Online Appendix for

"The Renovation Rebalance:
How Financial Intermediaries Affect Renter Housing Costs"

This document has accompanying material for the paper "The Renovation Rebalance: How Financial Intermediaries Affect Renter Housing Costs".

D Additional Implications of Improving Quality

This section describes additional implications of improving quality referenced in Appendix C of the paper.

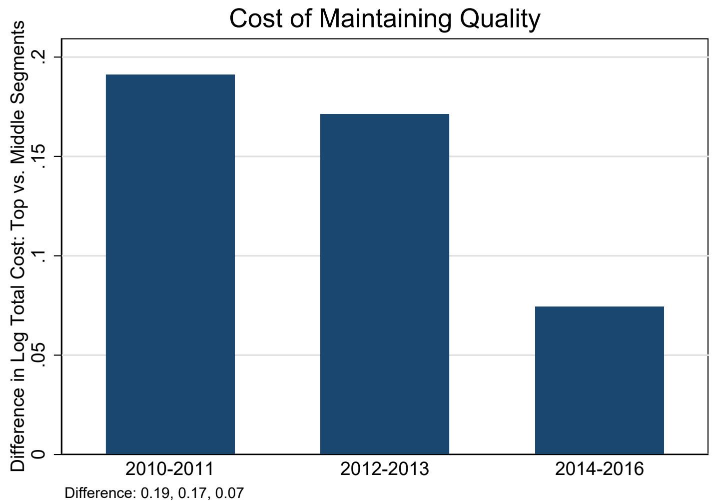
D.1 Implications for Innovation in Property Management

This extension presents stylized evidence that the routine costs of maintaining a unit's quality have fallen, coinciding with innovation in property management. The price-relevant dividend stream should, in principle, account for maintenance costs. Figure C11 plots the log difference in the median total cost of maintaining a unit in the top quality segment, measured by the MBA/CREFC rating, and that of a unit in the middle two segments.[82] The costs of maintaining a top quality unit are greater than those of average quality ones, which is fairly intuitive and can be understood as the result of geometric depreciation. This quality cost premium fell from around 13% over 2010-2011 to 7% over 2014-2016. The decline may be related to the contemporaneous decline in the quality rent premium and increase in the share of top quality units, described in the context of Figure 1 from the paper. For example, owners of top quality units compete until rent falls within a required margin over maintenance costs.

While it is difficult to infer too much from this decline, the 2010-2016 period also saw a burst of innovation in property management. Figure C12 documents this fact, measuring innovation by patent issuance and private investment in property management technology firms. One conjecture is that declining quality maintenance costs are related to technological advances in property management through endogenous innovation. Applying the Acemoglu and Linn (2004) logic to the housing sector, it a growing market for quality housing may have stimulated innovation in property management.

[82] Total costs include operating and capital expenses.

Figure C11: Cost Ratio of Managing Higher to Lower Quality Units



Note: This figure plots the difference in log median total cost of maintaining a unit in the top quality segment and that of a unit in the middle 2 quality segments. Quality segments are defined by the MBA/CREFC rating. Total costs include both operating and capital expenses. Data are from Trepp.

However, it is difficult to identify this channel in the current setting since I lack cross-industry variation that would typically be used.[83]

D.2 Implications for Official Rent Indices

This section relates the results from Section 3 of the paper to what one would obtain from an age adjustment procedure similar to that used by statistical agencies.[84] Following Gallin and Verbrugge (2007), I define the age adjustment regression as

$$\log\left(\text{Rent}_{i,t}\right) = \gamma\left(\text{Age}_{i,t}; X_{i,t}\right) + u_{i,t}, \tag{C14}$$

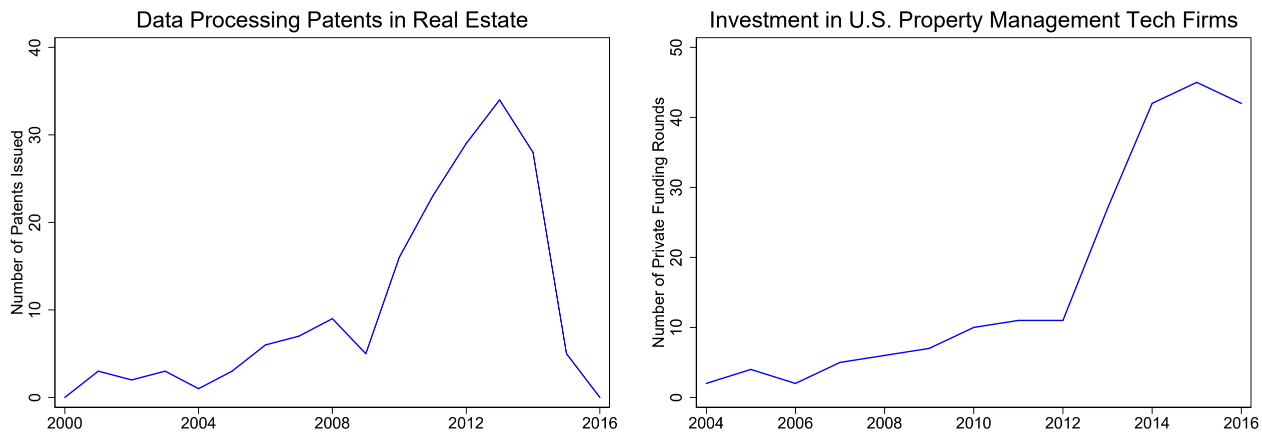
where $\text{Rent}_{i,t}$, $\text{Age}_{i,t}$, and $X_{i,t}$ are, respectively, a unit's rent, the age of the property, and a vector of structural features.[85] Then, one computes a unit's age adjusted rent as $\text{Rent}^A_{i,t} \equiv \text{Rent}_{i,t} e^{-\frac{\partial \gamma}{\partial \text{Age}_{i,t}}}$ and

[83] See Acemoglu and Linn (2004) or Jaravel (2018) for example research designs.

[84] Age is the primary attribute the Bureau of Labor Statistics (BLS) corrects for when computing the Rent of Primary Residence (Ptacek 2013). The other corrections pertain to the changes in the inclusion of parking or utilities in rent, and the addition of a new room or central air conditioning.

[85] The function $\gamma\left(\text{Age}_{i,t}; X_{i,t}\right)$ approximates that used by the BLS as closely as possible given a different dataset. It includes age, its square, and its interaction with: the number of units in the property and an indicator for whether the property is over 85 years old. Since I do not observe a unit's location and thus neighborhood features in the AHS data, I estimate (C14) as a panel regression and include a property fixed effect. When using the Trepp data, I weight

Figure C12: Technological Growth in Property Management



(a) Patent Issuance

(b) Property Management Technology

Note: This figure plots two measures of innovation in property management. Panel (a) plots the number of patents issued each year classified as data processing for business management in real estate, property management, and land development. Panel (b) plots the number of private capital fundraising rounds for property management technology firms. Data in panels (a) and (b) come from the USPTO and Crunchbase, respectively.

aggregates $\text{Rent}_{i,t}^{A}$ across units to produce an average rent π_t^A that is benchmarked to the reference period, similarly to the expression for the hedonic index.

$$\pi_t^A = \frac{\sum_{i \in \mathcal{I}} \text{Rent}_{i,t}^{A}}{\sum_{i \in \mathcal{I}} \text{Rent}_{i,t_0}}, \tag{C15}$$

The remainder of this section theoretically compares the structural and age adjusted rent indices π_t^S and π_t^A. Beginning with the setup described in the context of the structural index from Appendix C of the paper, consider the following exercise. Suppose a unit's equilibrium rent is a function of its quality $h_{i,t}$.

$$\begin{aligned} \log\left(\text{Rent}_{i,t}\right) &= a_i + P \log\left(h_{i,t}\right) + u_{i,t} \\ &= a_i + P \left[\log\left(\hat{h}_{i,t}\right) + \log\left(H_t\right)\right] + u_{i,t}, \end{aligned} \tag{C16}$$

where the notation is the same as in the paper with the addition of time subscripts, and $u_{i,t}$ is an iid shock. In particular, H_t is the highest quality in the market at t, which I call absolute quality, and $\hat{h}_{i,t} \equiv \frac{h_{i,t}}{H_t}$ is the relative quality of unit i. The parameter P is the equilibrium slope of the quality

observations in (C14) by number of units because the data are at the property-level.

