate in the ACCRA database.

The model that best measures total COL is

$$\begin{aligned}TOTCOL = {} & 96.548 + 0.835\ POPGTH^* + 0.00299\ DEN^* - 1.81(e\text{-}8)\ DENSQ^* \\ & + 1.221\ ELEC^* + 0.00312\ GCOST^* + 0.330\ NE - 8.504\ MA^* \\ & - 9.500\ SA^* - 11.291\ ESC^* - 15.069\ WSC^* - 9.049\ ENC^* \\ & - 9.800\ WNC^* - 6.469\ MTN^*\end{aligned}$$

(adjusted $R^2 = .787$, $F = 87.0$, *statistically significant at the 5 percent level).

This equation says that the total COL in an area (TOTCOL) is positively affected by population growth in the previous year (POPGTH) and the area's density of population (DEN and DENSQ). The density effect is nonlinear; higher density means a higher COL, but the effect increases at a decreasing rate with higher densities.[4] Electric rates (ELEC) and higher government costs (GCOST) also cause higher COL, as had been hypothesized. The last eight variables are the dummy variables for the regions of the country.[5] The Pacific region is used as the standard of comparison (and so is omitted from the model), and all other areas except New England had lower costs than the Pacific region on average, as reflected in the negative signs on all their variables.

The asterisks indicate that nearly all variables in this equation are statistically significant at the 5 percent level; in fact, they are significant at the 2 percent level. The only exception is the dummy variable for the New England area; this implies that the average costs in New England are not significantly different from costs in the Pacific region. The adjusted coefficient of determination is .787,

and the *F* statistic is 87.0, which is significant beyond the 0.00 percent level. Goldfeld-Quandt tests indicate that heteroskedasticity is not a problem, either.

But a question arises as to whether the model, which was fitted to a sample of 303 urban areas, is appropriate for estimating COL for the sixty-seven counties of Pennsylvania. First there is the issue of "urban areas" versus "counties." The ACCRA data are for urban areas, which need not include the whole county. Indeed, for the smallest counties, they certainly do not. Throughout this article, we assume that cost patterns that apply to an urban area of a county also apply more generally to the county as a whole. This seems a reasonable assumption, especially in cases where the explanatory variables are related to "rural-ness," such as population or density. However, it should be explicitly recognized that it is an assumption, since we have no direct data on COL in the rural areas. Given the data, it would be an interesting study in itself to look at urban versus rural COL within a county.

And how representative are the 303 urban areas of Pennsylvania's counties? The average 1997 population of the 303 areas in the database is 319,281, which is far greater than the population of most of Pennsylvania's counties. The urban areas in the sample ranged from 12,263 to 9,116,506 in population, while Pennsylvania's counties ranged from 4,965 to 1,450,683. Only three of the sixty-seven counties had populations smaller than the smallest area in the fitting sample. Density ranged from 3.8 to 71,690 people per square mile in the set of 303 urban areas; Pennsylvania's counties ranged from 11.5 to 10,171, well within the range of the wider set.

A better way to answer this question is to ask whether the final estimating equation fits equally well for large and small areas within the sample. A good way to test this is by using the Chow test,[6] fitting the estimating equation to a subset of the areas and then comparing the residuals of this subset with the residuals from the whole sample to see if they are significantly different. If they are not different, then the equation fits as well for the subset as for the whole sample.

To apply this test, we sorted the sample from largest to smallest area. We considered the smallest twenty-five areas in the sample, with 1997 population from 12,263 to 42,135, and the smallest fifty areas, which included areas with population up to 59,033. The Chow test resulted in small *F* statistics for both the smallest twenty-five areas and the smallest fifty areas, meaning that the estimation errors for the smallest areas were relatively small; the estimating equation from the whole sample of areas fit the small areas well. The conclusion that we can draw is that the estimating equation works for the smaller areas as well as for the whole sample. Since we are concerned with deriving estimates for the rural areas of the state, which have smaller populations, this is reassuring.[7]

Similarly, Chow tests were performed using density as the sorting variable. The least dense twenty-five areas in the sample had 1997 densities ranging from 31.9 down to 3.8 people per square mile, and the fifty least dense areas had density up to 60.0 people per square mile. Pennsylvania's counties ranged in density from 10,170.9 down to 11.5 people per square mile, well within the range of the sample areas. Again, the Chow test resulted in small *F* statistics for both the least dense

twenty-five areas and the least dense fifty areas. This implies that the estimating equation works well for the least dense areas, which should include Pennsylvania's rural areas.

Data for the independent variables were then applied to the estimating equation to calculate the COL for each of Pennsylvania's sixty-seven counties, which are presented in Table 2 and Figure 1.

For the total budget, the COL averaged 101.5 for Pennsylvania's sixty-seven counties, which is about 1.5 percent above the average for the 321 urban areas that participated in the *ACCRA Cost of Living Index* for the third quarter of 1997. The COL index numbers ranged from a low of 99.7 in McKean, Clarion, Elk, Warren, and Cameron counties (in northwest and north-central PA) to a high of 127.6 in Philadelphia in the southeast. This means that it cost about 28 percent more to live in Philadelphia, overall, than in the state's least expensive counties in 1997.

Figure 1 shows that COL tended to be highest in the eastern part of the state, especially in the southeastern portion. Philadelphia and its three neighboring counties make up four of the five highest-cost counties of the state, with Pittsburgh's Allegheny county rounding out the top five. The western frontier counties also tended to be higher cost, although not nearly as high cost as the southeast. The northern tier of the state, with the exceptions of the far east and west, were all low-cost places to live.

