Exploratory Analysis that Redefined the Parameter of a Structure's Year Built in Consumer Price Index Housing Age November 2024

Alice Yu, Ayme Tomson, Ben Houck, & Chun Wing Tse
U.S. Bureau of Labor Statistics, 2 Massachusetts Ave., NE,
Washington, DC 202129

Abstract

Rent and owner's equivalent rent in Consumer Price Index (CPI) uses building's age bias and structural change factors. These factors are calculated each December, then run in January, and put into production use for the following 12 months. A multivariable linear regression model with 35 independent variables and one dependent (unit's rent) variable is used to calculate these factors. One of the independent variables is the product of the building's age (the current year minus the year a structure is built) and a binary variable, "old". The "old" binary variable is defined as if the unit is built in 1919 or earlier, or if the unit is built in 1920 or later. This "old" parameter is defined in a static manner that underperforms in statistical significance and restricts CPI methodology. An alternative, less restrictive, definition of the "old" parameter which performs with higher statistical significance is found in this research.

Key Words: CPI shelter, building's age bias, exploratory analysis, python

This paper provides a summary of research results. The information is being released for statistical purposes, to inform interested parties, and to encourage discussion of work in progress. The paper does not represent an existing, or a forthcoming new, official BLS statistical data product or production series.

1. Introduction

The Consumer Price Index (CPI), produced by the U.S. Bureau of Labor Statistics (BLS), is a measure of the average change over time in the prices paid by urban consumers for a market basket of consumer goods and services. Indexes are available for the U.S. and various geographic areas in the nation. The CPI for shelter or the service that housing unit provides its occupants, is about one-third of the CPI market basket. Most of the changes in the shelter cost consumers receive from their primary residences are from two CPI indexes: 1) rent of primary residence (Rent) and 2) owners' equivalent rent of primary residence (OER). Rent estimates are based on contract rents, or all services provided by the landlord to the tenant in exchange for rent money as provided in the lease. OER measures the change in the amount a homeowner would pay in rent or earn from renting his or her home in a competitive market. [1]

In the context of the CPI, adjusting for aging or accounting for potential biases related to a building's year of construction is crucial for accurately measuring housing costs. As buildings age, their quality, functionality, and desirability change. For example, as newer buildings are added to the housing market and older ones are removed or renovated, the overall quality of available rental properties

may change. The year in which a building was constructed can influence its rental or ownership value, owing to differences in design, materials, regulations, and amenities. This paper will describe the methods and efforts that were behind the recent improvements in the CPI's housing age-bias adjustment factor.

2. Background Information on CPI's Housing Age-Bias Adjustment Factor

In 1988, there was a need to correct the shelter indexes and allow the physical depreciation of a housing unit over time to be accounted for in quality adjustment. This was an effort to keep the quality of consumer items in the CPI market basket constant over the period for which price change is observed, and the introduction of the age-bias adjustment factor for CPI housing units.

This originated with William C. Randolph using hedonic regression methods to estimate average rental housing physical depreciation for areas of the United States and he estimated a downward quality bias of the CPI caused by rental-housing physical depreciation. [8]

> Hedonic quality adjustment is a technique often used to adjust for quality change in the CPI. [7] The central idea is to model an item's cost as the sum of the prices of its individual features. When an existing item is replaced by a new one, hedonic quality adjustment estimates the price value of any differences in features between the old and new items.
>
> This adjustment is necessary when replacing an older item with a new one to ensure that the CPI accurately reflects changes in prices rather than changes in quality. Essentially, BLS economists estimate what the price of the new item would have been if it had been part of the CPI sample when the old item was last available.
>
> Although item replacements do not typically occur in the CPI Housing Survey, hedonic quality adjustments are still employed to calculate constant-quality price changes in shelter indexes. These adjustments are used in specific situations where there are changes to the physical structure or features of housing units.

The hedonic regression model used to estimate age-bias adjustments in the CPI was first developed in the 1980s and serves as the basis for the process still in use today. The BLS created the model by integrating data from the 1980 Census of Housing with information from approximately 10,000 housing units collected in the CPI Housing Survey. This model included variables such as structural characteristics, neighborhood, services included in the rent, and the age of the dwelling, providing a basis to estimate the effect of housing unit depreciation. [4]

The hedonic regression model was updated in 1999 due to changes in the CPI Housing Survey over time and to accommodate a new housing sample based on the 1990 census. This update also involved using census data by zip code instead of census tract, and units built after 1990. A critical need, to increase the accuracy of year-built data, was addressed in the early 2000. Since many renters did not know when their housing units were built, data collectors and national office staff made a strenuous effort to learn the year built from other sources. As a result of these efforts, the proportion of units with year-built data increased from 45 to 89 percent.[3]

Additional modifications were made in 2007, where BLS economist added new variables, like tenant length of occupancy and zip-code-level income data from the Internal Revenue Service (IRS). Furthermore, certain units were removed from the hedonic regression model, specifically rent-controlled units and units built before 1900. Rent-controlled units were excluded because they typically do not reflect market rents, while units built before 1900 had unique depreciation effects that affected model accuracy. [2] These changes improved the model's fit and shifted the estimation of age-bias factors from area averages to individual unit levels.

In 2020, there was a higher level need to adapt to the changing market conditions due to the economic impact of the COVID-19 pandemic. The rate of growth in shelter indexes decelerated, partly due to the increase of housing units with steep decline in rent. Some tenants were unable to pay rent, causing landlords to record a $0, a condition that can distort calculations. BLS excluded these extreme rent declines, a zero or near zero, when estimating age-bias factor in 2021 and thereafter, recognizing that they did not represent typical market-based rent prices.

