

Reanalysis of Multifamily Housing Permit Activity Shows No Reduction After Measure ULA

Greg Bonett*

Madeline Wander†

2026-04-28

Summary

This paper is a partial reanalysis of the data used by Jason Ward and Shane Phillips in a report released in April 2025 by the UCLA Lewis Center for Regional Policy Studies titled “Taxing Tomorrow: Measure ULA’s Impact on Multifamily Housing Production and Potential Reforms.”¹ This reanalysis examines Ward and Phillips’ claim to have shown a “robust causal linkage” between Measure ULA and a reduction in multifamily housing construction. Specifically, Ward and Phillips claim that Measure ULA had “caused” a nearly 50% reduction in the number of housing units permitted in projects with 20 or more units, amounting to a loss of at least 1,910 units annually. That conclusion has been widely cited in media accounts, by elected officials, in proposed legislation, and in a proposed constitutional amendment.² **In contrast, our reanalysis finds that the building permit data used by Ward and Phillips do not show any meaningful change in the number of multifamily housing units permitted on land sold after Measure ULA took effect.** In fact, there was a slight, albeit statistically insignificant, increase in the number of multifamily units permitted per month for projects with 10 or more units on land sold after Measure ULA.

A key flaw found through our reanalysis is that Ward and Phillips did not actually analyze the change in the number of multifamily units permitted. Rather, they tested whether the average *size* of multifamily projects had changed after Measure ULA took effect. Their findings leap from an analysis of the typical size of a project to a conclusion about the annual rate of housing construction, essentially assuming that the number of projects permitted each month did not change, when in fact it did. The data used by Ward and Phillips show that on land sold after Measure ULA took effect, the number of projects per month roughly doubled. Therefore, when we examine the total number of multifamily units permitted each month (rather than the size of individual projects), we find no change in the post-ULA period.

In addition to finding no meaningful change in the number of multifamily units permitted each month, our reanalysis also finds that the change in average project size observed by Ward and Phillips is more strongly associated with other factors, such as changes in the federal funds rate. When the federal funds rate is added to the regression model used by Ward and Phillips, the Measure ULA indicator is no longer statistically significant and has a value near zero. That is, statistically, the change in the federal funds rate is a much better predictor of the change in the size of projects than the timing of Measure ULA taking effect. In reality, the change in average project size is likely explained by many interacting factors, such as borrowing costs, construction costs, and zoning changes, that are not captured in a simple regression model.

*Senior Policy Counsel, Public Counsel Community Development Project

†Assistant Professor, Urban & Environmental Policy, Occidental College

¹Ward, J., & Phillips, S. (2025). *Taxing Tomorrow: Measure ULA’s Impact on Multifamily Housing Production and Potential Reforms*. UCLA: The Ralph and Goldy Lewis Center for Regional Policy Studies. We are grateful to the authors for making their data and statistical code available for reanalysis.

²Raman, Nithya V. (2026), Motion for Ballot Measure to Amend ULA, Los Angeles City Council; Local Taxpayer Protection Act to Save Proposition 13, Initiative 25-0006 Amdt. 1; Khouri, A. (2025, April 11), *Measure ULA reducing apartment development in city of L.A., study says*, Los Angeles Times; Wagner, D. (2025, April 11), *Mansion tax pitfalls: Study finds Measure ULA hits apartment development*, LAist; *LA multifamily permitting falls 18% after Measure ULA*, (2025, April 21), The Real Deal.

Ward and Phillips’ claim of a “robust causal linkage” between Measure ULA and declines in multifamily housing production is not reflected in the data they used. To the contrary, their data show no meaningful change in the number of multifamily units permitted per month after Measure ULA took effect. And the observed change in average project size is better explained by broader macroeconomic factors, such as changes in the federal funds rate.

Background

Measure ULA is a City of Los Angeles real estate transfer tax adopted by voter initiative that imposes a 4% tax on most real estate sales over \$5.3 million and a 5.5% tax on sales over \$10.6 million.³ Revenue from Measure ULA pays for affordable housing construction and preservation, as well as rental assistance, income support, and eviction defense programs for low-income residents. As of February 2026, Measure ULA has raised over \$1.1 billion for these programs.