Most important, the data confirm that the cost of living tends to be lower in Pennsylvania's "rural" counties than in its "urban" counties.[8] The average for Pennsylvania's twenty-five "urban" counties was 103.0, 2.4 percent more expensive than the average of its forty-two "rural" counties.[9]

However, the averages cited above are simple means of the index values for the counties; they simply sum the index values for all sixty-seven counties and divide by sixty-seven. These simple averages do not reflect the fact that the high-cost urban areas are by their very nature much more heavily populated than low-cost rural areas. To determine the COL experienced by the average urban resident as compared to the average rural resident, it is necessary to calculate a weighted average COL, with the weights being the population in each county. Table 3 presents the results.

These numbers indicate that the average resident of Pennsylvania experiences a COL index of approximately 105.5, or about 5.5 percent higher than the average of the 321 urban areas that participated in the *ACCRA Cost of Living Index* for the third quarter of 1997. It is not surprising that the weighted average is higher than the simple average, since it reflects the fact that the very high COL of Philadelphia and its environs applies to nearly a million and a half people, while the lower costs of the state's rural areas apply to much smaller numbers of people. Since the higher-cost counties generally tend to be more populous than the lower-cost counties, it is not surprising that the weighted averages are higher than the simple averages, even for the rural areas.

TABLE 2. Total Cost of Living (COL) in Pennsylvania's Counties Ranked from Lowest to Highest

County	*Rural?*	*COL Index*	*County*	*Rural?*	*COL Index*
1. McKean	R	99.7	35. Lawrence	R	100.6
2. Clarion	R	99.7	36. Washington		100.7
3. Elk	R	99.7	37. Mifflin	R	100.7
4. Warren	R	99.7	38. Lackawanna		100.7
5. Cameron	R	99.7	39. Juniata	R	100.9
6. Wyoming	R	99.8	40. Westmoreland		100.9
7. Venango	R	99.9	41. Forest	R	101.0
8. Indiana	R	100.0	42. Carbon		101.0
9. Schuylkill	R	100.0	43. Erie		101.0
10. Northumberland	R	100.0	44. Beaver		101.0
11. Clinton	R	100.1	45. Centre		101.1
12. Lycoming		100.1	46. Wayne	R	101.2
13. Greene	R	100.1	47. Franklin	R	101.2
14. Sullivan	R	100.1	48. Fulton	R	101.2
15. Armstrong	R	100.2	49. Perry	R	101.3
16. Huntingdon	R	100.2	50. Lebanon	R	101.4
17. Cambria		100.2	51. Dauphin		101.5
18. Potter	R	100.2	52. Butler	R	101.7
19. Jefferson	R	100.2	53. Cumberland		101.7
20. Columbia	R	100.3	54. Adams	R	101.9
21. Tioga	R	100.4	55. Berks		102.0
22. Luzerne		100.4	56. York		102.1
23. Bedford	R	100.4	57. Lancaster		102.3
24. Clearfield	R	100.4	58. Northampton		102.5
25. Somerset	R	100.5	59. Lehigh		103.1
26. Bradford	R	100.5	60. Chester		103.1
27. Union	R	100.5	61. Monroe	R	103.2
28. Susquehanna	R	100.5	62. Pike	R	103.2
29. Fayette	R	100.5	63. Bucks		103.5
30. Crawford	R	100.6	64. Allegheny		104.6
31. Blair		100.6	65. Montgomery		105.0
32. Montour	R	100.6	66. Delaware		108.4
33. Snyder	R	100.6	67. Philadelphia		127.6
34. Mercer		100.6			

	Descriptive Statistics		
	Total	*Rural*	*Urban*
Mean	101.5	100.6	103.0
Standard deviation	3.6	0.8	5.4
Minimum	99.7	99.7	100.1
Maximum	127.6	103.2	127.6

Note: 100.0 = average of 321 areas participating in *ACCRA Cost of Living Index*, third quarter, 1997.

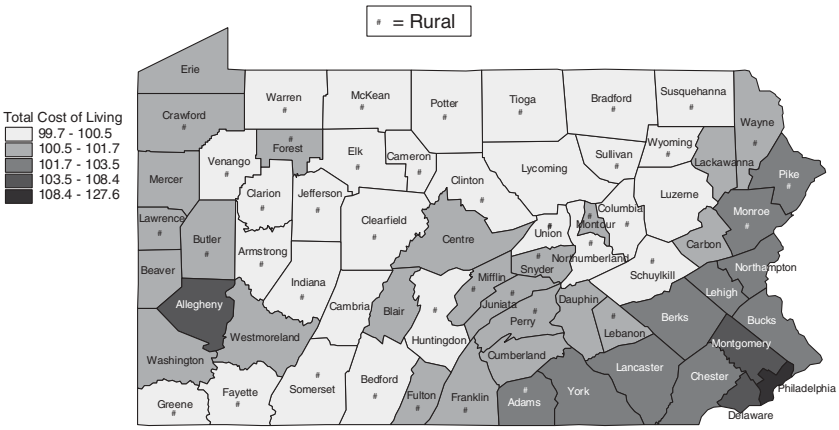

FIGURE 1. Estimated Cost of Living in Pennsylvania's Counties

TABLE 3. Comparison of Simple and Population-Weighted Average Cost of Living Index Values for Rural and Urban Counties

	Simple Average	*Population-Weighted Average*
All counties	101.5	105.5
42 rural counties	100.6	100.7
25 urban counties	103.0	106.7
Difference between urban and rural	2.4%	6.0%

A key conclusion of this study is that the average urban resident in Pennsylvania experiences an overall cost of living that is about 6 percent higher than that of the average rural resident.