ladder, or price of quality. A unit's relative quality is also a function of its age,

$$\log\left(\hat{h}_{i,t}\right) = -\delta \text{Age}_{i,t} + v_{i,t}, \tag{C17}$$

where δ is the rate of natural depreciation, and $v_{i,t}$ is not necessarily iid. The intuition is similar for more complicated depreciation schedules than (C17). For the sake of argument, suppose all parameters in (C16)-(C17) are known, but quality $h_{i,t}$ is not observed. Let $\tilde{\mathbb{E}}$ denote a cross-sectional expectations operator: $\tilde{\mathbb{E}}\left[z_{i,t}\right] = \mathbb{E}\left[z_{i,t}|t\right]$. Then the age adjusted index (C15) can be rewritten

$$\pi_t^A = e^{\delta P} \frac{\tilde{\mathbb{E}}\left[\text{Rent}_{i,t}\right]}{\tilde{\mathbb{E}}\left[\text{Rent}_{i,t-1}\right]}. \tag{C18}$$

The relationship between π_t^S and π_t^A is described by the following proposition.

Proposition D.1 (Bias in Age Adjusted Rent) *Suppose the total number of housing units is held fixed. Then age adjusted rent growth π_t^A is biased upward compared to the structural rent index π_t^S according to*

$$\frac{\pi_t^A}{\pi_t^S} = \left(\frac{H_t}{H_{t_0}}\right)^P \times \frac{\tilde{\mathbb{E}}\left[Rent_{i,t}\right]}{\tilde{\mathbb{E}}\left[Rent_{i,t} \times e^{-P\Delta v_{i,t}}\right]}. \tag{C19}$$

The result in (C19) decomposes bias in an age adjustment into two terms.[86] The intuition for the first term in (C19) is that an age adjustment accounts for relative quality, not absolute quality. A top-tier unit 2010 may have lost no or very little quality by 2014, but it will still rent at a discount compared to a unit renovated to top-tier standards in 2014 if there is growth in absolute quality, $H_{2014} > H_{2010}$. The intuition for the second term is that age is an imperfect proxy for quality in the

[86] The proof of Proposition D.1 is as follows:

Proof Let $\bar{R}_{t_0}$ and $\bar{U}$ denote aggregate rent expenditure and welfare at t_0. Let $\bar{R}_t$ denote the minimized cost function in t from Lemma G.2 when target utility is $\bar{U}$. By definition, $\bar{R}_t$ is the aggregate expenditure required to maintain welfare at $\bar{U}$. Since the number of units is unchanged, $\bar{R}_t$ equals aggregate expenditure at t when holding each unit's quality fixed at its t_0 level,

$$\begin{aligned}\bar{R}_t &= \tilde{\mathbb{E}}\left[R_{i,t} e^{P\left[\Delta \log\left(\hat{h}_{i,t}\right) + \Delta \log(H_t)\right]}\right] \\ &= \tilde{\mathbb{E}}\left[R_{i,t} e^{P\left(\delta - \Delta v_{i,t} - \Delta \log(H_t)\right)}\right],\end{aligned} \tag{C20}$$

where the second line uses (C17). Define $\pi_t \equiv \frac{\bar{R}_t}{\bar{R}_{t_0}}$. From the proof of Proposition 1 from the paper, $\pi_t^S = \pi_t$. Dividing each side of (C20) by $\bar{R}_{t_0}$ gives expression in (C19),

$$\frac{\pi_t^A}{\pi_t^S} = \left(\frac{H_t}{H_{t_0}}\right)^P \times \frac{\tilde{\mathbb{E}}\left[\text{Rent}_{i,t}\right]}{\tilde{\mathbb{E}}\left[\text{Rent}_{i,t} \times e^{-P\Delta v_{i,t}}\right]}$$

presence of improvement activity. To see this, first note from the depreciation process (C17) that an improvement would generate a large disturbance term $\Delta\nu_{i,t} > 0$. Furthermore, because improvements move a unit up the quality ladder, there is positive covariance between $\Delta\nu_{i,t}$ and $\text{Rent}_{i,t}$. Together, these two features would make the second term in (C19) greater than 1, leading to upward bias.

Recall that this exercise assumes knowledge of all parameters. If δP is unknown but estimated through OLS, then classical measurement error from the fact that age is an imperfect measure for relative quality would bias the estimated δP toward zero. This attenuation bias would partially offset the upward bias in (C19).

Proposition D.1 lays out two avenues for addressing upward bias in conventional age adjustments. First, collecting and incorporating information on improvement activity could reduce the measurement error from proxying quality with age. This could be accomplished by interviewing landlords, as opposed to the current practice of interviewing tenants. Second, addressing the bias from growth in absolute quality requires additional methodological tools. Measuring changes in absolute quality is not this paper's core contribution, but several papers in the price adjustment literature (e.g. Redding and Weinstein 2018) are making progress on this margin.

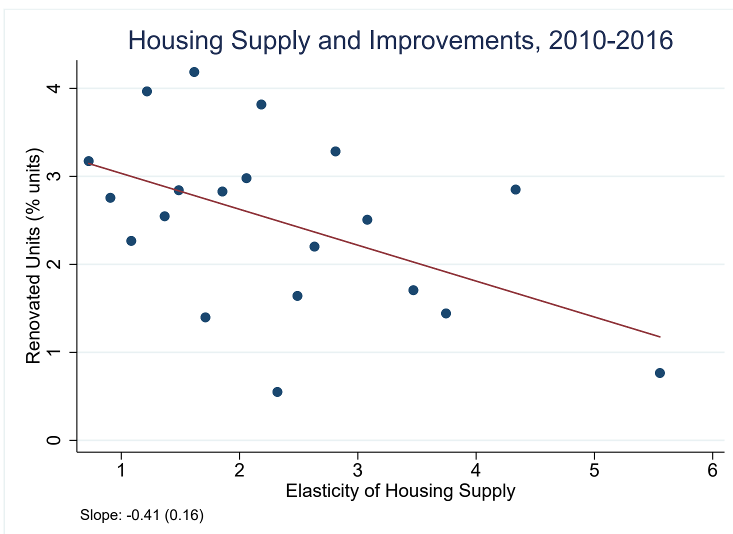
E Additional Facts About Improving Quality

This section describes additional facts about improving quality referenced in Section 2 of the paper.

E.1 Improvements in the Cross-Section

Figure C13 asks where improvements have occurred. It plots the annualized share of multifamily units in an MSA that were renovated over 2010-2016 against the MSA's Saiz (2010) elasticity of housing supply. Low values of this elasticity capture natural or regulatory constraints that make it difficult to build new housing units. Improvement activity is more intense in such MSAs, consistent with real estate investors substituting from construction to improvement projects. Section 4 studies a credit, in contrast to physical, supply shock that also induced such a substitution.

Figure C13: Improvements by Housing Supply



Note: This figure plots the average yearly probability of renovation for multifamily rental units in an MSA over 2010-2016 against the MSA's Saiz (2010) elasticity of housing supply. The plot is binned. Data are from Trepp.

There are other interpretations of Figure C13, which I explore in Table C15. I regress the share of renovated units in an MSA against the MSA's log average income, college education share, and an indicator for whether rent control or stabilization practices are in place.[87] The significance of the relationship between Renovation Probability$_m$ and the MSA's elasticity of housing supply is weaker than in Figure C13. Moreover, the significance of other variables in the table suggests that supply

[87] All variables are normalized to have unit variance.