Additionally in 2021, structural characteristics variables were investigated that would improve the hedonic regression model. Postestimation analysis was used to recognize essential variables for the model to function and to retain variables because of their highly statistically significant power to explain variations in rent. There was also research that began to explore new data that could be introduced into the model. The current hedonic regression model has 35 variables (excluding the dummy variables for location within a specific CPI pricing area). Refer to *Attachment 1* for list of variables description. [5]

$$\log(\text{rent}) = f[\text{17 structural characteristics variables, occupancy timeframe and income variables (bath, bathrooms}^2\text{, bed, bedrooms}^2\text{, other\_rms, electric heat, gas heat, central air, window air, other air, detached, mobile home, senior com, elevator, new tenant, length of occ, adj avg inc),}$$

$$\text{Dummy variables (for all CPI pricing areas included within the census region or division),}$$

$$\text{9 neighborhood characteristics variables (density, age 65pl, large buildings, mobile homes, renter occ, school age, high poverty, age 3+college, age25+college}$$

$$\text{3 variables for services provided with rent (heat incl, electric incl, extra changes),}$$

$$\text{6 depreciation variables (age, age}^2\text{ age}^3\text{, age\_allrm, age\_detach, ageold)]} \quad (1)$$

$$\text{+a random error term.}$$

2.1 Focus on Year-Built Value

Year-built value is periodically evaluated to deal with increases in records of missing values since it is crucial for calculating age-bias adjustment. In the early 2020s, BLS initiated a special round of data collection to obtain year-built values for housing units where this information was previously unknown. This special data collection increased accuracy and reduced the level of imputation required to fill in missing data. The collected information for 7,100 housing units allowed BLS to better estimate year-built values, increasing the percentage of units with known year-built data from 57% to 75% over a 4-month period at the end of 2020.

In 2024, the new threshold for “old” age of a housing unit is a 40-year dynamic window. The age of a housing unit structure is the current year minus the year-built value. This paper will focus on how the 40-year threshold was decided and the impact it has on the age-bias adjustment factor.

3. CPI’s Housing Age-Bias Adjustment Factor

The age-bias adjustment factor is the partial derivative of equation (1) with respect to age. The 1/12 in equation (2) is to convert the units of age, measured in years, to the units of depreciation, measured in months.

$$Age - Bias\ Adjustment\ Factor\ (or\ F) = depreciation = \frac{1}{12} x\ \frac{\partial \log(rent)}{\partial age} \quad (2)$$

Two age-bias factors are calculated for each housing unit: one using the census region regression parameter estimates and the other using the census division regression parameter estimates. The two factors are then averaged, a process introduced in [2]. Both factors are calculated similarly, as shown in (3).

$$F_{k,i} = \frac{-(\gamma_{1,k}+2\gamma_{2,k}age_i+3\gamma_{3,k}age^2{}_i+\gamma_{4,k}all\_rm_i+\gamma_{5,k}detached_i+\gamma_{6,k}old_i)}{12} \quad (3)$$

Where, F_k= area k (either a census region or a census division) age-bias adjustment factor

$\gamma_{1,k}$ thru $\gamma_{6,k}$ are the six depreciation variables (age, age^2 age^3, age_allrm, age_detach, ageold) parameter estimates, or marginal effects of age, for each housing unit in area k

age_i= current year minus year-built value (whether known or imputed) for housing unit i

$age^2{}_i$= age_i squared

all_rm_i= total number of rooms in housing unit i

$detached_i$= binary, 1 for single, detached structure (0 otherwise)

old_i= binary (current definition) 1 for age_i > or = 40 years ago (0 otherwise),

(previous definition) 1 for year-built value between 1900 and 1919 (0 otherwise).

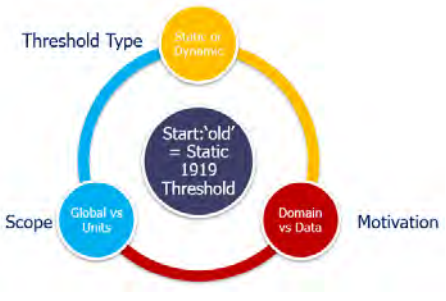
old_i (or "old") is the variable that we have redefined. Equation (3) above displays how "old" is related to age-bias adjustment factors.

3.1 Issue with Previous Definition of "old"

In 1988, when age bias was first introduced, "old" referred to housing unit built before 1900. Since 2007, "old" referred to housing units aged 100 and older. It has been a long time since this definition has been reviewed.

Attachments 2, 3, 4, and 5 show the age-bias adjustment factors with respective to age for each housing unit in that census region or census division. Issues with old housing units can be observed. All charts display instances where housing units exhibit negative values, which is unexpected. A potential explanation for this anomaly could be significant renovation the housing units. Some charts also appear to have a U-shaped trend, which indicates a possible understatement on depreciation effect on newer housing unit over time and an overstatement of rent reducing effect of age for older units. Separating the older housing units could improve the fit of the BLS model.