Jason Ward and Shane Phillips’ “Taxing Tomorrow” report claims to establish a “robust causal linkage” between Measure ULA and a reduction in multifamily housing production. The report includes two separate analyses: one examining real estate transactions and one examining multifamily building permit activity. In the first analysis, Ward and Phillips use a difference-in-differences model to test for a change in the rate of multifamily-zoned land sales in the City of Los Angeles compared with ten other jurisdictions in Los Angeles County before and after Measure ULA was enacted. They report a decrease in transactions of multifamily-zoned parcels in the City of Los Angeles relative to comparison jurisdictions after the implementation of Measure ULA. That finding is not the subject of this reanalysis, but previous work has raised questions about its validity, including the plausibility of satisfying the parallel trends assumption necessary for a difference-in-differences analysis. For example, that analysis assumes that the changes in factors such as borrowing costs, construction costs, and zoning are identical in the City of Los Angeles and the comparison jurisdictions, and that these changes will impact multifamily transactions in exactly the same way. That assumption is not realistic in this context.⁴

In the second analysis, which is reanalyzed here, Ward and Phillips use a “pre/post” model to estimate changes in multifamily building permit activity among projects on land that sold before or after Measure ULA took effect. That analysis uses building permit data from the City of Los Angeles and does not include a comparison with permit data from other jurisdictions. The authors extrapolated from that estimated change to conclude that Measure ULA is “causing” a reduction in multifamily permitting amounting to “3,820 units in its first two years, or 1,910 units annually.”⁵ That estimate has been widely reported in the media, appearing dozens of times across numerous publications. As shown below, the data used by Ward and Phillips do not support their conclusion. In fact, the data show no meaningful change in the number of multifamily units permitted on land sold after Measure ULA took effect.

Reanalysis

The Ward and Phillips analysis of building permit data examined the *size* of multifamily projects, not the *rate* of multifamily permitting, leading to an incorrect conclusion.

Ward and Phillips analyzed multifamily building permits issued within one year of a sale, for land sold before and after Measure ULA took effect. After testing for changes in permit activity at four separate building size thresholds, they concluded that Measure ULA was associated with a nearly 50% reduction in the number of housing units permitted in projects with 20 or more units.⁶ They then extrapolated from that estimate to

³The value thresholds are increased annually with inflation. See Measure ULA FAQ.

⁴Bonett, G., Breidenbach, J., Cummings, S., Dreier, P., Freer, R., Ling, J., Sharma, D., Tilly, C., & Wander, M. (2025). *Major Research Flaws Undermine Authors’ Bold Claims: Unpacking the Debate on Measure ULA*.

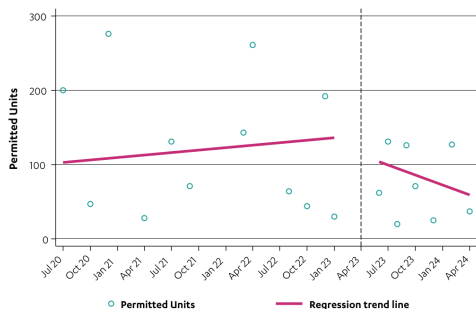
⁵Ward and Phillips, p. 34.

⁶Ward and Phillips found no statistically significant change at the three other size thresholds tested, including thresholds of 10+ units, 15+ units, and 30+ units.

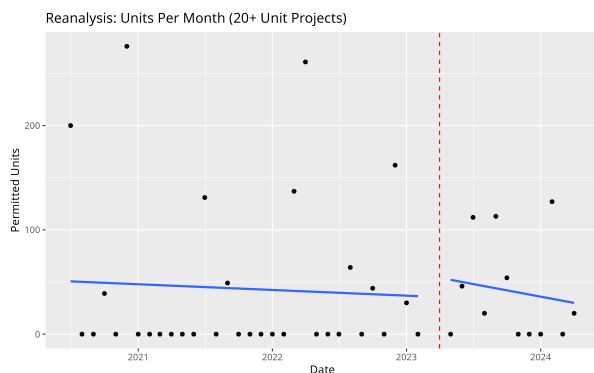
conclude that Measure ULA has reduced construction of new multifamily housing by at least 1,910 units per year. Our reanalysis finds that their conclusion is not supported by their data.

