

Nonresponse Imputations and Related Measurement Issues in the CPI for Shelter¹

Lara Loewenstein¹, Hugh Montag², and Randal Verbrugge¹

¹Federal Reserve Bank of Cleveland

²US Bureau of Labor Statistics

April 14, 2026

Shelter is the largest component of U.S. CPI, so the accuracy of shelter inflation is critical for the accuracy of overall inflation. Nonresponse in the BLS Housing Survey, which underpins CPI shelter measurement, has increased over time and now represents roughly 40 percent of observations. Missing rents are currently imputed using a class-mean approach based on rent tier, likely leading to upward-biased imputations. We study alternative simple imputation methods built upon variables correlated with both nonresponse and rent growth, including structure type and tenure length, and explore regression-based approaches in the appendix. While a simple model demonstrates that different methods could yield sharply different index biases, we find that in practice, these alternative methods yield similar shelter inflation indexes, suggesting that any index bias may be modest.

¹ The views expressed in this paper are solely those of the authors and do not necessarily reflect the opinions of the Federal Reserve Bank of Cleveland, the Federal Reserve System, or the BLS. Disclaimer: 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

Accurately measuring rental prices is essential to the federal government’s ability to construct price indices and, in turn, to measure inflation. The BLS Housing Survey provides the underlying data for CPI tenant rent and owner-equivalent rent, which together account for the bulk of CPI shelter inflation. Shelter represents more than 30 percent of headline CPI and nearly 20 percent of PCE, with even larger shares in their ex-food-and-energy measures. As a result, even small inaccuracies in the construction of shelter indices can translate into economically meaningful distortions in measured inflation.

That nearly 40 percent of observations in the BLS Housing Survey are currently imputed—and that most of these (more than 30 percent of all observations) reflect item nonresponse—is a cause for concern. This concern is amplified by the fact that the imputation rate has been steadily rising, effectively doubling since 2000 (see Figure 2).

If the imputed responses have the same inflation rate as the non-imputed responses, then the imputation procedures will have minimal effect on the aggregate inflation rate. However, we find that the inflation rates do vary by imputation status, as shown in Figure 3. Moreover, the imputed inflation is quite volatile and exerts a noticeable effect on the overall inflation rate.

For accuracy of the CPI shelter index, what matters is the accuracy of the *average rent change* estimate. In this context, it is less of a concern that the imputation method maintain the variance of the true underlying data (or even that unit-level rent change imputations be accurate). Some imputation methods, such as regression imputation (without added noise), could result in underestimates of variance, but this is not relevant to the construction of the index per se. In other contexts, such as regression analysis, preserving the distribution of the

underlying data via the use of methods like stochastic regression imputation is a much more important consideration. Since variance maintenance is of second order importance in this context, we focus mostly on class-means imputation methods. Use of such methods is common practice in inflation measurement (Graf, 2020) and our alternatives are similar enough to existing practice that implementation is feasible.

We make several contributions. First, we draw attention to an important issue in CPI shelter inflation measurement: a growing problem with survey nonresponse, resulting in a smaller effective sample size and driving a growing imputation rate. The BLS Housing Survey is not the only government survey affected by nonresponse and other forms of missing data. See Figure 1 for a comparison of household survey response rates. This project contributes to a broader economics literature on missing data and measurement in official statistics, including (Hokayem et al., 2015), who show that income nonresponse in the CPS March Supplement induces a 1 percent bias in official U.S. poverty rates, and (Heffetz & Reeves, 2021), who find that measurement issues (including nonresponse) may have generated a bias of up to 1.5 percent in either direction in the unemployment rate during the COVID19 pandemic. Other BLS surveys, such as the Consumer Expenditure Survey, have likewise adopted imputation methods to address missing data. (Groves & Peytcheva, 2008) provides a meta-analysis of methodological studies examining the conditions under which nonresponse bias emerges and when it is quantitatively meaningful.

Second, we analyze nonresponse in the BLS housing survey. Using observables available in the data, we assess whether nonresponse is correlated with rent level, length of tenure, and property type. We find nonresponse is slightly higher for higher rent level units, substantially lower for units in multifamily buildings, and is potentially much lower for newly

moved in tenants. Having said that, a nonresponse implies that the surveyor does not know whether the non-respondent is newly moved in or not, so there is a fair amount of uncertainty as to the tenancy of these tenants. Since shelter inflation rates also vary across these categories, any imputation strategy that does not account for these correlations can result in biased estimates of rent changes for nonresponses.

Third, we detail the current imputation method. It is a class means imputation approach based exclusively on three variables: city, month, and tercile of the lagged rent level of a unit. Our analysis reveals that rent level exhibits weak correlations with both nonresponse patterns and rental inflation rates. Notably, the existing method fails to account for structure type and tenancy length—the two covariates we identify as most strongly correlated with nonresponse and, crucially, with rent growth. Our evaluation demonstrates that this approach generally results in a positive bias to unit-level inflation estimates, since relative to the population, the method overweights the large jumps in rent associated with a new tenancy.

Fourth, we assess other class-means imputation strategies, including the use of tenure length and structure type. We examine the extent to which changing the covariates used for class means imputation reduce this bias and enhance accuracy. As noted above, exclusion of tenancy length in the imputation procedure appears to lead to upward-biased imputations in practice. However, tenancy length for nonrespondents is also not known in real time. At the extreme, all nonrespondents might be new tenants, or all might be continuing tenants: there is no way to know. The degree of imputation bias (and thus, the amount of measured inflation in the index) depends on which of these extremes is more accurate. Using an outside data source, we propose a probabilistic method for determining the average share of

nonrespondents that are new tenants. Our results indicate that the implied bias is much smaller than its theoretical maximum.

Fifth, we examine the relationship between unit-level imputation errors and index accuracy. In Appendix A, we present a simple model that captures many of the features of our data. This model demonstrates that minimizing unit-level imputation errors does not necessarily lead to index accuracy; instead, what is critical is to accurately estimate the average rent change among nonresponding units *that exit the sample*. Given that one-sixth of the sample exits annually, imputation inaccuracies could introduce substantial bias into shelter indices. However, because the relevant imputation errors concern nonresponding exiters—a group for which data are unavailable—a full quantitative assessment of index accuracy presents significant challenges that we reserve for future research. Nevertheless, we can observe that the similarity in rental inflation rates across our various imputation methods strongly suggests that at present, there is only minimal bias being induced in the shelter indices.

Finally, we provide some evidence about the extent to which the imputation process itself contributes to noise in the overall shelter index.

We contribute to the literature on imputation and inflation measurement. The vacancy imputation in the BLS housing survey has garnered much attention (see, for example, (Crone et al., 2004)), even though nonresponse imputation is currently a far more frequent occurrence in the BLS Housing Survey. Accordingly, this paper does not study vacancy imputation, but instead focuses on nonresponse imputation.

2 DATA

Our main dataset is the BLS Housing Survey. The housing survey underpins both the rent and owners' equivalent rent components of the Consumer Price Index. Typically, it tracks about 40,000 renter-occupied housing units. In drawing this sample, first, neighborhoods are selected from within a city; said neighborhoods almost always coincide with Census blocks or block groups. Then a handful of rental units are selected from each neighborhood. The survey is structured as a panel, where the housing units are surveyed (i.e., prices are collected) every six months. Accordingly, the sample is divided into six subsets called panels. Panel 1 consists of units that are surveyed in January and July, panel 2 consists of units that are surveyed in February and August, and so on. Note that neighborhoods are assigned to panels at random, subject to the restriction that each panel is intended to contain a weighted nationally representative set of units.² Thus, the month that a housing unit is surveyed should be uncorrelated with other housing characteristics.

We analyze the economic rent for each unit, which is the rent on the lease with the relevant adjustments that render the rent comparable from one period to the next. The housing survey contains detailed questions about the lease and housing unit in order to be able to accurately compare rents over time. For example, many leases specify that landlords are responsible for some or all utilities, but those utility responsibilities could change over time.³ Likewise, some tenants perform some in-kind labor for their lease, an agreement that may

² Likewise, within a city, each panel is representative of all housing units within the city.

³ In any case, this information is necessary in order to allow the BLS to adjust the rent for use in the OER index, which explicitly excludes utility costs.

also change over time. Most housing units experience a slight decrease in quality each year as they age and depreciate, although some involve more major renovations such as the addition of a room.⁴

The BLS does not collect information on the tenant, such as demographics or economic characteristics. It does collect the move-in date, but once a tenant moves out of the unit, the survey does not follow up with them or determine where they move to. One piece of relevant information which is collected is the date they moved into the housing unit, from which housing tenure is derived.

An important innovation relating to sample exit occurred in 2012. Prior to 2012, the BLS surveyed the same set of housing units indefinitely.⁵ Thus, many of the housing units we observe prior to 2012 are in the sample for a significant period of time. However, starting in 2012, the BLS began regularly rotating the sample of housing units. After entry into the sample, housing units are now surveyed for six years, after which they exit the sample. Each year, one-sixth of the housing units rotate out and are replaced. This has significant implications for the non-response and imputation rates post-2012. As we discuss below,

⁴ See the [factsheet on OER](#) and rent for more details.

⁵ Prior to 2010, the BLS routinely added additional housing units to the sample, with sampling based upon the decennial census data. In 2010, the CPI began initiating a plan that would continuously update the sample of rented housing units. Eventually, one-sixth of the rented housing unit sample every year would be replaced, with sample updating based on the basis of the latest available U.S. Census Bureau data. Ultimately, this new method reduces the age of the sample, more accurately reflects new construction and changes in where people live, and results in less sample attrition. Systematic sample retirement began in late 2012. See <https://www.bls.gov/opub/mlr/2013/article/updating-the-rent-sample-for-the-cpi-housing-survey.htm>, and the section on ([design in the Handbook of Methods](#)) for more information.

nonresponse imputation is chiefly important for housing units that do not respond to the final survey before they exit the sample. And sample rotation increased the exit rate dramatically.

The BLS typically uses all observations in its CPI calculation, whether or not field economists were able to obtain a response. There are three possible states for a housing observation: collected, non-response, or vacant. A housing unit is deemed “collected” if it responded to the survey, even if it did not complete all of the questions in the survey. A unit is a nonresponse if no one could be contacted, or if the respondent refused to fill out the survey.⁶ A housing unit is vacant if it is empty but has not been converted to owner-occupied housing. As there are two types of missing data, nonresponse and vacancies, there are two different imputation procedures that are applied.

We are interested in survey variables that may be correlated with nonresponse status or rent changes. One of these is tenancy length. We use the reported move-in date of the tenant, so that we can derive the length of the tenancy for the current tenant. Unfortunately, movein dates for nonresponding housing units are often unknown. In the event that the field economist cannot contact anyone at the unit, or the landlord refuses to participate, it is likely that the field economist cannot even determine if the previous tenant is still in the unit. In practice, the last known move in date is carried forward and applied to nonresponse observations. Faced with this uncertainty, we will implement several approaches to estimate the tenancy length of non-responding units, which we discuss below. A second variable that may be correlated with nonresponse status (and that is known to be correlated with rent

⁶ Survey respondents can be tenants, landlords, or property managers. In practice, the field economist attempts to contact the same respondent every survey period.

change; see (Adams & Verbrugge, 2025)) is the structure type of the housing unit.⁷ We make use of internal variables for the economic weight of housing units, so that we can construct national statistics in the same way that the BLS does.

Our dataset is the set of housing units in the BLS Housing Survey microdata from 1999-2024. Table 1 displays summary statistics for our dataset by imputation type averaged over the entire sample. The vast majority of observations are successfully collected, although a significant fraction of the sample are nonresponses. Note that as displayed above, the share of nonresponses is not static over this time period. As we demonstrate below, given current imputation procedures, this trend causes the average nonresponse rent in our sample to be larger than the average collected rent. Nonresponse observations have longer tenancy lengths than collected observations, indicating that nonresponse probabilities rise with tenancy lengths. (This could be for behavioral reasons, and/or survey attrition.) Vacant observations have short tenancies, as housing unit do not tend to be vacant for extended periods of time.

We supplement the BLS Housing Survey with mobility rates drawn from the 5-year American Community Survey (ACS) from 2010-2024. These statistics give us a measure of the share of tenants who have moved into their units within the past 12-months by year and CBSA.

⁷ Indeed, (Adams & Verbrugge, 2025) shows that rent changes for apartments and single family detached houses can vary dramatically, even within the same neighborhood.

3 CURRENT BLS PROCEDURES

3.1 The Current Nonresponse Imputation

The current BLS imputation procedure is a class means procedure based on city, month, and relative rent level (low, medium, or high relative to other rents in that city in that time period). The methodology has changed over time. Prior to 2010, the BLS considered within-city location instead of rent level. Cities were partitioned into multiple strata, roughly corresponding to a downtown area (split into two parts), then a northern quadrant, southern quadrant, etc. Nonresponding units were imputed using data from collected units in the same stratum. The present sample is considered too small to account for both rent level and within-city location simultaneously. We do not revisit the 2010 decision, and continue to focus on procedures that do not consider within-city location.

