Nonresponse Imputations and Related Measurement Issues in the CPI for Shelter*

Lara Loewenstein[1], Hugh Montag[2], and Randal Verbrugge[1]

[1]Federal Reserve Bank of Cleveland
[2]US Bureau of Labor Statistics

January 31, 2026

Shelter is the largest component of U.S. CPI, so the accuracy of shelter inflation is critical for the accuracy of overall inflation. Nonresponse in the BLS Housing Survey, which underpins CPI shelter measurement, has increased over time and now represents roughly 40 percent of observations. Missing rents are imputed using a class-mean approach based on a limited set of covariates. We evaluate this method and find it likely biases measured shelter inflation, and therefore overall CPI inflation, upward. We propose an alternative imputation strategy that incorporates structure type and tenure length, both of which are correlated with nonresponse and rent growth.

*The views expressed in this paper are solely those of the authors and do not necessarily reflect the opinions of the Federal Reserve Bank of Cleveland, the Federal Reserve System, or the BLS. Disclaimer: This paper provides a summary of research results. The information is being released for statistical purposes, to inform interested parties, and to encourage discussion of work in progress. The paper does not represent an existing, or a forthcoming new, official BLS statistical data product or production series.

1 INTRODUCTION

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

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

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

For accuracy of the CPI shelter index, what matters is the accuracy of the *average rent change* estimate. In this context, it is less of a concern that the imputation method maintain the variance of the true underlying data (or even that unit-level rent change imputations be accurate). Some imputation methods, such as regression imputation (without added noise), could result in underestimates of variance, but this is not relevant to the construction of the index per se. In other contexts, such as regression analysis, preserving the distribution of the underlying data via the use of methods like stochastic regression imputation is a much more important consideration. Since variance maintenance is of second order importance in this context, we focus on class-means imputation, which is standard practice in inflation measurement (Graf 2020).

We make several contributions. First, we draw attention to an important issue in CPI shelter inflation measurement: a growing problem with survey nonresponse, resulting in a smaller effective sample size and driving a growing proportion of observations being imputed. The BLS Housing Survey is not the only government survey affected by nonresponse and other forms of missing data. See Figure 1 for a comparison of household survey response rates. This project contributes to a broader economics literature on missing data and measurement in official statistics, including Hokayem et al. (2015), who show that income nonresponse in the CPS March Supplement induces a 1 percent bias in official U.S. poverty rates, and Heffetz and Reeves (2021), who find that measurement issues (including nonresponse) may

have generated a bias of up to 1.5 percent in either direction in the unemployment rate during the COVID-19 pandemic. Other BLS surveys, such as the Consumer Expenditure Survey, have likewise adopted imputation methods to address missing data. Groves and Peytcheva (2008) provides a meta-analysis of methodological studies examining the conditions under which nonresponse bias emerges and when it is quantitatively meaningful.

Second, we analyze nonresponse in the BLS housing survey. Using observables available in the data, we assess whether nonresponse is correlated with rent level, length of tenure, property type, and inflation rates. We find nonresponse is slightly higher for higher rent-level units, substantially lower for units in multifamily buildings, and is potentially much lower for newly moved in tenants, although a non-response implies that the surveyor does not know whether the non-respondent is newly moved in or not, so there is a fair amount of uncertainty as to the tenancy of these tenants. Since shelter inflation rates also vary across these categories, any imputation strategy that does not account for these correlations can result in biased estimates of rent changes for nonresponses.

Third, we detail the current imputation method. It is a class means imputation approach based solely on city, month, and tercile of the lagged rent level of a unit. However, we find that rent level is neither highly correlated with either nonresponse nor with rental inflation rates. The current imputation method does not control for structure type or tenancy length, the two covariates we determine are most highly correlated with nonresponse and which are much more highly correlated with rent growth. Our assessment indicates that this procedure generally results in a positive bias to shelter inflation estimates.

Fourth, we assess class-means imputation strategies including tenure length and structure type. We examine the extent to which small changes in this procedure, namely changing the covariates used for class means imputation, enhance accuracy and reduce this bias. We find that the exclusion of tenancy length in the imputation procedure can theoretically create an upward bias in the rent inflation rate. However, tenancy length for nonrespondents is also not known in real time. It is equally theoretically possible that all nonrespondents are new tenants, or that they are all the same tenants as the previous respondent. Any bias on rent inflation depends on which of these extremes is more accurate. We propose a probabilistic method for determining the average share of nonrespondents that are new tenants and find that the implied bias is much smaller than its theoretical maximum.

Finally, we document the extent to which the imputation process itself contributes to noise in the overall shelter index.

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

instead focuses on nonresponse imputation.

2 DATA

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

The housing survey contains detailed questions about the lease and housing unit in order to enhance accuracy in the the shelter indices. For example, many leases specify that landlords are responsible for some or all utilities. However, utility responsibility could change over time.[2] Likewise, many tenants perform some in-kind labor of their lease.[3] 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. 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 BLS field economists survey a specific housing unit rather than a tenant, so there is a paucity of information about tenant characteristics. Once tenant moves out of the unit, the survey does not follow up with them or determine where they move to. One piece of relevant information which is collected is the date they moved into the housing unit, from which housing tenure is derived.

Prior to 2012, the BLS surveyed the same set of housing units indefinitely. [4] Thus, many

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

[2] In any case, this information is necessary in order to allow the BLS to adjust the rent for use in the OER index

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

[4] Prior to 2010, the BLS routinely added additional housing units to the sample, with sampling based upon the decennial census data. In 2010, the CPI began initiating a plan that would continuously update the sample of rented housing units. Eventually, one-sixth of the rented housing unit sample every year would be replaced, with sample updating based on the basis of the latest available U.S. Census Bureau data. Ultimately, this new method reduces the age of the sample, more accurately reflects new construction and changes in where people live, and results in less sample attrition. Systematic sample retirement began in

of the housing units we observe prior to 2012 are in the sample for a significant period of time. However, starting in 2012, the BLS began regularly rotating the sample of housing units. After entry into the sample, housing units are now surveyed for six years, after which they exit the sample. Each year, one-sixth of the housing units rotate out and are replaced. This has significant implications for the non-response and imputation rates post-2012. As we discuss below, non-response imputation is chiefly important for housing units that do not respond to the final survey before they rotate out.

The BLS uses observations in its CPI calculation even if field economists were unable to obtain a response. There are three possible states for a housing observation: collected, non-response, or vacant. A housing unit is collected if it responded to the survey, even if it did not complete all of the questions in the survey. A unit is a non-response 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, non-response and vacancies, there are two different imputation procedures that are applied.

We are interested in survey variables that may be correlated with non-response status or rent changes. 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, field economists do not routinely record move-in dates for nonresponding housing units. 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. We will implement several approaches to estimate the tenancy length of non-responding units, which we discuss below. We also use information about the structure type for the housing unit, since as Adams and Verbrugge (2025) notes, rent changes for apartments and single family detached houses can vary dramatically, even within the same city. Lastly, we use internal variables for the economic weight of housing units so that we can construct national statistics.