Using the data and statistical code provided by Ward and Phillips, we were able to reproduce the regression results in “Taxing Tomorrow,” but found that these results do not support the authors’ conclusion. Ward and Phillips tested for an association between the number of units in a building permit issued for a multifamily-zoned parcel and whether that parcel was sold before or after Measure ULA took effect. That test was asking a question about the *size* of the project, not the *rate* at which multifamily units were being permitted.⁷ Figure 1 below replicates the data presented by Ward and Phillips. Panel (a) is a screenshot of the original figure from Ward and Phillips, and Panel (b) is a scatter plot of the data used by Ward and Phillips to estimate changes in building permit activity for projects with 20 or more units.⁸ To better illustrate the rate of housing production, Panel (b) includes points for months in which the Ward and Phillips data did not include any land sales resulting in a building permit for 20 or more units. **The original trend lines in Panel (a) did not account for months with zero multifamily units permitted.⁹ Including these months with zero permitted units shifts the trend lines down and noticeably changes the slopes.**

Figure 8.
Permitted Units Associated with Sales of Parcels Zoned for High-Density Multifamily Development (20+ Units) in Los Angeles



(a) Original figure from Ward and Phillips



(b) Reanalysis including zero-project months

Figure 1: Comparison of Figure 8 from Ward and Phillips and reanalysis. The red dashed line indicates when Measure ULA took effect. The blue lines are linear trend lines for the pre- and post-ULA periods. The original trend lines did not account for months with zero units.

Ward and Phillips used the following linear regression model to estimate the change in the number of units per project (*not* the number of multifamily units permitted) after Measure ULA took effect:¹⁰

$$\log(\text{ProjectUnits}_{it}) = \beta_0 + \beta_1 \text{Post}_t + \beta_2 \text{LandImprovementRatio}_i + \beta_3 \text{MultiFamily}_i + \varepsilon_{it}$$

The regression model described above was estimated using sampling weights equal to the inverse of the number of sales per month in which building permits were issued within one year of the sale. The results are

⁷More precisely, Ward and Phillips tested for an association between the log of the number of units in a building permit for 20 or more units issued within one year of the property being sold for properties sold before or after Measure ULA took effect, controlling for the zoning and value of improvements on the property, and giving greater weight to permits issued in months with fewer permits.

⁸See Ward and Phillips, Figure 8.

⁹Note that the original figure in Panel (a) includes two non-zero data points not included in Panel (b). These points are included in Panel (a) because of a coding error that included months with 20 or more units, rather than projects with 20 or more units. Panel (b) includes exactly the data used by Ward and Phillips to estimate the change in multifamily permitting of projects with 20 or more units.

¹⁰Note that the regression model presented in Ward and Phillips on page 32 is a simplified form of the model above. The results reported in Ward and Phillips are the coefficients from the regression included here, not the simplified form presented in their paper. The simple regression specified in Ward and Phillips on page 32 does not result in a statistically significant coefficient on the ULA indicator term. See Appendix Table A1.

reported in column (3) of Table A.3 in the Ward and Phillips report and are replicated in Table A2 of the appendix of this paper.

Interpreting the results of the regression model above is not straightforward. The regression tests whether land being sold after Measure ULA took effect is associated with a change in average project size, controlling for certain property characteristics, and weighting the observations by the inverse of the number of permitted projects per month to give equal weight to each month in the data. Ultimately, this regression analysis is answering a question about the *size* of individual projects—not about the *rate* of housing production. To illustrate this point, consider a hypothetical dataset in which there are a small number of 100-unit projects in the “pre” period and dozens of smaller 50-unit projects in the “post” period: the regression model used by Ward and Phillips would estimate a large negative coefficient on the “pre/post” indicator term, even though the rate of unit production in this hypothetical “post” period is much higher than in the hypothetical “pre” period.

The data used by Ward and Phillips show that, on land sold *after* Measure ULA took effect, the number of multifamily housing projects permitted per month roughly doubled.