The current procedure proceeds as follows. The most recent survey collection is partitioned by city. For each city, all housing units are sorted by the rent at the last collection period, which occurred six months prior. (Note that this 6-month-lagged rent may itself have been imputed.) Housing units are then partitioned by lagged rent into tertiles: low, medium and high rent. Next, vacancy imputations are applied to vacant units. (We explain the vacancy imputation procedure below.) Henceforth, these vacancy-imputed values are treated as data.

Next, each rent level cell (i.e., the subset of valid observations in each rent level in that city) is evaluated to ensure that there are a sufficient number of observations for imputation. If the BLS has more than 5 observations in the rent cell, then any nonresponding observations in the corresponding rent level are imputed using the average rent change from this comparison pool. If the BLS has fewer than 5 respondents in this comparison pool, it

increases the comparison group by relaxing the rent-level criteria. For example, if there are insufficient observations in a low-rent cell, then they may combine the low- and medium-rent cells. Finally, if this still does not result in a sufficient comparison pool, the BLS combines all collected and imputed vacant observations for that time period in the city.

The formula for the imputation is as follows. Let $H_{c,r,t}$ denote the set of collected and imputed vacant observations in city c , cell r , and time t . The imputed rent change in percent is given by:

$$NII_{c,r,t} = \frac{\sum_{i \in H_{c,r,t}} \omega_{i,t} \text{rent}_{i,t}}{\sum_{i \in H_{c,r,t}} \omega_{i,t} \text{rent}_{i,t-6}}$$

For each non-response observation in the city-date (c,t) , the rent is thus imputed as:

$$\text{rent}_{i,t}^* = NII_{c,r,t} \text{rent}_{i,t-6}$$

3.2 Vacant Units and Their Imputation

To understand why vacancy imputations are treated as data and included when imputing nonresponses, it is necessary to understand the rationale for imputing vacancy rents at all.

The potential for downward bias in rent indices due to vacant unit attrition is well documented in the literature. As (Crone et al., 2010) and others observe, rent inflation may be mismeasured when vacant units exit the sample before their new-tenant rent is recorded. If vacant units simply drop out of the sample without an accurate final imputation—one that reflects the fact that a new tenant is presumably going to move into the unit—this would lead to downward bias. Consider an illustrative extreme case. Suppose that all rents always remain constant throughout all tenancies, and that rent increases only occurred when a new tenant moved in. In that case, dropping exiting vacant units without an imputation would cause the

index to miss a substantial portion of actual rent increases. To address this issue, the BLS applies a “vacancy adjustment” imputation to units that become vacant, ensuring that rent changes are captured even if the unit exits the sample before a new tenant’s rent can be directly observed.

The BLS applies a class means imputation procedure to vacant units, effectively assuming that all newly vacant units receive a new tenant. As before, observations are partitioned by city and month. Observations are then separated by whether their tenancy length is greater than 6 months or not (this coincides with the definition of a “new tenant” in (Adams et al., 2024) and in this paper). Collected observations from new tenants are used to impute rent changes for vacancies. Long-tenancy collected units are used to impute rent changes for continuing tenancies—or continuing vacancies. In short, new vacancies are treated as if a new tenant just moved in, while subsequently, vacant units are adjusted as if they are inhabited by a continuing tenant. As above, if there are insufficient collected units in the short-tenancy or long-tenancy cells within a city-date, then the cells are combined until there is a sufficient comparison group.

Let the set $H_{c,v,t}$ denote the collected in city c , tenancy length cell v , and date t . The imputed rent change in percent is

$$VIM_{c,v,t} = \frac{\sum_{i \in H_{c,v,t}} \omega_{i,t} \text{rent}_{i,t}}{\sum_{i \in H_{c,v,t}} \omega_{i,t} \text{rent}_{i,t-6}}$$

For each vacant observation in city-date (c,t) , the rent is imputed as

$$\text{rent}_{i,t}^* = VIM_{c,v,t} \text{rent}_{i,t-6}$$

Because vacancy inflation rates are typically much higher than continuing tenant inflation rates, their implicit relative weight in the class means procedure must be accurate, or the imputation will be biased.

3.3 Measuring Rent Inflation

The CPI rent index resembles a Young price index for each geographic area.⁸ Since housing units are sampled every six months in a panel rotation, the set of housing units in consecutive months is not overlapping for a specific geographic area. But since the BLS wishes to construct a monthly index, its current practice is to calculate a six-month price relative for each city and convert it to a one-month price relative. For city c and time t , the six-month rent price relative is

$$REL6_{c,t} = \frac{\sum_i \omega_{i,t} \text{rent}_{i,t}^*}{\sum_i \omega_{i,t} \text{rent}_{i,t-6}^*}$$

$$REL_{c,t} = REL6_{c,t}^{\frac{1}{6}}$$

where $\omega_{i,t}$ denotes the weight on a housing unit and $\text{rent}_{i,t}^*$ denotes the economic rent after imputations.⁹ We aggregate these one-month price relatives across geographic areas to obtain a national one-month price relative and construct the resulting index from that.¹⁰ In practice, when we construct inflation rates using subsets of the housing units (e.g. only for

⁸ The geographic areas represent individual metropolitan statistical areas in the case of large cities. Smaller cities are not self-representing and are aggregated by Census Division.

⁹ We abstract from the aging bias adjustment applied to the denominator.

¹⁰ See <https://www.bls.gov/opub/hom/cpi/calculation.htm> for additional details about the CPI rent index construction.

detached houses) or by implementing alternative imputation methodologies, we will attempt to follow the same procedure above.

4 COVARIATE ANALYSIS

In this section, we assess whether nonresponse and/or rental inflation rates are correlated with observable characteristics.¹¹ If the nonresponse rates do not vary across observable characteristics, then independent of whether or not inflation rates vary by those characteristics, collected rents should still be representative and the class-means approach will result in an unbiased index. Similarly, if the inflation rates are uncorrelated with observable characteristics, then the price index will be valid even if nonresponse rates are correlated with that characteristic. It is only in the event that both the inflation rate and the nonresponse rate are correlated with an observable characteristic that a poor choice of imputation method may bias the result.

We focus on three covariates in the survey: rent level, tenancy length, and structure type. The construction of rent level is discussed in the Section 3 above and reflects the tercile of the lagged rent of the unit relative to the lagged rent other units in the same panel and city. The tenancy length of an observation refers to how long a tenant has been living in the unit. As (Gallin et al., 2024) notes, continuing tenants with long tenancies face infrequent rent changes and may obtain a substantial rent discount relative to new tenants. Finally, the structure type of a housing unit is easily observable and is invariant over time. We divide

¹¹ For the definition of missing at random and a discussion of imputation theory, see (van Buuren, 2018).

structure types into multi-family (apartment), single family detached, and single family attached units.

We first construct the national share of observations each period that are nonresponses. These shares are weighted averages, with the weights reflecting both geographic weights and lower-level within-city economic weights. We calculate these non-response shares by the covariates of interest, namely rent level, structure type, and tenancy length, and plot the results over time. The objective is to get an overall picture about whether the data is missing at random with respect to these variables, whether this is changing over time, and to observe the patterns of nonresponse. We then calculate the national inflation rate using only responding and vacant imputed rents for each level of our covariates separately using the price index formula described in Section 3.3.

Nonresponse and inflation rates by our three observable characteristics are on the left and right panels of Figure 4. Figure 4a and 4b are our plots of nonresponse and inflation rates by rent level, respectively. Similarly, Figures 4c and 4d are plots of nonresponse and inflation rates by structure type, and Figures 4e and 4f are plots of nonresponse and inflation by tenancy length based on the last date collected.

As described in Section 3, since 2010 nonresponses have been imputed based upon rent level. However, nonresponse rates do not differ substantially by rent level, especially when compared to the differences by structure type and measured tenancy. Early in our sample, when aggregate nonresponse rates were lower, the lowest rent units had slightly higher nonresponse rates on average. Since 2010, the highest rent units have had the highest nonresponse rates. This variation in nonresponse rates across rent levels is never more than a few percentage points.

By contrast, the *inflation* rates for low, medium and high rent units differ persistently over time. Over the entire sample period, low rent units have consistently experienced higher rent inflation, and high rent units have consistently experienced much lower rent inflation. This is potentially driven by mean reversion ((Verbrugge et al., 2017))—some units with low rent may be overdue for a large upward rent reset, while some units with high rent may soon experience much smaller (or even negative) rent resets. Regardless of its cause, however, it is a striking feature of the data, and one that may be consequential for imputation.

Across structure type, nonresponse differences are more stark. Nonresponse rates for single family detached units are consistently the highest, followed by single family attached, with multifamily last. Recently, about *half* of single family units have not been responding, about double the rate for multifamily units. Rent inflation rates also vary across structure type, though less notably. One possible cause for the nonresponse discrepancy is that multifamily unit respondents may be more likely to be landlords or property managers, while single family detached respondents may be tenants. The gap in nonresponse grows over time, reaching a peak during the COVID pandemic when the nonresponse rate for single family detached units was over 20 percent higher than that for multifamily units.

The BLS is well aware that different structure types can have different inflation rates (Adams & Verbrugge, 2025; Gallin et al., 2024), and now accounts for this in its calculation of OER. As can be seen in Figure 4d, multi-family units (over this period) experienced a systematically higher inflation rate than single-family detached units. Differences between single-family detached units and small multi-family units are less stark, especially earlier in the sample. Some of these differences may be explained by location, but not all (Adams & Verbrugge, 2025). Although large multifamily units feature higher turnover (and thus more

frequent rent resets toward new-tenant rents: see (Gallin et al., 2024)), this cannot explain rent inflation differences over such long horizons. An alternative explanation is that multifamily units are better maintained or depreciate more slowly, and that this is not currently captured by BLS aging adjustments.¹²

In Figures 4e and 4f we have parallel figures by tenure length, where tenure length is defined using the move-in dates in the BLS Housing Survey, which are assumed to be unchanged from the previous observation for nonrespondents. (In reality, the tenure lengths for nonrespondents is unknown.) In 4e there is a clear delineation in nonresponse rates between observations with tenancies of one year or less, and those with longer tenancies. Even after one year, there appears to be a clear monotonic relationship between non-response rates and tenancy lengths over time.

The pattern of nonresponse by tenure length could partly be an artifact of how the movein dates are carried forward. Suppose that there is a given rental unit which did not respond during the last collection period. In the next collection period, a new tenant may have moved in, but if there continues to be no response to the survey, the field economist has no way to know that. In that case, BLS procedures imply that the move-in date will remain unchanged, so that this response will be associated with a tenure length that is incorrectly too large.

However, to the extent that the relationship between tenure length and nonresponse in the figure represents reality, it is of concern, because as shown in Figure 4f (and discussed in detail in (Adams et al., 2024) and (Gallin et al., 2024)), inflation rates vary substantially by tenure length: new tenants are much more likely to receive a larger rent increase relative to the previous tenant, compared to tenants that remain in their unit. Failing to control for this

¹² Conversely, a higher proportion of the rent on detached units relates to a non-depreciating asset, land.

rent change differential, whether implicitly or explicitly, seems likely to result in biased imputations.

A measurement challenge is that it is difficult to determine tenure length for a unit that does not respond. In Section 6 we discuss how we can probabilistically estimate the tenure length of nonrespondents using local renter mobility rates from the ACS.

5 NON-RESPONSE SPELLS, ERROR PROPAGATION, AND EXIT

Before we turn to alternative imputation methodologies, let us consider the duration of nonresponse spells. It seems a priori unlikely that nonresponse status for a housing unit is distributed independently across across a tenure cycle. If a tenant actually declines to respond to the housing survey one month, then it seems plausible that they may well again decline to respond the next time they are contacted. In practice, this would mean that rent changes would end up being imputed for the same housing unit multiple times in a row, possibly compounding errors.

Figure 5 depicts the duration of imputation spells for both vacancy and non-response imputations. A spell is defined as the number of consecutive periods that a housing unit's rent is not collected in the survey, either through non-response or vacancy. Over 80 percent of vacancy imputation spells are of length 1; that is, receiving more than one vacancy imputation in a row is relatively uncommon. Conversely, 40 percent of non-response imputations are part of a longer non-response spell. Putting this differently, compared to vacancy, non-response is much more likely to be followed by another non-response.

Over extended periods, imputation method choice becomes inconsequential provided that each nonresponse is ultimately followed by a valid rent observation, which effectively

cancels out any prior imputation error. To illustrate, consider a hypothetical housing unit across three collection periods $t = 0, 1, 2$.¹³ The unit responds successfully in periods 0 and 2 but not in period 1, requiring imputation. Denote the true rent in period t as r_t and the imputed rent as \tilde{r}_t . The price relative between periods s and t be denoted by $rel_{s,t}$.

