

NBER WORKING PAPER SERIES

NONRESPONSE IMPUTATIONS AND RELATED MEASUREMENT ISSUES IN
THE CPI FOR SHELTER

Lara Loewenstein
Hugh G. Montag
Randal Verbrugge

Working Paper 35250
<http://www.nber.org/papers/w35250>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2026

We thank our discussant, Boaz Abramson, for his very helpful comments. We also want to extend our appreciation to the participants at the 2026 CRIW Conference on Measurement of Housing and the Housing Sector for their feedback and suggestions. 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, the BLS, or the National Bureau of Economic Research. 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.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2026 by Lara Loewenstein, Hugh G. Montag, and Randal Verbrugge. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Nonresponse Imputations and Related Measurement Issues in the CPI for Shelter
Lara Loewenstein, Hugh G. Montag, and Randal Verbrugge
NBER Working Paper No. 35250
May 2026
JEL No. C81, E31, R31

ABSTRACT

Shelter is the largest component of US consumer price index (CPI) inflation; therefore, the accuracy of shelter inflation is critical for the accuracy of overall CPI inflation. Nonresponse in the BLS Housing Survey, which underpins the measurement of CPI shelter inflation, has increased since 2000 and now represents roughly 40 percent of total observations. Missing rent data are currently imputed using a class-mean approach based on rent tier, potentially resulting in biased imputations, as we find that nonresponse is correlated with factors beyond rent tier. We study alternative simple imputation methods based on variables correlated with both nonresponse and rent growth, including structure type and tenure length. A simple model demonstrates that alternative methods could yield sharply different index biases. However, in practice, we find that these alternative methods yield similar shelter inflation indexes, suggesting that any index bias may be modest.

Lara Loewenstein
Federal Reserve Bank of Cleveland
lara.loewenstein@clev.frb.org

Randal Verbrugge
Federal Reserve Bank of Cleveland
Randal.Verbrugge@clev.frb.org

Hugh G. Montag
Bureau of Labor Statistics
Montag.Hugh@bls.gov

1 INTRODUCTION

Accurately measuring rental prices is essential to the federal government’s construction of price indices and, in turn, for measuring inflation. To this end, the BLS Housing Survey provides the underlying data for CPI tenant rent and owner-equivalent rent, which together account for over 90 percent of CPI shelter inflation. Shelter represents more than 30 percent of headline CPI inflation and nearly 20 percent of personal consumption expenditures (PCE) inflation and comprises even larger shares in their respective core inflation measures, which exclude food and energy. As a result, even minor inaccuracies in the construction of shelter indices can translate into economically meaningful distortions in overall inflation.

The fact that nearly 40 percent of surveyed housing units in the BLS Housing Survey are currently imputed—and that the majority of these are due to nonresponse as opposed to unit vacancy—is concerning. This concern is amplified by the fact that the imputation rate has been increasing over time, driven by an increase in nonresponse (see Figure 1).

If rent changes for imputed housing units are the same as those for non-imputed housing units, then the imputation procedures will have a minimal effect on the aggregate inflation rate. However, we find that the inflation rates do vary by whether units are imputed or not, as shown in Figure 2. Moreover, the imputed inflation is quite volatile and exerts a noticeable effect on the overall inflation rate.

For the CPI shelter index to be accurate, the only factor that matters is the accuracy of the *average rent change* estimate. It is of less concern in our context that the imputation method maintains the variance of the underlying data. Some imputation methods, such as regression imputation, could result in underestimates of variance, but this is not relevant to the construction of the index. In other statistical contexts, such as regression analysis, preserving the distribution of the underlying data through methods like stochastic regression imputation is a more important consideration. Since variance maintenance is of secondary importance in our context, we focus mostly on class-mean imputation methods. Using such methods is a common practice in inflation measurement as documented in Graf (2020) and the alternative class-mean methods are similar enough to existing practice that their implementation is feasible.

We make several contributions. First, we highlight an important issue in the measurement of CPI shelter inflation: the growing problem with survey nonresponse, which results in a smaller effective sample size. The BLS Housing Survey is not the only government survey affected by nonresponse and other forms of missing data; other surveys, such as the Current Population Survey and the Consumer Expenditure Survey, face similar challenges. In Figure 3, we compare household survey response rates across these surveys.

Second, we analyze nonresponse in the BLS Housing Survey. Using unit and tenant char-

acteristics available in the data, we assess whether nonresponse is correlated with rent level, structure type, and length of tenure. We find that nonresponse rates are very slightly higher for higher-rent units, substantially lower for multifamily buildings relative to single-family units, and possibly much lower for newly-moved-in tenants relative to continuing tenants. However, nonresponse leaves the tenant’s move-in status indeterminate; there is hence considerable uncertainty regarding the tenure status of nonresponding 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 nonrespondents.

Third, we detail the imputation method currently used by the BLS. It is a class-mean imputation approach based exclusively on three variables: city, month, and tercile of a unit’s lagged rent level. Our analysis shows that the rent level exhibits weak correlations with both nonresponse patterns and rental inflation rates. In particular, the existing method fails to account for structure type and tenancy length, which are the two covariates that we identify as most strongly correlated with nonresponse and, most importantly, with rent growth. Our evaluation demonstrates that this approach results in an upward bias in unit-level inflation estimates because the method overweights, relative to the population, the large rent increases associated with a new tenancy.

Fourth, we assess other class-mean imputation strategies, including those that use tenure length and structure type, to determine whether alternative covariates can reduce bias. As already noted, the exclusion of tenancy length in the imputation procedure can lead to upward-biased imputations in practice. However, for the current CPI production cycle, tenancy length for nonrespondents is not known at the time of imputation, though it may be determined at subsequent surveys. At one extreme, all nonrespondents might be new tenants; at the other extreme, all might be continuing tenants. The degree of imputation bias—and consequently, the measured inflation rate—depends on which of these extremes more accurately reflects the actual composition of nonrespondents. Using the American Community Survey (ACS), we propose a probabilistic method for determining the average share of nonrespondents that are new tenants. Our results show that the bias implied by our ACS-based estimates is much smaller than its theoretical maximum. The implied year-on-year inflation rate based on this method is up to 25 basis points below that of the current method.

Fifth, we examine the relationship between unit-level imputation errors and index accuracy (also measured as inflation accuracy). In Appendix A, we present a simple model that captures many features of the BLS Housing Survey. We use the model to demonstrate that minimizing unit-level imputation errors does not necessarily lead to index accuracy. Instead, what is critical is accurately estimating the average rent change among nonresponding units *that exit the sample*. The current rotating sample selection procedure results in one-sixth of

the sample exiting annually. Inaccurately imputing nonresponses among these units could introduce substantial bias into the shelter price indices. However, because the relevant imputation errors concern nonresponding exiting units, for which data are unavailable, a full quantitative assessment of index accuracy presents challenges that we reserve for future research. Nevertheless, we observe that our various imputation methods produce similar rental inflation rates. This similarity indicates that only minimal bias is introduced in the shelter indices.

Finally, we provide evidence on how much the imputation process contributes to noise in the overall shelter index. While the previous contributions focus on imputation bias, the imputation process can also introduce measurement noise—random fluctuations that increase the volatility of the published indices without systematically biasing them in one direction. We assess the magnitude of this imputation-driven volatility and its impact on the reliability of month-to-month shelter inflation measurements.

We contribute to multiple strands of the literature. First, we contribute to research on missing data and measurement in official statistics, including Hokayem et al. (2015), who show that income nonresponse in the Current Population Survey (CPS) March Supplement induces a 1 percentage point bias in official US poverty rates, and Heffetz and Reeves (2021), who find that measurement issues, including nonresponse, may have generated a bias of up to 1.5 percentage points positively or negatively in the unemployment rate during the COVID-19 pandemic. Second, Groves and Peytcheva (2008) provide a meta-analysis of methodological studies examining the conditions under which nonresponse bias emerges and when it is quantitatively meaningful. Third, we contribute to the literature on imputation in inflation measurement; while vacancy imputation in the BLS Housing Survey has garnered much attention (for example, see Crone et al. (2004)), nonresponse imputation is currently a far more frequent occurrence and is the focus of this paper.