4. Decision Architecture



The existing "ageold" variable (see *Attachment 1* for definition) used in the age-bias adjustment factor is based on a static threshold, derived from the "old" variable, but this approach lacks robust justification. To improve the motivation behind the "ageold" variable, it is crucial to re-examine the threshold year that defines "old". This investigation should explore several factors: whether the threshold should be static or dynamic, the balance of between domain-driven insights versus purely data-driven methods for setting the threshold year, and the CPI pricing

area (PSU) scope to which the threshold should apply. A more accurate and contextually relevant "ageold" variable can help achieve better age-bias adjustment across diverse geographical units.

4.1 Testing Data

Most of the input data are derived directly from the CPI Housing Survey collection. Census data are from the American Community Survey (ACS) 5-year data at the block group level. Income data are from the IRS's most recently published statistics and include the total adjusted gross income and number of returns by zip-code level [9]. The primary data used is 2022 data. The secondary data used to confirm initial findings is 2014 data.

The primary (2022) and the secondary (2014) data can be divided into smaller geographical units based on Census Regions and Divisions. There are four regions, nine divisions, and 75 (for 2022) / 87 (for 2014) primary sampling units (PSUs). The geographical size of each unit from largest to smallest is Census Region, Census Division, and PSU. While the number of geographical units is consistent across the primary and secondary data, there are variations in how the units are defined due to changes in methodologies over time.

The sample distributions for the regions and divisions across the primary and secondary data sets can be seen in *Table 1*.

Table 1: Sample distribution by Geographic Unit

	size	mean	std	min	max
2022 Regions	4	5195.75	1920.87	4113	8067
2014 Regions	4	10876.75	2476.34	8139	14061
2022 Divisions	9	2309.22	1359.2	897	4828
2014 Divisions	9	4834.11	1967.96	2385	7195

5. Domain-Driven Insights versus Data-Driven Method

5.1 Domain Knowledge

Recent advancements in collecting 'Year built' data include the use of tax assessment data and other resources by data collectors. However, a key concern is that the "year built" data may not fully capture depreciation due to major improvements, which are not reflected in this variable. For example, significant improvements such as projects requiring permits do not alter the 'Year built' data, and therefore, are not accounted for in the age bias as well. Only structural changes like the addition or removal of bathroom, bedroom, and other room are considered in the current model.

Previously, units built in 1919 or earlier are classified as "old". Analysis of major home improvements from 1920 to the present reveals several factors influencing property value that are not reflected in the previous definition. For instance:

- Plumbing: Around 1950, the first PVC pipes were laid in the US which provided any alternative and more affordable plumbing improvement work. [10]
- Electrical: In 1913, Fred W. Wolf invented the first home electric refrigerator. By the 1940s, having a compressor refrigerator became the standard for almost all home kitchens. [11]

- HVAC: By the late 1960s, most new homes had central air conditioning. Window air conditioners were more affordable than ever, fueling population growth in hot weather states. [12]

Other no-permit required projects includes:

- Windows: Double-pane windows are the most widely used, but they weren't widely available to residential homes until the 1970s. [13]
- Painting: In 1978, lead paint in homes were banned. [14]

The U.S. Department of the Interior's National Park Service maintains and possesses the National Register of Historic Places. This is the official list of the Nation's historic places worthy of preservation. There are many requirements for a property to be considered and one of the requirements is the property must be more than fifty years old. This fifty-years standard was established by National Park Service historians in 1948. [15]

Based on domain knowledge, it is recommended that the definition of "old" be updated to include units built in 1978 or earlier. This change would eliminate all the units that may have used "old" home improvement criteria. Additionally, to keep the new definition current, it is suggested to adopt a rolling time frame, aligned with the National Park Service's method, rather than a fixed year.

5.2 Coding / Modeling

The objective of this analysis is to conduct an exploratory investigation of the "ageold" variable within the CPI Housing age-bias adjustment factor models. This data-driven exploration utilizes Python code and regression analysis, analogous to the existing CPI Housing age-bias regression analysis performed in SAS. Specifically, two Python package are utilized:

1. pandas' read_sas() method, which is used to import the SAS dataset for analysis.
2. statsmodels.api, which can generate result summaries with many useful regression statistics, e.g., coefficient estimates, p-values, adjusted R-squared, F-statistics etc.

The p-values for the "ageold" variable in the regression results of each model were used to assess the "best fit" for the threshold years. The goal is to identify the optimal threshold year for "old" within the time range from 1919 to 2014. Each year within this range was tested as a potential threshold year. For each threshold year, the "old" variable is defined as 1 if the year-built is equal to or before the threshold year, and equal to 0 if the year-built is after the threshold year.

5.2.1 Primary (2022) Data

For the primary (2022) data, regressions were conducted across all four census regions and nine census divisions to obtain the p-values for "ageold" variable. The data sample distributions are shown in *Table 1* (refer to section **4.1 Testing Data**). For each threshold year, p-values for the "ageold" variable were examined by regions and divisions. Using the p-value results, descriptive statistics were performed to explore variations in p-values over time and across geographic units. The Agile, iterative project methodology facilitated the use of gathered domain knowledge to refine the years of interest to windows 50, 45, 40, 35, 30, 25, and 20 years prior (e.g. a 50-year window for the 2022 data corresponds to a threshold at the year 1972) – as illustrated by the vertical lines in *Figures 1 and 2*. The horizontal line represents the baseline p-value (using the previous definition of "old") for comparison.