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 non-responses. Note that as displayed above, the share of non-responses is not static over this time period. As we demonstrate below, given current imputation procedures, this trend causes the average non-response rent in our sample to be larger than the average collected rent, since non-responses are overrepresented in recent

late 2012. See , and the section on (design in the Handbook of Methods) for more information.

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

periods. Non-response observations have longer tenancy lengths than collected observations, which suggests that nonresponse probabilities rise with tenancy lengths, for behavioral reasons and/or survey attrition. Vacant observations have short tenancies, as housing unit do not tend to be vacant for extended periods of time.

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

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 rent level (low, medium, or high). 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, North, East, etc. Non-responding 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 simple procedures that do not consider within-city location.

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

Next, each rent level cell (i.e., the subset of valid observations in each rent level in that city) is evaluated to ensure that there are a sufficient number of observations for imputation. If the BLS has more than 5 observations in the rent cell, then any nonresponding observations in the corresponding rent level are imputed using the average rent change from this comparison pool, irrespective of its composition (between continuing tenants, new tenants, and vacancy imputations). If the BLS has fewer than 5 respondents in this comparison pool, it increases the comparison group by relaxing the rent-level criteria. For example, if there are insufficient observations in a low-rent cell, then they may combine the low- and medium-rent cells. Finally, if this still does not result in a sufficient comparison pool, the BLS combines all collected and imputed vacant observations for that time period in the city.

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

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

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

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

3.2 Vacant Units and Their Imputation

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

Omitting vacancies can potentially lead the price index to mismeasure rents, as Crone et al. (2010) and others note. We know that rent increases tend to be much larger when a new tenant moves in, relative to the rent change experienced by continuing tenants (Adams et al. 2024). If vacant units exit the sample before a subsequent tenant responds to the survey, then the housing survey would miss the (typically large) rent increase associated with the new tenant. Prior to a unit being dropped from the survey, "vacancy adjustment" is necessary in order to ensure that the index is unbiased. An extreme hypothetical example will explain the logic. Suppose that rents never change while a tenant occupied a unit; instead, we suppose that all rent adjustments occur only when a unit becomes vacant, before the new tenant moves in. Now suppose that units only exit after a tenant moves out, "just before" the unit was due for its rent increase. In that case, the BLS would miss a significant share of rent increases. For this reason, the BLS imputes a "vacancy adjustment" to units that become vacant, in case they exit before a new rent can be collected.

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 collapsed.

Let the set $H_{c,v,t}$ denote the collected in city c, tenancy length cell r, 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

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

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

3.3 Measuring Rent Inflation

The CPI rent index resembles a Young price index for each geographic area.[6] Since housing units are sampled every six months in a panel rotation, the set of housing units in consecutive months is not overlapping for a specific geographic area. But since the BLS wishes to construct a monthly index, the current practice is to calculate a six-month price relative for each city and convert it to a one-month price relative. For city c and time t, the six-month rent price relative is

$$REL6_{c,t} = \frac{\sum_i \omega_{i,t} \text{rent}^*_{i,t}}{\sum_i \omega_{i,t} \text{rent}^*_{i,t-6}}$$
$$REL_{c,t} = REL6_{c,t}^{\frac{1}{6}},$$

where $\omega_{i,t}$ denotes the weight on a housing unit and $rent^*_{i,t}$ denotes the economic rent after imputations.[7] 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.[8] 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.

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

[7] We abstract from the aging bias adjustment applied to the denominator.

[8] See https://www.bls.gov/opub/hom/cpi/calculation.htm for additional details about the CPI rent index construction.

4 COVARIATE ANALYSIS

In this section, we study whether observations in the housing survey are missing at random relative to observable characteristics and whether the rental inflation rates are similar by characteristic.[9] If the non-response rates do not vary across observable characteristics, then our subsample of collected rents should still be representative with the appropriate sampling weights. Even if the inflation rates vary by characteristics considered in the class means, the resulting price index would be accurate. Similarly, if the inflation rates are uncorrelated with observable characteristics (for example, the average rent change for single-family detached houses is the same as the average rent change for multi-family units), then the price index will be valid even if some types of units are more likely to non-respond than others. It is only in the event that both the inflation rate and the non-response rate are correlated with an observable characteristic that a poor choice of imputation method may bias the result.

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

We first construct the national share of observations each period that are non-responses. These shares are weighted averages, with the weights reflecting both geographic weights and lower-level within-city economic weights. We calculate these non-response shares by the covariates of interest, namely rent level, structure type, and tenancy length, and plot the results over time. The objective to get an overall picture about whether the data is missing at random with respect to these variables, whether this is changing over time, and to observe the patterns of nonresponse.

Even if the data is not missing at random, it would only affect the rent index if the inflation rate is also correlated with the observable characteristic. To test this, we 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 **??**.

Nonresponse and inflation rates by our three observable characteristics are on the left and right panels of Figure 4. Figure4a 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

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

rates by structure type, and Figures 4e and 4f are plots of of nonresponse and inflation by measured tenants length.

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

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

Across structure type, nonresponse differences are more stark. Response 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 multi-family 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 have different inflation rates (Adams and Verbrugge 2025; Gallin et al. 2024), and now accounts for this in its calculation of OER. As can be seen in Figure 4d, multi-family units have a systematically higher inflation rate than single-family detached units. Differences between single-family detached units and small multi-family units are less stark, especially earlier in the sample. Some of these differences may be explained by location, but not all (Adams and Verbrugge 2025). Although large multifamily units feature higher turnover (and thus more frequent rent resets toward new-tenant rents: see Gallin et al. (2024)), this cannot explain rent inflation differences over such long horizons. An alternative explanation is that multi-family units are better maintained or depreciate more slowly, and that this is not currently captured by BLS aging adjustments.[10]

In Figures 4e and 4f we have parallel figures by raw tenure length. In these figures, tenure

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

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 non-response rates between observations with tenancies shorter than one year, and those with longer tenancies. Even after one year, there appears to be a clear monotonic relationship between non-response rates and tenancy lengths over time.

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

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

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

5 NON-RESPONSE SPELLS, ERROR PROPAGATION, AND EXIT

Before we turn to alternative imputation methodologies, let us consider the duration of non-response spells. It seems a priori unlikely that non-response status for a housing unit is distributed independently across time. 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 rents would end up being imputed for the same housing unit multiple times in a row, possibly compounding errors.

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

non-response is much more likely to be followed by another non-response.

Over longer periods of time, the choice of the imputation method would not matter at all, if every non-response were to be followed by a valid rent collection in a subsequent period. Consider a hypothetical housing unit over three collection periods $t = 0, 1, 2$.[11] In periods 0 and 2 the housing unit is surveyed successfully, and in period 1 the housing unit does not respond, and its rent is imputed. Denote the actual rent in period t as r_t, and the rent entered in the housing survey $\tilde{r}_t$. Let the price relative between periods s and t be denoted by $rel_{s,t}$.