2 DATA

Our main dataset is the BLS Housing Survey. The housing survey underpins both the rent and owners' equivalent rent (OER) 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 6 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.¹ 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 responsibilities could change over time.² Likewise, some tenants perform some in-kind labor for their lease, an agreement that may 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.³

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 that is collected is the date the tenant moved into the housing unit, from which housing tenure is derived.

An important innovation related to sample exit occurred in 2012. Prior to 2012, the BLS surveyed the same set of housing units indefinitely.⁴ 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 nonresponse and imputation rates post-2012. As we discuss below, 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, nonresponse, 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

¹Likewise, within a city, each panel is representative of all housing units within the city.

²In 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.

³See the factsheet on OER and rent for more details.

⁴Prior to 2010, the BLS routinely added additional housing units to the sample, with sampling based upon the decennial Census data. In 2010, the CPI initiated a plan that would continuously update the sample of rented housing units. Eventually, one-sixth of the rented housing unit sample would be replaced every year, with sample updating based on the latest available US 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 Ptacek (2013), and the section on CPI design in the BLS Handbook of methods.

unit is a nonresponse if no one could be contacted, or if the respondent refused to fill out the survey.⁵ 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, two different imputation procedures 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, move-in 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 change; see Adams and Verbrugge (2025)) is the structure type of the housing unit.⁶ 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 units 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 core-based statistical area (CBSA).

⁵Survey respondents can be tenants, landlords, or property managers. In practice, the field economist attempts to contact the same respondent every survey period.

⁶Indeed, Adams and Verbrugge (2025) show 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 6 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 the BLS 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 nonresponse 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 occur when a new tenant moves 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 units in city c , tenancy length cell v , and date t . The imputed rent change in percent is

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

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

$$rent_{i,t}^* = VIM_{c,v,t} rent_{i,t-6}$$

Because vacancy inflation rates are typically much higher than continuing-tenant inflation rates, their implicit relative weight in the class means the 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.⁷ Since housing units are sampled every 6 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 6 month price relative for each city and convert it to a one-month price relative. For city c and time t , the 6 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.⁸ 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.⁹ In practice, when we construct inflation rates using subsets of the housing units (e.g., only for 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.¹⁰ 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

⁷The 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.

⁸We abstract from the aging bias adjustment applied to the denominator.

⁹See <https://www.bls.gov/opub/hom/cpi/calculation.htm> for additional details about the CPI rent index construction.

¹⁰For the definition of missing at random and a discussion of imputation theory, see van Buuren (2018).

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 Section 3 and reflects the tercile of the lagged rent of the unit relative to the lagged rent of 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. (2025) note, 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 structure types into multifamily (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 nonresponse 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 are missing at random with respect to these variables and 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. Figures 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 on 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 and Verbrugge 2025; Gallin et al. 2025), and now accounts for this in its calculation of OER. As can be seen in Figure 4d, multifamily units (over this period) experienced a systematically higher inflation rate than single-family detached units. Differences between single-family detached units and small multifamily units are less stark, especially earlier in the sample. Some of these differences may be explained by location, but not all (Adams and Verbrugge 2025). Although large multifamily units feature higher turnover (and thus more frequent rent resets toward new-tenant rents: see Gallin et al. (2025)), 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.¹¹

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 nonresponse rates and tenancy lengths over time.

The pattern of nonresponse by tenure length could partly be an artifact of how the move-in dates are carried forward. Suppose that there is a given rental unit that 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

¹¹Conversely, a higher proportion of the rent on detached units is related to a non-depreciating asset, land.

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. (2025)), 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 rent change differential, whether implicitly or explicitly, seems likely to result in biased imputations.

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

5 NONRESPONSE 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 nonresponse 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 nonresponse 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 nonresponse imputations are part of a longer nonresponse spell. Putting this differently, compared to vacancy, nonresponse is much more likely to be followed by another nonresponse.

Over extended periods, the choice of imputation method 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: $t = 0, 1, 2$.¹² 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 is $rel_{s,t} = \frac{r_s}{r_t}$.

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: $\frac{\tilde{r}_1}{r_0} \neq \frac{r_1}{r_0}$. 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 $rel_{1,2}$:

¹²For simplicity, we examine a single unit and abstract from the 6 month panel structure.

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

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 nonresponse 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 that 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 sample unless they transitioned to owner-occupancy. However, over this earlier period, a relatively large fraction of these units exited during a nonresponse spell. This is partly because nonresponse 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 nonresponses 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.

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, 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. Although it would be ideal if field economists could assess whether the same tenant

inhabits a housing unit, even if the tenant does not 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.

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 do not respond for multiple surveys. The move-in carryforward assigns these units the same tenant throughout their entire nonresponse 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 that of the raw move-in dates 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 the 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 6 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.¹³ 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 the ACS except for $x_{c,t}^{Nonresponse}$, the share of new tenants in nonresponses. In theory, $x_{c,t}^{Nonresponse} \in [0, 1]$, although in practice $x_{c,t}^{Nonresponse}$ can lie outside of these bounds for cities

¹³We ignore vacant units in this derivation.

and dates where the share of nonresponses 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 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 nonresponders, 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 multifamily 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 were 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. 2025). In the raw tenancy case, we aggregate new-tenant and continuing-tenant cells 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 $x_{c,t}^{Nonresponse}$. 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.¹⁴

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 that obtained with the unconditional approach. This is somewhat surprising, given that rent changes vary noticeably by rent level and structure 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 the 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 percentage points lower than the current rent cell (rent tier) and the unconditional methodologies for most periods. The effect rises to 0.2 percentage points 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 percentage points less than the current rent cell method. In Figures 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 is disproportionately made up of tenants with longer measured tenures. Those

¹⁴This represents a small departure from current BLS methods.

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 “nonresponses” 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 measure assumes a high rate of tenant turnover, and avoids combining new and continuing-tenant cells in the case of insufficient observations, instead choosing to use additional months of data. In contrast, the lower-bound tenancy imputation makes assumptions about tenancy status similar to those of 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 publication of ACS data.

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 Appendix A, we develop a stylized model showing that the choice of 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, the choice of imputation method 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 and Pakes (2011) regression imputation that uses both tenure and structure type. The second is a richer nonresponse 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 Appendix 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 measuring aggregate inflation. 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.¹⁵ We then impute rent $x_{i,c,t}^I$ 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.¹⁶

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.

In Figure 9a we include 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

¹⁵As the rent level is defined using lagged rent, we do not need to be concerned about redefining the rent levels.

¹⁶A jackknife procedure in this context may produce potentially unreliable results if the cause of nonresponse is significantly correlated with changes in rent 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 changes in rent 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. 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 C generate indexes that are very similar to those derived from the class mean methods.

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 percent. The average rent-level cell error is slightly higher, at 0.20 percent, also statistically significant. Both imputation methods yield notable *median* errors of 1.03 percent and 0.77 percent, 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 percent and 0.34 percent 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. It does illustrate that there is substantial heterogeneity in rent changes that neither method is able to capture.

The imputation errors for units that resume responding after a period of nonresponse display similar behavior, although at a greater scale. Figure 11 depicts a histogram of rent-level-cell-based imputation errors for these units. These errors possess a similar rightward drift. The median error is 1.83 percent and the mean error is 1.17 percent. The latter is statistically significant. However, two issues 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.¹⁷

9 CONCLUSION

In this paper, we investigate nonresponse imputation in the BLS Housing Survey, the source of data underlying CPI shelter inflation. This topic has grown increasingly important as nonresponse rates have risen steadily over the past decade. We document substantial heterogeneity in survey response rates across characteristics collected in the survey. Because inflation rates also differ across these characteristics, the choice of imputation method may materially affect measured inflation. This concern is reinforced by a stylized but reasonably realistic model presented in Appendix A, which shows that different imputation approaches can generate economically meaningful index biases under plausible data-generating processes.