Table C15: MSA Correlates with Improvement Activity

Outcome:	Renovation Probability$_m$
$\log(\text{Saiz Elasticity}_m)$	-0.120*
	(0.070)
$\log(\text{Income}_m)$	0.108*
	(0.056)
Rent Control$_m$	-0.071**
	(0.029)
College Education$_m$	-0.044
	(0.075)
R-squared	0.048
Number of Observations	211

Note: Subscript m denotes MSA. Renovation Probability$_m$ is the share of multifamily units that were renovated between 2010 and 2016. Saiz Elasticity$_m$ is the elasticity of housing supply as estimated by Saiz (2010). Income$_m$ is average real income per capita over 2010-2016. College Education$_m$ is the share of inhabitants with a bachelor's degree in 2010. Rent control$_m$ indicates if the MSA has rent control or stabilization policies. All variables are normalized to have unit variance. Observations are MSAs weighted by number of multifamily units over 2010-2016. Heteroskedasticity robust standard errors are in parentheses. Data are from Trepp and other data sources described in Appendix B.

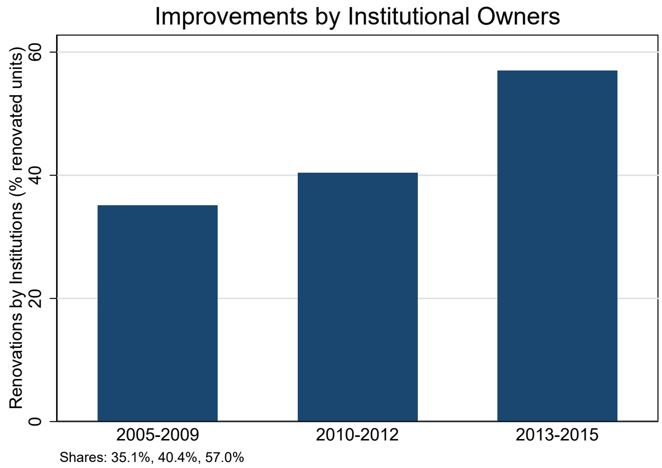
elasticities absorb other sources of variation which affect improvement activity.

For example, Table C15 shows how the renovation share is lower in MSAs with rent control. This correlation is quite intuitive, since rent control directly counteracts investors' reward for making improvements, and it is related to the Saiz (2010) insight that supply elasticities depend regulatory as well as physical constraints. In addition, there is a positive, though somewhat weak correlation between income and improvement activity. This correlation is consistent with the results of Section 3.3, specifically the finding that higher income households have a greater willingness to pay for quality. As pointed out by Davidoff (2016), supply elasticities are correlated with measures of local demand like income.

Collectively, one way to interpret the correlations in Figure 1 and Table C15 is that investors face a tradeoff between construction and improvements. Where it is more difficult to build, they tilt their activity towards improvements, provided policies like rent control do not make this strategy unattractive. Moreover, if the marginal cost of quality is either flat, then investors make improvements in higher income areas where the price of quality is higher, as estimated in Section 3.3.[88] However, this interpretation is only suggestive and it is beyond the scope of this paper to explore these effects in an equilibrium model.

[88] A flat marginal cost of quality is consistent with the cost curve of countertop installation. This figure is available upon request.

Figure C14: Improvements by Institutional Owners



Note: This figure plots the percent of renovated rental units owned by REITs, real estate corporations, partnerships, or LLCs over the indicated periods. Data are from the RHFS.

E.2 Improvements by Institutional Owners

Figure C14 studies who have performed improvements. The figure plots the percent of renovated rental units owned by for-profit institutions, obtained from the Rental Housing Finance Survey (RHFS).[89] The large and growing role of institutional landlords is consistent with Lambie-Hanson, Li and Slonkosky (2018).

[89]For-profit institutions are defined as REITs, real estate corporations, partnerships, or LLCs.

F Econometric Details

This appendix has econometric details related to the structural rent index. The structural index is described at a high level in Section 3 of the text and more fully in Appendix C. All material in this appendix is referenced in Appendix C.

F.1 Calculation of Structural Rent Index

Calculating the rent index π_t^S in (C5) requires two pieces of information: (1) a ranking variable to partition the sample into segments $\{\hat{h}\}$, and, relatedly, an identifier for which units are top-end in period t and thus have quality H_t; and (2) the preference parameter σ. I obtain these objects from the Trepp dataset and compute π_t^S over 2010-2016. As described in Appendix B, this dataset covers a geographically representative sample of multifamily properties, and it is particularly appealing for this exercise due to its detailed information on property upgrades, inspection ratings, and renovations which are collected as part of the multifamily mortgage servicing process.

First, to rank properties I use the MBA/CREFC property inspection rating. This rating captures a property's quality relative to a newly built unit and is obtained as part of the standard multifamily mortgage servicing protocol. The resulting index is robust to alternative partitioning variables such as the property's effective age. Importantly, because $\hat{h}$ does not directly appear in (C5), this rating only needs to identify a property's rank and does not need to be accurate in the cardinal sense.

To classify top-end properties I use a combination of renovation and inspection data. I classify a unit as having quality H_t in period t if it is in a property that was newly built or renovated and first on the market in year t, and if it also was ranked in the top MBA/CREFC quality segment after construction or renovation.[90] Under the assumption that top-end units in year $t-1$ retain their absolute quality H_{t-1} through at least year t, one can use the number of units with quality H_t (i.e. in segment 1 at t) and H_{t-1} (i.e. in segment 1_0 at t) and their respective revenue shares in year t to compute GQ_t according to the expression in (C5).[91] To account for the possibility that renovations and construction only occur in certain areas in year t, I also compute GQ_t within each zip code-year

[90] There is typically a 1-year lag between completion of construction and renovation and being rent-ready.

[91] Specifically, I compute ΔGQ_s for all $s = 1, ..., t$ and then take their product to obtain GQ_t. That is, I chain growth in absolute quality. When taking (C5) to the data, I will use the theoretically equivalent expression for growth in absolute quality, $GQ_t = \left(\frac{\text{Expend}_{1,t}}{\text{Expend}_{1_0,t}}\right)^{\sigma}\left(\frac{\text{Share}_{1,t}}{\text{Share}_{1_0,t}}\right)^{1-\sigma}$, where $\text{Expend}_{\hat{h},t}$ is the share of aggregate rent expenditure on segment $\hat{h}$. Doing so reduces measurement error from unit level rent. Feenstra (1994) also relies on expenditure shares when possible, since they are subject to less measurement error.

bin and then average across zip codes that year, which yields very similar results.[92]

Second, I must estimate σ, which I do using three methodologies: a property-level credit supply shock using idiosyncratic variation in payment timing similarly to Section 4.2, a zip code level version of the property-level shock, and the Feenstra (1994) GMM estimator. I now describe these three methodologies. Before doing so, Table E1 summarizes the estimated σ from each of them. The average estimate is 6.5, and Figure E1 helps interpret this magnitude by performing an introspective exercise.

Table E1: Estimated Preferences for Quality

	Estimated σ		
Specification:	Property IV	Zip Code IV	GMM
Estimate	6.5	6.9	6.1
Confidence Interval	[2.8, 18.2]	[1.2, 19.6]	[4.1, 9.4]

Note: This table shows the estimated elasticity of substitution σ for various methodologies. Property IV, Zip Code IV, and GMM are discussed in Sections F.1.1, F.1.2, and F.1.3 below. Bootstrapped 95% confidence intervals are shown in brackets. Data are from Trepp.

Table E2 summarizes the estimates after partitioning the sample into income cohorts as described in Appendix C.3. The highest income zip codes have the lowest value of σ (4.9) and thus the highest willingness to pay for quality.[93]

Table E2: Estimated Preferences for Quality by Income

	Estimated σ		
Specification:	Property IV	Zip Code IV	GMM
Full Sample	6.5	6.9	6.1
By Real Income 2010-2016:			
Bottom 30%	5.7	14.9	8.6
Middle 35%	17.9	25.0	5.0
Upper 35%	4.7	3.7	6.2

Note: This table shows the estimated elasticity of substitution σ for various methodologies and income groups. Property IV, Zip Code IV, and GMM are discussed in Sections F.1.1, F.1.2, and F.1.3 below. Data are from Trepp.