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

$$rel_{0,2} = \frac{r_2}{r_0} = \frac{\tilde{r}_2}{\tilde{r}_1}\frac{\tilde{r}_1}{\tilde{r}_0} = \tilde{rel}_{0,1}\tilde{rel}_{1,2}.$$

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

In Figure 6 we plot the share of the rental units exiting that housing survey that leave during a non-response spell (i.e., the share which did not respond during the collection period prior to exit). Prior to 2012, relatively few housing units left the survey each year, since units were kept in the sample unless they transitioned to owner-occupied. However, over this earlier period, a relatively large fraction of these units exited during a non-response spell. This is partly because non-response becomes more likely, the longer a unit has been in the survey. Once the housing rotation began, the fraction of exits jumped quickly, plateauing around 18-19 percent. However, since the start of the housing rotation, the share of non-responses among exits has grown markedly, from 15 to 40 percent. Therefore, the housing survey currently has an unprecedented combination of a relatively high exit rate and a high share of non-responses among exits, which means that any error in the imputation methodology has more severe consequences now than in previous times.

[11] For the sake of simplicity we focus on one housing unit and abstract away from the six-month panels in this exercise.

6 ESTIMATING TENANCY LENGTH FOR NONRESPONSES

As discussed 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 a high amount of uncertainty about the rent change experienced by specific units. 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.

We conduct a series of imputations using four measures of tenancy length to approximate or establish bounds on tenancy's effect on the rent price index.

First, we use the nonresponse move-in date information in the Housing Survey, which is principally calculated through a carryforward imputation. As shown below, many housing units non-respond for multiple surveys. The move-in carryforward assigns these units the same tenant throughout their entire non-response spell, with a long tenancy that can be years long. Throughout the paper, we refer to this as the housing survey movein (or raw movein). Although imperfect, this measure provides a necessary benchmark and may still be useful.

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

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

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

More specifically, let w denote the share of responses in the housing survey that are non-

responses and let x denote the share of units in a subsample that are new tenants.[12] 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 the BLS Housing Survey. All of the variables can be explicitly calculated from the housing survey or ACS except for $x_{c,t}^{Nonresponse}$, the share of new tenants in non-responses. In theory, $x_{c,t}^{Nonresponse} \in [0, 1]$, although in practice $x_{c,t}^{Nonresponse}$ can lie outside of these bounds for cities and dates where the share of non-responses is small. We truncate $x_{c,t}^{Nonresponse} \in [0, 1]$.

7 ALTERNATIVE IMPUTATIONS AND POTENTIAL BIAS

We estimate multiple alternative inflation rates utilizing alternative non-response imputation methods. Broadly speaking, these imputation methods employ the same class means approach utilized in the CPI rent index imputation, but differ based on the specific observable variable(s) used for observation partitioning. 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 can influence subsequent rent imputations, as well as the rent tercile classification of the unit in the following period. This process potentially compounds errors for consistent non-responders, as imputed rent adjustments are applied to rent levels derived from previously imputed rent changes. Throughout this paper, we refer to this process as "chaining" the imputation method, reflecting the iterative nature of our computational approach.

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

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

Third, we perform an unconditional imputation. In this scenario, we calculate the average rent change by city-date among all responding observations and use that as the rent

[12] We ignore vacant units in this derivation.

change for the nonresponders. This imputation method would be valid if the data was truly missing-at-random. It provides a benchmark against which to compare the other imputation methods.

Fourth, we partition by tenure, calculating imputation rates using our four different measures of tenancy length: the raw recorded value that is carried forward, our upper- and lower-bound scenarios described in Section 6, and our probabilistic estimate that uses the ACS mobility rate. For the first three measures, we partition responding observations into two bins: those who are newly moved in and therefore have a tenancy length shorter than 6 months, and those who have been in their unit for longer, as this distinction encapsulates most of the variation in rent changes across tenancy lengths (Gallin et al. 2024). In the raw tenancy case, we aggregate new tenant and continuing tenant cells together in the event of insufficient observations within a city-date. In the upper- and lower-bound cases, we pool observations across multiple months in the same city.

Finally, in an alternative tenure-based method, we use the ACS-implied mobility rates; this obviates the need to estimate tenure lengths. We instead impute the unobserved rent change using a weighted average of the local new- and continuing-tenant rent changes, where weights derive from our estimates of $x_{c,t}^{Nonresponse}$. 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. (This represents a small departure from current BLS methods.)

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 cells and structure type yield an inflation rate almost identical to the unconditional approach. This is somewhat surprising, given that rent changes vary noticeably by rent level and structure 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.

The year-on-year rent inflation rate using the raw-tenancy-based imputation method is about 0.05 p.p.% lower than the current rent cell and the unconditional methodologies for most periods. The effect rises to 0.1 p.p.% in 2020 and 2021 as the inflation rate rapidly rises. This discrepancy compounds fairly quickly: a rent index that started in January 2000 and used the tenancy imputation method would have grown by 2% p.p. less than the current

rent cell method. Figure 7b displays the indices.

The rationale for the magnitude of the effect is straightforward. The set of units that do not respond are disproportionately made up of tenants with longer measured tenures. Since new tenants often have larger and more volatile rent changes than continuing tenants, the non-responding units drag down the year-on-year inflation rate. However, for the typical city-month, only a handful of the successfully surveyed housing units contain a new tenant. In the case of a insufficient number of new tenant observations, our tenancy imputation approach combines the new tenant cell with the continuing tenant cell. In short, we impute new tenant "non-responses" with the average rent change across all units within a city-date often, mitigating the impact of the tenancy imputation.

As noted above, the raw tenancy measure may not accurately measure the share of non-responses 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 find that the upper-bound tenancy imputation is noticeably higher than the raw tenancy and lower-bound tenancy imputations. We ascribe the difference to the fact that the upper-bound measures assumes a high rate of tenant turnover and avoids combining new and continuing tenant cells together in the case of insufficient observations, instead choosing to use additional months of data. In contrast, the lower-bound tenancy imputation makes similar assumptions about tenancy status as the raw tenancy comparison.

Figure 8b compares the ACS-based tenancy index against the rent level 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 ends up a couple percentage points lower–a difference that is noteworthy. We think it likely that this differential stems from the rent cell methodology inadvertently overweighting rent changes observed in units with tenant turnover, given the correlation between response rates and tenure duration.

In summary, we find that the current rent cell approach is little different than an unconditional (city-date) imputation. At best, the rent level cell approach improves the accuracy of the rent index only marginally. Conversely, the raw tenancy-based imputation method yields slightly lower inflation rates, likely driven by the not-missing-at-random nature of the data. Extreme assumptions about tenure yield notably different inflation rates. The lower- and upper-bound approaches provide bounds for what the tenancy-based imputation could be if we had accurate tenancy information. An ACS-based approach allows one to condition on tenure, without the need to make strong assumptions about unmeasured tenure. Of the alternatives to the rent level method currently in use, this is our preferred approach. Note that it may need to be modified to account for lagged ACS data publication.

8 IMPUTATION ACCURACY AND PRECISION

In the previous section we determined that a tenancy-based imputation yields different results than a rent level cell-based imputation. This section addresses two subsequent questions: which method more accurately imputes value, and how precise are the estimates?

8.1 Accuracy

It is not obvious how one might estimate the accuracy of an imputation, because we do not observe the actual rent change for the units that don't respond. Hence, it is unclear what we would measure the an imputation against. We use two approaches to approximate the accuracy estimation.