We study several simple imputation methods. Both the current rent-level cell imputation approach and an unconditional imputation—which applies the average rent growth of all responding units—produce upward-biased imputations if response rates are higher for new

¹⁷Indeed, this conjecture may explain some of the positive unit-level rent change outliers in our data more generally.

tenants, who tend to experience much larger rent changes.¹⁸ These two approaches yield very similar indexes. Imputing by tenancy status reduces measured inflation. However, it presents challenges since information such as tenancy length cannot be collected when there is no response. Therefore, we develop an approach that estimates the average share of new tenants among nonrespondents using external survey data, circumventing the need for unit-level tenancy measurement. All methods have statistically significant unit-level imputation errors. Yet, as our Appendix model demonstrates, such errors need not be translated automatically into index bias because the index depends on aggregate averages, not individual unit accuracy. The relationship between unit-level imputation accuracy and aggregate index accuracy is indirect and subtle. Our central empirical finding is that the five imputation methods produce remarkably similar shelter inflation indexes. This similarity indicates that the current imputation procedure is not inducing substantial bias into published shelter indexes in practice; if it were, the different methods would produce divergent results.

There are several directions for future research. First, imputations for vacant units warrant deeper investigation. Although vacancies constitute a small share of the survey, the current procedure assumes that newly vacant units should receive a rent change associated with a new tenant, and therefore these units receive large imputed rent changes that can substantially influence the rent index. This sharp increase in rent upon a new tenant occupying the unit should depend upon the tenancy length of the previous tenant, since longer-tenured tenants typically have larger rent gaps relative to market rates, whereas the current imputation method abstracts from any such considerations. Second, the current nonresponse imputation procedure may generate a larger upward bias for OER compared to the rent index. 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 and Verbrugge (2025). Given that detached units now carry greater weight in the OER index, imputing by structure type could result in substantially different inflation rates, particularly in recent years. Third, we have only begun the exploration of hedonic modeling (see Appendix C), which would allow tenure, structure type, and rent level to be controlled for simultaneously through regression, potentially improving the accuracy of unit-level imputation. Also, machine learning methods offer new avenues for imputation. In short, assessing price index accuracy remains challenging, but important work.

¹⁸Having 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.

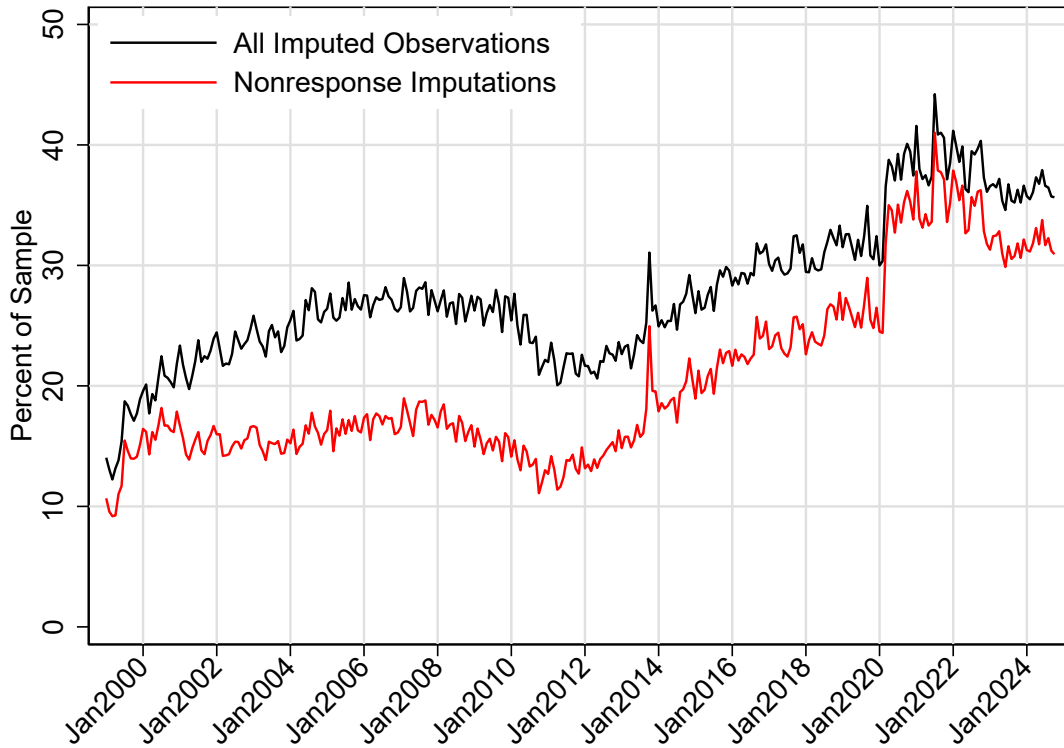
References

- Adams, Brian, Lara Loewenstein, Hugh Montag, and Randal Verbrugge (2024). “Disentangling rent index differences: Data, methods, and scope.” *American Economic Review: Insights*, 6(2), p. 230–45. doi:10.1257/aeri.20220685.
- Adams, Brian and Randal Verbrugge (2025). “Location, location, structure type: Rent divergence within neighborhoods.” *Journal of Housing Economics*, 69, p. 102,081. doi:10.1016/j.jhe.2025.102081.
- Ashley, Richard and Randal Verbrugge (2025). “The intermittent Phillips curve: Finding a stable (but persistence-dependent) Phillips curve model specification.” *Economic Inquiry*, 63(3), pp. 926–944. doi:10.1111/ecin.13281.
- Crone, Theodore M, Leonard I. Nakamura, and Richard Voith (2004). “Hedonic estimates of the cost of housing services: Rental and owner-occupied units.” *Federal Reserve Bank of Philadelphia Working Paper 04-22*. doi:10.21799/frbp.wp.2004.22.
- Crone, Theodore M, Leonard I Nakamura, and Richard Voith (2010). “Rents have been rising, not falling, in the postwar period.” *The Review of Economics and Statistics*, 92(3), pp. 628–642. doi:10.1162/REST_a_00015.
- Erickson, Tim and Ariel Pakes (2011). “An experimental component index for the CPI: From annual computer data to monthly data on other goods.” *American Economic Review*, 101(5), p. 1707–38. doi:10.1257/aer.101.5.1707.
- Gallin, Joshua, Lara P Loewenstein, Hugh Montag, and Randal J Verbrugge (2025). “Rent setting for new and continuing tenants.” *Work in progress*.
- Graf, Brian (2020). “Consumer price index manual, 2020: Concepts and methods.” In *Consumer Price Index Manual, 2020*. International Monetary Fund. doi:10.5089/9781484354841.069.
- Groves, Robert M and Emilia Peytcheva (2008). “The impact of nonresponse rates on nonresponse bias: a meta-analysis.” *Public Opinion Quarterly*, 72(2), pp. 167–189. doi:10.1093/poq/nfn011.
- Heffetz, Ori and Daniel Reeves (2021). “Measuring unemployment in crisis: effects of COVID-19 on potential biases in the CPS.” Technical report, National Bureau of Economic Research. doi:10.3386/w28310.

- Hokayem, Charles, Christopher Bollinger, and James P Ziliak (2015). “The role of cps non-response in the measurement of poverty.” *Journal of the American Statistical Association*, 110(511), pp. 935–945. doi:10.1080/01621459.2015.1029576.
- Houck, Ben (2023). “A review of recent improvements to the CPI’s housing age-bias adjustment.” *Monthly Labor Review*. doi:10.21916/mlr.2023.18.
- Ptacek, Frank (2013). “Updating the rent sample for the cpi housing survey.” *Monthly Lab Rev*, 136, p. 1.
- U.S. Bureau of Labor Statistics (2026). *BLS Handbook of Methods*. U.S. Department of Labor. URL <https://www.bls.gov/opub/hom/>, accessed: 2026-05-12.
- van Buuren, S. (2018). *Flexible Imputation of Missing Data. Second Edition*. CRC Press, Boca Raton, FL.
- Verbrugge, Randal, Alan Dorfman, William Johnson, Fred Marsh III, Robert Poole, and Owen Shoemaker (2017). “Determinants of differential rent changes: mean reversion versus the usual suspects.” *Real Estate Economics*, 45(3), pp. 591–627. doi:10.1111/1540-6229.12145.
- Verbrugge, Randal and Saeed Zaman (2024). “Post-COVID inflation dynamics: Higher for longer.” *Journal of Forecasting*, 43(4), pp. 871–893. doi:10.1002/for.3070.