Figure 1: 2022 Averaged p-value for Regions

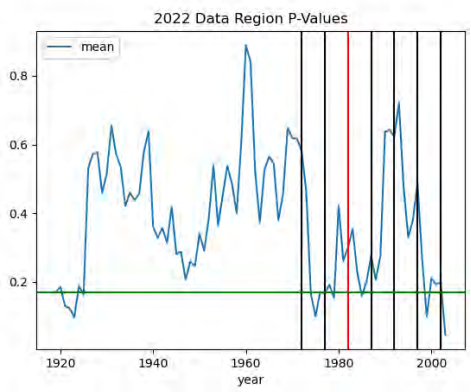
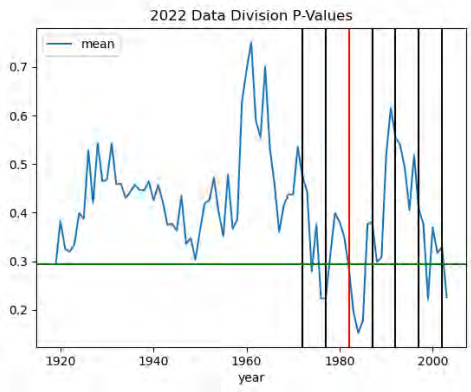


Figure 2: 2022 Averaged p-value for Divisions



As indicated by the spikes in the *Figures* above, the primary data exhibits significant noise. However, when the noise is disregarded, the 45-year, 40-year, and 35-year windows consistently show the lowest mean p-values.

5.2.2 Secondary (2014) Data

To assess if similar results could be obtained with a different dataset, the same regression methodology was applied to the secondary (2014) data. The data sample distributions are shown in *Table 1* (refer to section **4.1 Testing Data**). Every year between 1919 and 2014 was tested as the threshold year, with regressions performed for each census region and census division. This analysis allowed for the evaluation of p-values across cut-off years for each region and division using the 2014 data. A different set of geographical dummies (PSU dummies) was used, as the geographical classification in the 2014 data differs from that in the 2022 data. The previously defined windows (50, 45, 40, 35, 30, 25, and 20) were applied to obtain descriptive statistics and p-value difference comparisons for the years of interest. *Figures 3 and 4* depict these specific years of interest with vertical lines, and the comparison p-value from 1919 with a horizontal line.

Figure 3: 2014 Averaged p-value for Regions

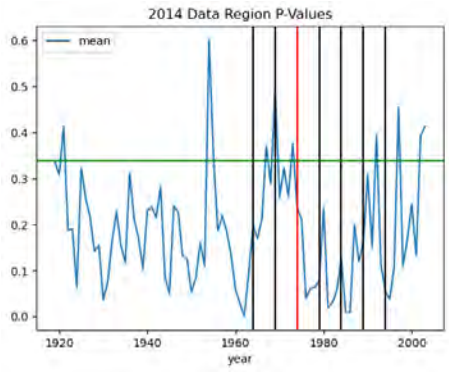
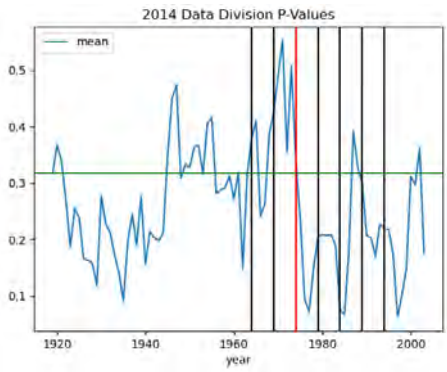


Figure 4: 2014 Averaged p-value for Divisions



Similarly, as indicated by the spikes in the *Figures* above, the secondary data exhibits significant noise. However, when the noise is disregarded, the 35-year, and 30-year windows consistently show the lowest mean p-values.

5.2.3 CPI Pricing Area (PSU) Scope

To validate the findings from the primary and secondary results, the data was further disaggregated by PSU (Primary Sampling Unit), providing the most granular geographic analysis. Every year between 1919 and 2014 was used as the threshold year, and regression were conducted for each PSU. For the primary 2022 data, 75 models (one for each PSU) were run for each year of interest. For the secondary 2014 data, 87 models were run for each year of interest. Unlike regressions for regions and divisions, PSU dummies were not included in the regressions. The previously defined windows (50, 45, 40, 35, 30, 25, and 20) were applied to obtain descriptive statistics and p-value difference comparisons for the years of interest. *Figures 8 and 9* illustrate the average p-values over time for the 75 PSU models from 2022 and the 87 PSU models from 2014, respectively. Vertical lines represent the specific years of interest, and a horizontal line represents the comparison p-value from 1919. The shaded area on both figures represents the standard deviation.

Figure 5: 2022 Averaged p-values for PSU's

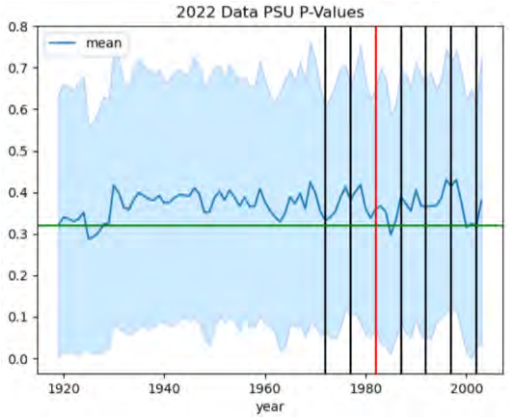
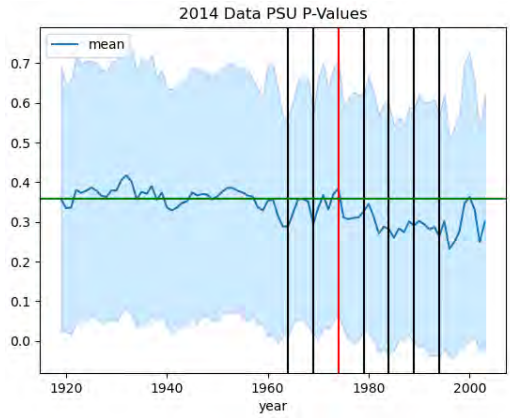


Figure 6: 2014 Averaged p-values for PSU's



The standard deviation reflects substantial variation in p-values across PSUs for both the primary and the secondary datasets. Furthermore, no consistent upper or lower bounds are observed when comparing the mean p-values between the primary data and the secondary data.