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.[13] We then impute rent $x^I_{i,c,t}$ 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.[14]

Second, we focus on housing units that exit the sample for a period of time due to nonresponse, and then re-enter the sample later. In these cases, we can impute their rent for several consecutive dates and then evaluate the accuracy of their imputations when their rent is collected by the survey again.

Figure 9a displays two histograms for the jackknife imputation errors across cities and dates. The left-hand histogram depicts the imputation error using the rent level cell method, while the right-hand histogram shows the imputation error with the ACS-tenancy imputations. Given the heterogeneity associated with the rental market, and the fact that class

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

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

mean imputations impute the same rent change for all units within a given rent level-city-date combination, it is unsurprising that imputation errors have such wide variance.

Both imputation methods yield imputations that are biased and with notable variance. On average, the ACS-tenancy method overestimates rent changes by a statistically significant 0.18%. The average rent level cell error is slightly higher, at 0.20%, also statistically significant. Both imputation methods yield notable median errors of 1.03% and 0.77%, suggesting that there is significant asymmetry in the errors. Tails are also quite large. Further, as figure 10 displays, there is significant time variation in the average imputation errors. Both methods yield highly correlated average imputation errors; the ACS-tenancy errors tend to be somewhat more volatile. The average error over time is about 0.42% and 0.34% for the ACS-tenancy and rent cell approach, respectively. The jackknife approach to measuring accuracy does not conclusively indicate that one method is superior to the other, but it does suggest that there is substantial rent change heterogeneity that is neither captured by tenancy status nor rent level.

The imputation errors for re-entered observations display similar behavior, although at a greater scale. Figure 11 depicts a histogram of rent-level-cell-based imputation errors for re-entered units. These errors possess a similar rightward draft. The median error is 1.83% and the mean error is 1.17%. The latter is statistically significant. However, there are two issues that complicate the interpretation of these results. First, there is some heteroskedasticity to these re-entered imputation errors, since different units may non-respond for different durations. Second, we are not absolutely certain whether the same tenant was present before the non-response spell as at the end; this means that there may some tenant turnover, and new tenant rents.

9 CONCLUSION

This paper investigates the imputation process underlying the BLS Housing Survey. We demonstrate that the share of non-responses in the survey has been rising steadily over the past decade, which means that this is an increasingly important topic. We find that there are substantial differences in response rates and inflation rates by observable covariates, highlighting the necessity for careful imputation. While the current rent level cell imputation approach yields similar results to an unconditional imputation, we show that imputing by tenancy status reduces the subsequent rent inflation rate. However, as the tenancy length of nonresponding units is also unobserved, performing an imputation by tenancy status is a non-trivial task. We provide a tenancy-based method that does not require accurate measurement of tenancy at the unit level. In two separate tests of imputation accuracy, we find that the imputation methods yield statistically significant imputation errors. (Recall that what matters for index accuracy is the average error.)

There are three lines of future research that we hope to explore related to this topic. First, vacancy imputations have not yet been studied in depth. Although vacancies make up a small share of the survey, new vacancies can have large imputed rent changes that are important drivers of the rent index. Second, the non-response imputation for owners' equivalent rent could yield possibly different results. Like the rent index, the price index for owners' equivalent rent is derived from the BLS Housing Index. However, structure type plays a more significant role in OER, as noted in Adams and Verbrugge (2025). It is possible that imputing by structure type could affect the OER index, especially over the more recent period, since detached units now carry greater weight in the OER index. Third, we have not explored imputations based upon hedonic modeling, a method which would (in principle) allow one to simultaneously control for tenure, structure type, and rent level. It may also lead to imputations featuring less bias.

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. URL https://www.aeaweb.org/articles?id=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.

Crone, Theodore M., Leonard I. Nakamura, and Richard Voith (2004). "Hedonic estimates of the cost of housing services: Rental and owner-occupied units." Technical report, Federal Reserve Bank of Philadelphia. URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=610942.

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.

Gallin, Joshua, Lara P Loewenstein, Hugh Montag, and Randal J Verbrugge (2024). "Sticky continuing-tenant rents." *Work in progress*.

Graf, Brian (2020). "Consumer price index manual, 2020: Concepts and methods." In *Consumer Price Index Manual, 2020*. International Monetary Fund.

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.

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.

Hokayem, Charles, Christopher Bollinger, and James P Ziliak (2015). "The role of cps nonresponse in the measurement of poverty." *Journal of the American Statistical Association*, 110(511), pp. 935–945.

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.

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
Non-response Obs	383,929	21	1192	973	70	49
All	1,805,042	100	1030	838	52	29