COL SUBINDEXES

The *ACCRA Cost of Living Index* also provides data for the six subindexes that compose the overall index: groceries, housing, utilities, transportation, health care, and miscellaneous goods and services. The estimation process discussed above was applied to each of these subindexes as well. Table 4 shows the estimating equations that resulted, and COL estimates for each county for each index are presented in Table 5.

The subindexes tended to follow the general spatial pattern of the overall COL. Table 6 shows that all of the subindexes were positively correlated with the overall cost of living, although to differing degrees. Within the subindexes, housing and

(text continues on p. 108)

TABLE 4. Regression Results: Cost of Living Estimating Equations

Variable	*Expected Sign*	*Total*	*Groceries*	*Housing*	*Utilities*	*Transportation*	*Health Care*	*Miscellaneous*
Intercept		+96.548*	+97.145*	+123.170*	–27.787	+87.093*	+104.292*	+97.929*
Population 1997	(+)						+4.59E-06*	
Population squared							–7.06E-13*	
Population growth 1996-97	+	+0.835*						+0.508*
Density 1997	(+)	+0.00299*	+0.000423*	+0.00492*		+0.00214*		
Density squared		–1.81E-08*				–2.72E-08*		
Income per capita 1997 squared							+2.03E-08*	+6.38E-09*
Income per capita growth	+							–0.168*
Aggregate income 1997	+		+3.48E-08*		+8.37E-08*			
Government cost per worker 1992	(+)	+0.00312*		+0.00689*		+0.208*		
Government cost squared						–5.33E-05*		
Electric rate 1997	+	+1.221*	+0.928*		+4.750*	+1.241*		+0.707*
Gas rate 1997	(+)				+19.336*		+1.350*	
Gas rate squared					–1.107*			
Unemployment rate 1997	–			–0.976*				
New England		+0.330	+0.289	+1.943	+36.369*	–5.546	–12.426*	+0.012
Middle Atlantic		–8.504*	–5.045*	–18.207*	17.252*	–7.859*	–29.569*	–4.699*
South Atlantic		–9.498*	–4.784*	–24.971*	16.491*	–9.072*	–32.815*	–2.839*
East South Central		–11.291*	–5.721*	–30.786*	9.363*	–7.397*	–34.749*	–3.254*
West South Central		–15.069*	–12.643*	–34.968*	8.650*	–5.326*	–31.300*	–4.872*
East North Central		–9.049*	–4.336*	–22.006*	9.788*	–5.533*	–29.130*	–4.928*
West North Central		–9.801*	–6.828*	–25.815*	6.761*	–4.951*	–27.972*	–4.122*
Mountain		–6.469*	–1.501	–13.952*	5.554	0.840	–16.901*	–4.482*
Adjusted R^2		.787	.570	.765	.606	.434	.688	.341
F statistic		87.0	37.4	90.6	39.7	18.8	56.4	14.0
Probability (*F* statistic)		.000	.000	.000	.000	.000	.000	.000

Note: Expected signs: (+) indicates that the variable is expected to exert a positive effect on cost of living, although the combination of signs may vary when both the linear and quadratic variables are included in the model.

*Significant at the 5% level.

TABLE 5. Cost of Living Estimates, Total and All Subindexes, for Pennsylvania Counties, 1997

County	*Rural?*	*Total*	*Groceries*	*Housing*	*Utilities*	*Transportation*	*Health Care*	*Miscellaneous*
Adams	R	101.9	101.0	101.4	120.9	100.0	95.8	100.8
Allegheny		104.6	102.9	109.5	123.9	105.2	108.2	101.3
Armstrong	R	100.2	100.9	98.3	120.9	98.6	94.5	98.6
Beaver		101.0	101.1	102.3	121.1	100.7	96.4	98.5
Bedford	R	100.4	100.9	98.2	120.8	96.8	92.2	98.4
Berks		102.0	101.3	103.1	121.5	101.3	101.3	101.1
Blair		100.6	101.0	100.9	121.0	98.0	95.1	98.7
Bradford	R	100.5	100.9	100.0	120.8	99.0	93.3	99.4
Bucks		103.5	101.8	105.9	122.2	105.0	107.6	102.3
Butler	R	101.7	101.0	101.5	121.1	100.2	96.9	100.9
Cambria		100.2	101.0	98.5	121.0	98.1	94.6	98.8
Cameron	R	99.7	100.8	98.7	120.8	97.9	96.1	98.2
Carbon		101.0	100.9	98.8	120.8	98.6	94.9	99.2
Centre		101.1	101.0	102.9	121.0	100.4	95.5	99.4
Chester		103.1	101.6	104.9	122.1	104.4	118.2	105.7
Clarion	R	99.7	100.9	99.1	120.8	96.8	93.8	97.8
Clearfield	R	100.4	100.9	97.5	120.9	96.9	93.7	99.1
Clinton	R	100.1	100.9	98.0	120.8	101.7	93.0	98.6
Columbia	R	100.3	100.9	98.7	120.9	98.0	94.1	99.0
Crawford	R	100.6	100.9	100.1	120.9	99.3	93.9	99.0
Cumberland		101.7	101.2	104.3	121.2	101.2	102.0	101.6
Dauphin		101.5	101.2	104.0	121.3	101.9	102.1	100.7
Delaware		108.4	102.6	115.3	122.1	108.5	107.1	102.0
Elk	R	99.7	100.9	99.5	120.8	99.1	97.2	99.0
Erie		101.0	101.2	101.2	121.3	100.2	97.1	99.6
Fayette	R	100.5	101.0	97.8	121.0	96.8	93.7	98.7
Forest	R	101.0	100.8	95.3	120.8	97.8	91.3	98.8
Franklin	R	101.2	101.0	101.3	121.0	98.4	96.6	99.7