Table 1. Summary Statistics

Imputation Type	Obs. (#)	Share (%)	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
Nonresponse Obs	383,929	21	1192	973	70	49
All	1,805,042	100	1030	838	52	29

**Figure 1.** SHARE OF OBSERVATIONS THAT ARE IMPUTED. *Note:* Observations are imputed due to nonresponse or vacancy. *Source:* Authors' calculations using the BLS Housing Survey.

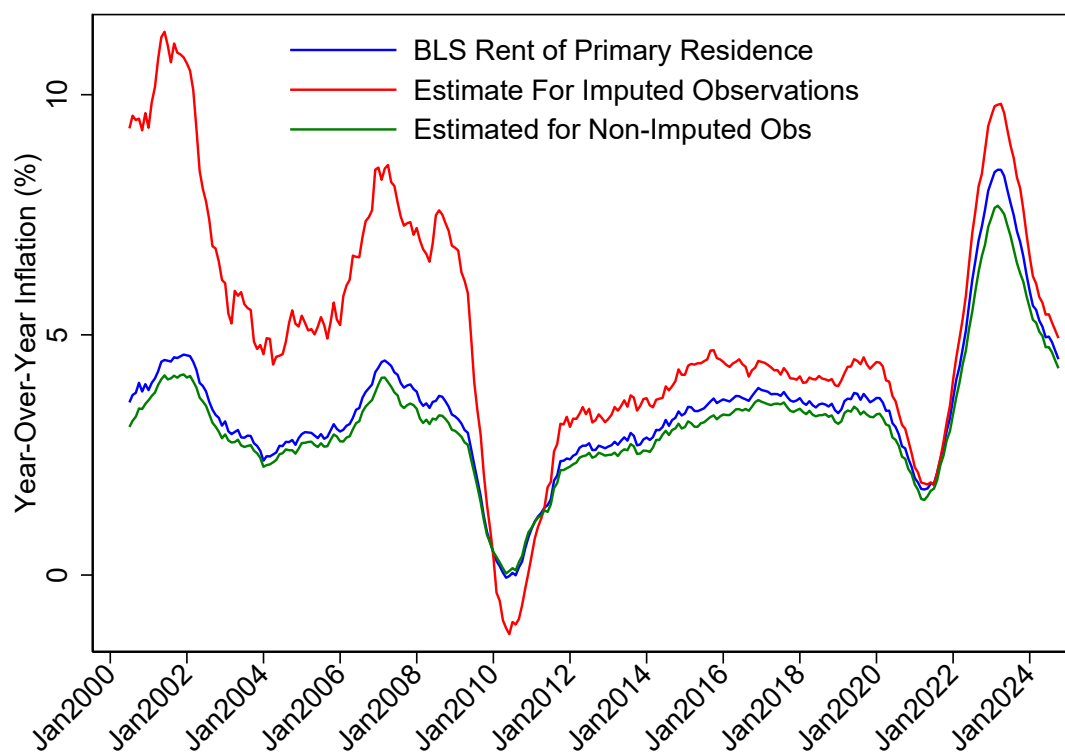


Figure 2. INFLATION RATES FOR IMPUTED AND NON-IMPUTED OBSERVATIONS. *Note:* Inflation is calculated using the method for official CPI shelter inflation, but limiting the observations to imputed and non-imputed observations. *Source:* Authors' calculations using BLS Housing Survey.

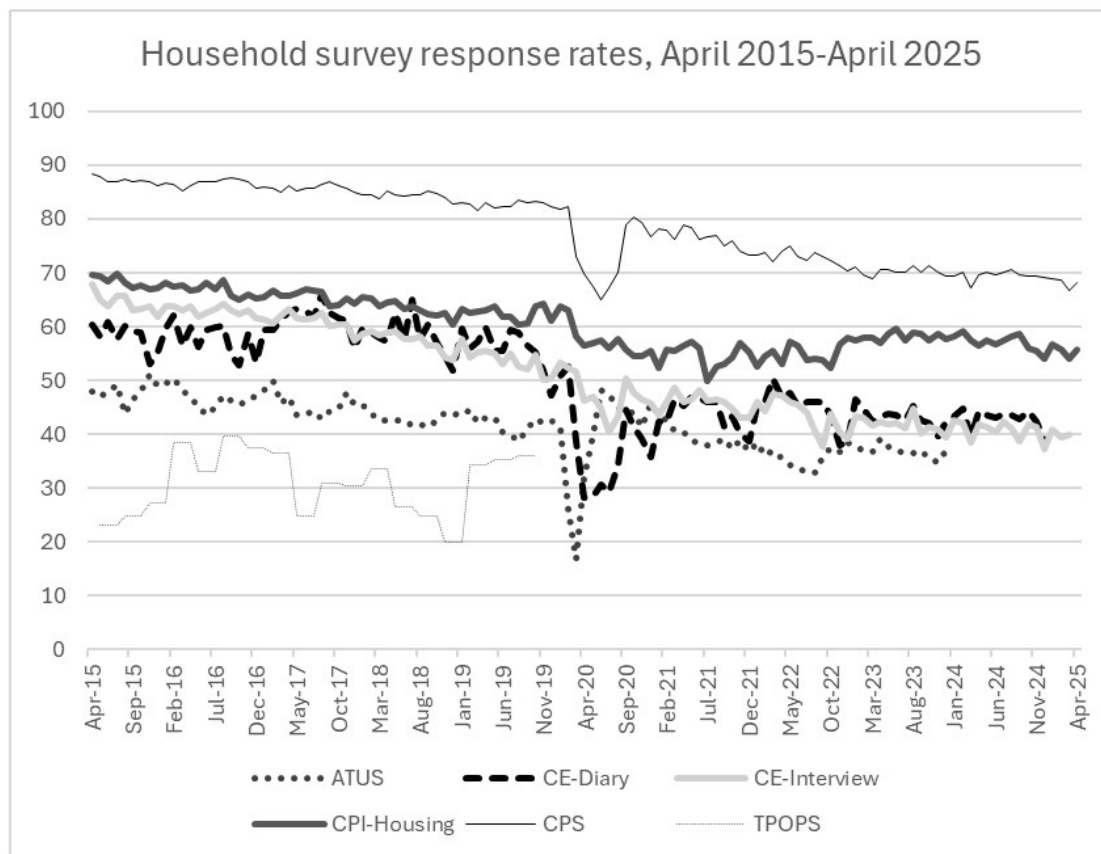
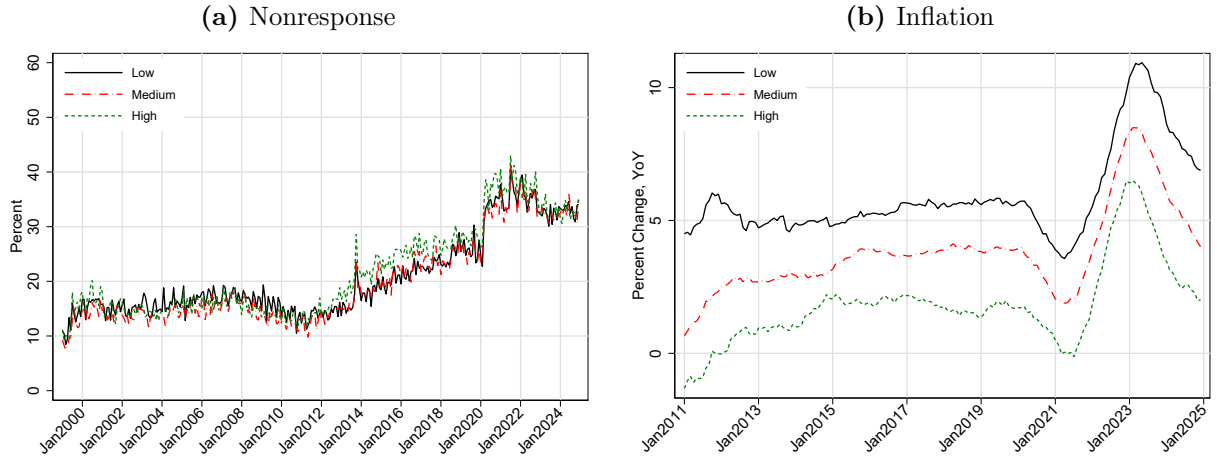
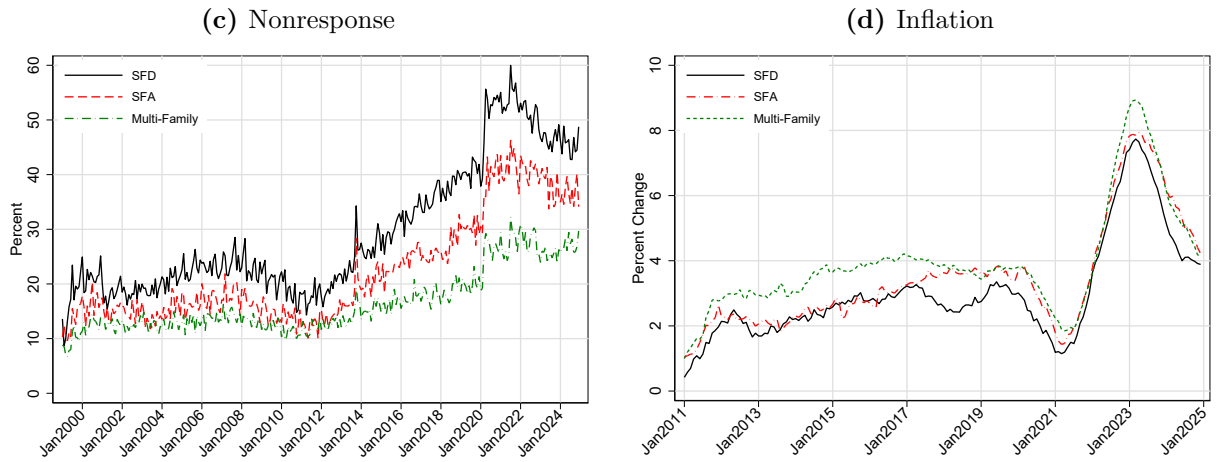