As discussed above, the estimate in Ward and Phillips is misleading if building permits were issued more frequently on land sold after Measure ULA took effect. In fact, of the 27 projects with 20 or more units used to estimate the impact of ULA on multifamily housing production, the average number of projects per month is notably higher on land sold after Measure ULA: there are almost **twice** as many projects per month on parcels sold after Measure ULA took effect compared with before.¹¹ Therefore, **simply looking at the change in the average size of a project does not give an accurate picture of the change in the number of units permitted each month.**

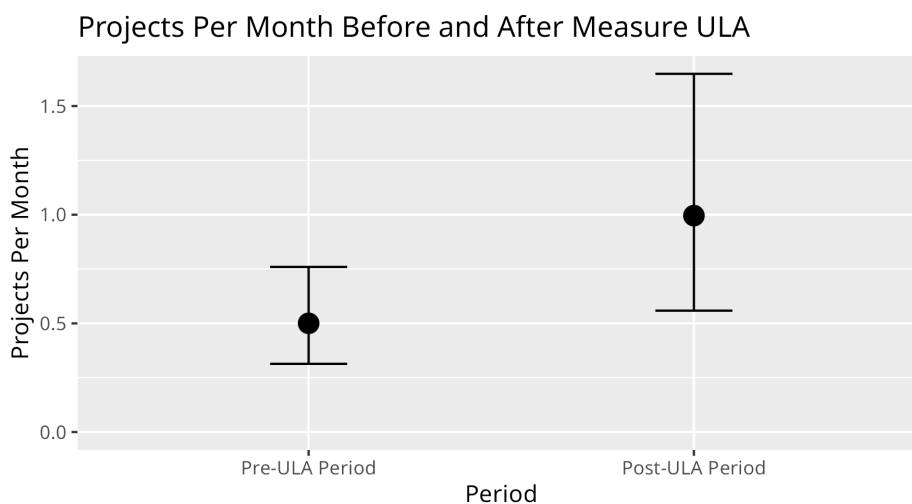


Figure 2: Projects per month before and after Measure ULA. The pre-ULA period spans from July 2020 to February 2023, and the post-ULA period spans from May 2023 to April 2024. N=27 projects with 20+ units. 90% confidence intervals calculated by modeling the issuance of a building permit as a Poisson process.

¹¹If we model the issuance of building permits as a Poisson process and compare the rates, we can say with 90% confidence that the rate after Measure ULA took effect is at least 1.11 times higher than the pre-ULA rate.

The dataset shows no meaningful change in the number of multifamily units permitted per month on land sold after Measure ULA took effect.

As discussed above, the analysis in Ward and Phillips did not estimate the change in the number of multifamily units permitted per month on land sold before or after Measure ULA took effect. Rather, their analysis estimated a change in the size of projects and did not account for the possibility that projects occurred more frequently after the measure went into effect. In fact, in their dataset, as we stated above, projects were about twice as frequent after Measure ULA. When we examine the total number of units per month before and after Measure ULA, we find no meaningful change. This is the case for all project size thresholds, including among projects with 20 or more units. As shown in Table 1, for projects with 10 or more units, there is a slight, but statistically insignificant, increase in the number of units permitted per month in the post-ULA period. To test for statistical significance, which tells us how likely it is that the difference in means occurred by chance, we ran t-tests for the four project-size thresholds that Ward and Phillips use. As Table 1 shows, since the p-values are very large (greater than 0.05), there is a high probability we would see differences this large purely by chance, and we cannot conclude that there is any real difference between the “pre” and “post” averages. This is not surprising given the small sample size of projects analyzed here.

Table 1: Permitted units by project size threshold before and after Measure ULA.

Unit threshold	Avg. units/mo. pre	Avg. units/mo. post	Statistically significant?	N
10+ units	48.6	48.9	No (p=0.989)	45
15+ units	46.0	47.8	No (p=0.942)	37
20+ units	43.5	41.0	No (p=0.919)	27
30+ units	40.6	33.6	No (p=0.778)	19

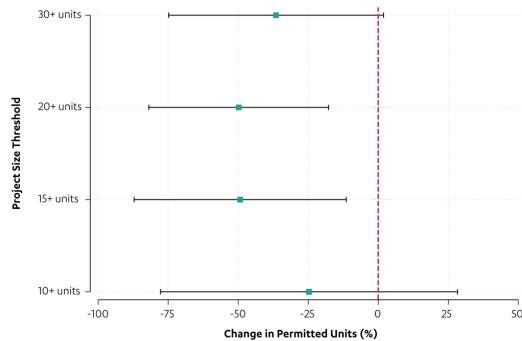
Another method for testing whether there was a meaningful change in the number of units permitted each month on land sold after ULA took effect is to use a regression model similar in form to the model presented by Ward and Phillips on page 32 of their report. But rather than using the log of units per project as the outcome variable, we use the log of total units per month—including months with zero units. While Ward and Phillips’ unit of analysis is the project, our unit of analysis is the month.