(continued)

TABLE 5 (continued)

County	*Rural?*	*Total*	*Groceries*	*Housing*	*Utilities*	*Transportation*	*Health Care*	*Miscellaneous*
Fulton	R	101.2	100.8	97.7	120.8	96.4	92.7	97.7
Greene	R	100.1	100.9	96.8	120.8	98.1	92.0	98.7
Huntingdon	R	100.2	100.9	95.2	120.8	97.3	91.5	98.6
Indiana	R	100.0	100.9	97.8	120.9	97.4	93.8	99.2
Jefferson	R	100.2	100.9	98.6	120.8	98.3	94.2	99.2
Juniata	R	100.9	100.9	97.6	120.8	96.6	92.7	100.3
Lackawanna		100.7	101.2	100.6	121.1	101.6	97.4	99.3
Lancaster		102.3	101.4	104.6	121.7	102.3	100.3	101.2
Lawrence	R	100.6	101.0	100.3	120.9	98.6	94.6	98.7
Lebanon	R	101.4	101.0	103.5	121.0	99.1	97.3	99.8
Lehigh		103.1	101.5	105.0	121.4	103.9	102.8	101.4
Luzerne		100.4	101.2	99.9	121.3	99.5	97.7	99.0
Lycoming		100.1	100.9	99.3	120.9	99.4	94.7	99.4
McKean	R	99.7	100.9	98.9	120.8	97.4	94.8	98.4
Mercer		100.6	101.0	100.9	120.9	98.3	94.6	99.2
Mifflin	R	100.7	100.9	98.8	120.8	99.7	92.4	98.9
Monroe	R	103.2	101.0	98.5	121.0	101.5	95.5	101.4
Montgomery		105.0	102.4	109.0	123.1	107.0	121.8	106.8
Montour	R	100.6	100.9	102.1	120.8	111.8	101.8	101.3
Northampton		102.5	101.3	103.7	121.3	101.9	99.5	100.6
Northumberland	R	100.0	101.0	99.3	120.9	97.7	94.4	98.9
Perry	R	101.3	100.9	101.8	120.8	98.0	94.0	99.8
Philadelphia		127.6	106.3	148.6	123.5	119.2	101.7	99.7
Pike	R	103.2	100.9	100.4	120.8	101.0	94.2	102.0
Potter	R	100.2	100.8	98.0	120.8	97.5	94.0	98.3
Schuylkill	R	100.0	101.0	98.7	121.0	98.4	95.3	98.9
Snyder	R	100.6	100.9	101.0	120.8	98.5	97.4	100.0
Somerset	R	100.5	100.9	98.3	120.9	97.1	93.6	100.2

Sullivan	R	100.1	100.8	98.4	120.8	97.1	92.4	100.0
Susquehanna	R	100.5	100.9	98.5	120.8	96.8	93.2	100.4
Tioga	R	100.4	100.9	98.8	120.8	96.7	93.2	98.6
Union	R	100.5	100.9	101.5	120.8	99.3	94.1	101.4
Venango	R	99.9	100.9	99.0	120.9	98.1	97.8	99.1
Warren	R	99.7	100.9	100.2	120.8	98.5	96.1	99.9
Washington		100.7	101.1	101.1	121.2	99.4	99.3	99.1
Wayne	R	101.2	100.9	97.4	120.8	100.1	93.5	100.4
Westmoreland		100.9	101.3	101.5	121.5	101.2	98.8	99.8
Wyoming	R	99.8	100.9	97.4	120.8	99.0	93.9	99.5
York		102.1	101.3	103.3	121.5	101.6	99.4	101.5

TABLE 6. Correlation of Estimated Subindexes

	Total	*Groceries*	*Housing*	*Utilities*	*Transportation*	*Health Care*	*Miscellaneous*
Total	1.000						
Groceries	.958	1.000					
Housing	.963	.965	1.000				
Utilities	.703	.851	.757	1.000			
Transportation	.802	.825	.843	.746	1.000		
Health Care	.384	.522	.488	.786	.655	1.000	
Miscellaneous	.285	.327	.321	.575	.535	.830	1.000

TABLE 7. Rural versus Urban Costs, Unadjusted for Population

	Overall	*Rural*	*Urban*	*Urban % Higher*
Total	101.5	100.6	103.0	2.4
Groceries	101.2	100.9	101.6	0.7
Housing	101.3	99.0	105.2	6.3
Utilities	121.1	120.8	121.6	0.7
Transportation	100.0	98.7	102.3	3.6
Health Care	96.9	94.4	101.1	7.1
Miscellaneous	99.9	99.4	100.6	1.2

groceries tended to most closely follow the overall pattern, with miscellaneous goods and health care being less closely related.

The groceries subindex exhibited relatively little variation across counties in Pennsylvania. This is also true for the nation as a whole; the groceries and miscellaneous goods subindexes had the least amount of variation across the 321 areas in the *ACCRA Cost of Living Index* for the third quarter of 1997. Given the easy portability of these kinds of goods, it is not surprising that they would be quickly transported from lower- to higher-cost places, tending to reduce spatial price differences.