Assume that the imputed rent for period 1 differs from the actual rent, $r_1 \neq \tilde{r}_1$. Then the period 1 survey rent relative calculated in the survey will not equal the actual rent relative: $\frac{\tilde{r}_1}{\tilde{r}_0} \neq \frac{r_1}{r_0}$. However, once the housing unit responds to the survey in period 2, then the error will correct itself in that $rel_{0,2}$ will be correct despite the inaccuracy in $rel_{0,1}$ and $rel_{1,2}$:

$$rel_{0,2} = \frac{r_2}{r_0} = \frac{r_2 \tilde{r}_1}{\tilde{r}_1 r_0} = \tilde{rel}_{0,1} \tilde{rel}_{1,2}$$

In practice, what this means is that if a), the sample of housing units is static, and b), housing units eventually respond to the survey again, then any errors induced by imperfect imputation methods will eventually be corrected in the index. Unfortunately, the BLS Housing Survey does not have these two features. Indeed, even abstracting from other causes of sample exit, 6-year sample rotation alone (initiated in 2012 to ensure that the rental sample remains similar to the rental stock) implies that about a sixth of the sample exits the survey every year. As we show next, a quite sizable fraction of these exiting rental units feature a non-response in their final survey.

In Figure 6 we plot the share of the rental units exiting the survey during a nonresponse spell (i.e., the share which did not respond during the collection period prior to exit). Prior to 2012, relatively few housing units left the survey each year, since units were kept in the

¹³ For simplicity, we examine a single unit and abstract from the six-month panel structure.

sample unless they transitioned to owner-occupancy. However, over this earlier period, a relatively large fraction of these units exited during a non-response spell. This is partly because non-response becomes more likely the longer a unit has been in the survey. Once the housing rotation began, the fraction of exits jumped quickly, plateauing around 18-19 percent. However, since the start of the housing rotation, the share of nonresponses among exits has grown markedly, rising from 15 to 40 percent. Therefore, the housing survey currently has an unprecedented combination of a relatively high exit rate and a high share of non-responses among exits, which means that any error in the imputation methodology has potentially increased in severity.

6 ESTIMATING TENANCY LENGTH FOR NONRESPONSES

As mentioned above, we wish to consider imputations based upon tenancy length, but tenancy length is not measured for non-responding housing units. Unlike structure type or location, which are time invariant characteristics, a field economist may have no information about the move-in date if no one in the housing unit responds to inquiries.¹⁴

The lack of information about the tenure length for nonrespondents creates uncertainty about how to impute their rent changes. In the event that a housing unit does not respond for several years in a row, it is possible that the same tenant lived in the unit for the entire time. But it is also possible that a new tenant has moved in every 6 months during the period,

¹⁴ Although it would be ideal if field economists could assess whether the same tenant inhabits a housing unit, even if they non-respond, that task may be infeasible if a landlord no longer wishes to participate in the survey or if no one can be contacted at the unit.

generating multiple short-tenure tenants. Given that rent changes depend on tenure, these two scenarios could result in very different inflation rates if tenure is used in the imputation process.

We conduct a series of imputations using four measures of tenancy length to establish bounds on tenancy's effect on the rent price index. First, we use the nonresponse move-in date information populated in the Housing Survey, which is principally calculated through a carryforward imputation. As shown below, many housing units non-respond for multiple surveys. The move-in carryforward assigns these units the same tenant throughout their entire non-response spell.

Second, we estimate the move-in dates using the extreme assumption that the same tenant lives in the housing unit the entire time it is non-responding. This scenario provides a result similar to the raw movein above. We refer to the tenancy length derived from this as the lower-bound tenancy length, so-termed because it still features low turnover and hence slower rent growth.

Third, we repeat the previous exercise, but assume that a new tenant moves into a housing unit every 12 months. This scenario approaches maximum turnover (assuming that annual leases are norm) and generates more new tenants. We refer to this tenancy length as the upper-bound tenancy length, as it features high turnover and thus faster rent growth.

Fourth, we estimate the *share* of nonresponses that are new tenants, rather than focusing on the tenancy length of any specific unit. We calculate the share of tenants that moved into their unit by city in the last six months from the 5-year ACS. The ACS gives us the share of tenants that moved in within the past year. We convert this to a 6-month mobility rate by assuming a constant monthly hazard rate of moving out. We then estimate the share of

nonrespondents that moved in within the last 6 months by assuming that the aggregate share of the BLS sample should match the ACS mobility rate.

More specifically, let w denote the share of responses in the housing survey that are nonresponses and let x denote the share of units in a subsample that are new tenants.¹⁵ Then for city c and year t ,

$$x_{c,t}^{ACS} = w_{c,t}^{Collected} x_{c,t}^{Collected} + w_{c,t}^{Nonresponse} x_{c,t}^{Nonresponse}$$

All of the variables on the right-hand-side of the equation are values from the BLS Housing Survey. All of the variables can be explicitly calculated from the housing survey or ACS except for $x_{c,t}^{Nonresponse}$, the share of new tenants in non-responses. In theory, $x_{c,t}^{Nonresponse} \in [0,1]$, although in practice $x_{c,t}^{Nonresponse}$ can lie outside of these bounds for cities and dates where the share of non-responses is small. We truncate $x_{c,t}^{Nonresponse} \in [0,1]$.

7 ALTERNATIVE IMPUTATIONS AND POTENTIAL BIAS

7.1 Imputation Methods

We estimate multiple alternative inflation rates using alternative imputation methods. Broadly speaking, these imputation methods employ the same class means approach utilized in the official CPI rent index imputation, but differ based on the specific observable variable(s) used for observation partitioning. (We discuss more complex imputation methods below.) For each non-responding housing unit at time t , we impute a rent using each alternative method. Note

¹⁵ We ignore vacant units in this derivation.

that any imputation is subsequently carried forward to the subsequent survey period. Consequently, the imputation methodology will influence subsequent rent imputations, as well as the rent tercile classification of the unit in the following period. This process potentially compounds errors for consistent non-responders, as imputed rent adjustments are applied to rent levels derived from previously imputed rent changes. Throughout this paper, we refer to this process as “chaining” the imputation method, reflecting the iterative nature of our computational approach.

First, we replicate the standard BLS CPI methodology of imputing by rent level. We separate observations into low, medium and high lagged rent cells, and follow the procedure for collapsing cells with insufficient collected observations. We then apply the price index formula of Section 3.3 and aggregate to a national series. Note that this and the subsequent inflation rates are calculated using all units, not simply the imputed ones.

Second, we partition the data solely by structure type. We divide units into single family detached, single family attached, and multi-family cells. When there are insufficient responding single family attached or detached units, we collapse to a single family cell. If necessary, we further collapse to all units within a city and time period.

Third, we perform an unconditional imputation. In this scenario, we calculate the average rent change by city-date among all responding observations and use that as the rent change for the nonresponders. This imputation method would be valid if the data was truly missing-at-random. It provides a benchmark against which to compare the other imputation methods.

Fourth, we partition by tenure, calculating imputation rates using our four different measures of tenancy length: the raw recorded value that is carried forward, our upper- and lower-bound scenarios described in Section 6, and our probabilistic estimate that uses the

ACS mobility rate. For the first three measures, we partition responding observations into two bins: those who are newly moved in and therefore have a tenancy length shorter than 6 months, and those who have been in their unit for longer, as this distinction encapsulates most of the variation in rent changes across tenancy lengths (Gallin et al., 2024). In the raw tenancy case, we aggregate new tenant and continuing tenant cells together in the event of insufficient observations within a city-date. In the upper- and lower-bound cases, we pool observations across multiple months in the same city.

Finally, in an alternative tenure-based method, we use the ACS-implied mobility rates; this obviates the need to estimate tenure lengths. We instead impute the unobserved rent change using a weighted average of the local new- and continuing-tenant rent changes, where weights derive from our estimates of $\chi_{c,t}^{Nonresponse}$. In case of cell insufficiency, something more common in the ACS case, we use a different method. In particular, in cities and dates where there are insufficient new tenant collected observations, we pool across multiple (lagged) months first, and across similar cities second, to obtain a reliable imputed rent change.¹⁶

7.2 *Impact on Inflation Measurement*

In Figure 7a we plot the year-on-year change of the rent index using alternative imputation methods, including the raw tenancy method. The imputations based on rent tier and structure type yield an inflation rate almost identical to the unconditional approach. This is somewhat surprising, given that rent changes vary noticeably by rent level and structure

¹⁶ This represents a small departure from current BLS methods.

type. The data are almost missing-at-random with respect to the rent level, which explains the irrelevance of that imputation method. It is, however, somewhat surprising that the imputation by structure type does not yield rent change estimates that are more dissimilar. One feature of the data that plays a role in these and subsequent time-series comparisons is that exiting-during-nonresponse rates over the first part of the period were quite low, leading to small inflation differentials. We would not expect notable inflation differentials prior to 2013.

In Figure 7c we plot the difference between year-over-year inflation rate implied by the rent cell approach and the inflation rates based on the alternative imputation methods. The year-on-year rent inflation rate using the raw-tenancy-based imputation method is about 0.05 p.p.% lower than the current rent cell (rent tier) and the unconditional methodologies for most periods. The effect rises to 0.2 p.p.% in 2020 and 2021 as the inflation rate rapidly rises. This discrepancy compounds fairly quickly: a rent index that started in January 2000 and used the raw tenancy imputation method would have grown by 1.5% p.p. less than the current rent cell method. In Figure 7b and 7d we plot the indexes and the difference between the rent cell index and the alternative measures respectively.

The rationale for the magnitude of the effect is straightforward. The set of units that do not respond are disproportionately made up of tenants with longer measured tenures. Those units receive smaller imputed rent changes. Further, regarding those units with shorter tenures, for the typical city-month, only a handful of the successfully surveyed housing units contain a new tenant. In the case of an insufficient number of new tenant observations, our tenancy imputation approach combines the new tenant cell with the continuing tenant cell.

In short, we impute new tenant “non-responses” with the average rent change across all units within a city-date often, mitigating the impact of the new tenancy imputation.

As noted above, the raw tenancy measure may not accurately measure the share of nonresponses in the housing survey. We plot the inflation rates associated with tenure-based imputation using raw, upper-bound, lower-bound, and ACS-based year-on-year inflation rates in Figure 8a. We also plot the difference in inflation rates relative to the rent cell approach in Figure 8c. We find that the upper-bound tenancy imputation is noticeably higher than the raw tenancy and lower-bound tenancy imputations. We ascribe the difference to the fact that the upper-bound measures assumes a high rate of tenant turnover, and avoids combining new and continuing tenant cells together in the case of insufficient observations, instead choosing to use additional months of data. In contrast, the lower-bound tenancy imputation makes similar assumptions about tenancy status as the raw tenancy comparison.

Figure 8b compares the ACS-based tenancy index against the rent tier cell, structure type, and other inflation rates. The ACS-approach is quite similar to the current rent cell methodology until the post-pandemic period, when it diverges slightly. This differential stems from the rent tier methodology inadvertently overweighting rent changes observed in units with tenant turnover, given the correlation between response rates and tenure duration.

We now summarize the impact of imputation method on measured inflation. First, we find that the current rent cell approach is little different than the naive method, an unconditional (city-date) imputation. Hence, relative to the unconditional imputation, at best the rent level cell approach improves the accuracy of the rent index only marginally. The raw tenancy-based imputation method yields slightly lower inflation rates, likely driven by the not-missing-at-random nature of the data. Extreme assumptions about tenure yield notably different

inflation rates. The lower- and upper-bound approaches provide bounds for what the tenancy-based imputation could be, if we had accurate tenancy information. An ACS-based approach allows one to condition on tenure, without the need to make strong assumptions about unmeasured tenure. Of the alternatives to the rent level method currently in use, this is our preferred approach. Note that it may need to be modified to account for lagged ACS data publication.

To explore the extent to which the implied inflation rates truly differ across imputation methods, in Appendix B we report results from estimating linear and nonlinear Phillips curve models using the 12-month inflation rates associated with five of the methods. Results are remarkably similar across methods, indicating that these indexes are not very different from one another.

A central finding of this study is that while different imputation methods do yield non-negligibly different inflation rates when viewed over long periods, the differences are notably smaller than they could have been. In the Appendix A, we develop a stylized model showing that imputation method choice has the potential to substantially affect measured inflation rates. Examining three approaches-unconditional imputation, structure-type stratification, and a method analogous to our ACS-based approach-the model demonstrates potential index biases as large as 0.8 percentage points under plausible parameters. That our empirical analysis finds alternative methods yielding relatively similar inflation estimates is therefore a meaningful result, not a foregone conclusion. The model makes clear that under different response patterns, rent dynamics, and turnover and exit rates, imputation method choice could materially alter measured shelter inflation.

The group-means imputation approach above is restricted to imputing on one covariate at a time. In Appendix C, we consider imputations based upon regression-based methods that incorporate multiple variables. We study two variants. The first is a simple (Erickson & Pakes, 2011) regression imputation that uses both tenure and structure type. The second is a richer non-response imputation whose specification mimics the regressions that the Bureau of Labor Statistics performs annually in order to estimate the depreciation of housing units. In both of these cases, we find that multi-variate regression-based imputation methods yield rent inflation rates that are quite similar to the current approach.