$$\log(\text{TotalUnitsPerMonth}_t) = \beta_0 + \beta_1 \text{Post}_t + \varepsilon_t$$

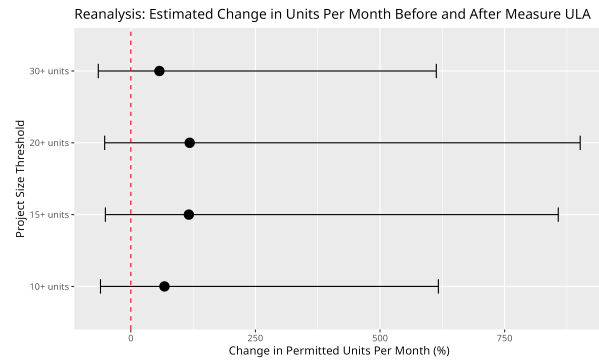
The results of this regression model show that the indicator term for Measure ULA going into effect is associated with an **increase** in the number of units per month for each of the four project-size thresholds used by Ward and Phillips (although this increase was not statistically significant at any threshold). To illustrate these results in terms of a percent change, the exponentiated estimates and confidence intervals based on our model are presented in Figure 3, Panel (b).¹²

¹²Note that even at the 30+ unit threshold, the log-transformed estimate for the change in units permitted per month is positive, suggesting that the difference in monthly averages was driven by the small number of large projects on land sold in the “pre-ULA” period. A log transformation decreases the impact of a small number of large projects on the estimate and often helps shape data to more closely adhere to the assumptions of linear regression analysis. We also tested an untransformed version of our outcome variable (total units per month) and found no statistically significant change in the number of units permitted per month after Measure ULA took effect for any of the four project-size thresholds.

Figure 9.
Estimated Change in Permitted Units Associated with Sales of Parcels Zoned for High-Density Multifamily Development in Los Angeles After Implementation of Measure ULA



(a) Original figure from Ward and Phillips



(b) Reanalysis using log of monthly total units

Figure 3: Comparison of Figure 9 from Ward and Phillips and our reanalysis. The figure on the right shows the change in the log of monthly total units with 95% confidence intervals, showing no statistically significant change for any project size threshold and positive change estimates at all project size thresholds.

The results above illustrate the error in the Ward and Phillips analysis. Ward and Phillips did not test for a change in the **rate** of multifamily unit production after Measure ULA took effect—only whether the average number of units in an individual project had changed. To arrive at their estimate that Measure ULA decreased multifamily housing production by at least 1,910 units per year, the authors multiplied the estimated percentage change in units per project by the average number of units permitted each year in projects with 20 or more units. But that extrapolation leaps from an analysis of project *size* to a conclusion about the *rate* of housing construction—it assumes that the number of projects per month did not change. That assumption is not supported by the data: in the dataset analyzed by Ward and Phillips, the average project size may have decreased, but this decrease was almost exactly offset by an increase in the number of projects per month. When the total number of units permitted per month is analyzed directly, there is no evidence of a reduction in multifamily housing construction after Measure ULA took effect. In fact, using the 10+ unit threshold, which includes all of the units in projects with higher thresholds, we find a slight, albeit statistically insignificant, increase in the number of units permitted per month in the post-ULA period.

The change in average project size is better explained by other factors.

What if we really do care about whether Measure ULA is associated with a reduction in the average size of projects, even if projects occurred more frequently? Even on this point, the Ward and Phillips analysis does not make a valid case.

The Ward and Phillips analysis estimates the change in project size on land sold after Measure ULA took effect, controlling for certain property characteristics, and giving each month equal weight. It does not control for any other potentially confounding factors that may have changed around the same time as Measure ULA took effect. Ward and Phillips justify that design choice by pointing to their difference-in-differences analysis of real estate sales which, as noted above, requires very strong and unrealistic assumptions.