Figure 3. IMPUTED SHARE OF HOUSEHOLD SURVEYS. *Note:* Observations are imputed due to nonresponse or vacancy. ATUS is the American Time Use Survey, CE is the Consumer Expenditure Survey. CPS is the Current Population Survey. TPOPS is the Telephone Point of Purchase Survey. *Source:* BLS

By Rent Level



By Structure Type



By Tenure Length

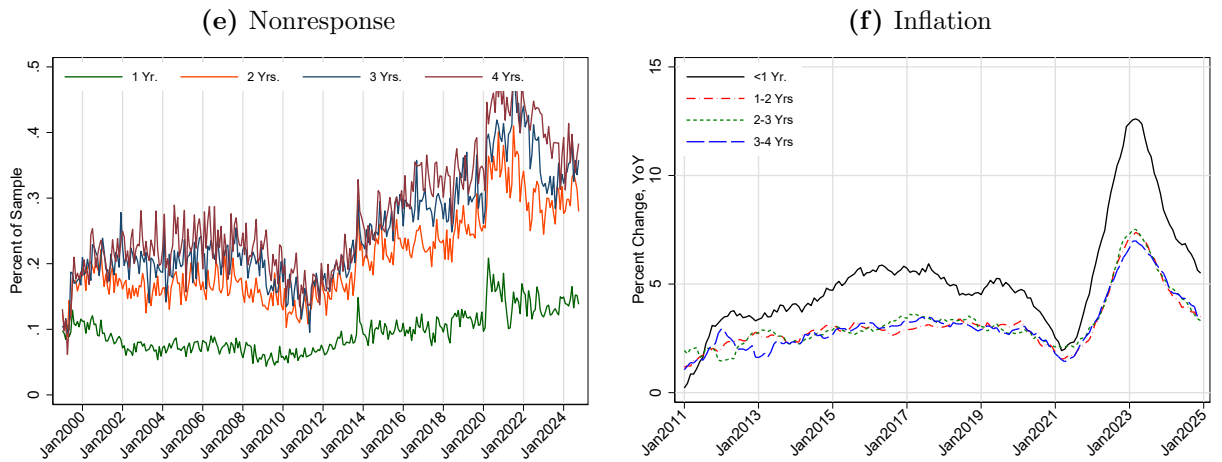


Figure 4. NONRESPONSE AND INFLATION RATES BY CATEGORY. *Note:* Inflation is calculated using all observations, collected and imputed, within a given category. *Source:* Authors' calculations using the BLS Housing Survey.

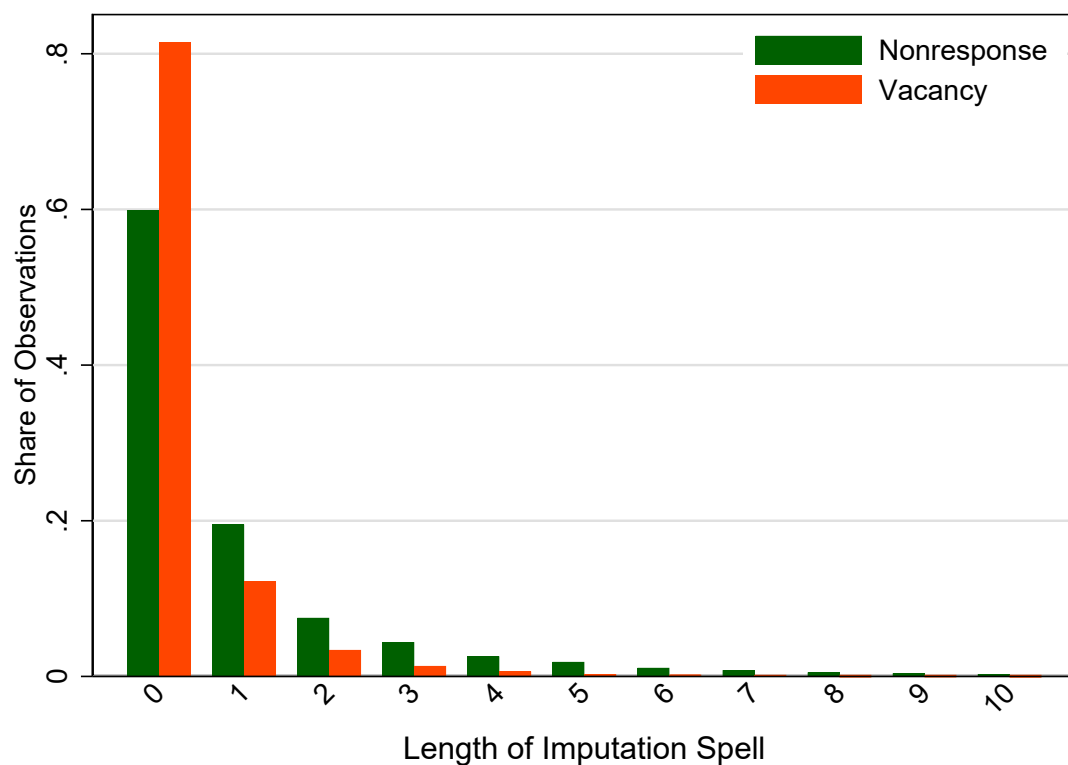


Figure 5. LENGTH OF IMPUTATION SPELLS. *Note:* An imputation spell is defined as the number of consecutive survey periods that a housing unit's rent is not collected in the survey. The x-axis denotes the number of surveys not collected. *Source:* Authors' calculations using the BLS Housing Survey.