[92] Growth in the rent index π_t^S is 1.5% as opposed to 1.7% under the baseline.

[93] There is an apparent non-monotonicity in σ with respect to income bracket. This likely reflects the fact that σ is both the inverse willingness to pay for quality and the market level substitutability across segments. The estimated σ in very low income zip codes may be small if the data reveal limited movement across segments. By analogy, with CRRA preferences there is not a distinction between the coefficient of relative risk aversion and the inverse intertemporal elasticity of substitution.

Figure E1: Example Willingness to Pay for Quality

(a) Top Tier: \$1,600

(b) Above Average General Market: \$1,448

(c) Below Average General Market: \$971

(d) Bottom Tier: \$625

Note: This figure shows the willingness to pay for quality assuming the rent of a unit in the bottom quality segment is \$625 per month, the elasticity of substitution across quality segments is $\sigma = 6.5$, and the properties only differ in their structural quality. The photographs are from the website of a large real estate investor in the Dallas, TX and San Antonio, TX markets. Panels (a)-(d) show properties corresponding to the 4 bins of the MBA/CREFC rating used in this paper, respectively. Moving from the top segment to the bottom segment entails consecutive reductions of 0.6, 2.6, and 2.9 log points of relative quality, respectively. Given the average estimated σ of 6.5, this implies a willingness to pay of 10% for "highest current market standards" versus "above average" (i.e. panel (a) vs. panel (b)); 40% for "minimal" versus "general" wear and tear (i.e. panel (b) vs. panel (c)); and 44% for "no" to "some" or "multiple" life safety violations (i.e. panel (c) vs. panel (d)).

F.1.1 Estimating σ: Property-Level Credit Supply Shock

This methodology estimates σ insofar as it is the inverse marginal willingness to pay for quality. Using Lemma G.2, $\frac{1}{\sigma}$ is the elasticity of rent with respect to quality, holding the distribution of units across quality segments constant, $\frac{1}{\sigma} = \frac{\partial \log(\text{Rent}_{\hat{h},t})}{\partial \log(\hat{h})}$.

As described in Section 4.2, the structure of most multifamily mortgage contracts generates spikes in improvement activity. Combining this institutional feature with the effectively exogenous variation in their due date established in panel (b) of Figure C7, I construct an instrument for the change in log relative quality $\log(\hat{h})$. Suppose now that the quality of units in property i evolves according to

$$\log(h_{i,t}) = \log(h_{i,t-1}) + \log\left(\text{Improvements}_{i,t}\right) - \delta_{i,t}, \tag{E1}$$

where $\delta_{i,t}$ is a depreciation shock. As discussed above and shown in Table C7, having an impending loan due reduces the probability of making a quality improvement, lowering $\text{Improvements}_{i,t}$ and thus $\Delta \log(h_{i,t})$ in (E1). Reviewing the discussion from Section 4.2, this is because most multifamily mortgages are balloon loans which require renewal at the end of every loan term, with a modal term of 10 years. Moreover, refinancing is generally not an option and must be done through a process of defeasance.[94] Because of the possibility of cheaper borrowing costs after renewal, one would expect that having an impending loan due covaries negatively with a unit's change in quality.

Mapping to a regression equation, I estimate the system

$$\log(\text{Rent}_{i,z,t}) = \frac{1}{\sigma}\Delta \log\left(\text{Quality}_{i,z,t}\right) + \beta_0 \log\left(\text{Quality}_{i,z,t-1}\right) + a_i + a_{z,t} + u_{i,z,t} \tag{E2}$$

$$\Delta \log\left(\text{Quality}_{i,z,t}\right) = \tilde{\beta}_0 \text{Impending}_{i,z,t} + \beta_1 \log\left(\text{Quality}_{i,z,t-1}\right) + \tilde{a}_i + \tilde{a}_{z,t} + \tilde{u}_{i,z,t}, \tag{E3}$$

where i, z, and t index property, zip code, and year, and $\text{Impending}_{i,z,t}$ indicates if the property's loan is due in t or $t{+}1$.[95] Relative quality $\hat{h}_{i,z,t}$ is denoted $\text{Quality}_{i,z,t}$ and measured using the MBA/CREFC rating, as in Section 4.2. The second-stage equation is (E2), and its first stage (E3) is similar to that used in (C10) from Section 4.2.[96] However, I do not include information about the lender because my interest is on average improvement activity, not its dependence on whether the

[94] Defeasance is a fairly complicated process in which the borrower must exchange the loan for another security of equal maturity, such as a Treasury.

[95] I weight observations in (E2)-(E3) by number of units because the Trepp data are at the property-level.

[96] Relative to (C10), equation (E3) restricts $\tau \leq 0$ and $\text{Impending}_{i,z,t} = \max\{\text{Due}_{i,z,t}, \text{Due}_{i,z,t+1}\}$.

Table E3: Substitutability Across Quality Rungs with Property-Level IV

	First Stage	Second Stage
Outcome:	$\Delta \log\left(\text{Quality}_{i,t}\right)$	$\log\left(\text{Rent}_{i,t}\right)$
	(1)	(2)
$\text{Impending}_{i,t}$	-0.208**	
	(0.035)	
$\Delta \log\left(\text{Quality}_{i,t}\right)$		0.153**
		(0.072)
$\log\left(\text{Quality}_{i,t-1}\right)$	-0.805**	0.132**
	(0.008)	(0.058)
Estimator	OLS	2SLS
Property FE	Yes	Yes
Zip Code-Year FE	Yes	Yes
First Stage F		34.661
Number of Observations	67210	67210

Note: Subscripts i and t denote property and year. Columns 1 and 2 estimate (E3) and (E2), respectively. $\text{Quality}_{i,t}$ is relative quality based on the MBA/CREFC property inspection rating. $\text{Impending}_{i,t}$ indicates if the investor has a mortgage due in year t or year $t+1$. The estimator in column 2 is 2SLS, and the instrument for $\Delta \log\left(\text{Quality}_{i,t}\right)$ is $\text{Impending}_{i,t}$. Observations are property-years weighted by number of units. The sample period is 2010-2016. Standard errors are in parentheses. Data are from Trepp.

lender was affected by HVCRE regulation.

The zip code-year fixed effect $a_{z,t}$ absorbs local demand effects that would otherwise affect rent. Thus, any violation of the exclusion restriction due to expectations of future growth would need to require sub-zip code variation in demand. Also, as shown in panel (b) of Figure C7, having an impending loan due is uncorrelated with interest rate spreads or other measures of credit risk. The property fixed effects a_i absorb amenities.

Column 1 of Table E3 has the results of the first stage regression (E3). Like in Figure C7 of the text, having an impending loan leads to a deterioration in relative quality. The second stage in column 2 implies $\sigma = 6.5$, based on the point estimate of 0.15 on $\Delta \log\left(\text{Rel Quality}_{i,t}\right)$.[97]

F.1.2 Estimating σ: Zip Code Level Credit Supply Shock

If CES market demand is a poor approximation, one might be concerned that the previous strategy does not identify the appropriate parameter because it relies on a highly misspecified functional form. To address concerns about functional form, I propose a second strategy which estimates σ insofar

[97] When computing the bootstrapped standard errors in Table E1 and the heterogeneous preference parameters in Table E2, I estimate (E2) through a constrained optimization such that the implied value of σ lies in the interval [1, 25]. This remark also applies when using the zip code level credit supply instrument and the Feenstra (1994) GMM estimator.

as it is the aggregate elasticity of substitution across quality segments, and which thus obtains identification through a different functional form that is nonetheless consistent with the CES market demand structure. The source of variation is similar to the previous strategy, and the instrument used is a zip code level share of property owners with an impending loan due.