Housing was the subindex that varied the most across the state, from a low of 95.2 in Huntingdon county to 148.6 in Philadelphia. Again, urban counties had a higher cost than rural counties. Housing costs in Pennsylvania's urban counties was 6.2 percent higher than in its rural counties, on average. In fact, the housing index for rural counties (99.0) was just below the average (100.0) for the 321 areas in the *ACCRA Cost of Living Index*, which is the standard of comparison for this study. After weighting the average for population, housing costs were 12.7 percent higher for the average urban Pennsylvanian than for the average rural Pennsylvanian.

A key question to be asked is whether the same rural-urban patterns exist for each of the subindexes as for the overall cost of living. As Tables 7 and 8 show, the

TABLE 8. Rural versus Urban Costs, Adjusted for Population

	Overall	*Rural*	*Urban*	*Urban % Higher*
Total	105.5	100.7	106.7	6.0
Groceries	102.1	100.9	102.5	1.5
Housing	109.3	99.4	112.0	12.7
Utilities	122.0	120.9	122.3	1.1
Transportation	104.1	98.6	105.6	7.1
Health Care	102.2	94.7	104.1	9.9
Miscellaneous	100.9	99.5	101.1	1.7

answer is yes. Urban areas tend to be more expensive for all of the subindexes, as well as for the overall cost of living, although the differential varies by type of good.

Conclusions and Policy Implications

The key question asked in this research is whether rural counties are less expensive places to live than urban counties. The short answer, at least for Pennsylvania, is yes. COL estimates for the rural counties in Pennsylvania tend to be lower overall and for each of the six subindexes than for urban counties.[10]

However, not all rural places are less expensive than all urban places in the state. For example, urban Lycoming (Williamsport), Cambria (Johnstown), and Luzerne (Wilkes-Barre) counties tended to have lower costs than many rural counties, and eastern-tier Pike and Monroe counties tended to have higher costs despite their rural classification. In other words, "urban-rural" is not the only factor affecting COL differences through space.

How much more expensive are urban counties than rural counties? Table 7 summarizes the results. Using simple averages of the counties' COL data, urban areas are about 2.4 percent more expensive overall than rural areas. But as explained earlier, more people live in urban than rural areas, so if we wish to compare the cost of the average urban resident with that of the average rural resident, it is necessary to calculate population-weighted averages. Table 8 does this. Virtually all of the numbers in Table 8 are higher than in Table 7. Using the population-weighted averages, the COL differences between urban and rural residents are larger than for the unadjusted averages. For the overall cost of living, urban residents of the state pay about 6.0 percent more on average than do rural residents. Similar patterns apply for all of the subindexes, although to differing degrees.

Why are these results important for rural areas? A number of policy issues are related. First, a low COL may represent a significant competitive advantage of rural areas, which can be used as part of rural areas' economic development campaigns. Employers may be attracted to low COL areas since they may be able to pay lower wages in such areas, and thus be more competitive in national and world markets,

without having their employees suffer a lower real standard of living. Of course, not all firms would respond to this location factor, but those firms that do not need close access to urban agglomerations may be interested in low-cost rural locations.

Another target group to consider is households. Most simply, a low COL means that their dollars buy more in rural areas than they would elsewhere. But lower-cost places also tend to be lower-wage places (DuMond, Hirsch, and Macpherson 1999). This might make it difficult to sell low COL as a reason for moving to a place, when the new resident's earnings will also be lower. However, a low COL would be an attractive attribute for people who bring their income with them, and that leads to some possible economic development approaches.

Who falls into the category of "bringing their income with them"? There are two types that come to mind: those who are not in the labor force and those who are in the labor force but can earn the same income regardless of their location. The first group includes retirees. They have earned the bulk of their lifetime income already and are living primarily off the earnings of past savings or retirement checks, neither of which vary with location of the retiree. A lower COL will make the nest egg or retirement check go further, so a campaign to attract retirees might be fruitful.[11] Areas that can combine a low COL with other attributes that may be attractive to retirees, such as low crime rates, access to health care, and appropriate recreational activities, may be especially successful. If an area succeeds in creating a community of retirees, the demand they would generate would make it possible for activities that cater to senior citizens to spring up or move in, adding to the attractiveness of the area for more retirees. In other words, a local area may spawn an "industry cluster" that centers on the demands of retirees.

And a program focusing on low cost of living for retirees may also help prevent current residents from moving away to places that they may not realize have a significantly higher COL. For example, Flagstaff, Arizona, had a COL of 113.9 in the third quarter of 1997, with housing costs 36.9 percent above the average for the 321 places in that issue of the *ACCRA Cost of Living Index*. Sarasota and West Palm Beach, Florida, each had a COL about 7 percent above average. A COL-focused campaign might sell itself to current residents who are investment-conscious by saying that staying in the low-cost local area results in an immediate increase in the value of their wealth by *x%* compared to selected high-cost areas.

A second group of people who bring their income with them would be those whose income is not tied to a local firm, perhaps entrepreneurs who work from home or telecommuters who work for a company located elsewhere. Those who live by the modem, the fax, and the satellite are relatively footloose and may benefit from a low-cost location. The development campaign could also stress amenities that appeal to this group, and an examination of their demographics would be necessary. For example, if they tend to be interested in arts and cultural activities or professional sports, low-cost counties that are relatively close to big cities would have the most success with this type of campaign. This is also true of those who need access to a major airport for visits to clients. If a substantial portion of tele-

commuters tend to be interested in wilderness hiking and outdoor sports, more remote rural counties could target this group for economic development purposes. And the smaller the county in terms of population, the fewer high-income residents it would take to make a difference in local spending and income levels.