More sophisticated imputation methods-including predictive mean matching and machine learning techniques such as gradient boosting, random forest models, or neural networks-could potentially improve unit-level prediction accuracy by exploiting additional covariates or complex relationships in the data. The data exhibit spatial and temporal correlation, in addition to correlation by housing unit attributes, which creates additional challenges for deploying machine learning algorithms. While exploring these methods would be valuable, available sample sizes are relatively small, raising concerns about overfitting and the curse of dimensionality. Moreover, our model in A reveals that the connection between unit-level imputation accuracy and index precision is neither direct nor mechanical, suggesting that enhanced unit-level performance may not guarantee commensurate improvements in aggregate inflation measurement. Exploring more sophisticated approaches is a task we leave for future work.

8 IMPUTATION ACCURACY

While unit-level imputation accuracy does not directly translate into improved index accuracy, it remains an informative metric in its own right. Having established in the previous

section that tenancy-based and rent level cell-based imputation methods yield somewhat different results, we now address a natural follow-up question: which method more accurately imputes unit-level values?

It is not obvious how one might estimate the accuracy of an imputation for a nonresponding unit that exits, because we do not observe its actual rent change. We use two approaches to approximate estimation accuracy.

First, we perform a jackknife resampling analysis on the subset of collected observations. For housing unit i in city c at date t , we treat the rent $x_{i,c,t}$ as a missing observation and perform a class means imputation using the remaining collected observation.¹⁷ We then impute rent $x_{i,c,t}^j$ and repeat the process for the remaining collected units. Since we do know the true rent for collected units, we can evaluate the accuracy of the imputation process. We restrict this analysis to observations between January 2019 and December 2024. We perform this analysis using the current rent level imputation method and the chained tenancy method.¹⁸

¹⁷ As the rent level is defined using lagged rent, we do not need to be concerned about redefining the rent levels.

¹⁸ A jackknife procedure in this context may produce potentially unreliable results if the cause of nonresponse is significantly correlated with rent change patterns. We contend that most nonresponse causal factors—such as respondent time constraints, age demographics, survey length, declining public trust in institutions, or increased telephone call screening—are likely uncorrelated with rent change. However, certain nonresponse determinants, including structure type and survey fatigue (which correlates with tenure duration), do demonstrate correlation with rent change patterns. While income is a probable nonresponse factor that correlates with absolute rent levels—and rent levels correlate strongly with rent changes—our analysis indicates that rent level itself does not strongly correlate with nonresponse probability. The rent level imputation methodology controls for absolute rent values, while the ACS-tenancy imputation accounts for tenure factors.

Second, we focus on housing units with nonresponse spells that end with a response. In these cases, we can impute their rent for several consecutive dates and then evaluate the accuracy of their imputations when their rent is collected by the survey again.

Figure 9a displays two histograms for the jackknife imputation errors across cities and dates. The left-hand histogram depicts the imputation error using the rent level cell method (the present method), while the right-hand histogram shows the imputation error with our preferred ACS-tenancy imputations. Given the heterogeneity associated with the rental market, and the fact that class mean imputations impute the same rent change for all units within a given rent level-city-date combination, it is unsurprising that imputation errors have such wide variance.

Both imputation methods yield imputations that are biased and with notable variance. On average, the ACS-tenancy method overestimates rent changes by a statistically significant 0.18%. The average rent level cell error is slightly higher, at 0.20%, also statistically significant. Both imputation methods yield notable *median* errors of 1.03% and 0.77%, suggesting that there is significant asymmetry in the errors. Tails are also quite large. Further, as Figure 10 displays, there is significant time variation in the average imputation errors. These two methods yield highly correlated average imputation errors; the ACS-tenancy errors tend to be somewhat more volatile. The average error over time is about 0.42% and 0.34% for the ACS-tenancy and rent cell approach, respectively. The jackknife approach to measuring accuracy does not conclusively indicate that one method is superior to the other.

A comprehensive hedonic procedure incorporating controls for rent level, tenure duration, and structure type could generate imputations effectively uncorrelated with rent change patterns. However, the hedonic models studied in the Appendix generate indexes that are very similar to those derived from the class mean methods.

It does illustrate that there is substantial rent change heterogeneity that neither method is able to capture.

The imputation errors for re-responding observations display similar behavior, although at a greater scale. Figure 11 depicts a histogram of rent-level-cell-based imputation errors for re-responding units. These errors possess a similar rightward draft. The median error is 1.83% and the mean error is 1.17%. The latter is statistically significant. However, there are two issues that complicate the interpretation of these results. First, there is some heteroskedasticity to these imputation errors, since the length of the nonresponse spell differs across units. Second, it is possible that in some cases, a new tenant has moved in, yet this fact was not captured in our data. In other words, re-responding units may include some new tenant rents.¹⁹

9 CONCLUSION

In this paper, we investigate the imputation process underlying the BLS Housing Survey, a topic of growing importance given that nonresponse rates have risen steadily over the past decade. We document substantial heterogeneity in both response rates and inflation rates across observable characteristics, raising the possibility that imputation method choice may substantially affect measured inflation. A stylized but reasonably realistic model presented in the Appendix reinforces this concern, demonstrating that different imputation approaches

¹⁹ Indeed, this conjecture may explain some of the positive unit-level rent change outliers in our data more generally.

can generate economically meaningful index biases under plausible data-generating processes.

We study several simple imputation methods. The current rent-level cell imputation approach generates upward-biased imputations. An unconditional imputation—which simply applies the average rent growth of all responding units—is also upward biased in our data, driven by the fact that response rates are higher for new tenants, who tend to experience much larger rent changes.²⁰ These two approaches yield very similar indexes. Imputing by tenancy status, a more theoretically grounded method, reduces measured inflation. However, since tenancy length is unobserved for nonresponding units, implementing tenancy-based imputation presents challenges; we develop an approach that circumvents the need for unit-level tenancy measurement. Two separate validation tests reveal statistically significant unit-level imputation errors across methods. Yet as our Appendix model demonstrates, such errors need not translate mechanically into index bias—the relationship between unit-level imputation accuracy and aggregate index accuracy is indirect and subtle. Our central empirical finding is that the various imputation methods produce remarkably similar shelter inflation indexes. This observed similarity provides reassuring evidence that the current imputation procedure, despite its theoretical potential for upward bias, is not inducing material bias into published shelter indexes in practice.

Several promising directions for future research emerge from this work. First, vacancy imputations warrant deeper investigation. Although vacancies constitute a small share of the

²⁰ Having said that, the simple model in the Appendix generates both negative imputation bias, and negative index bias, from this method. Hence the bias from this method depends on details of the data generating process.

survey, newly vacant units often receive large imputed rent changes that substantially influence the rent index. Yet the jump in rent upon a new tenant occupying the unit should depend upon the tenancy length of the previous tenant, whereas the current method abstracts from any such considerations. Second, the nonresponse imputation process for owners' equivalent rent (OER) may yield different results. Like the rent index, OER is derived from the BLS Housing Survey, but structure type plays a more significant role in OER, as noted by (Adams & Verbrugge, 2025). Given that detached units now carry greater weight in the OER index, imputing by structure type could have more substantial effects, particularly in recent years. Third, we have only begun the exploration of hedonic modeling, which should allow simultaneous control for tenure, structure type, and rent level, potentially improving unit-level imputation accuracy. Machine learning methods also offer new avenues for imputation. Finally, assessing index accuracy directly remains a challenging but important undertaking. One promising approach involves developing a realistic calibrated model and computing index accuracy under various imputation methods. This work is in progress, and preliminary results align encouragingly with our conclusions thus far.

References

- Adams, B., Loewenstein, L., Montag, H., & Verbrugge, R. (2024). Disentangling Rent Index Differences: Data, Methods, and Scope. *American Economic Review: Insights*, 6(2), 230–245. <https://doi.org/10.1257/aeri.20220685>
- Adams, B., & Verbrugge, R. (2025). Location, location, structure type: Rent divergence within neighborhoods. *Journal of Housing Economics*, 69, 102081. <https://doi.org/10.1016/j.jhe.2025.102081>
- Ashley, R., & Verbrugge, R. (2025). The intermittent Phillips curve: Finding a stable (but persistence-dependent) Phillips curve model specification. *Economic Inquiry*, 63(3), 926–944. <https://doi.org/10.1111/ecin.13281>
- Crone, T. M., Nakamura, L. I., & Voith, R. (2004). *Hedonic Estimates of the Cost of Housing Services: Rental and Owner-Occupied Units*. Federal Reserve Bank of Philadelphia. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=610942
- Crone, T. M., Nakamura, L. I., & Voith, R. (2010). Rents Have Been Rising, Not Falling, in the Postwar Period. *Review of Economics and Statistics*, 92(3), 628–642. https://doi.org/10.1162/REST_a_00015
- Erickson, T., & Pakes, A. (2011). An Experimental Component Index for the CPI: From Annual Computer Data to Monthly Data on Other Goods. *American Economic Review*, 101(5), 1707–1738. <https://doi.org/10.1257/aer.101.5.1707>
- Gallin, J., Loewenstein, L. P., Montag, H., & Verbrugge, R. J. (2024). Sticky Continuing-Tenant Rents. *Work in Progress*.

- Graf, B. (2020). *Consumer Price Index Manual, 2020: Concepts and Methods*. International Monetary Fund. <https://doi.org/10.5089/9781484354841.069>
- Groves, R. M., & Peytcheva, E. (2008). The impact of nonresponse rates on nonresponse bias: A meta-analysis. *Public Opinion Quarterly*, 72(2), 167–189.
- Heffetz, O., & Reeves, D. (2021). *Measuring unemployment in crisis: Effects of COVID-19 on potential biases in the CPS*. National Bureau of Economic Research.
- Hokayem, C., Bollinger, C., & Ziliak, J. P. (2015). The Role of CPS Nonresponse in the Measurement of Poverty. *Journal of the American Statistical Association*, 110(511), 935–945.
- Houck, B. (2023). A review of recent improvements to the CPI's housing age-bias adjustment. *Monthly Labor Review*. <https://doi.org/10.21916/mlr.2023.18>
- van Buuren, S. (2018). *Flexible Imputation of Missing Data. Second Edition*. CRC Press.
- Verbrugge, R., Dorfman, A., Johnson, W., Marsh, F., Poole, R., & Shoemaker, O. (2017). Determinants of Differential Rent Changes: Mean Reversion versus the Usual Suspects. *Real Estate Economics*, 45(3), 591–627. <https://doi.org/10.1111/1540-6229.12145>
- Verbrugge, R., & Zaman, S. (2024). Post-COVID inflation dynamics: Higher for longer. *Journal of Forecasting*, 43(4), 871–893. <https://doi.org/10.1002/for.3070>

Table 1. Summary Statistics

Imputation Type	Obs (#)	Mean Rent (\$)	Median Rent (\$)	Mean Tenancy Length (months)	Median Tenancy Length (months)	
Collected Obs	1,294,897	72	994	816	51	27
Vacant Obs	126,216	7	905	701	6	0
Non-response Obs	383,929	21	1192	973	70	49
All	1,805,042	100	1030	838	52	29

