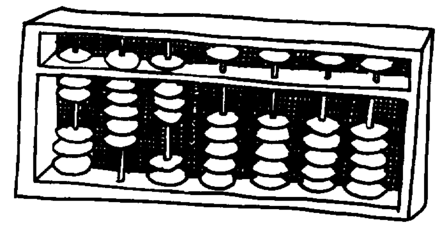
Technical Notes



Adjusting the CPI shelter index to compensate for effect of depreciation

WALTER F. LANE, WILLIAM C. RANDOLPH, AND STEPHEN A. BERENSON

Beginning with the Consumer Price Index (CPI) for January 1988, the Bureau of Labor Statistics (BLS) introduced an adjustment into the CPI shelter indexes for the phenomenon known as "age bias." The need to correct the shelter indexes (the rent index and the owners' equivalent rent index) arises from the need to keep the quality of consumer items in the CPI market basket constant over the period for which price change is observed.

The CPI measures price changes for urban consumers in the United States.[1] It does this by following the prices of a sample of consumer goods and services called the CPI market basket. Using longitudinal surveys, the CPI tracks the prices of consumer goods and services in the market basket, comparing their current prices to those previously observed. The percent change in the CPI is an average of these price changes. For this to be an accurate measure of price change, it is critically important that the quality of the goods and services be identical in both periods for which prices are observed.

The CPI measures price change for both renter-occupied and owner-occupied housing with a longitudinal survey of housing units. The residential rent index, which is the measure of the change in the cost of renter-occupied shelter, is based on the movement of rents in the survey. Since 1983, the Bureau has also used the changes in the rents of rental units in the survey to estimate the change in the owners' equivalent rent index, which measures the change in the cost of owner-occupied shelter.[2] Currently, the Bureau does this by imputing an implicit rent for each owner-occupied unit in the survey from an initial estimate obtained at the first interview with the unit's owner, and then periodically adjusting this estimate with the average rent change of similar renter-occupied housing units nearby. BLS agents ask for rent information from the rental units in the survey every 6 months. At each interview they obtain the rent and services provided for the rent. BLS computes the rent change for each sample unit and then averages the rent change. Following the same units—which provide services that rarely change—and computing the average rent change from the individual units' rent changes avoids most apparent problems with quality changes between periods of rent comparison.

Consequently, a large proportion of the sample appears to be of "constant quality." In other words, the services offered in exchange for rent are the same at successive interviews. However, the Bureau has procedures to estimate the value of and adjust for common changes in the services offered. For example, if an apartment's owner stops providing electricity as part of the rental services, the BLS estimates its value and adds it to the rent for the current month. The estimated value depends on the electrical appliances and size of the unit, and the climate and price of electricity in that area. Currently, BLS does not estimate the value of changes such as the addition of a room or a bathroom. When a rental unit in the sample experiences such a change, the Bureau leaves it out of the CPI calculation for one period.

Because quality changes of the types mentioned above occur quite rarely in the housing survey, there are relatively few problems with them. However, following the same units over time introduces the possibility of a subtle but potentially significant quality change. Even though BLS surveys the same set of housing units in successive periods, the quality of the units is not exactly the same because in the later period they are older and their quality is slightly lower. Of course, some housing units receive maintenance that slows and can even prevent them from wearing out. However, because most housing eventually wears out and is torn down or is closed down for complete rehabilitation, the housing stock as a whole—and therefore the average housing unit—clearly depreciates. Empirical estimates of the physical depreciation show that it is small but significant.

Walter F. Lane, William C. Randolph, and Stephen A. Berenson are economists in the Office of Prices and Living Conditions, Bureau of Labor Statistics.

Measuring the age bias

The question of how much housing depreciates during each period must be answered empirically. To do this, BLS researchers[3] used a hedonic regression model in which the logarithm of the rent of a housing unit is a function of (1) its structural characteristics—such as number of rooms, (2) its location, (3) services, such as utilities, included in the rent, (4) neighborhood characteristics, such as percent of the population with some college education, and (5) a set of six "depreciation" variables. The depreciation variables are the only variables in the model that depend directly on the age of the dwelling. The first two depreciation variables are (1) the age, (2) the age squared divided by 2. The others are variables interacting with age, namely: (3) age times a dummy for detached housing, (4) age times a rent control dummy, (5) age times the number of rooms, and (6) age times a dummy variable for very old units (built before 1900). (Table 1 lists all the variables, except the location variables, which are very numerous.)