Along with attracting new business, retention of existing companies is an important focus of economic development. It is not unusual for firms to consider moving out of the local area at key junctures in their life cycle. Young and inexperienced firms that are growing rapidly may be especially prone to overlook the issue of COL in making their location decisions. Since COL data are not widely available, it is easy to forget the whole issue. Local officials from low-cost areas would be doing their local firms a favor by informing them of the low cost in the local area and helping them make more informed choices.

Aside from economic development campaigns, COL has policy implications related to the spatial dimensions of growth (and urban sprawl), rural emigration, and real versus nominal income levels for private contracts and government programs.

As suggested above, low-cost rural counties that neighbor urban counties may be very attractive places for people who want or need to work downtown but still want to live in a low-cost, high-natural-amenity environment. This may mean that commuters will spread out beyond current urban boundaries into rural areas. This can be viewed in two ways. First, it makes sense for people to move to the low-cost rural fringe. They will experience a higher standard of living both from the lower costs and from the natural environment. And from the rural community's point of view, they will bring property and income tax revenues and disposable income for local retailers and service firms. They may be a route to development for the community.

On the other side, this would promote sprawl. Those who use the term "sprawl" often see uncontrolled development as a bad thing, since the people who leap the urban boundary will be driving more, using more fuel, and adding to pollution problems. They will also be adding development to a rural area and perhaps converting farmlands or other green space to residential development. And low-density residential development may be less efficient with regard to provision of public services such as roads, water, and sewer. Those in rural areas who are concerned about the negative aspects of sprawl, especially in low-cost counties near urban concentrations, should recognize the possibility that longer-distance urban commuters may be coming their way, and may wish to consider taking steps to enact land controls or other sprawl-defense systems. There are clearly two very different points of view on this opportunity/problem, but regardless of the viewpoint, it is necessary to recognize that cost of living issues will have an impact.

Another issues is that of depopulation. To the extent that rural areas suffer outmigration of current residents, a campaign that advertises the low COL in a place might help resolve a "brain drain" syndrome. It may be successful in dissuading some from leaving. People who are attracted to the higher incomes of urban

areas should be made aware that those incomes are often accompanied by significantly higher COL, which negates much of the benefit of the higher incomes. For example, someone who moves from rural Luzerne County, Pennsylvania, to Manhattan might expect housing costs to be higher, but they probably do not know that Manhattan's housing costs are more than four times as much as Luzerne's! Even within a state, rural residents may not realize that housing in the next big city may be 50 percent more expensive. Again, a campaign to inform residents of the COL differentials across places will help them make better decisions, both for themselves and for their home counties.

An economist would also point out that spatial differences in COL mean that labor contracts that specify wages that are identical through space result in different real compensation levels for workers doing the same job in different places. While labor negotiations have long recognized the need to adjust for cost of living changes through time (i.e., inflation), they have been less willing to adjust for similar cost differences across space. Firms and unions may wish to take this into account as they decide on wage levels.

Similarly, spatial differences in COL have an implication for government redistribution programs that focus on nominal amounts. Nord (2000) pointed out that nominal poverty thresholds that are invariant over space result in overstatement of the poverty problem in rural areas compared with that in urban areas. To the extent that government programs use the official poverty thresholds (which are stated in nominal terms) to determine eligibility for benefits, use of COL-adjusted values would change the number of people who qualify in various areas and, thus, change the distribution of federal funds. An alternative way of considering this would be that the current, spatially invariant thresholds have the unintended consequence of incorrectly excluding some families that are actually poor but happen to live in high-cost places, while counting as poor those that are better off but live in lower-cost places. This implies that the money is not going to the most needy. Of course, this makes the issue a political hot potato.

There are nearly eighty government programs that use means tests of some kind currently.[12] Several programs (such as Head Start, the Food Stamp Program, the National School Lunch Program, and the Low-Income Home Energy Assistance Program) use the "Poverty Guidelines" of the Department of Health and Human Services to determine eligibility rather than the official poverty thresholds. Although the Guidelines use different dollar amounts than the poverty thresholds, they also do not adjust for spatial COL differences, at least for the forty-eight contiguous states and the District of Columbia. They do, however, specify higher values for Hawaii and Alaska. It might be noted that separate—higher—guideline values for Hawaii and Alaska may be taken as tacit recognition by the federal government that COL does in fact vary over space, even though poverty thresholds and the guideline values for the forty-eight contiguous states do not incorporate it.[13]

And while this study has not explicitly tested for causality, it is interesting to note that the government efficiency variable was statistically significant in the

equations for the overall COL as well as the housing and transportation subindexes. A higher cost to taxpayers per government worker was associated with a higher overall COL for area residents as well as higher housing and travel costs. The implications for conscientious government officials are obvious.

All of these policy implications point out the importance of having good COL information. Perhaps it is time for the federal government to undertake creation of an official spatial price index?

NOTES

1. A word about terminology: in this article, we discuss spatial differences in prices. The tool for measuring these differences by comparing the price of a fixed basket of goods across space might most appropriately be called a spatial price index (SPI). This is in contrast with the more familiar temporal indexes that measure price differences through time, such as the Consumer Price Index. A true "cost of living index" would allow consumers to substitute items in the basket as relative prices change, either through space or time, as long as they maintain the same standard of living (utility level). However, the literature on spatial price differences frequently refers to a spatial price index as a "cost of living index," and this article will follow that usage.