Using Lemma G.2, the CES market demand curve can be written

$$\log\left(\frac{\text{Expend}_{\hat{h}}}{\text{Expend}_{\hat{h}_0}}\right) = \left(1-\frac{1}{\sigma}\right)\log\left(\frac{\text{Share}_{\hat{h}}}{\text{Share}_{\hat{h}_0}}\right) + \frac{1}{\sigma}\log\left(\frac{\hat{h}}{\hat{h}_0}\right), \tag{E4}$$

where, using the notation introduced in Appendix G, $\text{Expend}_{\hat{h}}$ is the aggregate share of rent expenditure on segment $\hat{h}$. I then estimate the following system through 2SLS,

$$\log\left(\frac{\text{Expend}_{\hat{h},z,t}}{\text{Expend}_{\hat{h}_0,z,t}}\right) = \left(1-\frac{1}{\sigma}\right)\log\left(\frac{\text{Share}_{\hat{h},z,t}}{\text{Share}_{\hat{h}_0,z,t}}\right) + \gamma X_{z,t} + a_{\hat{h}} + a_{m,t} + u_{\hat{h},z,t} \tag{E5}$$

$$\log\left(\frac{\text{Share}_{\hat{h},z,t}}{\text{Share}_{\hat{h}_0,z,t}}\right) = \tilde{\beta}_0 \text{Impending}_{z,t} + \gamma X_{z,t} + \tilde{a}_{\hat{h}} + \tilde{a}_{m,t} + \tilde{u}_{\hat{h},z,t} \tag{E6}$$

where z indexes zip codes, t indexes years, m indexes MSAs $\hat{h}$ indexes quality segments according to the MBA/CREFC score; $\hat{h}_0$ is the reference segment, which I set at the lowest quality segment on the MBA/CREFC rating scale; the segment fixed effects $\alpha_{\hat{h}}$ and $\tilde{\alpha}_{\hat{h}}$ absorb the second term in (E4); and $\text{Impending}_{z,t}$ is the fraction of units in zip code z whose owner has a loan due in t or $t+1$, and it is a zip code level average of the instrument from the property-level system (E2)-(E3).[98] The equation of interest is the second stage (E5), and its first stage is (E6).

The MSA-year fixed effect $a_{m,t}$ and its first-stage counterpart in (E6) restrict variation within MSA m in which z is located and year t. The zip code controls include measures of investors' financial condition and local demand.[99] Since all variation comes from within MSA-year bins, one can think of (E5)-(E6) as comparing the timing of when most investors in a zip code took out their loan. Thus, the instruments are predetermined as of time t and do not capture contemporaneous demand shocks. However, it is plausible that the timing of borrowing decisions and the resulting interest rate reflected expectations about future demand in a given zip code with an MSA. These expectations

[98] Specifically, $\hat{h}_0$ corresponds to a raw MBA/CREFC rating of 3, 4, or 5. I group these three segments together because they are the lowest collective segment observed in all zip codes and years. See Appendix B for full details on the interpretation of MBA/CREFC ratings.

[99] Financial controls are the average interest rate spread, securitization rate, and log term for loans on units in zip code z and year t. Demand controls are the log average income, log population, and fraction of households with social security benefits, capital gains, dividend income, and children in z and t, all based on IRS tax returns. See the footnote to Table E4 for more details on how these variables are proxied using the IRS data.

Table E4: Local Demand, Financial Condition, and the Timing of Loan Renewal

Outcome:	$\text{Impending}_{z,t}$
$\log\left(\text{Income}_{z,t}\right)$	0.002
	(0.005)
$\log\left(\text{Population}_{z,t}\right)$	-0.002
	(0.002)
$\text{Stock Ownership}_{z,t}$	-0.228
	(0.166)
$\text{Family Households}_{z,t}$	0.110
	(0.088)
$\text{Social Security Benefits}_{z,t}$	-0.060
	(0.060)
$\text{Capital Gains Income}_{z,t}$	0.209
	(0.190)
$\text{Rate Spread}_{z,t}$	0.117
	(0.145)
$\text{Securitized}_{z,t}$	0.000
	(0.004)
$\log\left(\text{Term}_{z,t}\right)$	-0.037**
	(0.003)
MSA-Year FE	Yes
R-squared	0.210
Number of Observations	9282

Note: Subscripts z and t denote zip code and year. $\text{Impending}_{z,t}$ is the fraction of units whose investor has a mortgage due in year t or t+1. $\text{Income}_{z,t}$ is average income per tax return. $\text{Population}_{z,t}$ is number of tax returns. $\text{Stock Ownership}_{z,t}$ is the fraction of households with dividend income. $\text{Family Households}_{z,t}$ is the fraction of returns with a child tax credit. $\text{Social Security Benefits}_{z,t}$ is the fraction of returns with social security income. $\text{Capital Gains Income}_{z,t}$ is the fraction of returns with capital gains. $\text{Rate Spread}_{z,t}$ is the average difference between the loan's current interest rate and the average loan interest rate the year of origination or renewal. $\text{Securitized}_{z,t}$ is the fraction of units whose loan was securitized within 3 months of origination. $\text{Term}_{z,t}$ is the average loan term in months. The sample period is 2010-2016. Standard errors are in parentheses.

would be reflected ex post in measures of local demand, or ex ante in the loan's rate spread or initial securitization status. To investigate this possibility, I project the instrument $\text{Impending}_{z,t}$ onto the control vector $X_{z,t}$. The results in Table E4 show that the only significant partial correlation with $\text{Impending}_{z,t}$ is the mechanical effect of having a shorter term. This finding suggests that, within the same MSA-year bin, $\text{Impending}_{z,t}$ does not reflect expectations about local demand.

Table E5 has the results of (E5)-(E6). Column 1 has the estimates from the first stage (E6). Consistent with the property-level specification, zip codes where more property owners have an impending loan due see fewer units in segments $\hat{h} > \hat{h}_0$, recalling that the reference segment $\hat{h}_0$ is the lowest on the MBA/CREFC rating scale. In these zip codes, there is a compositional shift toward lower quality units. The second stage results in columns 2-4 show how this compositional shift

Table E5: Substitutability Across Quality Rungs with Zip Code Level IV

	First Stage	Second Stage		
Outcome:	$\log\left(\frac{\text{Share}_{\hat{h},z,t}}{\text{Share}_{\hat{h}_0,z,t}}\right)$	$\log\left(\frac{\text{Expend}_{\hat{h},z,t}}{\text{Expend}_{\hat{h}_0,z,t}}\right)$		
	(1)	(2)	(3)	(4)
$\text{Impending}_{z,t}$	-0.538**			
	(0.173)			
$\log\left(\frac{\text{Share}_{\hat{h},z,t}}{\text{Share}_{\hat{h}_0,z,t}}\right)$		0.884**	0.857**	0.856**
		(0.143)	(0.138)	(0.146)
MSA-Year FE	Yes	Yes	Yes	Yes
Segment FE	Yes	Yes	Yes	Yes
Credit Controls	Yes	No	Yes	Yes
Demand Controls	Yes	No	No	Yes
First Stage F		9.596	10.677	9.683
Number of Observations	11574	11574	11574	11574

Note: Subscripts $\hat{h}$, z, and t denote quality segment, zip code, and year. Column 1 estimates (E6) and columns 2-4 estimate (E5). Quality segments are based on relative quality from the MBA/CREFC property inspection rating. Segment $\hat{h}_0$ is the lowest available in all zip codes and years. $\text{Impending}_{z,t}$ is the fraction of units whose property has a loan due in year t or $t+1$. $\text{Expend}_{\hat{h},z,t}$ and $\text{Share}_{\hat{h},z,t}$ are the aggregate share of rent expenditure and number of units in segment $\hat{h}$ within a given zip code-year. The estimator in columns 2-4 is 2SLS and the instrument for $\log\left(\frac{\text{Share}_{\hat{h},z,t}}{\text{Share}_{\hat{h}_0,z,t}}\right)$ is $\text{Impending}_{z,t}$. Credit and demand controls are those from Table E4. The sample period is 2010-2016. Standard errors are in parentheses. Data are from Trepp.

affected relative expenditure shares under the CES market demand curve (E4). The point estimate of 0.86 in column 4 implies σ of around 6.9, similar to the result of the property-level specification.[100]

F.1.3 Estimating σ: Feenstra GMM

Feenstra (1994) proposes an estimator for σ in the context of time-varying quality. It exploits the panel structure of the data to provide an identification condition. This approach is potentially problematic in my dataset because it requires a large number of time periods to produce consistent estimates. That said, I estimate σ using this method as well.