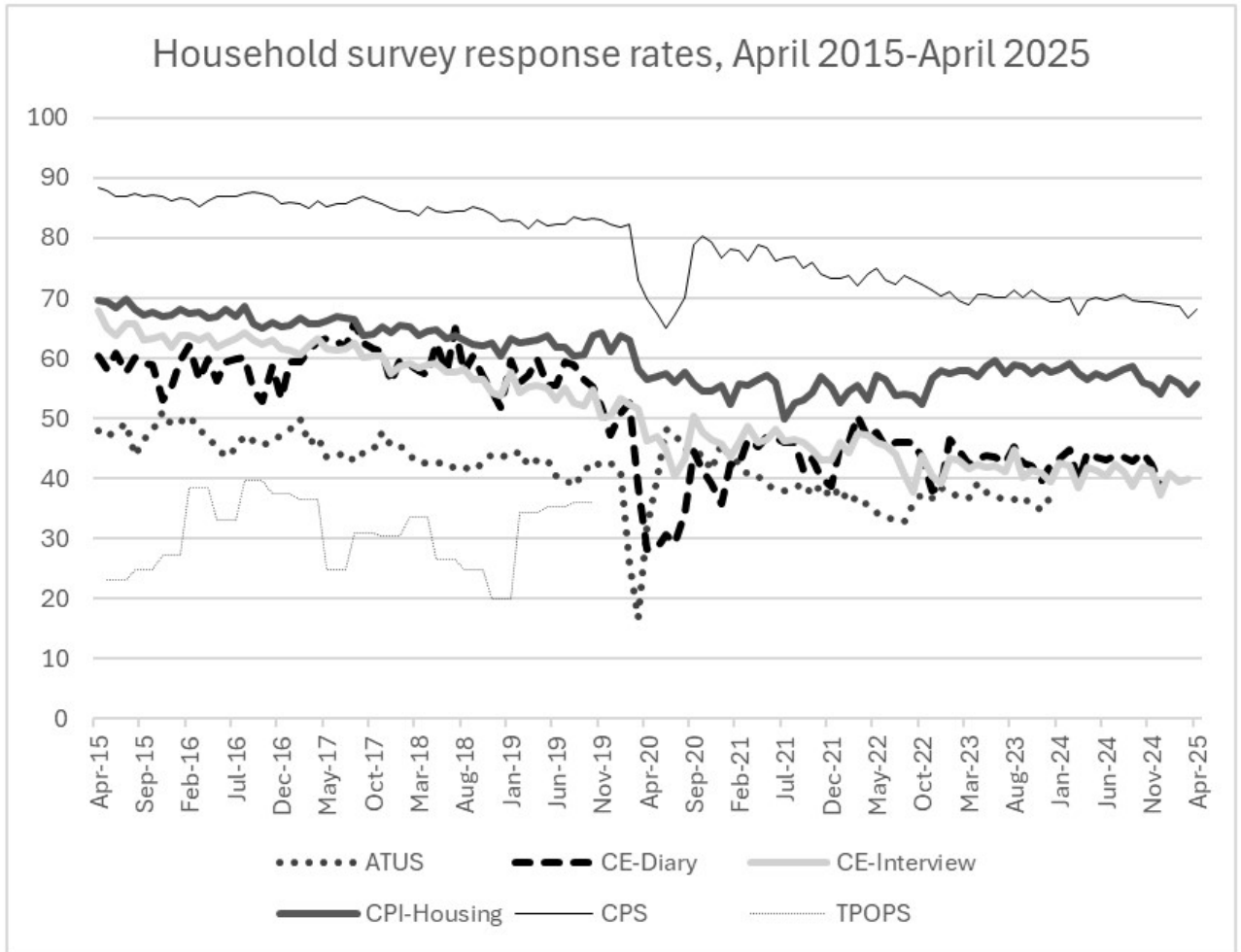


Figure 1. Imputed Share of Household Surveys. *Note:* Observations are imputed due to nonresponse or vacancy. *Source:* BLS, <https://www.bls.gov/osmr/response-rates/#chart1a>

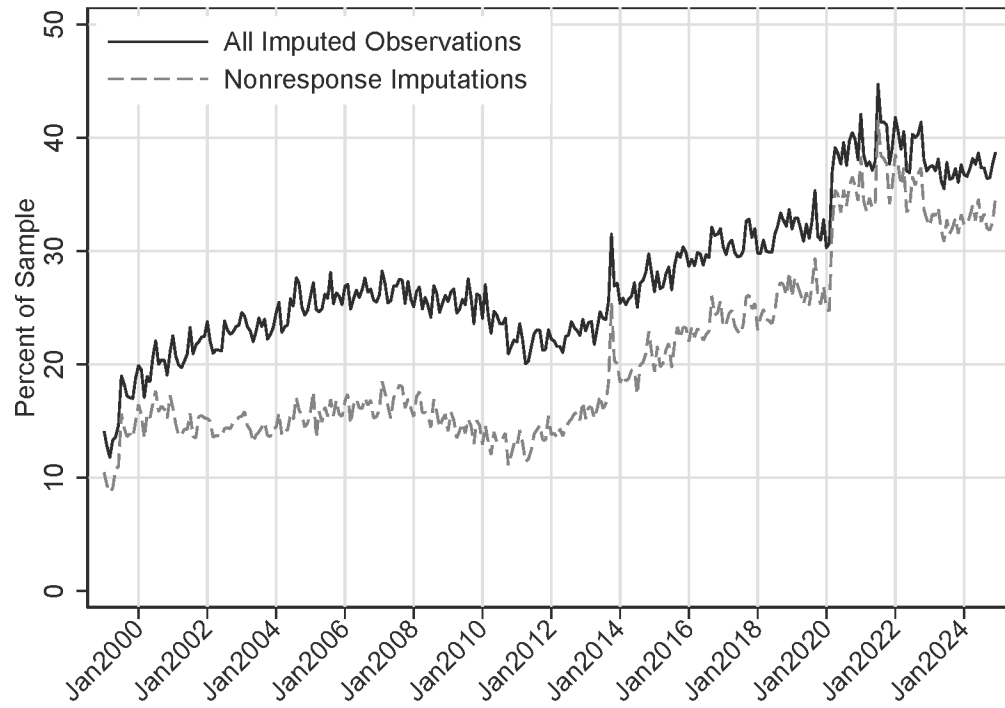


Figure 2. Share of Observations that are Imputed. *Note:* Observations are imputed due to nonresponse or vacancy. *Source:* Authors' calculations using the BLS Housing Survey.

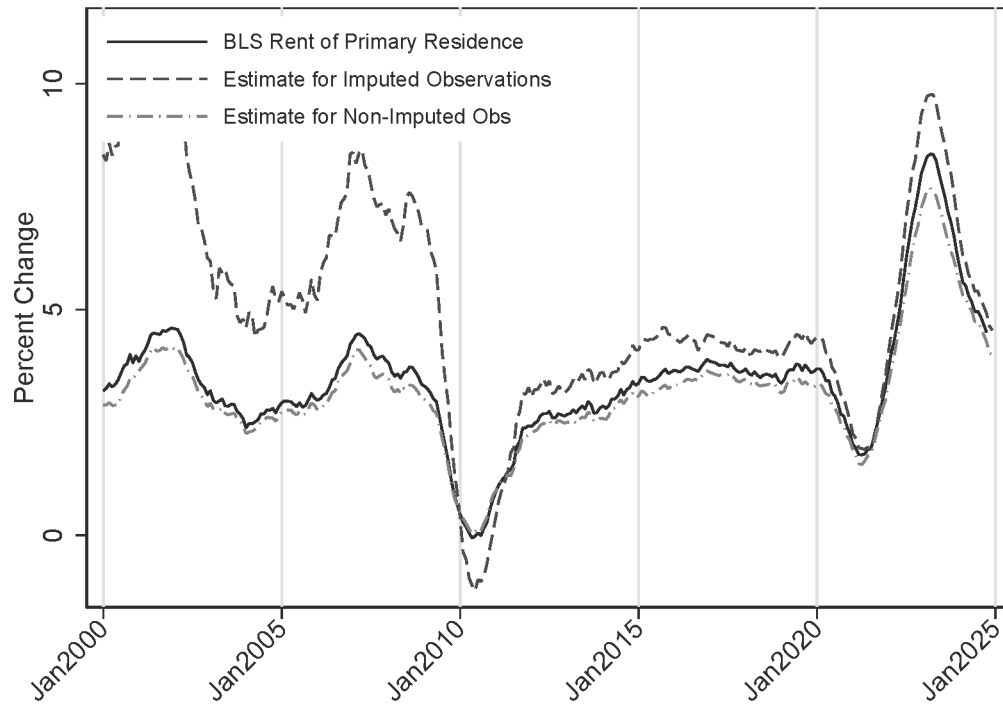
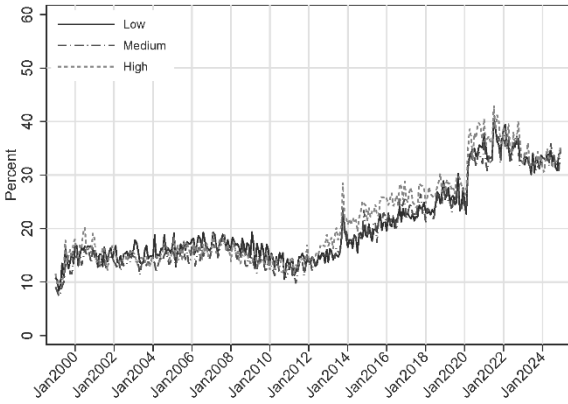


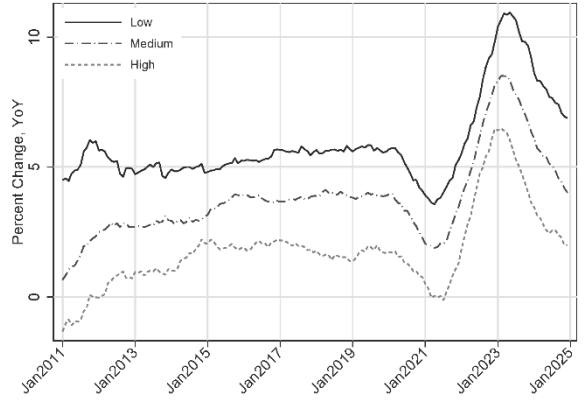
Figure 3. Inflation Rates for Imputed and Non-Imputed Observations. *Note:* Inflation is calculated using the method used for official CPI shelter inflation, but limiting the observations to imputed and non-imputed observations. *Source:* Authors' calculations using BLS Housing Survey.

By Rent Level

(a) Nonresponse

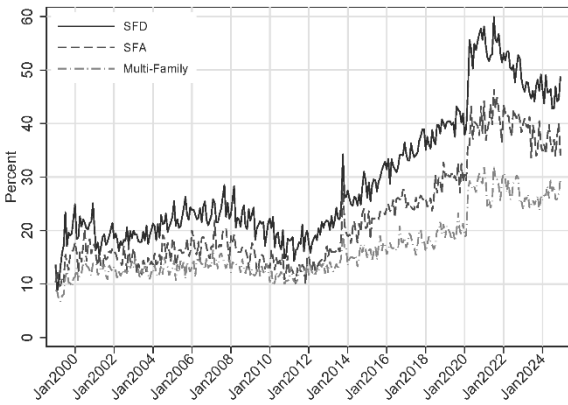


(b) Inflation

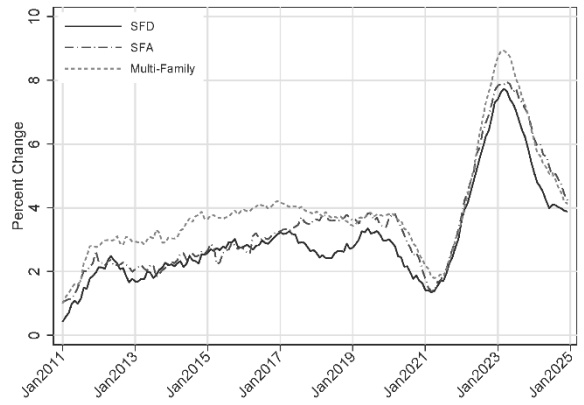


By Structure Type

(c) Nonresponse

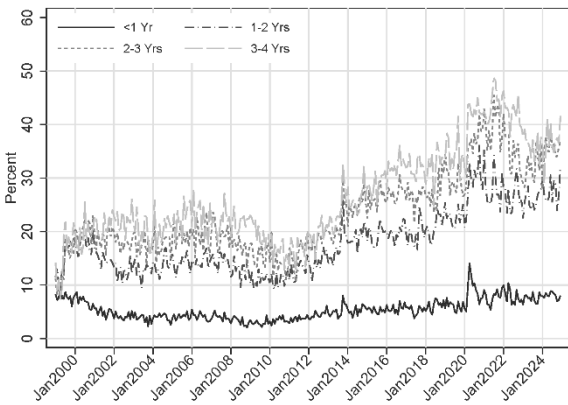


(d) Inflation



By Tenure Length

(e) Nonresponse



(f) Inflation

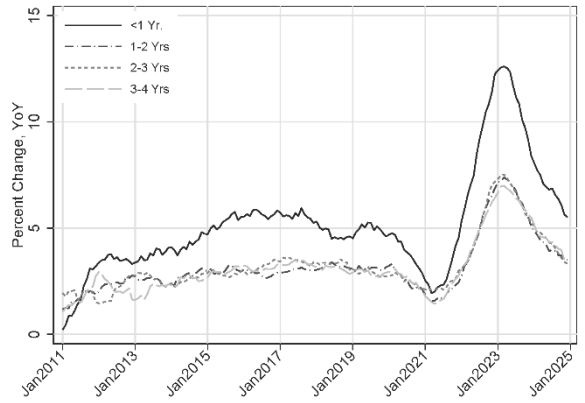


Figure 4. Non-Response and Inflation Rates By Category. *Note:* Inflation is calculated using all observations, collected and imputed, within a given category. *Source:* Authors' calculations using the BLS Housing Survey.