2. Complete details on the market basket and pricing procedures are presented in the *ACCRA Cost of Living Index Manual* (1997).

3. Kurre (1992, 2000) presented the literature review leading to the choice of these variables and more of the theoretical underpinnings.

4. Theoretically, a high enough density could actually mean a lower cost of living (COL) according to this formulation. However, this would not occur until a density of more than 23 million people per square mile, which is significantly higher than in any area in the sample. For comparison, the highest density in the 303 areas of the study occurs in New York county at 71,690 people per square mile, and the next highest is 10,958 in Suffolk, Massachusetts. The average density for these 303 urban areas is just 711 people per square mile, and for Pennsylvania's 67 counties it is 428.

5. The Census Bureau's standard regions are NE New England, MA Middle Atlantic (including Pennsylvania), SA South Atlantic, ESC East South Central, WSC West South Central, ENC East North Central, WNC West North Central, MTN Mountain, and PAC Pacific.

6. See Gujarati (1978, 305-6) and Quantitative Micro Software (1998, 347-50).

7. Full results of both the Goldfeld-Quandt and Chow tests are presented in Appendix C of Kurre (2000).

8. Pennsylvania's "rural" counties were identified by the Center for Rural Pennsylvania (http://www.ruralpa.org). A county is defined as rural if more than 50 percent of its population is defined as nonurban—living outside census-identified urbanized areas and places of twenty-five hundred or more. By these standards and using 1990 Census data, Pennsylvania has forty-two predominantly rural counties and twenty-five predominantly urban counties. The U.S. Census Bureau's urban/rural definitions are at http://www.census.gov/population/censusdata/urdef.txt.

9. More precisely, the COL was lower for urban areas in "rural" counties than for urban areas in "urban" counties. Of course, we do not have direct data on COL in rural portions of the "rural" counties.

10. The reader should remember that these estimates are based on data for urban areas, some of which are in rural counties, and not on direct rural data per se.

11. See Fagan and Longino (1993) for more ideas on retirees as the focus of economic development.

12. U.S. Congressional Research Service (1999).

13. Information on the Health and Human Services Guidelines is available at http://aspe.hhs.gov/poverty/97poverty.htm.

REFERENCES

Abraham, Katharine, Ernst Berndt, W. Erwin Diewert, Zvi Griliches, Charles Schultze, and Jack Triplett. 2000. Price index research in the coming decades. *Monthly Labor Review* 123 (9): 31-36.

American Chamber of Commerce Researchers Association. 1997. *ACCRA Cost of Living Index manual*. Rev. ed. Arlington, VA: American Chamber of Commerce Researchers Association.

Bartik, Timothy J. 1991. *Who benefits from state and local economic development policies?* Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Bishop, John A., John P. Formby, and Paul D. Thistle. 1994. Convergence and divergence of regional income distributions and welfare. *Review of Economics and Statistics* 76 (2): 228-35.

Borooah, Vani K., Patrick P. L. McGregor, Patricia M. McKee, and Gwyenth E. Mulholland. 1996. Cost-of-living differences between the regions of the United Kingdom. In *New inequalities: The changing distribution of income and wealth in the United Kingdom*, edited by John Hills, 103-32. Cambridge: Cambridge University Press.

Cebula, Richard J. 1989. The analysis of geographic living-cost differentials: A brief empirical note. *Land Economics* 65 (1): 64-67.

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

Cebula, Richard J., Gigi M. Alexander, and James V. Koch. 1992. A further note on determinants of geographic living-cost differentials. *Quarterly Review of Economics and Finance* 32 (4): 144-47.

Crawford, Ian. 1996. UK Household Cost-of-Living Indices, 1979-92. In *New inequalities: The changing distribution of income and wealth in the United Kingdom*, edited by John Hills, 76-102. Cambridge, UK: Cambridge University Press.

Deller, Steven C., and James C. McConnon Jr. 1990. Statistical estimates of Cost of Living Indices for selected Maine communities. Staff Paper ARE 409, Maine Agricultural Experiment Station, Department of Agricultural and Resource Economics, University of Maine, Orono.

Deller, Steven C., Martin Shields, and David Tomberlin. 1996. Price differentials and trends in state income levels: A research note. *Review of Regional Studies* 26 (1): 99-113.

Denslow, David. 1999. Personal correspondence to the author from David Denslow, Bureau of Business and Economic Research, University of Florida, 31 March 1999.

DuMond, J. Michael, Barry T. Hirsch, and David A. Macpherson. 1999. Wage differentials across labor markets and workers: Does cost of living matter? *Economic Inquiry* 37 (4): 577-98.

Eberts, R. W., and M. E. Schweitzer. 1994. Regional wage convergence and divergence: Adjusting wages for cost of living differences. *Federal Reserve Bank of Cleveland's Economic Review* 30 (2): 26-37.

Environmental Systems Research Institute, Inc. 1996. *ArcView GIS*. Redlands, CA: Environmental Systems Research Institute, Inc.

Fagan, Mark, and Charles F. Longino Jr. 1993. Migrating retirees: A source for economic development. *Economic Development Quarterly* 7 (1): 98-106.

Florida Department of Education. 1996-2001. *Florida Price Level Index*. Latest edition available from http://www.firn.edu/doe/bin00047/home0047.htm.

Ghelfi, L. M. 1988. About that lower cost of living in nonmetro areas. *Rural Development Perspectives* 5 (1): 30-34.