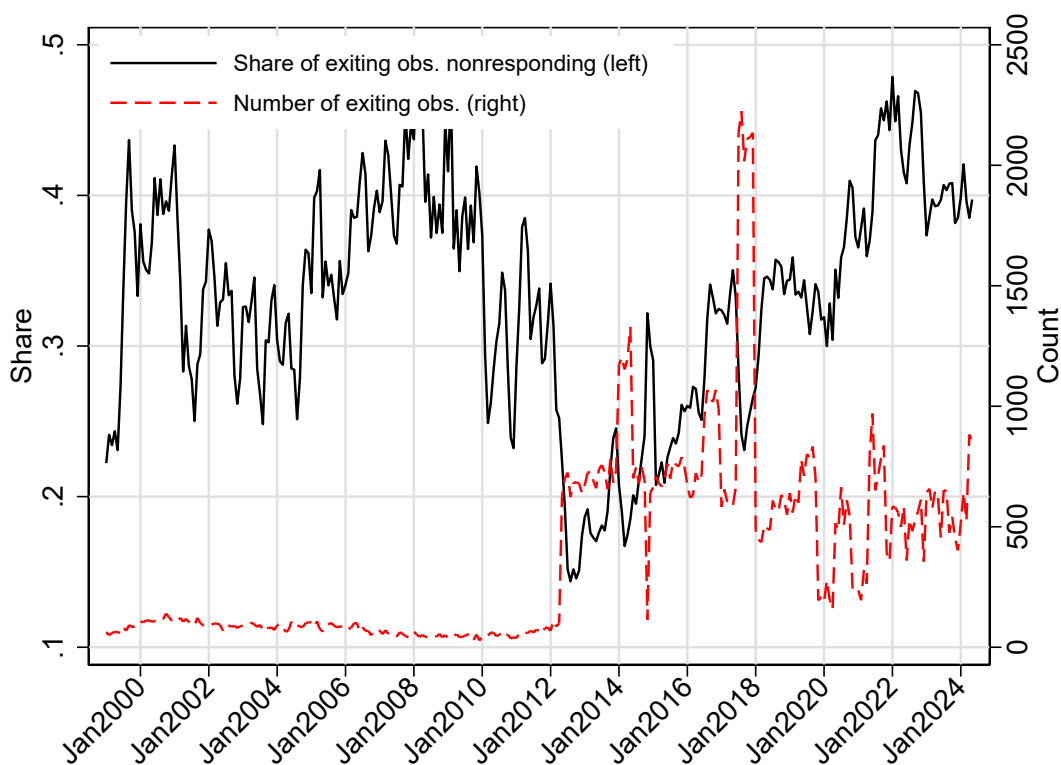


Figure 6. NONRESPONSE SHARE OF UNITS. *Note:* A unit is defined as exiting if it is surveyed for the last time before January 2025. The weighted share of exits that are nonresponses has been smoothed with a 3-month moving average. The unweighted count displays the total number of units exiting. *Source:*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

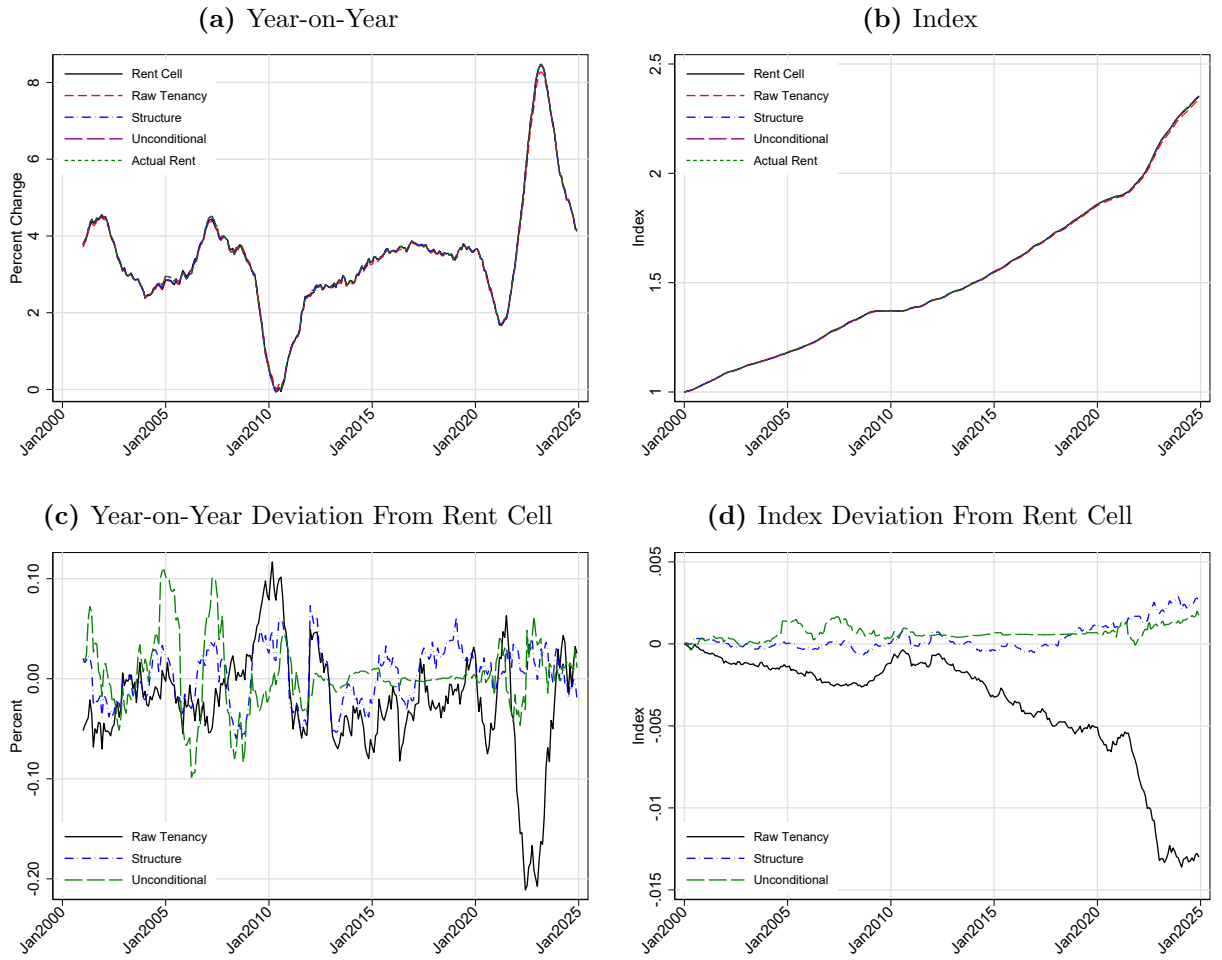


Figure 7. ALTERNATIVE IMPUTATION INFLATION RATES. *Note:* The time series are created by using a class means method and picking different covariates by which to partition observations. See text for more details. *Source:* Authors' calculations using the BLS Housing Survey.