BLS economists used data from the CPI housing survey and the 1980 Census of Housing to estimate the model. The data for the structural characteristics, location, services included in the rent, age, and other data needed for the depreciation variables came from the CPI survey. Neighborhood characteristics data came from the smallest published geographic levels of the 1980 census (usually the census block) containing the housing unit. The Census Bureau defined these areas to be relatively homogeneous with respect to population characteristics, economic status, and living conditions.[4] BLS accounted for location by letting all regression parameters vary over the four census regions and letting many of them vary across CPI pricing areas within regions.

Table 1. National averages of parameter estimates in age bias adjustment model for selected variables (multipled by 100)

Item	Average	Item	Average
Depreciation		**Continued—Services included in rent**	
Age	-0.48	Fuel oil	2.3
Age*Rent control	-.026	Parking	3.5
Age*Old	.06	Furnishings	-2.6
Age*Detached	-.37	Swimming pool	4.3
Age*Rooms	.05	Other recreation	1.7
Structural characteristics		**Neighborhood characteristics (by percent)**	
Detached	16.0	Renters	-.05
Bedrooms	14.6	White	.11
Other rooms	3.1	Large buildings	.08
Complete kitchen	7.3	Two or more autos	-.036
Dishwasher	11.0	Without complete plumbing	-.63
Washer/dryer	6.6	Air conditioned	-.065
Oil heat	-.21	Children age 6 to 18	-.13
Electric heat	.57	College students	-.04
Central air conditioning	10.9	Families below poverty level	-.15
Window air conditioning	4.2	Elderly over 65	.12
Extra bathroom	9.9	Mobile homes	-.17
Rent control	-7.5	Unemployment	-.32
Services included in rent		With college education	.46
Gas	4.2		
Electric	7.3		

The model was estimated both with maximum likelihood and with ordinary least squares regression methods. The maximum likelihood estimates are slightly more efficient; however, the ordinary least squares estimates are much easier to obtain and they are unbiased. For this reason, the least squares method was chosen for computing the aging adjustments for the shelter indexes.

To provide lower variance depreciation estimates for individual CPI areas, the Bureau uses a composite estimation procedure[5] that combines the regional depreciation estimate with that of the local area. Similar procedures are used elsewhere in the CPI—most notably in the calculation of base-period expenditure weights—to reduce variance.

The hedonic regression model can be summarized with the following expression:

(1) Log (rent) = f(13 structural characteristics variables, various location and survey variables,
13 neighborhood characteristics variables,
7 dummy variables for services provided with rent,
6 depreciation variables, and a random error term)

The function is linear in most of these variables. After BLS estimated the coefficients using the data and the hedonic regression techniques, this function became a formula that can project the log of rent for any unit for which the values of the explanatory variables are known. In addition, a coefficient estimated for a variable in such a semilog function (the log is taken of the dependent variable only) can be interpreted as an approximation of the expected percentage change in the dependent variable (here, the rent) that will result from a unit change in the explanatory variable. For example, the national average of the coefficients for the structural characteristic variable "more than one bathroom" is about 0.099, which implies that a housing unit with an additional bathroom would have a rent about 10 percent higher than another housing unit in which all else was the same. Table 1 provides estimates in percentage terms of the effect of the regression variables on rent.

Depreciation is the effect of aging on the quality of a housing unit. The partial derivative of the full equation with respect to age provides a formula for depreciation, which can be interpreted as the approximate percent change in rent (net of inflation) as the rental unit ages 1 year. Note that the only variables in the formula that contain age are the depreciation variables. When the partial derivative is taken, the other terms drop out, leaving only the depreciation terms. Keep in mind that the inclusion of other variables influences the estimates of the depreciation variables' coefficents.

We can obtain depreciation as:

(2)

$$\text{depreciation} = \frac{\partial \log (\text{rent})}{\partial \text{age}} = g\,(\text{depreciation variables})$$

Depreciation, or the amount of quality change lost because the sample does not keep the age constant, can be found from this simpler formula.