Gujarati, Damodar. 1978. *Basic econometrics*. New York: McGraw-Hill.

Isserman, Andrew M. 2001. Competitive advantages of rural America in the next century. *International Regional Science Review* 24: 38-58.

Johnston, Richard, Martin McKinney, and Tom Stark. 1996. Regional price level variations and real household incomes in the United Kingdom, 1979/80-1993. *Regional Studies* 30 (6): 567-78.

Knapp, John L., Patricia A. Gorski, and Robert W. Cox. 1988. *Predicting faculty salaries using an Estimated Inter-City Cost-of-Living Index*. Business and Economics Section, Center for Public Service, University of Virginia, Charlottesville.

Koo, Jahyeong, Keith R. Phillips, and Fiona D. Sigalla. 2000. Measuring regional cost of living. *Journal of Business and Economic Statistics* 18 (1): 127-36.

Kurre, James A. 1992. *The cost of living in rural Pennsylvania*. Erie: Center for Rural Pennsylvania, Pennsylvania State University–Erie.

———. 2000. *Differences in the cost of living across Pennsylvania's 67 counties.* Erie: Center for Rural Pennsylvania, Pennsylvania State University–Erie. Available from http://www.ruralpa.org.

Marshall, Ray. 2001. Rural policy in the new century. *International Regional Science Review* 24 (1): 59-83.

McHugh, Richard. 1990. Cost-of-living differences in urban versus rural areas: An analysis with expenditure data. Center for Agricultural and Rural Development Staff Report 90-SR 49, Iowa State University, Ames.

McMahon, Walter W. 1991. Geographical cost of living differences: An update. *AREUEA Journal* 19 (3): 426-50.

Nord, Mark. 2000. Does it cost less to live in rural areas? Evidence from new data on food security and hunger. *Rural Sociology* 65 (1): 104-25.

Pearce, Diana, and Jennifer Brooks (with Janice Hamilton Outtz). 1997. *The self-sufficiency standard for Pennsylvania*. Washington, DC: Wider Opportunities for Women, Inc.

Quantitative Micro Software. 1998. *EViews 3 user's guide*. Irvine, CA: Quantitative Micro Software.

Raper, Michael. 1990. Using the ACCRA Cost of Living Index to explain geographic differences in the cost of living. Paper presented at Southern Regional Science Association meetings, March, in Washington, DC.

Ravallion, Martin, and Dominique van de Walle. 1991. Urban-rural cost-of-living differentials in a developing economy. *Journal of Urban Economics* 29 (1): 113-27.

Runzheimer International. 1998. Over half of U.S. companies have pay differential policies; both salaried and hourly employees participate. News Release, 18 May. Available from http://www.runzheimer.com.

U.S. Bureau of Economic Analysis. 1999. *Regional economic information system, 1969-1997* [Compact disk]. Washington, DC: U.S. Bureau of Economic Analysis.

U.S. Census Bureau, Compendium of Government Finances. 1997a. *1992 census of government* Vol. 4, no. 5. Washington, DC: Government Printing Office. Available from http://www.census.gov:80/prod/2/gov/gc92-4/gc924-5.pdf.

———. 1997b. *1992 census of government*. Vol. 3, no. 2. Washington, DC: Government Printing Office. Available from http://www.census.gov:80/prod/2/gov/gc92-3/gc923-2.pdf.

U.S. Congressional Research Service, Library of Congress. 1999. Cash and noncash benefits for persons with limited income: Eligibility rules, recipient and expenditure data, FY1996–FY1998 (CRS Report for Congress—Order Code RL30401). 15 December. Washington, DC: U.S. Congressional Research Service, Library of Congress.

U.S. Department of Energy, Energy Information Administration. 1998a. *Electric sales and revenue 1997*. Washington, DC: U.S. Department of Energy, Energy Information Administration.

———. 1998b. *Natural gas annual, 1997*. Washington, DC: U.S. Department of Energy, Energy Information Administration.

U.S. Department of Labor, Bureau of Labor Statistics. 1982. Final report on family budgets: Cost increases slowed, Autumn 1981. *Monthly Labor Review*, July, pp. 44-46.

U.S. Office of Personnel Management. 1997. Cost-of-living allowances (nonforeign areas). *Federal Register* 62 (54, 20 March): 13354-55. (5 CFR Part 591; RIN 3206-AH51)

———. 1999. *1999 general schedule and locality pay tables*. Available from http://www3.opm.gov/oca/01tables/GSannual/index.htm.

Voicu, Alexander, and Michael L. Lahr. 1999. Expenditure Based Inter-Area Cost of Living Index. Paper presented at the 38th annual meeting of the Western Regional Science Association, February, in Ojai, CA.

Walden, Michael L. 1997. How much income variation really exists within a state? *Review of Regional Studies* 27 (3): 237-50.

———. 1998. Geographic variation in consumer prices: Implications for local price indices. *Journal of Consumer Affairs* 32 (2): 204-26.

Watts, Harold W. 1980. Special panel suggests changes in BLS Family Budget program. *Monthly Labor Review*, December, pp. 3-10.

Wojan, Timothy R., and Adam C. Maung. 1998. The debate over state-level inequality: Transparent methods, rules of evidence, and empirical power. *Review of Regional Studies* 28 (1): 63-80.

Zimmerman, Julie N., and Lorraine Garkovich. 1998. Meeting the monthly budget: Earnings and the cost of living in rural Kentucky. Paper presented at the 1998 annual meetings of the Southern Rural Sociological Association, Little Rock, AR.