Setting aside the critiques of the Ward and Phillips difference-in-differences analysis (which is not reanalyzed here), that logic is flawed. Ward and Phillips provided no explanation for why a change in sales volume would be tightly linked to a change in average project size. And, even if that metric were closely linked, the control jurisdictions in the difference-in-differences analysis also experienced a decline in sales volume—which presumably would be associated with a change in project size in control jurisdictions as well.

The linear regression model used by Ward and Phillips, shown above, omits several potentially confounding factors that coincided with the implementation of Measure ULA, including changes in borrowing costs, insurance rates, capitalization rates, construction costs, and local zoning. To test if one of these potentially

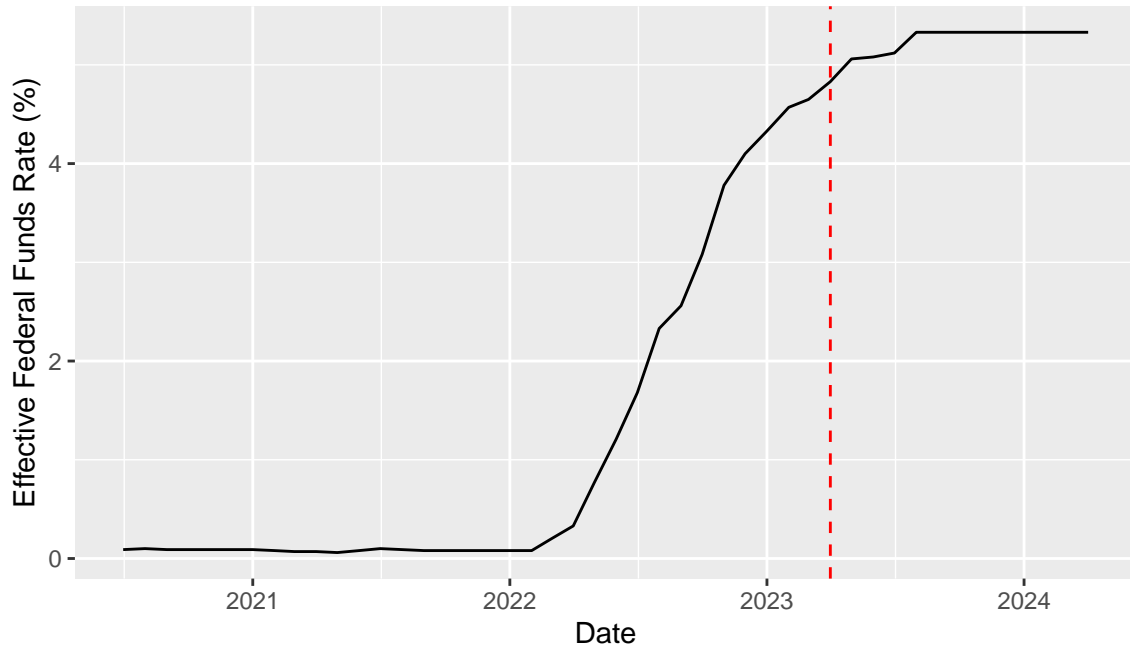


Figure 4: Effective Federal Funds Rate Over Time. Data from FRED, Federal Reserve Bank of St. Louis. Vertical dashed line indicates when Measure ULA took effect.

confounding factors could explain the change in project size observed by Ward and Phillips, we repeat their analysis and add the federal funds rate to the regression model as a control variable. As Figure 4 shows, the federal funds rate increased during the same time period in which Measure ULA went into effect. And indeed, with the federal funds rate included as a control, the Measure ULA indicator term is no longer statistically significant and has a value near zero.¹³ Exponentiating the coefficients, we could interpret these results as follows: after controlling for the federal funds rate, Measure ULA is associated with a statistically *insignificant* 2.9% decrease in average project size, while a one percentage point increase in the federal funds rate is associated with a statistically *significant* 16% decrease in average project size.¹⁴ That is, the change in the federal funds rate is a much better predictor of the change in the size of projects than the timing of Measure ULA taking effect.