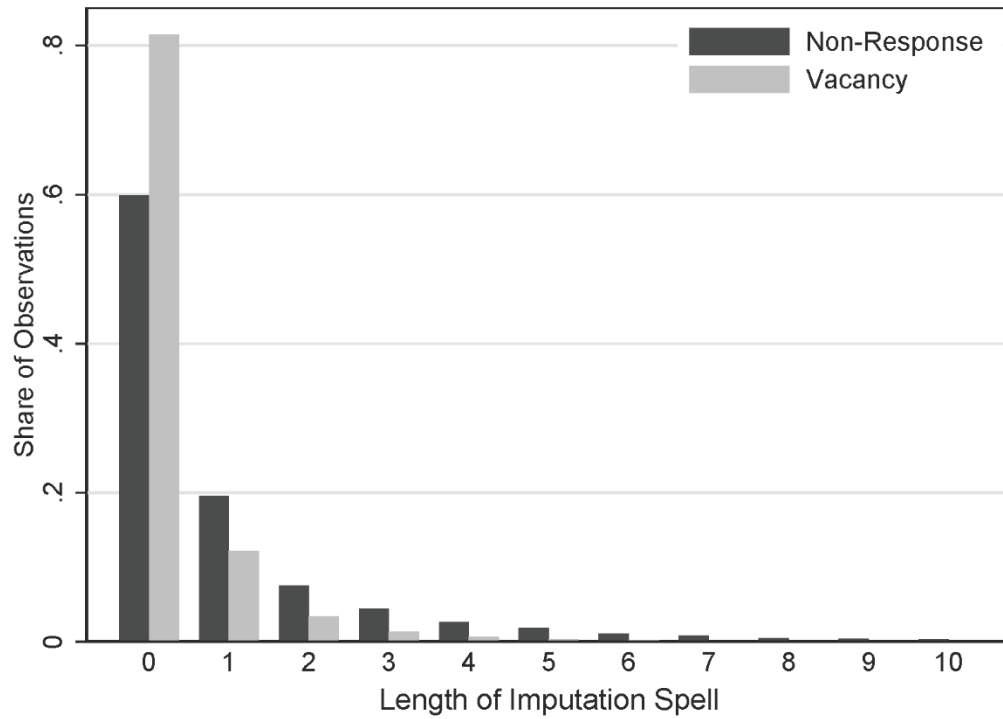


Figure 5. Length of Imputation Spells. *Note:* An imputation spells is defined as the number of consecutive survey periods that a housing unit is not surveyed. The x-axis denotes the number of surveys not collected. *Source:* Authors' calculations using the BLS Housing Survey.

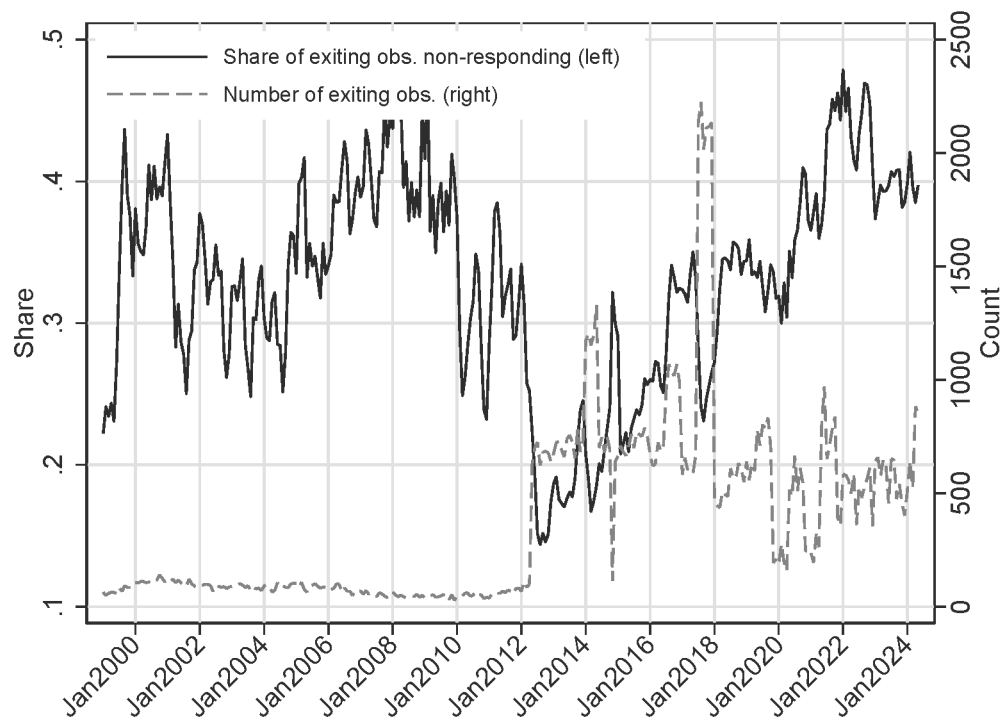


Figure 6. Non-response Share of Units. *Note:* A unit is defined as exiting if it is surveyed for the last time before January 2025. The weighted share of exits that are non-responses has been smoothed with a 3-month moving average. The unweighted count displays the total number of units exiting. *Source:* Authors' calculations using the BLS Housing Survey.

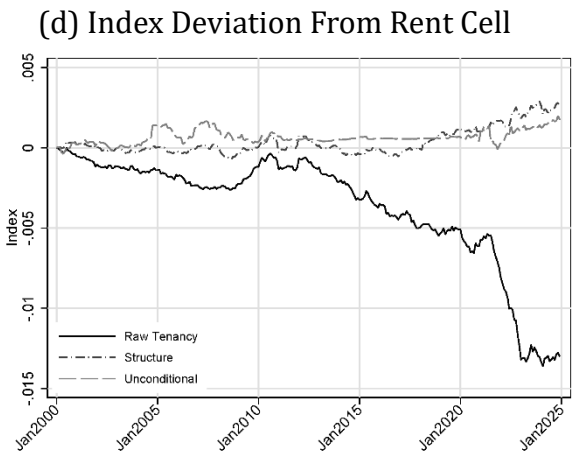
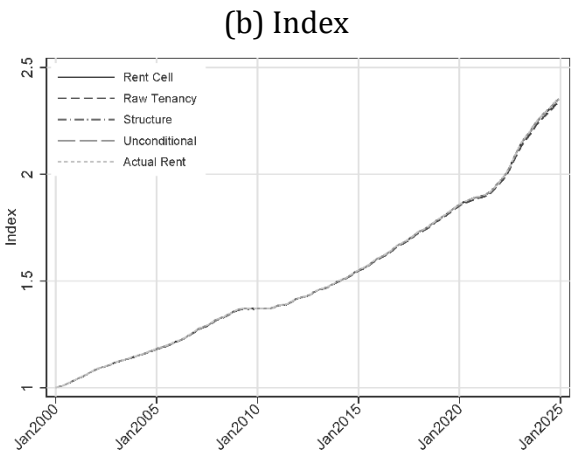
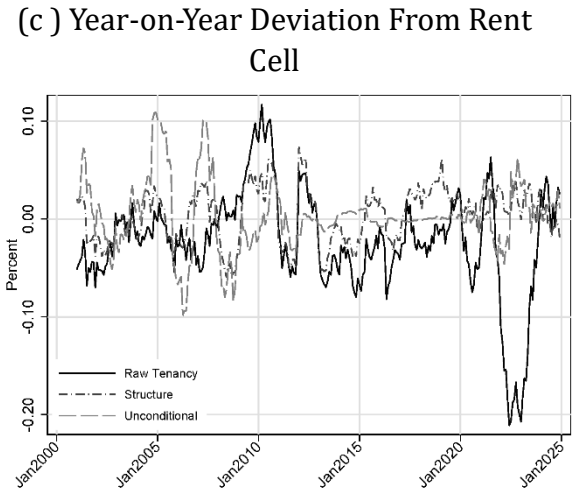
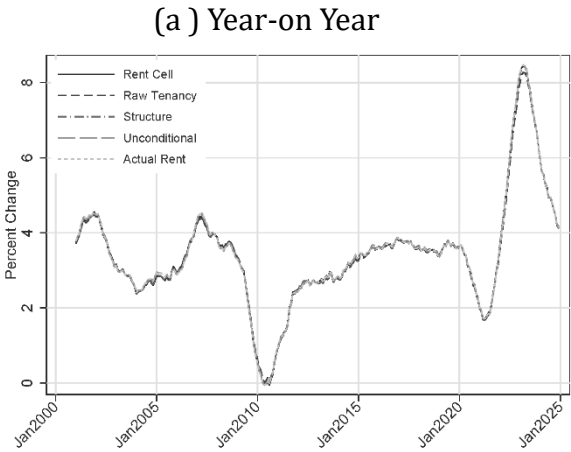
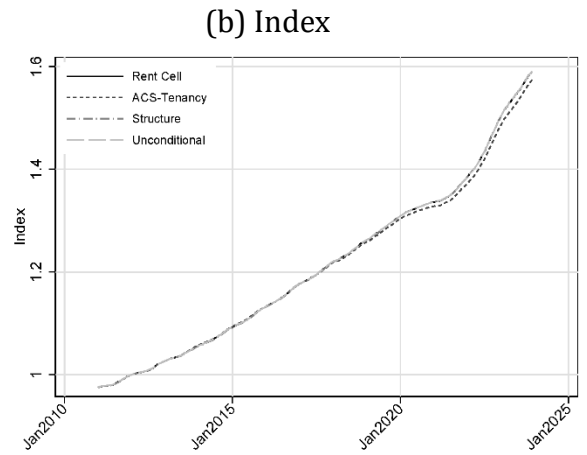
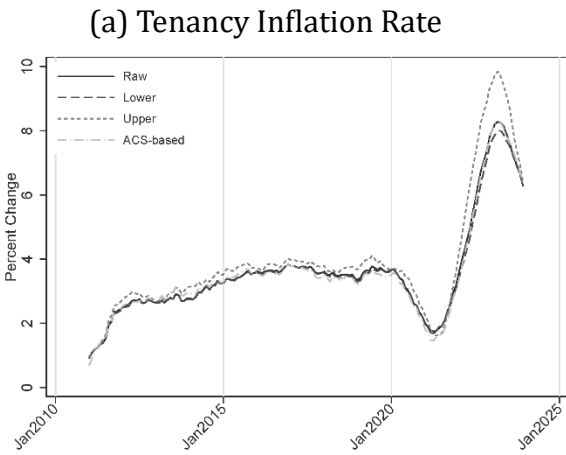


Figure 7. Alternative Imputation Inflation Rates. *Note:* The time series are created by using a class means method and picking different covariates to partition observations by. See text for more details. *Source:* Authors' calculations using the BLS Housing Survey.



(c) Inflation Rate Deviation From Rent Cell

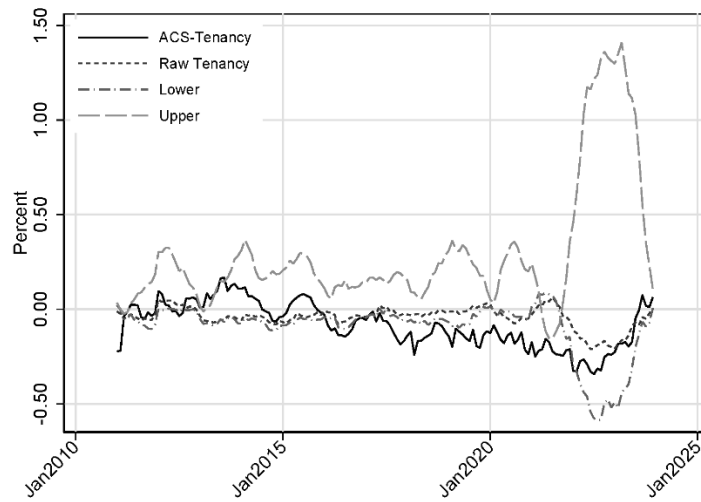


Figure 8. ACS-Tenancy Inflation Comparison *Note:* The time series are created by using a class means method using different measures of tenure length for nonresponses. *Source:* Authors' calculations using the BLS Housing Survey, ACS.

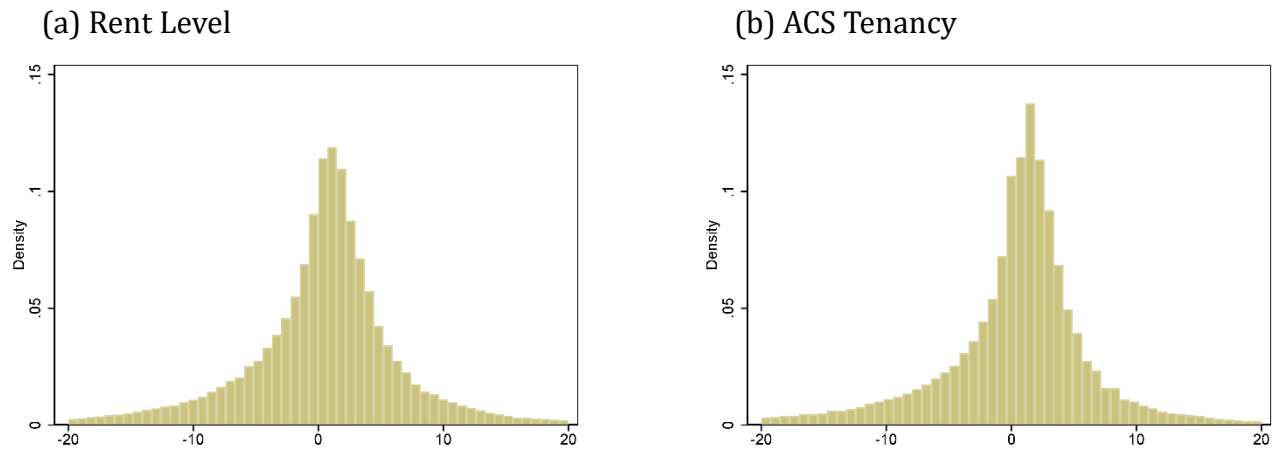


Figure 9. Jackknife Imputation Error by Imputation Method. *Note:* The imputation error is defined as the difference in percent between the imputed rent and collected rent. The histogram is truncated at -20% and 20%. Observations are unweighted. Observations are from January 2019 to December 2024. See text for more details. *Source:* Authors' calculations using the BLS Housing Survey.