To derive the depreciation formula explicitly, we can rewrite equation *(1)*, the hedonic regression for the logarithm of rent, as:

(3)

$$\log (\text{rent}) = a X + b_1 \text{age} + b_2 (\text{age}^2) + b_3 \text{age (number of rooms)} + b_4 \text{age (very old unit)} + b_5 \text{age (detached housing)} + b_6 \text{age (rent control)},$$

where X is a vector of the variables that do not measure depreciation and a is the vector of regression coefficients for the variables in vector X. The depreciation variables, which are all functions of the age of the housing unit, are listed individually. The derivative provides an explicit version of equation *(2)*:

(4)

$$\frac{\partial \log (\text{rent})}{\partial \text{age}} = b_1 + 2b_2 (\text{age}) + b_3 (\text{rooms}) + b_4 (\text{very old unit}) + b_5 (\text{detached}) + b_6 (\text{rent control})$$

Because none of the variables in vector X depends directly on the age of the unit, those variables do not appear in the derivative.

Vintage effects

The results from this approach to measuring the age bias are subject to possible error caused by vintage (or year built) effects. The regression coefficient estimates for the age variables (the depreciation variables) may reflect something other than the effect of aging. If the housing units built in a certain year are consistently better or worse than those built in years before or since, the regression would not be able to separate the effect of depreciation from the effect of vintage on the rent of a housing unit. For example, older homes that still survive may have been more soundly constructed, have more mature landscaping, or be in better locations than newer homes. On the other hand, newer homes may have better floor plans, insulation, appliances, equipment, wiring, and plumbing. To the degree that the market values newer houses more highly or less highly than older ones, the regression equation estimating the effect of age on the rent may be distorted. However, BLS research[6] on the subject indicates that the regression, by including structural and neighborhood variables, accounts for most vintage effects or that vintage effects that favor older housing are offset by other vintage effects favoring newer housing. The model used to adjust the shelter indexes for depreciation ultimately rests on the assumption that, when the regression contains such variables, vintage effects are negligible. In light of empirical evidence, such an assumption is preferable to assuming—as the CPI implicitly did before aging adjustments began—that housing does not depreciate and that depreciation effects can be ignored.

Results

Equation *(4)* provides a way to estimate the depreciation of housing units for the CPI housing index and also provides adjustments for the CPI rent and owners' equivalent rent indexes. To minimize the variance that this new procedure introduces, BLS uses the same adjustment for all housing units in a CPI area. To obtain the adjustment for each area, BLS estimates the area's average value for the depreciation variables with equation *(4)*. The Bureau then uses the composite estimation procedure that combines the adjustments estimated at the regional level with those from the local level to obtain the final local values. Table 2 gives the estimates of the corrections for the largest CPI areas. These are the annualized values, in percentage terms, of the corrections introduced in January 1988. They will be recomputed yearly using the same (or improved) methods with newer data. BLS makes the adjustments by adding the estimated percent change from depreciation to the percent change in rent for each shelter index in each of the 85 CPI pricing areas. For example, if an area has an annual adjustment of, say, 0.3 percent, one-twelfth that amount (0.025) would be added to the percent increase for the rent and owners' equivalent rent indexes in each month of the year. In the future, BLS may apply the corrections at smaller geographic levels, possibly as low as the sample unit level.

The adjustments in table 2 show relatively little variation in housing depreciation by region. This is a result of the composite estimation process described on the preceding

Table 2. Age bias adjustments for selected metropolitan areas
[In percent]

Pricing area	Annual adjustment	Pricing area	Annual adjustment
Northeast		**South**	
New York	0.36	Washington, DC	0.17
Philadelphia	.36	Dallas	.14
Boston	.36	Baltimore	.17
Pittsburgh	.36	Houston	.11
Buffalo	.35	Atlanta	.17
		Miami	.16
Midwest		**West**	
Chicago	.22	Los Angeles	.22
Detroit	.24	San Francisco	.23
St. Louis	.21	Seattle	.25
Cleveland	.24	San Diego	.21
Minneapolis	.21	Portland, OR	.24
Milwaukee	.22	Honolulu	.22
Cincinnati	.24	Anchorage	.19
Kansas City	.23	Denver	.24

page. Among the regions, the Northeast has the highest depreciation rates and therefore requires the largest adjustment. This may reflect the effects of the severe climate, but it also results from the prevalence of rent control and multi-unit housing in that region. The lower rates in the South result from the milder climate, more detached housing, and less rent control. □

———*FOOTNOTES*———

[1]For a complete description of the Consumer Price Index, see chapter 19 of the *BLS Handbook of Methods*, Bulletin 2285 (Bureau of Labor Statistics, April 1988).