Does this mean that the change in project size was caused by the change in the federal funds rate and that Measure ULA had no impact on project size? No. There are many other potentially confounding factors and potential interactions between factors. And, with such a small sample size, multicollinearity between the Measure ULA indicator and the federal funds rate could mask a potential association between Measure ULA and changes in project size. But it is clear that the simple pre/post analysis reported by Ward and Phillips does not provide compelling evidence that Measure ULA “caused” the change in project size observed in their data.

Additional errors and questionable choices in the Ward and Phillips building permit analysis.

The most significant findings of this reanalysis are that the Ward and Phillips building permit data do not show any meaningful change in the rate of multifamily housing production after Measure ULA took effect, and that the change in project size observed in their data is better explained by changes in other variables, such as the federal funds rate. In addition to these primary issues, our reanalysis identified a number of other errors or questionable analysis choices:

- As mentioned above, the regression model presented in Ward and Phillips’ report on page 32 is a

¹³See Table A4 in Appendix.

¹⁴The federal funds rate increased by over five percentage points between July 2020 and April 2024.

simplified form of the model that produced the results reported. There is no discussion in the text justifying the regression model that produced the reported results. The simple model presented on page 32 of “Taxing Tomorrow” does not produce a statistically significant coefficient on the Measure ULA indicator term (see Table A1 in the appendix of this paper).

- Ward and Phillips used the “delta method” to calculate the confidence interval for the change in average project size. That method is an approximation that often requires large sample sizes to be valid. In this case, choosing to use the delta method exaggerates the significance of their finding. A more appropriate and accurate statistical method produces wider confidence intervals that are shifted to the right, as shown in Figure A1 of the appendix.
- Ward and Phillips excluded *all* projects that used Executive Directive 1 (ED1) from their analysis. ED1 is indeed a confounding factor that coincided with the adoption of Measure ULA that should be accounted for in some way. But the total exclusion of ED1 projects likely excluded projects that would have received building permits even if ED1 had not been enacted. The total exclusion of ED1 projects is likely an over-correction. It reduces the number of units permitted in the “post-ULA” period, which may bias the results. For illustration, the number of units at each project-size threshold is shown in Table A5 of the appendix with ED1 projects included.

Conclusion

This partial reanalysis of the Ward and Phillips “Taxing Tomorrow” report finds that the conclusions regarding multifamily permit activity are not supported by the data it relied on. In fact, the data do not show any reduction in the number of multifamily units permitted on land sold after Measure ULA took effect. And changes in project size are better explained by changes in other factors, such as the federal funds rate, than by the timing of Measure ULA taking effect. Accordingly, the data do not support the claim that Measure ULA “caused” a decline in multifamily housing production.

Appendix

Replication of results from the Ward and Phillips regression model.

Tables A1 and A2 compare the simplified model described in Ward and Phillips with the fuller model that reproduces the “20+ unit” column of their Table A.3.

Table A1 reports the simplified regression specified on page 32 of Ward and Phillips. Standard errors are shown in parentheses.

Table A1: Simple Regression Predicting Project Size from Post-ULA Indicator

	<i>Dependent variable:</i>
	Log Units
Post ULA	-0.461 (0.297)
Constant	4.113*** (0.190)
Observations	27
Log Likelihood	-30.825
Akaike Inf. Crit.	65.650
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

The simple specification does not result in a statistically significant coefficient on the ULA indicator term.

Table A2 reports the actual regression used by Ward and Phillips to estimate the change in project size for projects with 20 or more units.

Table A2: Replication of Table A.3, 20+ Unit Projects

	<i>Dependent variable:</i>
	Log Units
Post ULA	-0.689** (0.307)
Land Improvement Ratio	0.00000* (0.00000)
Multi-Family Zoning	-0.750** (0.273)
Constant	4.306*** (0.283)
Observations	27
Log Likelihood	-25.630
Akaike Inf. Crit.	59.261
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

These results match the “20+ unit” column of Table A.3 in Ward and Phillips to three decimal places. This regression is testing for the change in project size on land sold after Measure ULA took effect, controlling for

certain property characteristics, and giving each month equal weight. It is not testing for a change in the monthly rate of housing production.

Regression results using the log of total units permitted per month as the outcome variable.

Table A3 reports the results of a regression model that uses the log of total units permitted per month as the outcome variable.