6. The Impact from New "old" Definition

The CPI Housing Survey are collected for the same housing unit every 6 months and, in most cases, represent the amounts tenants actually paid in the reference month to occupy the unit. Age-bias adjustments for Rent and OER indexes are estimated annually for the January index and remain constant for one calendar year.

6.1 Applying Age-Bias Adjustment Factor

This is the BLS equation for 1-month price relatives for index estimation:

$$REL_{t-1,t,s} = \left[\frac{\sum W_s R_{i,t}}{\sum W_s R_{i,t-6}}\right]^{-6} \tag{4}$$

where, W_s= rents weight (for either Rent or OER) in segment s;

For CPI Housing Survey, there are 75 primary sampling units (PSUs) broken down into small geographic areas by PSU, state, count, average rent (or rent level), and tract

called segment. An average of about five rental housing units are selected within each segment.

$R_{i,t}$= rent (either economic rent or pure rent) for housing unit i in time period t (month);

$R_{i,t-6}$ = rent for housing unit i in time period $t-6$ (month).

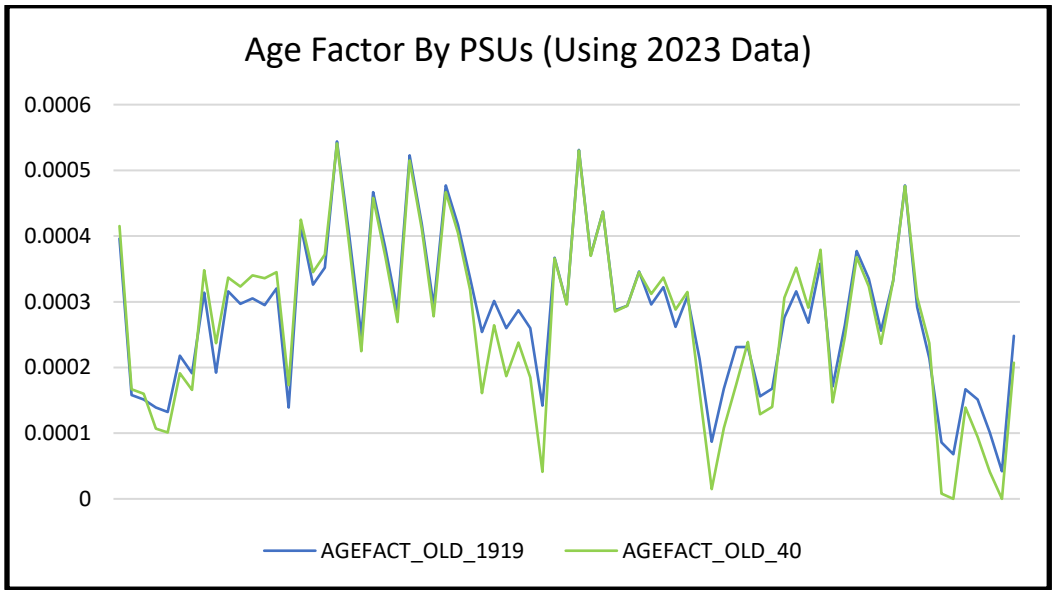
The age-bias adjustment factor affects only the previous rent, $R_{i,t-6}$ (highlighted in grey above). It accounts for the depreciation of housing stock within a PSUs in the 6 months between the collection of rent data. There are also facility adjustments, structural characteristics adjustments, and other adjustments, but for simplicity, those variables won't be mentioned in the following equation. *

$$R^{*}_{i,t-6} = R_{i,t-6} + \left[R_{i,t-6}(e^{-6*F}) - 1\right] \quad (5)$$

Where, F= age-bias adjustment factor.

Figure 9 illustrates the housing age-bias adjustment factor for the year 2023. The x-axis represents 75 PSUs (primary sampling units). The darker line corresponds to the previous "old" definition with 1919 threshold, while the lighter line corresponds the current "old" definition with 40-year window threshold. The consistent overlap of both lines indicates that there was no significant impact from the change in definition. Nonetheless, the update to the variable "old" was necessary and has been addressed.

Figure 9: Previous Definition vs New Definition of "old"



7. Conclusion

Previously, the definition of "old" housing unit was static and limited, encompassing units built between 1900 and 1919. As fewer units from that period are collected over time, the definition became increasingly obsolete. This team was tasked with updating the "old" definition, aligning it more closely with those used by other agencies and institutions.

Based on domain-driven insights, it is recommended to redefine "old" as units built at or before 1978. This adjustment would exclude units that may have used outdated home improvement criteria, such as lead paint. To ensure the definition remains relevant, it is also suggested to adopt a dynamic threshold approach like that of the National Park Service, which uses a length-of-year definition rather than a specific year cutoff.