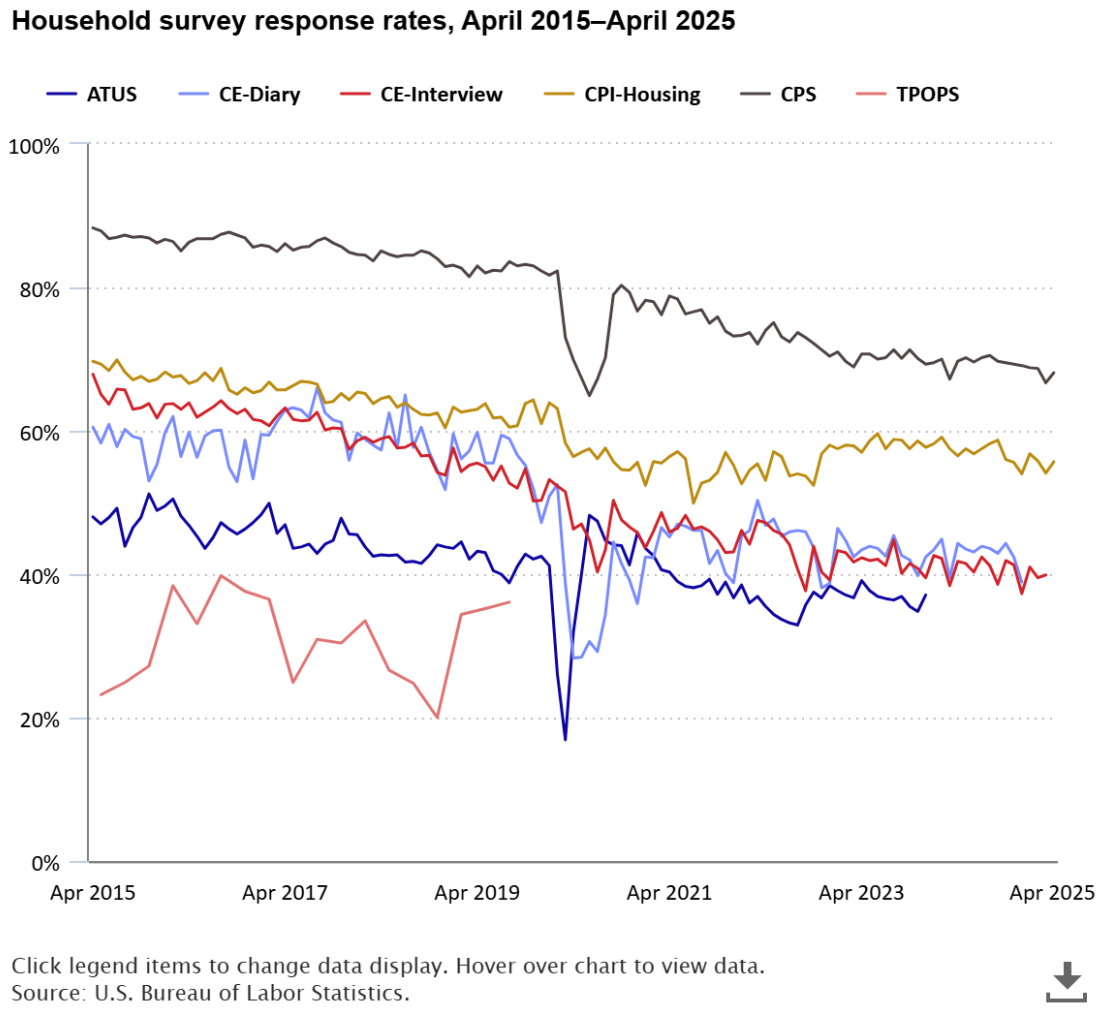


Figure 1. Imputed Share of Household Surveys. *Note:* Observations are imputed due to nonresponse or vacancy. *Source:* BLS, https://www.bls.gov/osmr/response-rates/#chart1a

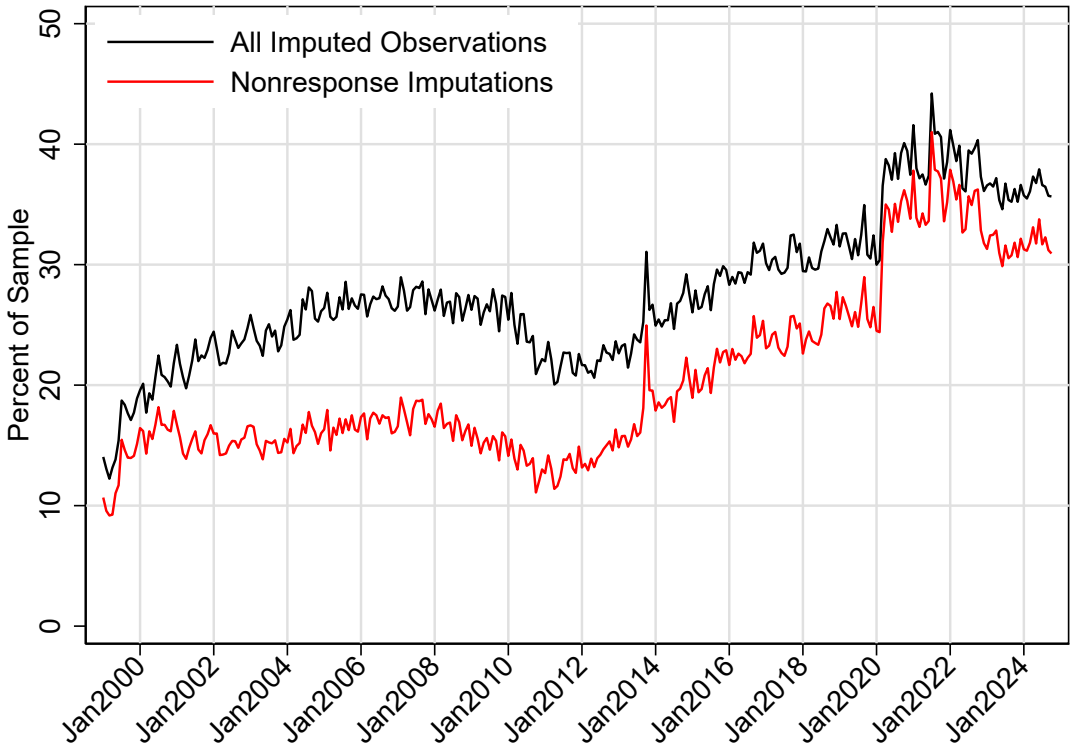


Figure 2. Share of Observations that are Imputed. *Note:* Observations are imputed due to nonresponse or vacancy. *Source:* Authors' calculations using the BLS Housing Survey.

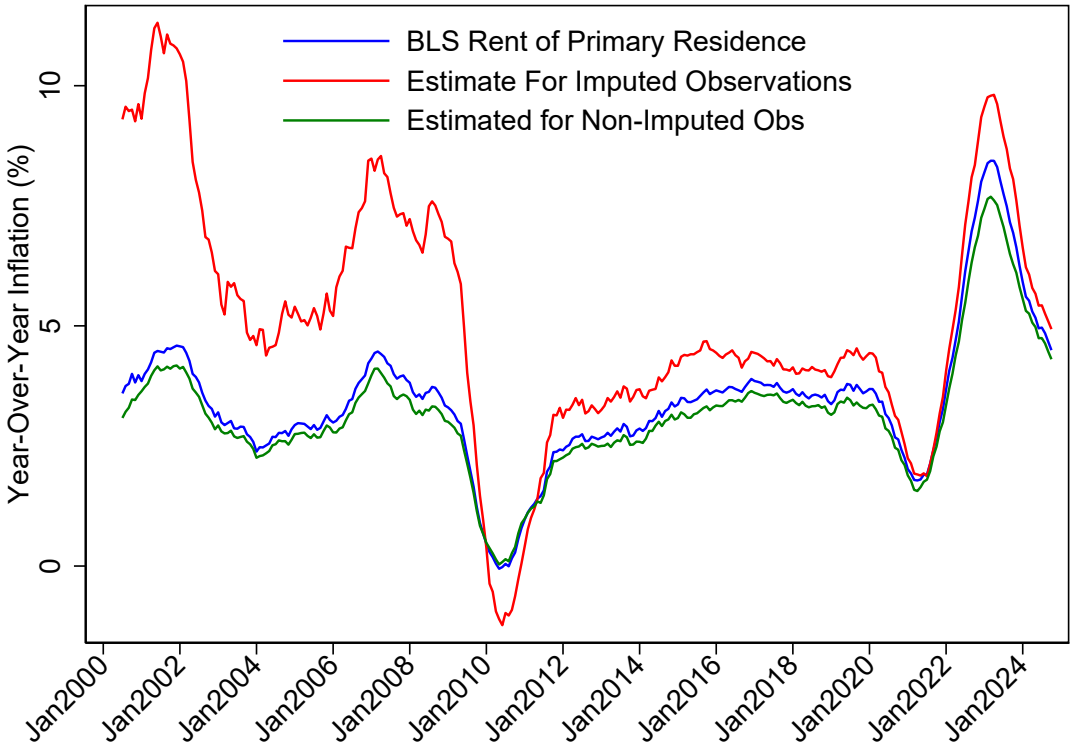


Figure 3. Inflation Rates for Imputed and Non-Imputed Observations. *Note:* Inflation is calculated using the method used for official CPI shelter inflation, but limiting the observations to imputed and non-imputed observations. *Source:* Authors' calculations using BLS Housing Survey.