[2]Rental equivalence is described in the *BLS Handbook of Methods*, and in more detail in "Changing the Homeownership Component of the CPI to Rental Equivalence," *CPI Detailed Report*, January 1983, pp. 7–11.

[3]The full development of the aging adjustment is described in William C. Randolph, "Housing Depreciation and Aging Bias in the Consumer Price Index," *Journal of Business and Economic Statistics*, July 1988, pp. 359–71.

[4]See *Census of Population and Housing Technical Documentation* (Bureau of the Census, 1982), p. 221.

[5]For a complete development of the composite estimation procedure, see William C. Randolph and Kimberly D. Zieschang, "Aggregation Consistent Restriction Based Improvement of Local Area Estimators," *Proceedings of the Business and Economics Section* (American Statistical Association, January 1988).

[6]For a full development of the vintage effect question, see William C. Randolph, "Estimation of Housing Depreciation: Short-Term Quality Change and Long Term Vintage Effects," *Journal of Urban Economics*, March 1988, pp. 162–78.

Establishment survey incorporates March 1987 employment benchmarks

JOHN B. FARRELL

With the release of data for May 1988, the Bureau of Labor Statistics introduced its annual revision of national estimates of employment, hours, and earnings from the monthly survey of establishments. The revision uses employment counts for March 1987 as a benchmark. As part of the annual benchmarking process, the Bureau also revised seasonally adjusted series for the past 5 years, and computed new seasonal adjustment factors.

Adjustment procedure

Monthly estimates from the Current Employment Statistics (CES) survey are based on information collected from a sample of approximately 300,000 establishments. To help improve their accuracy, the Bureau adjusts CES estimates each year to new benchmarks. Benchmarks are comprehensive counts of employment based primarily on mandatory unemployment insurance reports filed by employers with the State employment security agencies.

John B. Farrell is an economist in the Division of Monthly Industry Employment Statistics, Bureau of Labor Statistics.

The current revisions are based on March 1987 benchmarks and affect all unadjusted series from April 1986 forward. As is the usual practice with the introduction of new benchmarks, the Bureau has also revised the seasonally adjusted series from January 1983 forward and has published new seasonal adjustment factors to be used for the coming year.

Current revisions. The March 1987 benchmark level for total nonagricultural employment—100.4 million—was only 35,000, or 0.04 percent, below the corresponding sample-based estimate, resulting in one of the survey's smallest benchmark revisions. The pattern of revisions was mixed across industry divisions, with downward revisions in the goods-producing industries (−0.7 percent) being about offset by upward revisions in the service-producing industries (0.2 percent). Table 1 provides the revisions for March 1987 by industry division.

New estimates were computed for April 1987 forward, based on the new March 1987 benchmark levels and recomputed seasonal adjustment and bias factors. In addition, the sample was redistributed into estimating cells that reflect their March 1987 employment size, and sample reports were added that had been received since the original estimates were made. The combined effect of the new benchmark level, recomputed seasonal and bias factors, the resized sample, and added late reporters resulted in the new estimates generally showing larger over-the-month employment gains than previously reported. The cumulative effect on the survey estimate from April 1987 through February 1988 was the addition of

Table 1. Differences between nonagricultural employment benchmarks and estimates, by industry division, March 1987

[Numbers in thousands]

Industry	Benchmark	Estimate	Difference	
			Number	Percent
Total nonagricultural ..	100,427	100,462	−35	([1])
Total private	83,173	83,152	21	([1])
Mining	696	718	−22	−3.2
Construction	4,531	4,599	−68	−1.5
Manufacturing	18,810	18,897	−87	−.5
Transportation and public utilities	5,274	5,275	−1	([1])
Wholesale trade	5,763	5,725	38	.7
Retail trade	17,902	17,737	165	.9
Finance, insurance, and real estate	6,443	6,478	−35	−.5
Services	23,754	23,723	31	.1
Government	17,254	17,310	−56	−.3
Federal	2,916	2,916	0	0
State	4,050	4,036	14	.3
Local	10,288	10,358	−70	−.7

[1]Less than 0.05 percent.