Incorporating data-driven results, the years of interest were further narrowed to the 45-year, 40-year, 35-year, and 30-year windows. Due to the noise in the data, greater emphasis was placed on the higher year windows to align more closely with other agencies' practices. The regression model using a 40-year dynamic threshold demonstrated acceptable mean p-value across different data sources and is considered the most suitable replacement for the static 1919 threshold for the "old" variable. This revised definition would, at minimum, provide a better fit to our BLS model compared to the previous, outdated definition.

7.1 Further Research

The age-bias factor consists of six depreciation variables. To understand their impact on $'age'$, each variable was individually removed from the model. The most significant effect came from removing $'age^2'$. This variable serves as a multiplier in the calculation of a housing unit's age-bias factor, and its removal addressed the U-shaped curve issue (refer to section **3.1 Issue with Previous Definition of "old"**). However, age^2 is considered essential for the age-bias adjustment factor, so its removal is not a viable solution unless further research indicates otherwise.

Figure 7: U-shaped Curves That Use "age2" Variable Calculation

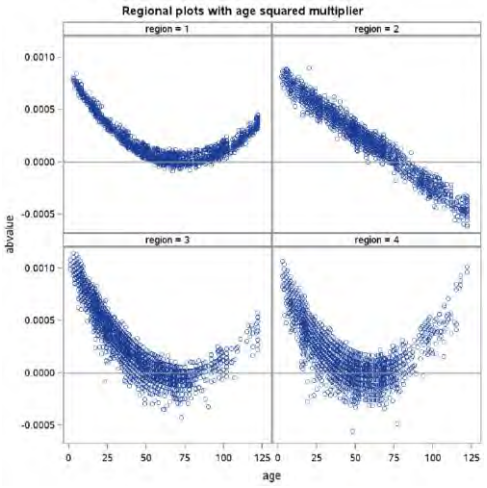
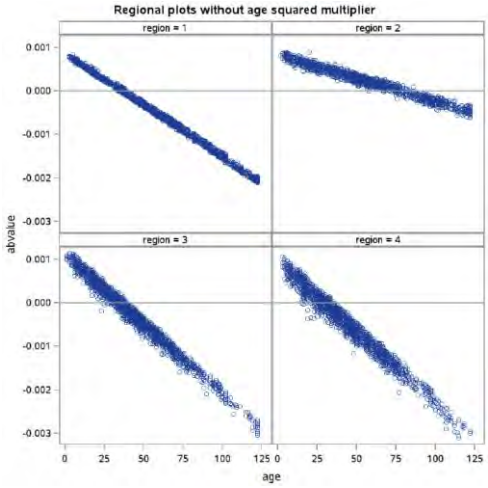


Figure 8: Curves With "age2" Removed From



The original U-shaped curves can be seen in *Figure 7*. The adjusted linear curves calculated without the "age2" variable can be seen in *Figure 8*.

References

[1] Bureau of Labor Statistics (2023, September) *Handbook of Methods: Consumer Price Index.* https://www.bls.gov/opub/hom/cpi/home.htm

[2] Gallin, J., & Verbrugge, R. (2007, October). *Improving the CPI's Age-Bias Adjustment: Leverage, Disaggregation and Model Averaging*. Bureau of Labor Statistics. https://www.bls.gov/osmr/research-papers/2007/ec070100.htm

[3] Campbell, L. L. (2006, November) *Updating the Housing Age-Bias Regression Model in the Consumer Price Index*. Bureau of Labor Statistics. https://www.bls.gov/cpi/quality-adjustment/updating-housing-age-bias.pdf

[4] Lane, W. F., Randolph, W. C., & Berenson, S. A. (1988, October) *Adjusting the CPI shelter index to compensate for effect of depreciation*. Bureau of Labor Statistics. https://www.bls.gov/opub/mlr/1988/10/rpt1full.pdf

[5] Houck, B. (2023, August) *Monthly Labor Review: A review of recent improvements to the CPI's housing age-bias adjustment*. Bureau of Labor Statistics. https://www.bls.gov/opub/mlr/2023/article/a-review-of-recent-improvements-to-the-cpis-housing-age-bias-adjustment.htm

[6] Ptacek, F. (2013, August) *Monthly Labor Review: Updating the rent sample for the CPI Housing Survey*. Bureau of Labor Statistics. https://www.bls.gov/opub/mlr/2013/article/updating-the-rent-sample-for-the-cpi-housing-survey.htm

[7] Bureau of Labor Statistics (2024, March) *Factsheets: Measuring Price Change in the CPI: Rent and Rental Equivalence.* https://www.bls.gov/cpi/factsheets/owners-equivalent-rent-and-rent.htm

[8] Randolph, W.C. (1988, July) *Housing Depreciation and Aging Bias in the Consumer Price Index*. Journal of Business & Economics, Vol. 6, No.3, pp. 359-371

[9] Internal Revenue Service (2022, December) *SOI tax stats – individual income tax statistics – SIP code data (SOI).* https://www.irs.gov/statistics/soi-tax-stats-individual-income-tax-statistics-zip-code-data-soi.