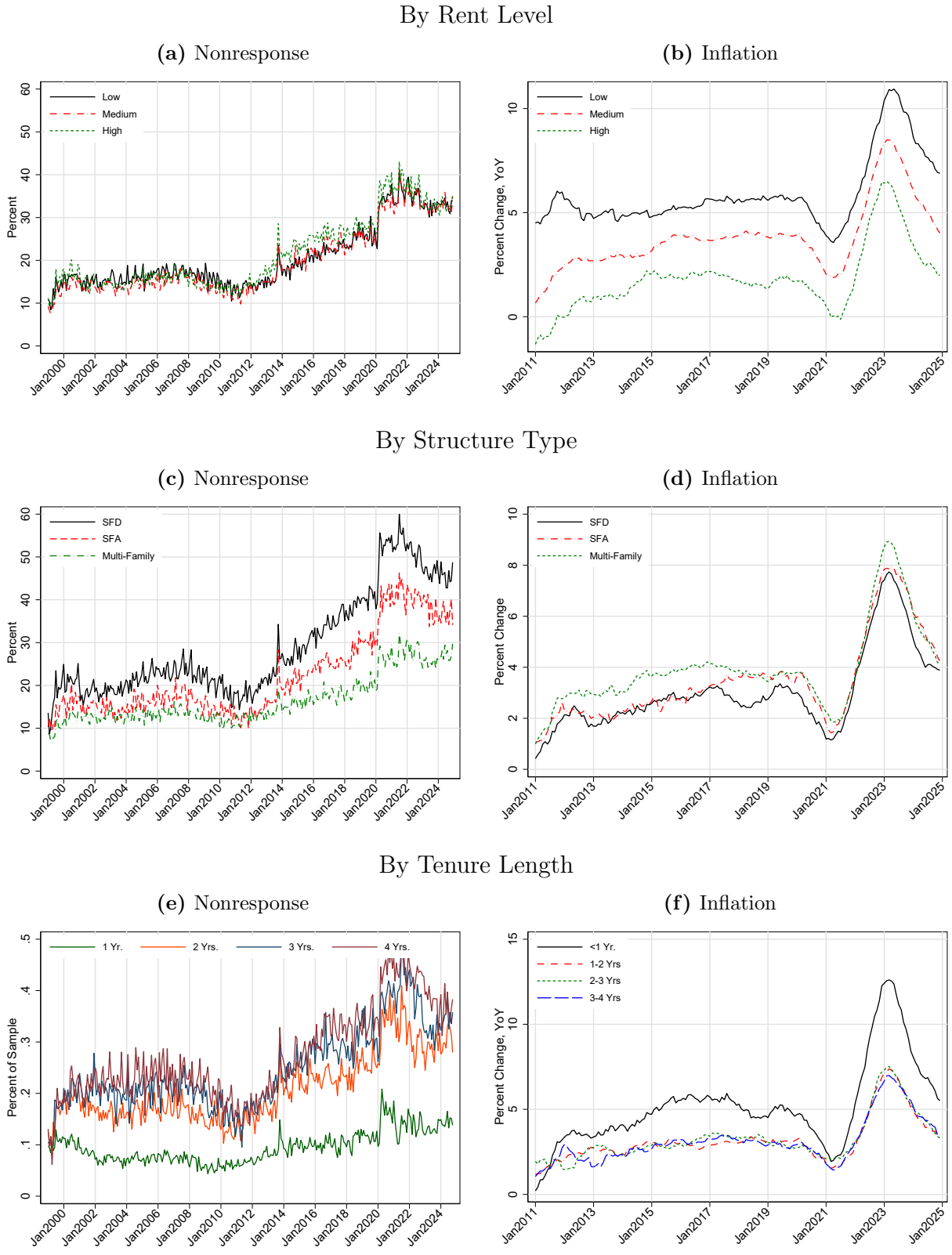


Figure 4. Non-Response and Inflation Rates By Category. *Note:* Inflation is calculated using all observations, collected and imputed, within a given category. *Source:* Authors' calculations using the BLS Housing Survey.

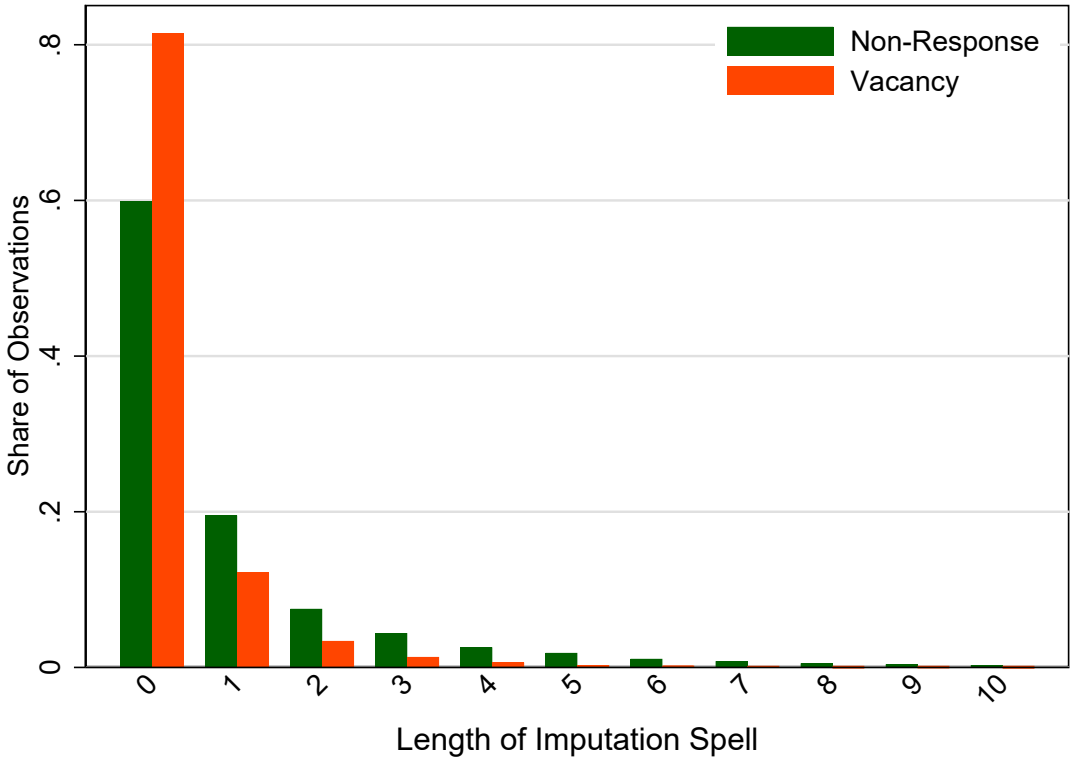


Figure 5. Length of Imputation Spells. *Note:* An imputation spells is defined as the number of consecutive survey periods that a housing unit is not surveyed. The x-axis denotes the number of surveys not collected. *Source:* Authors' calculations using the BLS Housing Survey.