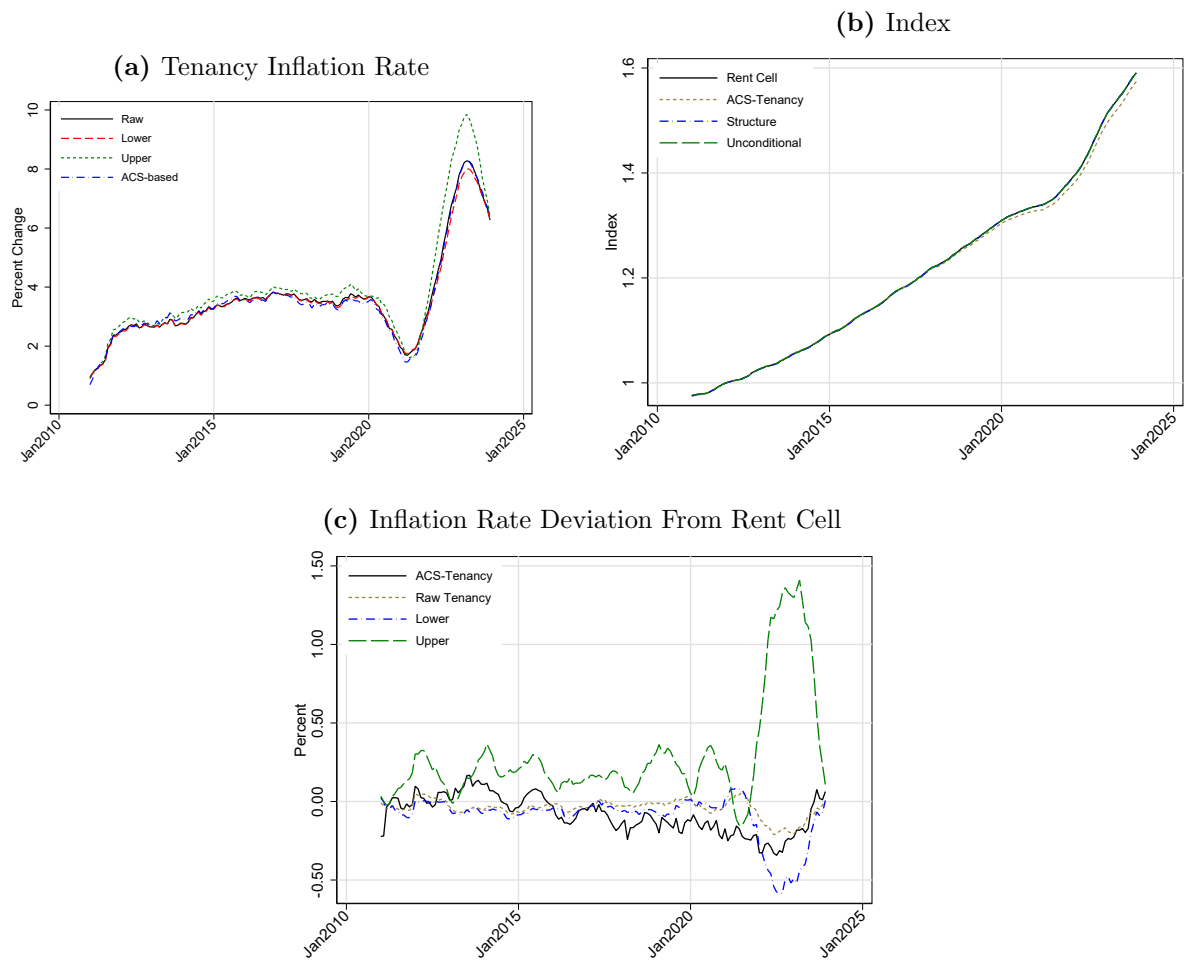


Figure 8. ACS-TENANCY INFLATION COMPARISON *Note:* The time series are created by using a class means method using different measures of tenure length for nonresponses. *Source:* Authors' calculations using the BLS Housing Survey, ACS.

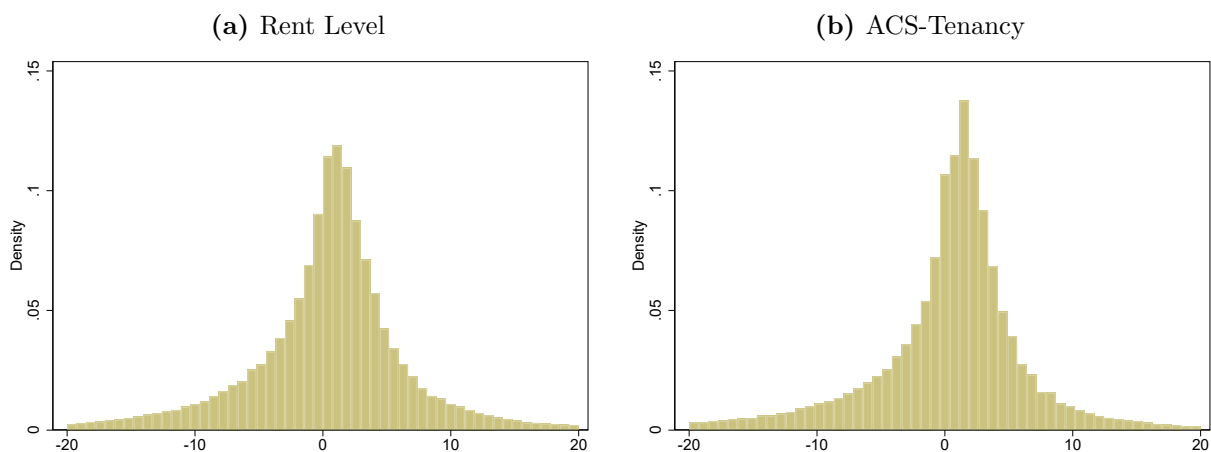


Figure 9. JACKKNIFE IMPUTATION ERROR BY IMPUTATION METHOD. *Note:* The imputation error is defined as the difference in percent between the imputed rent and collected rent. The histogram is truncated at -20% and 20%. Observations are unweighted. Observations are from January 2019 to December 2024. See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

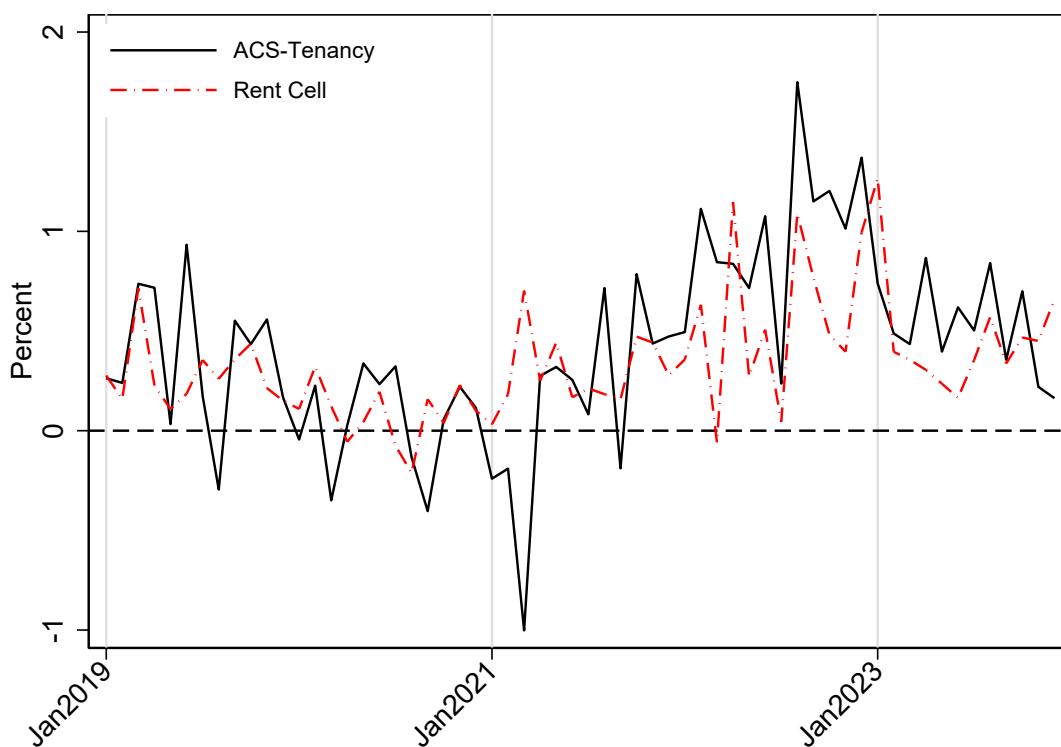


Figure 10. JACKKNIFE AVERAGE IMPUTATION ERROR. *Note:* The imputation error is defined as the difference in percent between the imputed rent and collected rent. Errors are averaged across cities using weights to obtain nationally representative series. *Source:* Authors' calculations using the BLS Housing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