[10] Walker, B. (2011, January) *North America's Cinderella Pipe Story: A Look at PVC Pipes' Climb to the Top.* Journal of ASTM International, Vol. 8, No. 7. p 3. https://dnr.wisconsin.gov/sites/default/files/topic/Wells/PVCStudyGroup/PVCPipesClimbtoTopPaper.pdf

[11] Whirlpool. *The history of the refrigerator: ancient origins to today*. https://www.whirlpool.com/blog/kitchen/history-of-the-refrigerator.html

[12] Lester, P. (2015, July) *History of Air Conditioning*. Department of Energy. https://www.energy.gov/articles/history-air-conditioning

[13] Hicks, W. (2023, June) *NREL Windows Research Clearly Making a Difference*. National Renewable Energy Laboratory. https://www.nrel.gov/news/features/2023/nrel-windows-research-clearly-making-a-difference.html

[14] Centers for Disease Control and Prevention. *About Lead in Paint*. https://www.cdc.gov/lead-prevention/prevention/paint.html?CDC_AAref_Val=https://www.cdc.gov/nceh/lead/prevention/sources/paint.htm#cdc_prevention_res-resources

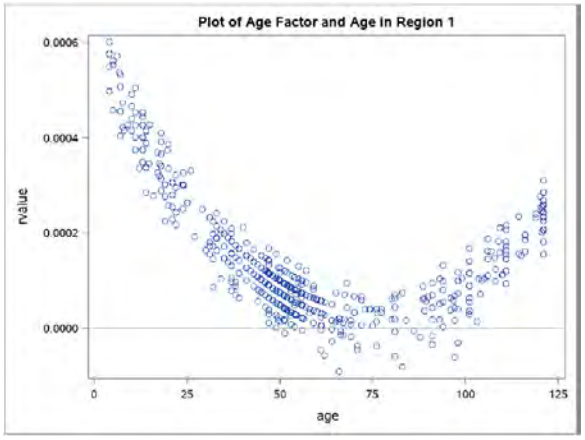
[15] Sprinkle Jr, J. (2007, Spring) *Of Exceptional Importance: The Origins of the Fifty-Year Rule in Historic Preservation.* The Public Historian, Vol. 29, No.2, pp. 81-103. https://doi.org/10.1525/tph.2007.29.2.81

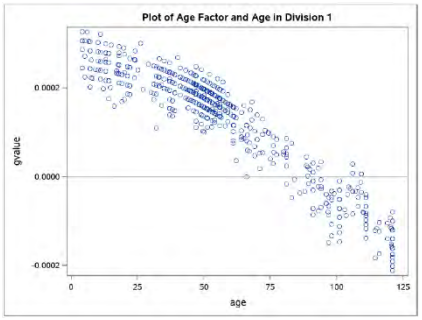
Attachment 1: List of Variables in the Hedonic Regression Model

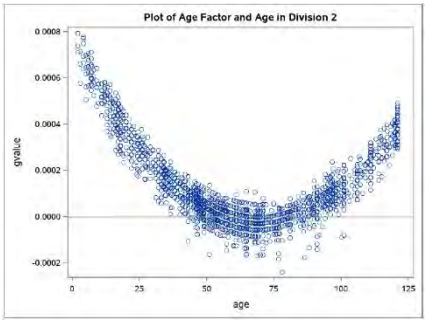
VARIABLE	SOURCE	VARIABLE DESCRIPTION
age	CPI HOUSING SURVEY	Current year less the year a unit was built (whether known or imputed)
age^2	CPI HOUSING SURVEY	Age squared
age^3	CPI HOUSING SURVEY	Age cubed
age_allrm	CPI HOUSING SURVEY	Age times the total number of rooms
age_detach	CPI HOUSING SURVEY	Age for single, detached structures (0 otherwise)
ageold	CPI HOUSING SURVEY	**Current: age of structure built more than 40 years ago (0 otherwise)** *Previous: age of structures built between 1900 and 1919 (0 otherwise)*
bath	CPI HOUSING SURVEY	Total number of full and half bathrooms
$bathrooms^2$	CPI HOUSING SURVEY	(Total number of full and half bathrooms) squared
bed	CPI HOUSING SURVEY	Number of bedrooms
$bedrooms^2$	CPI HOUSING SURVEY	(Number of bedrooms) squared
other_rms	CPI HOUSING SURVEY	Number of other rooms, such as living rooms, kitchens, and finished basements
central air	CPI HOUSING SURVEY	To indicate if a unit has central air-conditioning
window air	CPI HOUSING SURVEY	To indicate if a unit's air-conditioning is a window air-conditioning unit
other air	CPI HOUSING SURVEY	To indicate if a unit's air-conditioning is swamp cooler, heat pump, or through-the-wall air conditioner
electric heat	CPI HOUSING SURVEY	To indicate if a unit's heating is electric
gas heat	CPI HOUSING SURVEY	To indicate if a unit's heating fuel is natural gas
electric incl	CPI HOUSING SURVEY	To indicate if electricity is included in the rent
heat incl	CPI HOUSING SURVEY	To indicate if heating is included in the rent
detached	CPI HOUSING SURVEY	To indicate if a structure is single family, detached
mobile home	CPI HOUSING SURVEY	To indicate if a structure is a mobile home
elevator	CPI HOUSING SURVEY	To indicate if the structure has an elevator
extra changes	CPI HOUSING SURVEY	To indicate that a unit's rent includes any extra charges
new tenant	CPI HOUSING SURVEY	To indicate that the occupant(s) moved into the unit since the last time rent was collected
length of occ	CPI HOUSING SURVEY	Length of occupancy, or the number of months between the month a tenant moved into a unit and the most recent month of CPI Housing Survey data collection
senior com	CPI HOUSING SURVEY	To indicate that a unit is part of a senior-living residential community