Table A3: Regression of Log Total Units Per Month on Post-ULA Indicator (20+ Units)

	<i>Dependent variable:</i>
	Log Total Units Per Month
Post ULA	0.780 (0.755)
Constant	1.575*** (0.395)
Observations	44
R ²	0.025
Adjusted R ²	0.002
Residual Std. Error	2.232 (df = 42)
F Statistic	1.065 (df = 1; 42)
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

The coefficient on the post-ULA indicator term is positive but not statistically significant, indicating no meaningful change in the number of units permitted per month after Measure ULA took effect.

Regression results including the federal funds rate as a control variable.

Table A4 adds the effective federal funds rate to the Ward and Phillips project-size model to illustrate the importance of controlling for other factors that coincided with the implementation of Measure ULA.

Table A4: Predicting Project Size with Federal Funds Rate

	<i>Dependent variable:</i>
	Log Units
Post ULA	-0.029 (0.301)
Land Improvement Ratio	0.00000** (0.00000)
Multi-Family Zoning	-0.606** (0.222)
Effective Federal Funds Rate	-0.175** (0.064)
Constant	4.551*** (0.275)
Observations	27
Log Likelihood	-23.344
Akaike Inf. Crit.	56.688
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

The coefficient on the post-ULA indicator term is no longer statistically significant and is close to zero, while the coefficient on the effective federal funds rate is negative and statistically significant.

Including ED1 units increases monthly averages in the post-ULA period.

Table A5 reports the monthly unit-production results when projects permitted using the streamlined Executive Directive 1 (ED1) process are included.

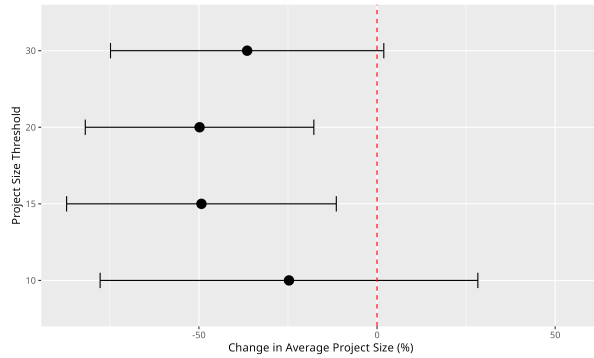
Table A5: Permitted units by project size threshold before and after Measure ULA, including ED1 projects.

Unit threshold	Avg. units/mo. pre	Avg. units/mo. post	Statistically significant?	N
10+ units	48.6	69.4	No (p=0.445)	48
15+ units	46.0	68.3	No (p=0.409)	40
20+ units	43.5	61.5	No (p=0.506)	30
30+ units	40.6	54.1	No (p=0.621)	22

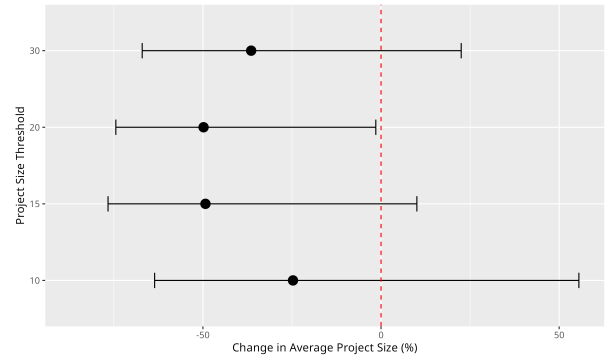
Including ED1 projects increases the post-ULA monthly unit averages, but the pre/post differences remain statistically insignificant at each project-size threshold.

Using the delta method to calculate confidence intervals exaggerates the result.

Figure A1 compares the delta-method confidence intervals used in the Ward and Phillips figure with confidence intervals produced by directly transforming the original log-scale confidence intervals. The direct transformation produces wider confidence intervals that are shifted to the right, reducing the apparent statistical significance of the original estimates.



(a) Recreation of Figure 9 using delta method confidence intervals



(b) Reanalysis using direct transformation of confidence intervals

Figure A1: Comparison of the original Figure 9 estimates with different methods for calculating confidence intervals. The figure on the left uses the delta method to estimate the standard error of the transformed estimate, which is not appropriate for this sample size and exaggerates the significance of the findings. The figure on the right directly transforms the confidence interval from the original estimate, showing wider and more accurate confidence intervals that are shifted to the right.