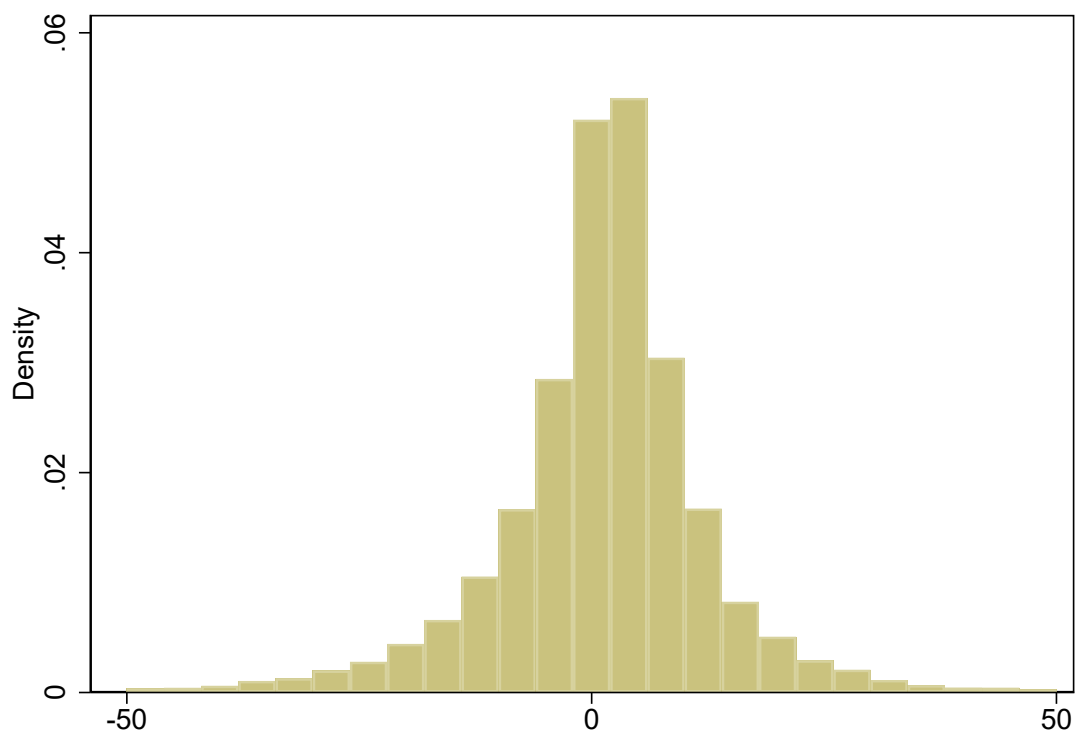


Figure 11. RE-ENTERED UNIT IMPUTATION ERRORS. *Note:* The imputation error is defined as the difference in percent between the imputed rent and collected rent. Observations are unweighted. Observations are from January 2019 to December 2024. See text for more details. *Source:* Authors' calculations 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 are related 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.¹⁹ Apartment tenants remain in their units for exactly 2 years, while tenants in detached units remain for 6 years. New-tenant rents grow by 5 percent annually. During tenancy, apartment rents increase by 4 percent annually (producing small rent gaps), while detached rents increase by 1 percent annually (producing large and growing rent gaps). Upon turnover, apartment rents increase by 6 percent to eliminate the rent gap, while detached rents increase by 25 percent. Both unit types thus experience average annual rent growth of 5 percent.

Each period, 10 percent of units exit the sample independent 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. New tenants respond with 90 percent probability. Responding units transition to nonresponse with probability $1/3$, while nonresponding units transition back to response with probability 0.1. In the 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 percent of units exit the sample. Under this timing, rent data for units in their final year of tenure reflect the old tenant’s response behavior, with new-tenant response rates affecting the following year’s data collection.

We compare three imputation methods:

1. 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

¹⁹For 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.

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.

2. 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.
3. 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 percent probability; the other half respond with 61 percent probability. Among detached units, one-sixth are new tenants who respond with 90 percent probability; response rates decline with tenure to 61 percent in year 2, 44.6 percent in year 3, 35.2 percent in year 4, 30 percent in year 5, and 27 percent in year 6. Overall, 62.3 percent of units respond in any given period, while 37.7 percent do not respond.

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

Since the true population rent change is 5 percent, 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 the 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 percent (reflecting a reweighted mass of 0.767 apartment units with an average 4.82 percent change and a reweighted mass of 0.480 detached units with an average 3.25 percent change; note that the detached average is reduced by the large proportion of continuing tenants experiencing a 1 percent 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 percent 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 percent, 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 (Percent)	Index Bias (Percent)	Imputed Value(s) (Percent)
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 the 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 percent 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 percent 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.²⁰

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 percent for detached units rather than 5.8 percent, Method 2 comes closer to the true average for nonresponding detached units, which include both many continuing tenants (1 percent increases) and some year-6 turnover units (25 percent

²⁰The 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 periods, etc. Suppose the average imputation error is +1 percent. 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 percent, 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.

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 percent) 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 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 nonresponse (and how it interacts with tenant turnover and rent-setting), so in this context it is possible to precisely impute rents for non-respondents. Conversely, in the real world the challenge is that we do not know this data-generating process.²¹ Owing to the paucity of information about exiting nonresponders (the set of tenants about whom 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.

²¹We 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 and 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 and 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:

$$\pi_{k,t}^{t+12,t} = \beta_0 + \beta_1 \text{ungap}_t + \beta_2 \pi_{k,t-12}^{t+12,t} + \beta_3 \pi_{k,t-24}^{t+12,t} + \epsilon_t$$

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 percent, ranging from -0.433 to -0.441. For β_2 , the largest gap was 7.4 percent (0.276 to 0.298), while for β_3 there was a 4.0 percent 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 percent (-2.286 to -2.389). R^2 values differed by at most 1.4 percent (0.804 to 0.815), and standard error estimates by 3.0 percent (0.691 to 0.712).

Across all specifications and sample periods examined, no coefficient estimate varied by more than 18 percent across imputation methods, and the vast majority varied by less than 5 percent. 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 across 30 cities, a significant minority of these units are not collected. 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 and 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 and Pakes (2011), for each city c and time period t , we regress

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

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 $\hat{\beta}_{c,t}$ and $\hat{\gamma}_{j,c,t}$ to predict the change in log-rents for all nonresponding units, and impute a rent on this basis.²² 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 and Pakes (2011), we test two variants: one where we only impute rents for nonresponders, and one where we impute rents for all units, both nonresponders and collected units.

Figure 12a compares the two Erickson and Pakes (2011) imputation inflation rates to

²²Note 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.

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 nonresponses. 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 Houck (2023). 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.²³ We restrict our replication to the 2020-2024 period, and use time-invariant 2025 Census tract ACS data as well as 2022 IRS zip-code-level income data. We perform this regression for each Census region and period, with city-level fixed effects to capture geographic heterogeneity. 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 non-responding 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.

²³We omit the length of occupancy variable, because we are unable to impute it for nonresponses.

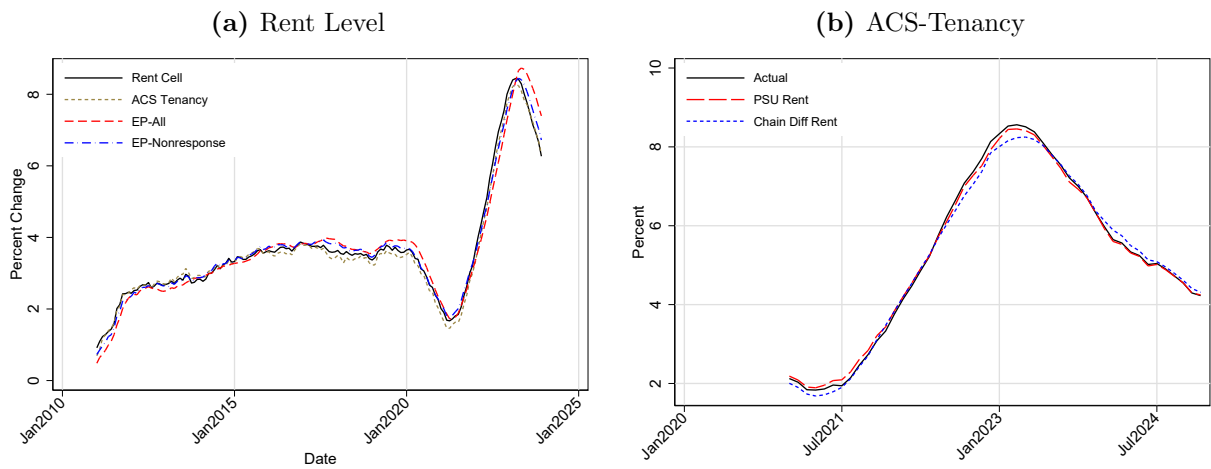


Figure 12. ERICKSON-PAKES AND AGE BIAS REGRESSION IMPUTATIONS. *Note:* The Erickson-Pakes imputation uses a regression-based approach that incorporates tenure and structure type. The age bias approach includes a wider array of covariates in the regression. See text for more details. *Source:* Authors' calculations using the BLS Housing Survey.