density	Census Bureau	The log of total block group population divided by land area
age 65pl	Census Bureau	The percentage of the block group's population that are ages 65 years and older
large buildings	Census Bureau	The percentage of the block group's housing units in structures with 50 or more units
mobile homes	Census Bureau	The percentage of the block group's housing units that are mobile homes or trailer units
renter occ	Census Bureau	The percentage of the block group's households that are renter households
school age	Census Bureau	The percentage of the block group's population that are ages 6 to 18
age25+college	Census Bureau	The percentage of the block group's population ages 25 years and older with at least some college attendance or completion
high poverty	Census Bureau	The percentage of the block group's population for whom poverty status is determined as under the poverty line
age 3+college	Census Bureau	The percentage of the block group's population ages 3 and older enrolled in undergraduate, graduate, or professional school
adj avg inc	IRS	The average adjusted gross income for a zip code expressed as a ratio of the zip-code average to the average for the CPI pricing area

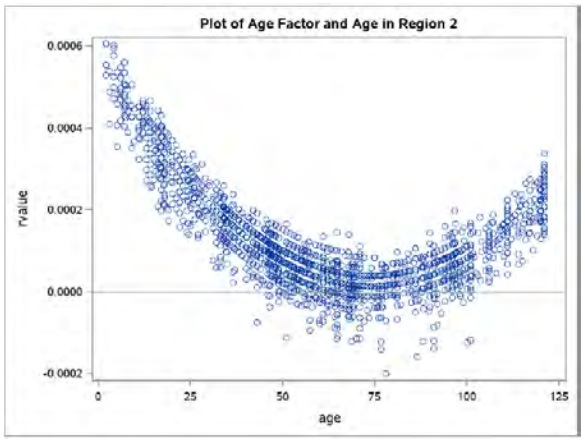
Attachment 2: Census Region 1 and their Census Divisions (1 & 2)

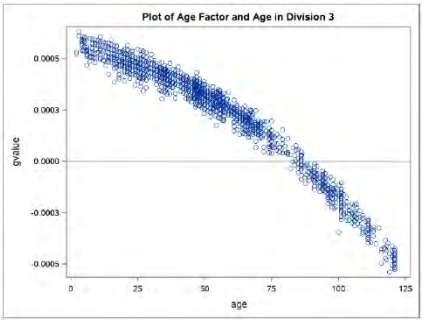


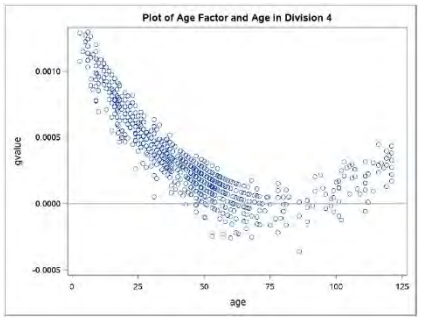




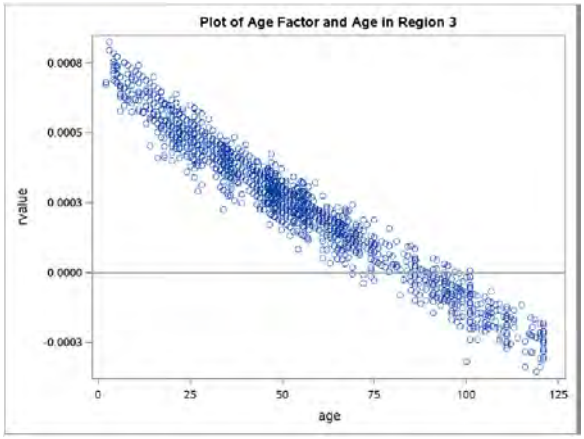
Attachment 3: Census Region 2 and their Census Division (3 &4)

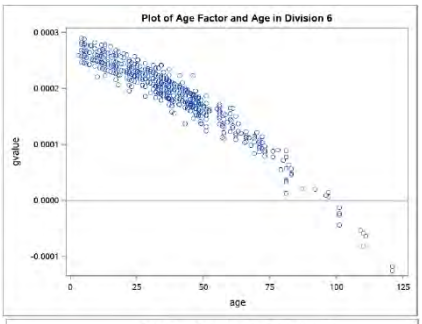


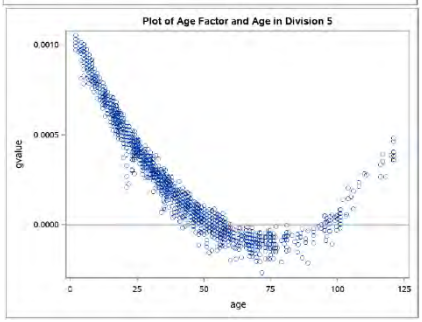


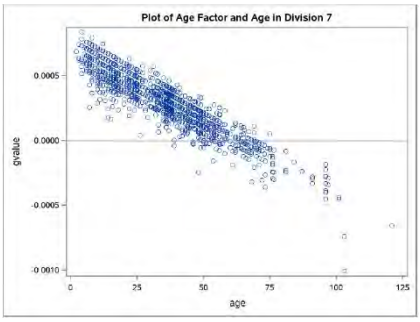


Attachment 4: Census Region3 and their Census Division (5, 6, &7)









Attachment 5: Census Region 4 and their Census Dvision (8 & 9)