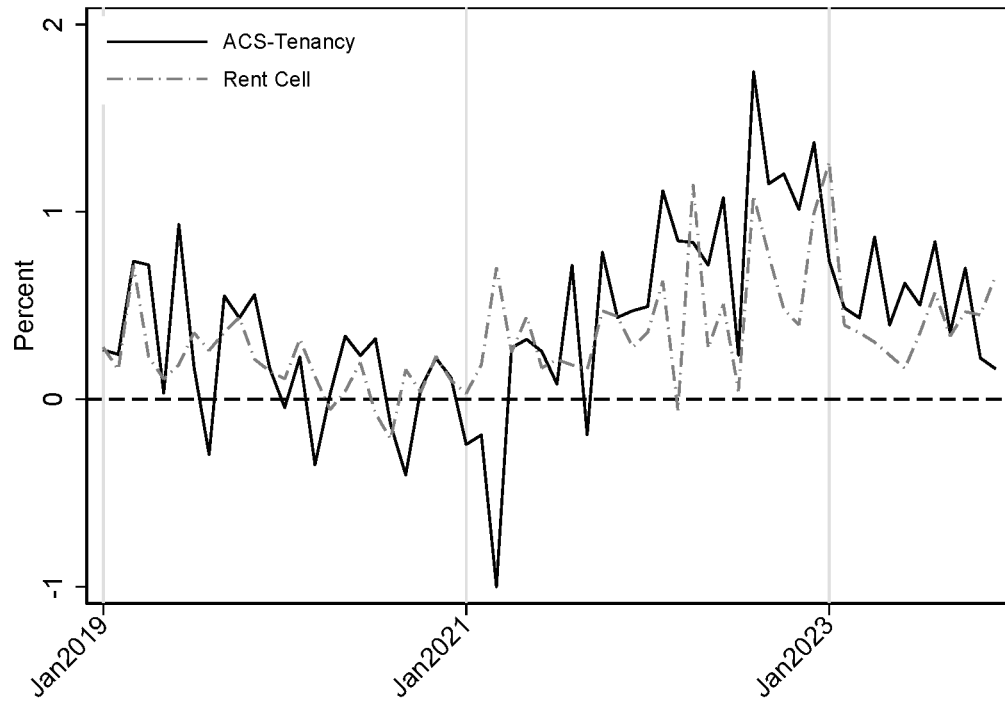


Figure 10. Jackknife Average Imputation Error. *Note:* The imputation error is defined as the difference in percent between the imputed rent and collected rent. Errors are averaged across cities using weights to obtain nationally representative series. *Source:* Authors' calculations using the BLS Housing Survey.

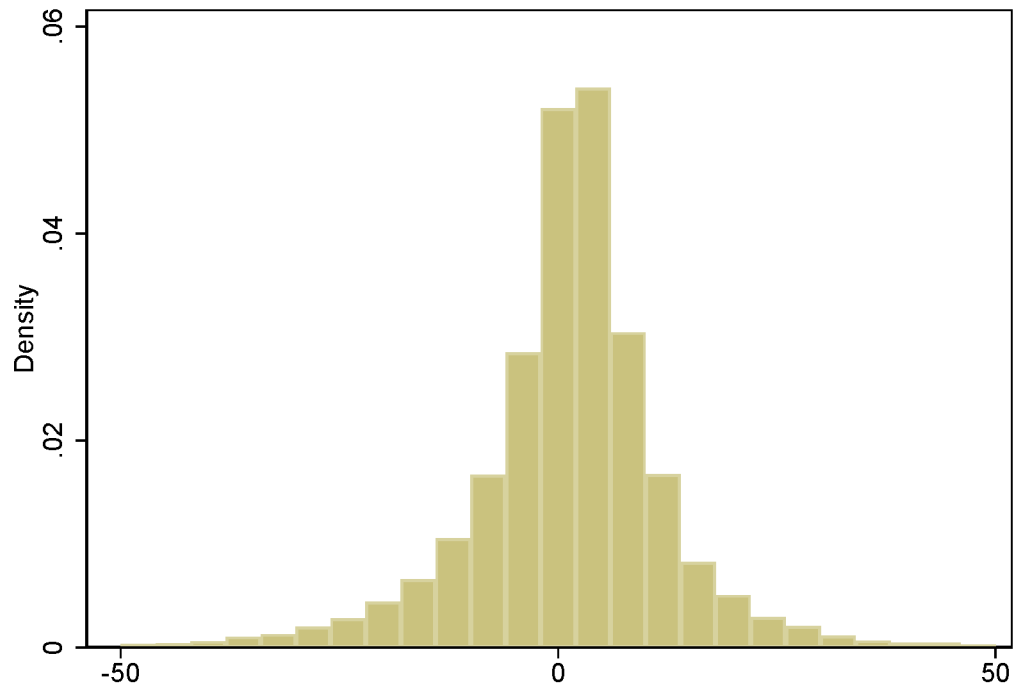


Figure 11. Re-entered Unit Imputation Error Histogram. *Note:* The imputation error is defined as the difference in percent between the imputed rent and collected rent. The histogram is truncated at -20% and 20%. Observations are unweighted. Observations are from January 2019 to December 2024. See text for more details. *Source:* Authors' calculations using the BLS Housing Survey.

Appendix

A A SIMPLIFIED MODEL OF RENTAL IMPUTATION BIAS

Model Setup

We develop a stylized model to illustrate the relationship between unit-level rent changes, imputation, nonresponse, and sample exit, and to demonstrate how unit-level imputation errors relate to aggregate index bias. The model highlights the key distinction between units that exit while in nonresponse, versus other units.

Our multiperiod model features heterogeneous tenure lengths, unit exit, and nonresponse dynamics that capture some of the key patterns observed in the data. The rental universe consists of equal proportions of apartment and detached units.²¹ Apartment tenants remain in their units for exactly 2 years, while detached tenants remain for 6 years. New tenant rents grow by 5% annually. During tenancy, apartment rents increase by 4% annually (producing small rent gaps), while detached rents increase by 1% annually (producing large and growing rent gaps). Upon turnover, apartment rents increase by 6% to eliminate the rent gap, while detached rents increase by 25%. Both unit types thus experience average annual rent growth of 5%.

Each period, 10% of units exit the sample independently of tenure or structure type; this captures the fact that in reality, almost all sample exit occurs via sample rotation, which is also independent of these characteristics. Nonresponse follows a two-state Markov process.

²¹ For simplicity, we reweight the sample every period so that despite sample exit, the total mass of apartment units is 1, and the total mass of detached unit is 1.

New tenants respond with 90% probability. Responding units transition to nonresponse with probability $1/3$, while nonresponding units transition back to response with probability 0.1. In steady state, this process generates substantial nonresponse, with response rates declining over tenure length.

The timing within each period is as follows: (1) nonresponse status is determined based on the previous period and transition probabilities; (2) the statistical agency collects rent data at mid-year, using actual values for responding units and imputed values for nonresponding units; (3) turnover occurs after data collection for units reaching the end of their tenure; and (4) 10% of units exit the sample. Under this timing, rent data for units in their final year of tenure reflects the old tenant's response behavior, with new tenant response rates affecting the following year's data collection.

We compare three imputation methods:

- Method 1 imputes a weighted average of the "vacancy jump rate" (the average rent increase among responding units experiencing turnover) and the continuing tenant rent change (the average change among continuing responding tenants), where the weight on the jump rate equals the true fraction of units receiving new tenants. This is similar to the ACS-based method studied in the main body of the paper.

- Method 2 stratifies by unit type, imputing the average rent change among responding units of the same type. This is similar to the structure-type imputation studied in the main body of the paper.

- Method 3 imputes the overall average rent change among all responding units to all nonresponding units. This is similar to the unconditional imputation studied in the main body of the paper.

A.1 Average Imputation Bias versus Index Bias

Response rates are central to understanding the results, not least because nonresponding units receive imputations, and imputations are estimated from responding units. Among apartments, half are new tenants who respond with 90% probability; the other half respond with 61% probability. Among detached units, one-sixth are new tenants who respond with 90% probability; response rates decline with tenure to 61% in year 2, 44.6% in year 3, 35.2% in year 4, 30% in year 5, and 27% in year 6. Overall, 62.3% of units respond in any given period, while 37.7% do not respond.

Under the specified parameters, Method 1 produces a vacancy jump rate of 12.1% and a continuing tenant change of 2.5%. Since 34.5% of units experience turnover, the imputed change applied to each nonresponding unit is 5.8%. Method 2 produces imputed changes of 4.8% for apartments and 7.1% for detached units, yielding a weighted average imputation of 6.4%. Method 3 produces an imputed change of 4.2%, the average among all responding units.

Since the true population rent change is 5%, all three methods generate average unit-level imputation errors: roughly +0.8 percentage points for Method 1, +1.4 percentage points for Method 2, and -0.8 percentage points for Method 3.

However, these unit-level errors do not translate directly into index bias. The key insight is that imputation errors are corrected when units return to response status. Only units that exit the sample while non-responding contribute permanent bias to the index. This mechanism both attenuates and alters the relationship between unit-level imputation accuracy and aggregate index accuracy.

To compute index bias, recall that the index uses actual rent changes for responding units and imputed changes for nonresponding units. In steady state, the index reflects a weighted average of these values across all units. For responding units, the weighted average rent change is 4.22% (reflecting a reweighted mass of 0.767 apartment units with an average 4.82% change and a reweighted mass of 0.480 detached units with an average 3.25% change; note that the detached average is reduced by the large proportion of continuing tenants experiencing a 1% rent increase). The mass of nonresponding apartment units is 0.233, and the mass of nonresponding detached units is 0.520, so that nonresponding units comprise 37.7% of the entire sample and receive imputed values. The impact of nonresponding units on the index varies by imputation method.

Under Method 1, the index growth rate is $(1.247 \times 4.22\% + 0.753 \times 5.8\%) / 2.0 = 4.81\%$, generating a bias of -0.19 percentage points. Under Method 2, the index growth rate is $(1.247 \times 4.22\% + 0.233 \times 4.8\% + 0.520 \times 7.1\%) / 2.0 = 5.04\%$, generating a bias of +0.04 percentage points. Under Method 3, the index growth rate is $(1.247 \times 4.22\% + 0.753 \times 4.22\%) / 2.0 = 4.22\%$, generating a bias of -0.78 percentage points. To achieve an unbiased index, the average nonresponse imputation would need to be 6.29%, associated with an average unit-level bias of +1.29 percentage points. The following table summarizes the results:

Table 2. Imputation Methods: Unit-Level Errors and Index Bias

Method	Unit-Level Error	Index Bias	Imputed Value(s)
Method 1	0.83%	-0.19%	5.8% (all units)
Method 2	1.38%	0.04%	4.8% (apt), 7.1% (detached)
Method 3	-0.78%	-0.78%	4.22% (all units)
Ideal Imputation	1.29%	0.00%	6.38% (all units)

A.2 Interpretation

This exercise demonstrates several important points about rental imputation. First, reasonable imputation methods can generate large average unit-level errors. All three methods produce average absolute errors of approximately 1 percentage point.

Second, the relationship between unit-level errors and index accuracy is indirect and subtle. Unit-level imputation errors do not translate mechanically into index bias, and minimizing unit-level imputation errors need not enhance index accuracy. In this exercise, the relationship between these two measures varies dramatically across methods. Method 1 exhibits a unit-level imputation error of +0.8 percentage points but an index bias of only 0.19 percentage points—note that the sign is reversed. Method 2 has the largest unit-level error (+1.4 percentage points) but the smallest index bias (+0.04 percentage points)—the magnitude is dramatically attenuated, with no change in sign. Method 3 demonstrates yet another pattern: its unit-level error of 0.78 percentage points translates directly into an index

bias of 0.78 percentage points with identical sign and magnitude. None of the methods produces an unbiased index, although the bias of Method 2 is quite small and that of Method 1 is modest.

What explains these patterns? Index errors stem from the interaction between nonresponse and exit on one hand, and imputation on the other. As noted earlier, a key insight is that imputation errors are corrected once a unit resumes responding, leaving only the selected subsample of exiting nonresponders to generate permanent bias. While exit itself is independent of tenure, the subset of units that are both nonresponding and exiting is comprised disproportionately of turnover units. Among apartments, the vast majority (39/49) of exiting-while-nonresponding units are in their second year and will experience turnover, while among detached units, 23% of exiting-while-nonresponding units are in their final year and will experience turnover. Hence, exiting nonresponders are disproportionately likely to experience large turnover-related rent increases that exceed the population average rent change.