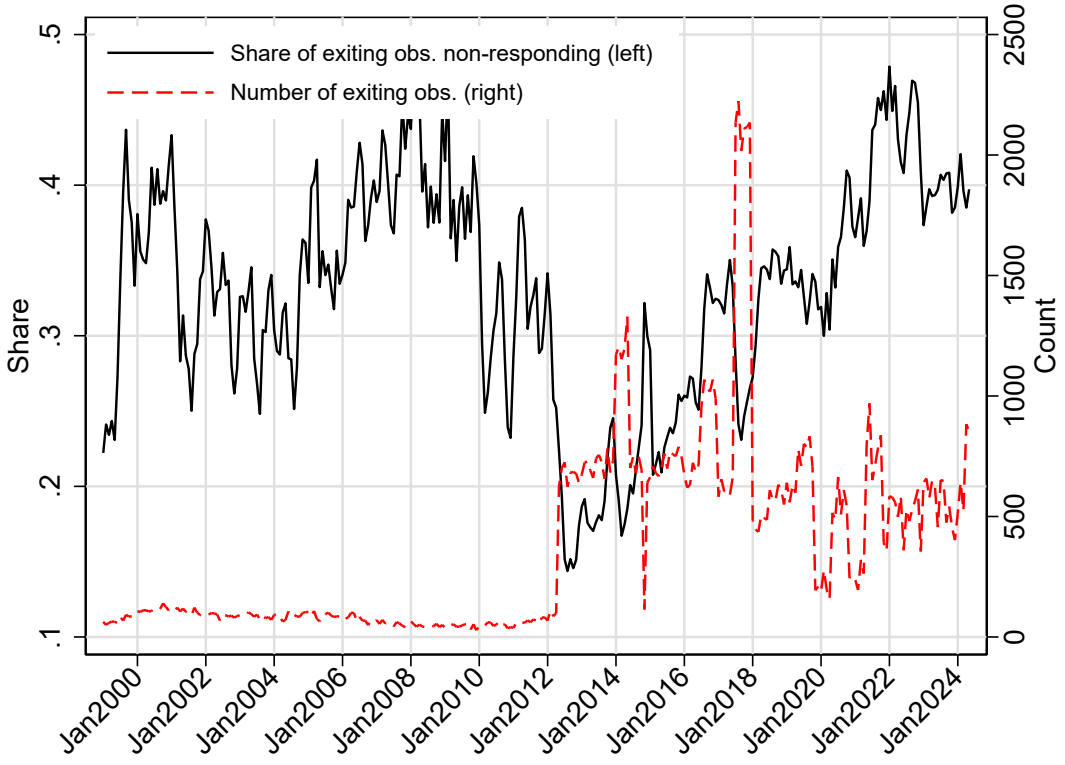


Figure 6. Non-response Share of Units. *Note:* A unit is defined as exiting if it is surveyed for the last time before January 2025. The weighted share of exits that are non-responses has been smoothed with a 3-month moving average. The unweighted count displays the total number of units exiting. *Source:*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

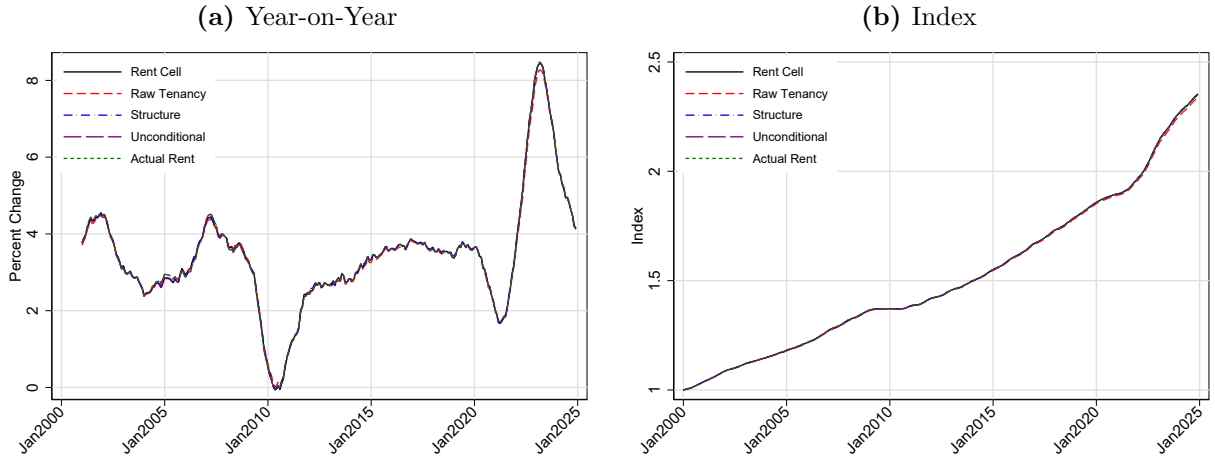


Figure 7. Alternative Imputation Inflation Rates. *Note:* The time series are created by using a class means method and picking different covariates to partition observations by. See text for more details. *Source:* Authors' calculations using the BLS Housing Survey.

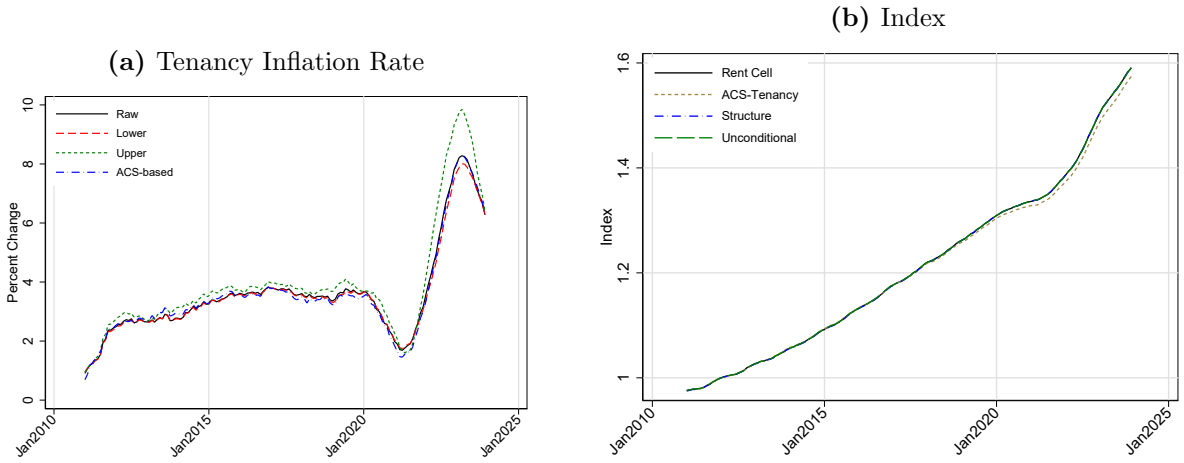


Figure 8. ACS-Tenancy Inflation Comparison *Note:* The time series are created by using a class means method using different measures of tenure length for nonresponses. *Source:* Authors' calculations using the BLS Housing Survey, ACS.