To summarize the methodology briefly, one begins with the market demand curve implied by Lemma G.2 and previously expressed in (E4),

$$\log\left(\frac{\text{Expend}_{\hat{h},t}}{\text{Expend}_{\hat{h}_0,t}}\right) = \frac{1}{\sigma}\log\left(\frac{\hat{h}}{\hat{h}_0}\right) + \left(1-\frac{1}{\sigma}\right)\log\left(\frac{\text{Share}_{\hat{h},t}}{\text{Share}_{\hat{h}_0,t}}\right) + \nu_{\hat{h},t} \tag{E7}$$

where, using the notation introduced in Appendix G, $\text{Expend}_{\hat{h}}$ is the aggregate rent expenditure share

[100] Explicitly, $\sigma = \frac{1}{1-0.86}$.

on segment $\hat{h}$ and $\nu_{\hat{h},t}$ is a demand shifter.[101] Like with the zip code credit supply methodology, $\hat{h}_0$ is the reference segment, which I set at the lowest quality segment on the MBA/CREFC rating scale.

Then, one specifies the following isoelastic supply curve for a representative property owner deciding how many units in quality segment $\hat{h}$ to provide. This representative property owner aggregates the improvement decisions of individual property owners, like those described in Section 2.3, giving rise to a supply curve which I express in terms of revenue, $\text{Expend}_{\hat{h},t}$,

$$\log\left(\frac{\text{Expend}_{\hat{h},t}}{\text{Expend}_{\hat{h}_0,t}}\right) = \alpha_0 + \alpha \log\left(\frac{\text{Share}_{\hat{h},t}}{\text{Share}_{\hat{h}_0,t}}\right) + \alpha_{\hat{h},t}. \tag{E8}$$

One takes differences of the demand and supply curves (E7) and (E8) to obtained the differenced shocks $\Delta\nu_{\hat{h},t}, \Delta\alpha_{\hat{h},t}$. These shocks give the moment condition

$$\mathbb{E}\left[\Delta\nu_{\hat{h},t}\Delta\alpha_{\hat{h},t}\right] = 0. \tag{E9}$$

Note that (E9) must apply to each zip code z. Therefore, rearranging (E9) gives the regression equation

$$\begin{aligned}\left[\Delta\log\left(\frac{\text{Expend}_{\hat{h},z,t}}{\text{Expend}_{\hat{h}_0,z,t}}\right)\right]^2 &= \theta_1\left[\Delta\log\left(\frac{\text{Expend}_{\hat{h},z,t}}{\text{Expend}_{\hat{h}_0,z,t}}\right)\Delta\log\left(\frac{\text{Share}_{\hat{h},z,t}}{\text{Share}_{\hat{h}_0,z,t}}\right)\right] + \ldots \\ &\ldots + \theta_2\left[\Delta\log\left(\frac{\text{Share}_{\hat{h},z,t}}{\text{Share}_{\hat{h}_0,z,t}}\right)\right]^2 + u_{\hat{h},z,t},\end{aligned} \tag{E10}$$

where the notation is the same as in previous specifications. Intuitively, equation (E10) expresses the relationship among the second moments of expenditure and unit shares, $\text{Expend}_{\hat{h},z,t}$ and $\text{Share}_{\hat{h},z,t}$. The coefficients θ_1 and θ_2 encode the elasticity of substitution σ and the supply elasticity α. In particular,

$$\sigma = \frac{1}{1+\theta_2\alpha}, \quad \alpha = \frac{\theta_1}{2} + \frac{1}{2}\sqrt{\theta_1^2 + 4\theta_2}. \tag{E11}$$

Since $u_{\hat{h},z,t}$ is a function of the differenced demand and supply shocks $\Delta\nu_{\hat{h},z,t}, \Delta\alpha_{\hat{h},z,t}$, one cannot

[101] The estimator I derive is slightly different than the original proposed by Feenstra (1994) because I reason on quantity (i.e. number of units) rather than price (i.e. rent). Reasoning on quantity is more appropriate in my setting because the share of units in each segment must sum to 1, and doing so reduces measurement error from the fact that I approximate rent as revenue per occupied unit. However, the setup is effectively the same after replacing "goods" with "quality segments".

estimate (E10) consistently. However, one can obtain consistent estimates by taking the average of (E10) across time periods, and estimating the resulting regression equation by weighted least squares. That is, $\text{plim}_{T\to\infty} \sum_{t=0}^{T} u_{\hat{h},z,t} = 0$. The resulting estimates of θ_1 and θ_2 imply $\sigma = 6.1$, shown in Table E1.

G Mathematical Details

This appendix has mathematical details related to the structural rent index. The structural index is described at a high level in Section 3 of the text and more fully in Appendix C. All material in this appendix is referenced in Appendix C.

G.1 Household Preferences: Structural Rent Index

This extension describes households' problem in greater detail. Reviewing the setup from Section C.2, household j selects a unit i and derives additive random utility from the unit's quality h_i according to the preferences in (C4),

$$u_{i,j} = \log(h_i) + \epsilon_{i,j}.$$

In the baseline case, $\epsilon_{i,j}$ follows a Gumbel, or type 1 extreme value distribution.[102] Moreover, quality is itself a composite of a unit's space s_i (e.g. square feet) and other amenities a_i (e.g. granite countertops) according to $\log(h_i) = \mu \log(s_i) + \log(a_i)$. For simplicity I assume that all units i in the same quality segment $\hat{h}$ have the same space $s_{\hat{h}}$. The next lemma describes how this preference structure gives rise to the discrete choice problem verbally articulated in Appendix C.2.

Lemma G.1 (Discrete Choice) *A household with preferences (C4) chooses her shelter according to*

$$\max_{i \in \mathcal{I}} \left\{ -\log(Rent_i) + \frac{1}{\mu}\log(h_i) + \frac{1}{\mu}\epsilon_{i,j} \right\}. \tag{F1}$$

Note that while households do not consume the numeraire, it is straightforward to allow for non-housing consumption. To do so, I follow Anderson, de Palma and Thisse (1992) and modify the baseline preferences (C4) as follows

$$u_{i,j} = \kappa \log(c_j) + \log(h_i) + \epsilon_{i,j}, \tag{F2}$$

where c_j is household j's consumption of the numeraire. Maximizing (F2) under the budget constraint

[102] The cumulative distribution function is $\Pr[\epsilon_{i,j} \leq \epsilon] = \exp\left[-\exp\left[-\left(\frac{\epsilon}{\tilde{\mu}} + \gamma^e\right)\right]\right]$, where $\gamma^e = 0.58$ is Euler's constant and $\tilde{\mu}$ is a scaling parameter. In particular, $\mathbb{E}[\epsilon_{i,j}] = 0$ and $\mathrm{Var}[\epsilon_{i,j}] = \tilde{\mu}^2 \frac{\pi^2}{6}$.

$y_j = y_j^h + c_j$, where y_j^h is income available for housing expenditure,

$$c_j = \frac{\kappa}{\kappa+1} y_j.$$

Thus, the analysis is effectively the same after replacing income y_j with income net of non-housing consumption, $y_j^h = \frac{1}{\kappa+1} y_j$.

The next lemma provides a useful aggregation result which enables Proposition C.1.

Lemma G.2 (Anderson, De Palma, and Thisse 1992) *If $\epsilon_{i,j}$ follows a type 1 extreme value distribution with scaling parameter 1, then the distribution of number units across segments $\{Share_{\hat{h}}\}$ behaves according to*

$$\min_{\{Share_{\hat{h}}\}} \sum_{\hat{h}\in\mathcal{H}} Share_{\hat{h}} \times Rent_{\hat{h}} \quad s.t. \quad \bar{U} = \left[H^{\frac{1}{\sigma}} \sum_{\hat{h}\in\mathcal{H}} \hat{h}^{\frac{1}{\sigma}} Share_{\hat{h}}^{1-\frac{1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \tag{F3}$$

where $\sigma = \mu + 1$ and $\mathcal{H} \subseteq [0,1]$ is the set of quality segments. Moreover, $\frac{\partial \log\left(Rent_{\hat{h}}\right)}{\partial \log\left(\hat{h}\right)} = \frac{1}{\sigma}$.