While Method 1 overestimates rent changes for most nonresponding units (producing positive unit-level errors), it substantially underestimates changes for year-6 detached units experiencing 25% turnover increases (though the imputation error is only a modest underestimate for exiting, nonresponding apartment units). For this method, an increase in the unit-level imputation would enhance index accuracy.²²

²² The sign reversal also contradicts a simple heuristic that would link average imputation errors to index error (weighted by exit-in-nonresponse probability). The heuristic argument proceeds as follows: Imputation errors matter only for units that never return to response status; we can therefore focus on units that exit and assess error using the average imputation error. Exits occur after a nonresponse spell lasting 1 period, 2 periods, 3

Method 2 performs best in this model because stratification by unit type partially captures heterogeneity in rent changes, and ends up reducing errors for units experiencing turnover. By imputing 7.1% for detached units rather than 5.8%, Method 2 comes closer to the true average for nonresponding detached units, which include both many continuing tenants (1% increases) and some year-6 turnover units (25% increases). The resulting index bias of +0.04 percentage points is modest.

Method 3 produces the most concerning results, despite having a unit-level imputation error similar in magnitude to Method 1. By imputing the average of responding units (4.22%) to all nonresponders, Method 3 causes the index to exactly equal the responding units' average. The attenuation mechanism that partially protects Methods 1 and 2 from bias is absent. Since responding units experience systematically lower rent growth than the population due to selection, Method 3's index bias of 0.78 percentage points is substantial.

Third, since index errors depend on the proportion of units that exit while in nonresponse status, and imputation errors for these units, index errors are likely increasing in both nonresponse rates and exit rates. Nonresponse rates have increased markedly over the past decade or two. Prior to the introduction of a rotating rental unit sample, exit rates from the

periods, etc. Suppose the average imputation error is +1%. Consider a unit that did not respond last period and does not respond this period. Its rent is imputed too high this period but was also imputed too high last period; these errors roughly cancel. Only the imputation error in the first nonresponse period matters. Hence, once a unit receives a faulty imputation of +1%, this error persists in the index. The overall index error would then equal the average imputation error multiplied by the overall probability that a unit exits while nonresponding. This argument can explain how an index has less bias than the average imputation error, but would suggest that the average unit-level imputation error is the key determinant, and cannot explain a sign reversal.

BLS sample were far smaller (and generally were treated with a vacancy imputation), limiting the scope for the introduction of index bias.

These results underscore the importance of response rates and of understanding the selection process governing which units become and remain nonresponders. They also demonstrate that minimizing unit-level imputation errors is not the primary objective. The critical challenge is not minimizing unit-level imputation errors across all nonresponders, but rather accurately capturing the rent change distribution among nonresponders that exit. When nonresponse correlates with tenure and tenure correlates with rent gap accumulation, imputation methods that ignore these interactions may generate substantial bias even when unit-level errors appear modest. Conversely, methods with larger unit-level errors may produce better index estimates if they avoid the systematic selection bias inherent in other approaches.

An important difference between this model and the real world is that, in the model, we know the data generating process of non-response (and how it interacts with tenant turnover and rent-setting), so in this context it is possible to precisely impute rents for nonrespondents. Conversely, in the real world the challenge is that we do not know this data generating process.²³ Owing to the paucity of information about exiting non-responders (the set of tenants about which the BLS collects no information), deducing index bias is challenging. In the present exercise, different imputation methods yielded notably different index growth rates. In the main body of the paper, we present evidence that different imputations methods yield similar index growth rates; this provides reason to hope that index bias may be limited in practice.

²³ We are grateful to our discussant, Boaz Abramson, for pointing this out.

B ROBUSTNESS OF PHILLIPS CURVE ESTIMATES ACROSS IMPUTATION METHODS

To assess whether differences in imputation methods affect macroeconomic inference, we estimate Phillips curve regressions using 12-month shelter inflation series constructed under each of five imputation methods: rent cell, rent level, structure type, unconditional, and ACS-based. We consider both a standard linear specification and a nonlinear (frequency-dependent) specification, following (Ashley & Verbrugge, 2025). The labor market variable is the unemployment gap, measured as the permanent component of the unemployment rate minus the CBO natural rate (itself adjusted to match the difference in means), following (Verbrugge & Zaman, 2024). We estimate the model over the pre-COVID period and over the full sample; conclusions are unchanged.

The baseline linear specification regresses the shelter inflation variable on the unemployment gap and on two annual lags of shelter inflation:

$$\pi_{k,t}^{t+12,t} = \beta_0 + \beta_1 ungap_t + \beta_2 \pi_{k,t-12}^{t+12,t} + \beta_3 \pi_{k,t-24}^{t+12,t} + \varepsilon_t$$

The nonlinear specifications extend this specification by decomposing the unemployment gap by frequency and, in one specification, allow each frequency gap term to enter asymmetrically. Here we provide more detailed results pertaining to two specifications, the linear specification (over the 2005-2019 period) and the most highly nonlinear specification (over the full sample).

Linear Specification: For the 2005-2019 sample, the largest percentage gap between β_1 estimates was 3.4%, ranging from -0.433 to -0.441. For β_2 , the largest gap was 7.4% (0.276 to 0.298), while for β_3 there was a 4.0% gap (-0.293 to -0.305). All coefficients remained highly significant across methods, with t-statistics varying by less than 10

Nonlinear Specification: This is estimated over the full 2005-2024 sample. In this specification, there are five more coefficients to estimate, and we might expect that minor variations between series would be highlighted by the frequency decomposition. Yet coefficient stability for this specification was even more striking. The biggest coefficient differential was for $\beta_1^{low,negative}$, the coefficient on the negative part of the low-frequency unemployment gap. But variation was a mere 4.4% (-2.286 to -2.389). R^2 values differed by at most 1.4% (0.804 to 0.815), and standard error estimates by 3.0% (0.691 to 0.712).

Across all specifications and sample periods examined, no coefficient estimate varied by more than 18% across imputation methods, and the vast majority varied by less than 5%. Statistical significance patterns were entirely consistent across methods. This high degree of stability demonstrates that the choice among these imputation approaches does not materially affect macroeconomic inference regarding shelter inflation dynamics.

C REGRESSION-BASED IMPUTATION

The number of observations required for a rigorous group means imputation grows linearly in the number of levels of a variable and exponentially in the number of variables. As noted in the main body, although the BLS Housing Survey follows around 40,000 housing units, a significant minority of these units are not collected, and the collected units are divided between over 30 cities and 6 panels. Even the official imputation procedure, computing class means from three relative rent levels, requires careful rules for cell collapsing owing to the frequent event that there are insufficient collected observations for a city-date-cell. As discussed in greater detail in the text, these sample-size restrictions prohibit any group means imputation with multiple variables or finer categories.

We test two regression-based approaches that permit us to include additional variables, at the cost of a stricter functional form (and in one case, broadening the sample). First, we follow (Erickson & Pakes, 2011) in regressing the change in log rents on two covariates: tenure and structure type. Second, we estimate a much richer regression-based approach that mimics the depreciation adjustments that the BLS performs. We study two variants of each approach.

Following (Erickson & Pakes, 2011), for each city c and time period t , we regress

$$\frac{\ln(\text{rent}_{i,c,t})}{\ln(\text{rent}_{i,c,t-6})} = \alpha_{c,t} + \beta_{c,t}T_{i,c,t} + \sum_{j=0}^2 \gamma_{j,c,t}S_{i,c,t} + \epsilon_{i,t}$$

where $T_{i,c,t} = 1[\text{tenure}_{i,c,t} < 6\text{months}]$ and $S_{i,c,t}$ denotes the structure type of the housing unit. We estimate this regression using housing units that were successfully sampled at time t with a valid *rent*; this excludes vacancies and nonresponses. We use the estimated coefficients $\widehat{\beta}_{c,t}$ and $\widehat{\gamma}_{j,c,t}$ to predict the change in log-rents for all nonresponding units, and impute a rent on this basis.²⁴ We estimate the regressions iteratively over the entire date range. Since tenure is unknown for all nonresponding units, we use the ACS data to impute a tenure for each such unit within a city-date, in parallel with the ACS-based method studied in the main body. Following the imputation, we aggregate the housing units using the weighted rent index formula described above to calculate a counterfactual national rent index. Inspired by (Erickson & Pakes, 2011), we test two variants: one where we only impute rents for

²⁴ Note that a unit may be collected at time t but a nonresponder at time $t-6$. In this case, its imputed rent is used for the $t-6$ observation, in parallel with all the methods considered in this paper.

nonresponders, and one where we impute rents for all units, both nonresponders and collected units.

Figure 12a compares the two (Erickson & Pakes, 2011) imputation inflation rates to the rent cell and ACS-tenancy approach. In keeping with our previous results, we find that these imputation methods do not produce substantially different results—although imputing rents for all units tends to yield a higher inflation rate. The two regression-based indexes appear to lag the rent cell index slightly. A potential explanation is that we use lagged ACS data, corresponding to the data available at production time, to impute tenure for non-responses. The lagged tenure imputations could generate a slightly sluggish inflation rate in a regression-based approach.

Our second approach is to follow the age bias adjustment calculations that the BLS routinely performs to address depreciation in housing unit quality related to aging. Our primary source for the methodology is (Houck, 2023). In particular, we regress change in log rent on the list of variables displayed in Table 2 in the article. These variables include housing unit-level attributes, such as age, structure type, and utility type, as well as geographic attributes, such as zip code income and the share of the block group’s population with at least some college. We follow the procedure for imputing the age of buildings for observations with no year-built information, and we impute the new tenant status for nonresponses as above.²⁵ We restrict our replication to 2020-2024 period, and use time-invariant 2025 Census Tract ACS data as well as 2022 IRS income zip-code-level data. We perform this regression for each Census region and period, with city-level fixed effects to capture geographic heterogeneity.

²⁵ We omit the length of occupancy variable, because we are unable to impute it for nonresponses.

To ensure that we have sufficient observations for each period, we perform the age-bias-style hedonic regression at a semi-annual frequency instead of a monthly frequency.

Figure 12b compares the rent inflation rate from the official Rent index against the indexes derived from two age-bias regression imputations. “PSU Rent” runs the semi-annual Census region regression with only city fixed effects, to provide a benchmark. This PSU Rent series should, in theory, more-or-less mimic the unconditional imputation in the main body of the paper, except using a regression-based approach instead of a group means one. The series “Chain Diff Rent” uses the full age bias regression and updates current and lagged nonresponding rents. We find that the age bias-style imputation produces noticeable differences in the overall rent inflation rate, but these differences remain of the same magnitude as those between the ACS-tenancy imputation and the official imputation.

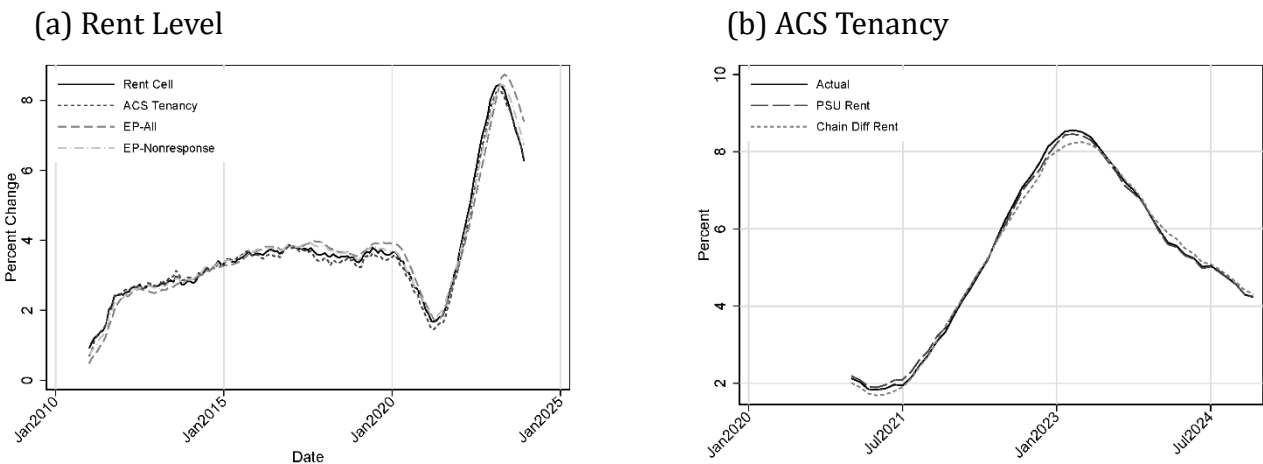


Figure 12. Erickson-Pakes and Age Bias Regression Imputations. *Note:* The Erickson-Pakes imputation uses a regression-based approach that incorporates tenure and structure type. The age bias approach includes a wider array of covariates in the regression. See text for more details. *Source:* Authors’ calculations using the BLS Housing Survey.