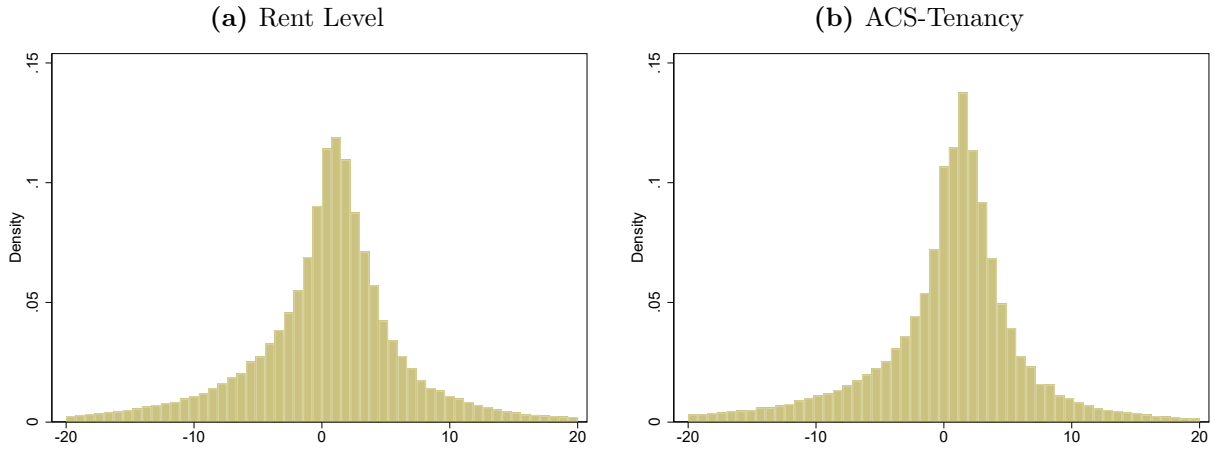


Figure 9. Jackknife Imputation Error by Imputation Method. *Note:* The imputation error is defined as the difference in percent between the imputed rent and collected rent. The histogram is truncated at -20% and 20%. Observations are unweighted. Observations are from January 2019 to December 2024. See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

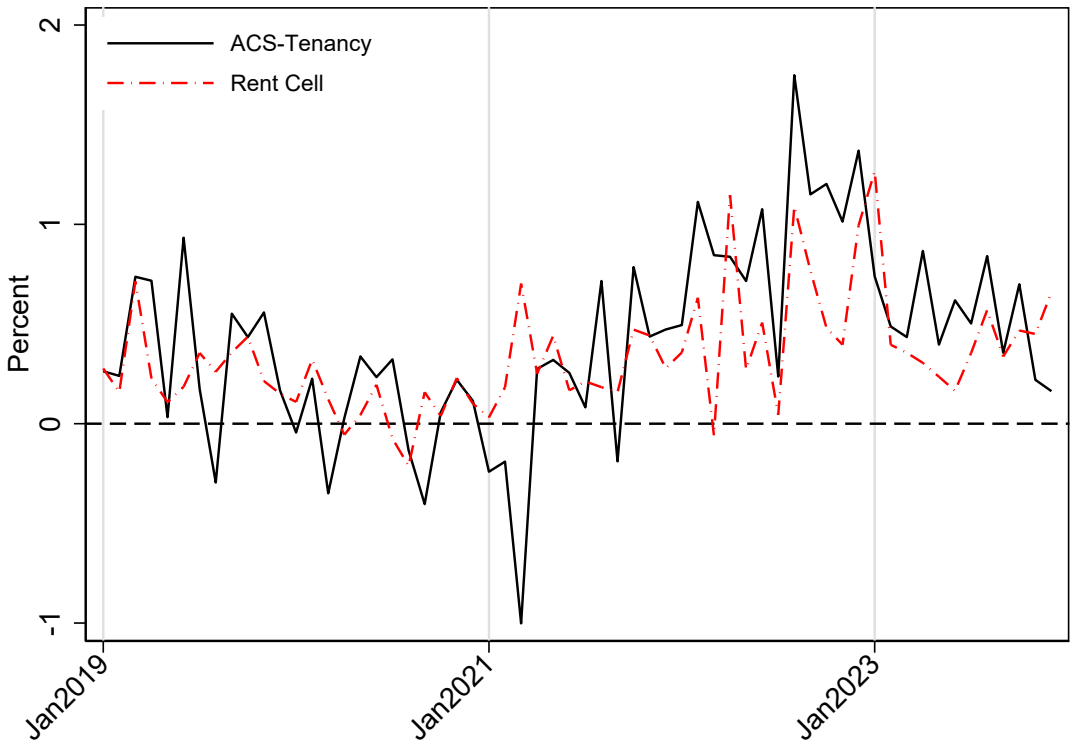


Figure 10. Jackknife Average Imputation Error. *Note:* The imputation error is defined as the difference in percent between the imputed rent and collected rent. Errors are averaged across cities using weights to obtain nationally representative series. *Source:* Authors' calculations using the BLS Housing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

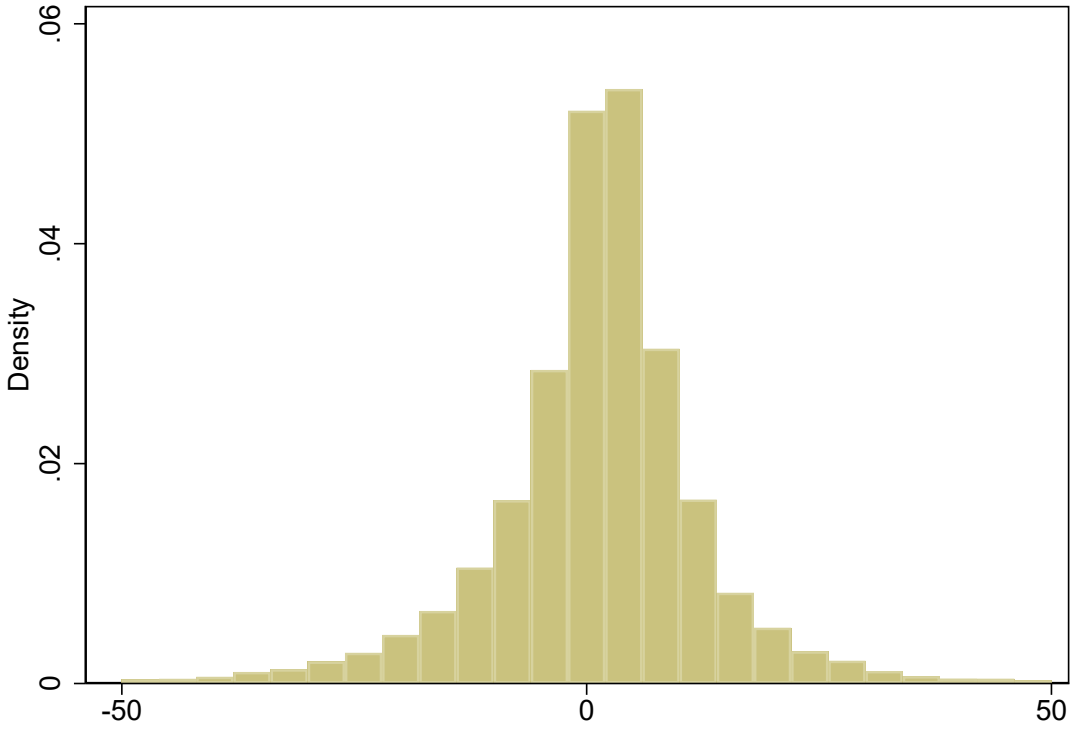


Figure 11. Re-entered Unit Imputation Error Histogram. *Note:* The imputation error is defined as the difference in percent between the imputed rent and collected rent. The histogram is truncated at -20% and 20%. Observations are unweighted. Observations are from January 2019 to December 2024. See text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