Finally, section 3.3 provides evidence that higher income households have a lower value of σ, and thus a lower μ, which governs preferences for space versus amenities. If households treat space as a necessity, one can suppose that μ is a decreasing function of income y_j. Thus, as y_j rises, households' relative preference for space versus amenities falls, so that high income households have low values of $\sigma = \mu + 1$, matching the estimates from Table E2.

G.2 Proofs

Proof of Lemma G.1

The preferences in (C4) are preserved when multiplying by $\tilde{\mu} \equiv \frac{1}{\mu}$. Therefore, since rent Rent_i is denominated in numeraire per housing unit, a household with y_j available to spend on housing has the following utility, based on (C4),

$$u_{i,j} = \log\left(\frac{y_j}{\text{Rent}_i/s_i}\right) + \tilde{\mu}\log(a_i) + \tilde{\epsilon}_{i,j} = \log(y_j) - \log(\text{Rent}_i) + \tilde{\mu}\log(h_i) + \tilde{\epsilon}_{i,j}, \tag{F4}$$

where the second equality uses $\log(h_i) = \mu \log(s_i) + \log(a_i)$, and $\tilde{\epsilon}_{i,j} \equiv \tilde{\mu}\epsilon_{i,j}$. From (F4) it follows that household j selects unit $i \in \mathcal{I}$ as the solution to

$$\max_{i \in \mathcal{I}} \left\{ -\log(\text{Rent}_i) + \frac{1}{\mu}\log(h_i) + \frac{1}{\mu}\epsilon_{i,j} \right\},$$

as in (F1).

Proof of Lemma G.2

First note that the problem in (F10) is equivalent to its corresponding primal problem,

$$\max_{\{\text{Share}_{\hat{h}}\}} \left[\sum_{\hat{h} \in \mathcal{H}} h^{\frac{1}{\sigma}} \text{Share}_{\hat{h}}^{1-\frac{1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \quad s.t. \quad \bar{y} = \sum_{\hat{h} \in \mathcal{H}} \text{Share}_{\hat{h}} \times \text{Rent}_{\hat{h}}, \tag{F5}$$

with

$$\bar{U} = \bar{y} \left(\left[\sum_{\hat{h} \in \mathcal{H}} h \times \text{Rent}_{\hat{h}}{}^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \right)^{-1} \equiv \bar{y}\bar{R}^{-1}, \tag{F6}$$

where $\bar{R}$ is the minimized cost function associated with the problem in (F10), which was referred to as "welfare relevant rent" in the text. In particular, using $h = \hat{h}H$, the solution implies that the share of aggregate expenditure on segment $\hat{h}$ is

$$\text{Expend}_{\hat{h}} = \frac{\hat{h}\text{Rent}_{\hat{h}}^{1-\sigma}}{\sum_{\hat{h} \in \mathcal{H}} \hat{h}\text{Rent}_{\hat{h}}^{1-\sigma}}. \tag{F7}$$

Continuing, it suffices to show that the aggregate demand generated by individual households solving (F1) behaves according to (F5) for some σ. I work with the normalized preferences in (F4) from the proof of Lemma G.1, where $\tilde{\epsilon}_{i,j} \equiv \tilde{\mu}\epsilon_{i,j}$ follows a Gumbel distribution with scaling parameter $\tilde{\mu} \equiv \frac{1}{\mu}$. Then, use equation (3.51) from Anderson, de Palma and Thisse (1992) for the case when there is the additional quality term $\tilde{\mu}\log(h)$, to write the probability a household chooses a unit in segment $h = \hat{h}H$ as

$$\varrho_{\hat{h}} = \frac{\hat{h} \times \text{Rent}_{\hat{h}}^{-\mu}}{\sum_{\hat{h} \in \mathcal{H}} \hat{h} \times \text{Rent}_{\hat{h}}^{-\mu}}. \tag{F8}$$

It follows that aggregate demand across households for units in segment $\hat{h}$ is

$$\text{Share}_{\hat{h}} = \frac{\bar{y}}{\text{Rent}_{\hat{h}}/s_{\hat{h}}} \cdot \varrho_{\hat{h}} \cdot \frac{1}{s_{\hat{h}}}, \tag{F9}$$

where $s_{\hat{h}}$ is the space afforded by units in segment $\hat{h}$. In particular, the three terms in (F9) are, respectively: (i) the space demanded by the average household, (ii) the share of households selecting a housing unit in segment $\hat{h}$, and (iii) the inverse space per housing unit.

Finally, using Proposition 3.8 from Anderson, de Palma and Thisse (1992), the aggregate demand system (F9) equals that of the representative household (F7) if and only if $\sigma = \mu + 1$. That is, the distribution of units across quality segments $\{\text{Share}_{\hat{h}}\}$ behaves according to the solution to the problem (F10). In addition, as pointed out by Anderson, de Palma and Thisse (1992), the value function associated with (F5) is a utilitarian welfare function.

To obtain the marginal willingness to pay, use the demand curve (F7) to write

$$\frac{\text{Rent}_{\hat{h}}}{\text{Rent}_{\hat{h}_0}} = \left(\frac{h}{h_0}\right)^{\frac{1}{\sigma}} \left(\frac{\text{Share}_{\hat{h}}}{\text{Share}_{\hat{h}_0}}\right)^{-\frac{1}{\sigma}}, \tag{F10}$$

for some reference segment $\hat{h}_0$, which gives $\frac{\partial \log(\text{Rent})}{\partial \log(h)} = \frac{\partial \log(\text{Rent})}{\partial \log(\hat{h})} = \frac{1}{\sigma}$ using $h = \hat{h}H$. This completes what needed to be shown.

Proof of Proposition C.1

By definition, π_t^S is the growth in the unit cost function $\bar{R}_t$ from t_0 to t, explicitly $\pi_t^S \equiv \frac{\bar{R}_t}{\bar{R}_{t_0}}$. Using (F6) from the proof of Lemma G.2,

$$\bar{R} = \left[\sum_{\hat{h}\in\mathcal{H}} \hat{h}H_t \times \text{Rent}_{\hat{h},t}{}^{1-\sigma}\right]^{\frac{1}{1-\sigma}}, \tag{F11}$$

where $h_t = \hat{h}H_t$ is the absolute quality of segment $\hat{h}$ at t. Therefore, write π_t^S as

$$\pi_t^S = \left[\sum_{\hat{h}\in\mathcal{H}} \hat{h} \times \text{Rent}_{\hat{h}}{}^{1-\sigma}\right]^{\frac{1}{1-\sigma}} \times \left(\frac{H_t}{H_{t_0}}\right)^{-\frac{1}{\sigma-1}} \equiv DQ_t \times GQ_t^{-\frac{1}{\sigma-1}}. \tag{F12}$$

Next, following Feenstra (1994), use the results of Diewert (1976), Sato (1976), and Vartia (1976) to rewrite DQ_t as

$$DQ_t = \exp\left[\sum_{\hat{h}\in\mathcal{H}} w_{\hat{h},t} \log\left(\frac{\text{Rent}_{\hat{h},t}}{\text{Rent}_{\hat{h},t_0}}\right)\right], \tag{F13}$$

where the Sato-Vartia weights are

$$w_{\hat{h},t} = \frac{\frac{\text{Expend}_{\hat{h},t} - \text{Expend}_{\hat{h},t_0}}{\log\left(\text{Expend}_{\hat{h},t}\right) - \log\left(\text{Expend}_{\hat{h},0}\right)}}{\sum_{\hat{h}\in\mathcal{H}} \frac{\text{Expend}_{\hat{h},t} - \text{Expend}_{\hat{h},t_0}}{\log\left(\text{Expend}_{\hat{h},t}\right) - \log\left(\text{Expend}_{\hat{h},t_0}\right)}}, \tag{F14}$$

and, as in (F7) from the proof of Lemma G.2, $\text{Expend}_{\hat{h},t}$ is the share of aggregate expenditure on segment $\hat{h}$ in t,

$$\text{Expend}_{\hat{h}} = \frac{\text{Share}_{\hat{h}} \times \text{Rent}_{\hat{h}}}{\sum_{\hat{h}\in\mathcal{H}} \text{Share}_{\hat{h}} \times \text{Rent}_{\hat{h}}}. \tag{F15}$$

Finally, Lemma G.2 implies that the market demand curve has a CES structure, and using (F10),

$$\frac{\text{Rent}_{1,t}}{\text{Rent}_{1_0,t}} = \left(\frac{H_t}{H_{t_0}}\right)^{\frac{1}{\sigma}} \left(\frac{\text{Share}_{1,t}}{\text{Share}_{1_0,t}}\right)^{-\frac{1}{\sigma}}, \tag{F16}$$

where H_{t_0} is absolute quality in t_0 and segment $1_0 \equiv \frac{H_{t_0}}{H_t}$ contains units that were in segment 1 in year t_0.[103] Rearranging (F16) gives growth in absolute quality,

$$GQ_t = \left(\frac{\text{Rent}_{1,t}}{\text{Rent}_{1_0,t}}\right)^{\sigma} \frac{\text{Share}_{1,t}}{\text{Share}_{1_0,t}}. \tag{F17}$$

Combining (F13) and (F17) gives the expression in (C5),

$$\pi_t^S = \exp\left[\sum_{\hat{h}\in\mathcal{H}} w_{\hat{h},t} \log\left(\frac{\text{Rent}_{\hat{h},t}}{\text{Rent}_{\hat{h},t_0}}\right)\right] \times \left[\left(\frac{\text{Rent}_{1,t}}{\text{Rent}_{1_0,t}}\right)^{\sigma} \frac{\text{Share}_{1,t}}{\text{Share}_{1_0,t}}\right]^{-\frac{1}{\sigma-1}} \qquad \equiv DQ_t \times GQ_t^{-\frac{1}{\sigma-1}}.$$

[103] While the setup does not feature depreciation, when taking π_t^S to the data I require that units in segment 1_0 retained their absolute quality through t, as discussed in Appendix F.1.

H Non-CES Rent Indices

This section uses detailed data on property upgrade activity to infer time variation in quality with minimal structural assumptions. While the CES aggregator is one of the most commonly used in economics, one might be concerned that the expenditure function from Lemma G.2 is highly misspecified. To address this concern, this section performs a quasi-hedonic quality adjustment to correct each segment's rent for time-varying quality. The corrected rent can then be used in any non-CES price index formula (e.g. Tornqvist, Paasche, Laspeyres).

Write the rent on unit i in year t as

$$\begin{aligned} \log\left(\text{Rent}_{i,t}\right) &= a_i + a_t + P\log\left(h_{i,t}\right) + u_{i,t} && \text{(G1)} \\ &= a_i + a_t + P\left[\log\left(\hat{h}_{i,t}\right) + H_t\right] + u_{i,t}, && \text{(G2)} \end{aligned}$$

where P is the equilibrium slope of the quality ladder, or price of quality. Let $\text{New}_{i,t}$ indicate if i is first on the market in t after renovation and is in the top quality segment. I maintain the assumption from the structural index in Appendix C of the paper that such units retain their absolute quality for at least one year. Then (G1) implies that rent growth conditional on being new ($\text{New}_{i,t} = 1$) or almost-new ($\text{New}_{i,t-1} = 1$) is

$$\Delta\log\left(\text{Rent}_{i,t}\right) = \Delta a_t + \underbrace{\beta_0}_{P\times\log(GQ_t)} \text{New}_{i,t} + \underbrace{\beta_1}_{-P}\left(\text{New}_{i,t}\times\log\left(\hat{h}_{i,t-1}\right)\right) + \Delta u_{i,t}, \qquad \text{(G3)}$$

where $GQ_t = \frac{H_t}{H_{t-1}}$ is growth in absolute quality, using the notation from Proposition 1. In words, the rent growth differential between new and almost-new units reflects new unit quality growth, after controlling for the previous quality of new units. To measure relative quality, I again use the MBA/CREFC property inspection rating.[104]

Using the Trepp dataset, I estimate (G3) year-by-year on units such that

$$\max\left\{\text{New}_{i,t}, \text{New}_{i,t-1}\right\} = 1$$

and extract the coefficients $\{\beta_{0,t}\}$.[105] These point estimates give a sequence of quality growth rates in units of log rent, $\{P\times(GQ_t)\}$. The identifying assumption is that the renovation decision $\text{New}_{i,t}$

[104]Since inspections are rare during the year of renovation, I proxy for $\hat{h}_{i,t-1}$ using the most recent rating prior to renovation. Similarly, I proxy for $\Delta\log\left(\text{Rent}_{i,t}\right)$ using the most recent available rent data and annualizing.

[105]I weight observations in (G3) by number of units because the Trepp data are at the property-level.

is orthogonal to unobserved changes in the rental market as contained in $\Delta u_{i,t}$. This assumption would be violated if, for example, renovations only occur in high-growth areas. To account for this difficulty, I estimate (G3) with MSA fixed effects and controls for property size. Thus, the comparison is strictly between units in new and almost-new properties of the same size and in the same MSA and year. While reducing bias, this approach substantially limits the available variation to estimate (G3) given the inclusion of so many covariates. Therefore, I use the James-Stein estimator, which optimally biases the point estimate toward 0.[106]

Given the estimated sequence of annualized growth rates $\{P \times \log(GQ_t)\}$, I correct each unit's rent according to

$$\text{Rent}_{i,t}^{GQ} = \text{Rent}_{i,t} \times \exp\left[-\sum_{\tau=t_0}^{t} P \times \log(GQ_\tau)\right]. \tag{G4}$$

Using (G4), one obtains the corrected rent in each quality segment $\{\text{Rent}_{\hat{h},t}^{GQ}\}$. Along with the appropriate data on aggregate expenditure shares $\{\text{Expend}_{\hat{h},t}\}$, one can compute effective rent using any price index formula. Some common formulae used in this paper are

$$\begin{aligned}
\pi_t^{\text{Tornqvist}} &= \exp\left[\sum_{\hat{h}\in\mathcal{H}} \frac{\text{Expend}_{\hat{h},t} + \text{Expend}_{\hat{h},t_0}}{2} \log\left(\frac{\text{Rent}_{\hat{h},t}^{GQ}}{\text{Rent}_{\hat{h},t_0}^{GQ}}\right)\right], \\
\pi_t^{\text{Paasche}} &= \exp\left[\sum_{\hat{h}\in\mathcal{H}} \text{Expend}_{\hat{h},t} \log\left(\frac{\text{Rent}_{\hat{h},t}^{GQ}}{\text{Rent}_{\hat{h},t_0}^{GQ}}\right)\right], \\
\pi_t^{\text{Laspeyres}} &= \exp\left[\sum_{\hat{h}\in\mathcal{H}} \text{Expend}_{\hat{h},t_0} \log\left(\frac{\text{Rent}_{\hat{h},t}^{GQ}}{\text{Rent}_{\hat{h},t_0}^{GQ}}\right)\right].
\end{aligned}$$

Excess-CPI growth in $\pi_t^{\text{Tornqvist}}$, π_t^{Paasche}, and $\pi_t^{\text{Laspeyres}}$ was between 0.1% and 0.2% over 2010-2016.

[106]The James-Stein estimator is $\hat{\beta}^{JS} = \max\left\{1 - \frac{c}{\text{F-statistic}}, 0\right\} \cdot \hat{\beta}^{OLS}$, where $\hat{\beta}^{OLS}$ is the OLS estimator. For $0 < c < \bar{c} < 2$ and at least three predictor variables, $\hat{\beta}^{JS}$ dominates $\hat{\beta}^{OLS}$ under the L^2 norm.