Online Appendix for "Political Consumerism and the Emergence of Rare Information on User-Generated Content Platforms"

Uttara M. Ananthakrishnan
Heinz College of Information Systems and Public Policy, Carnegie Mellon University, Pittsburgh, PA 15213
Email: uttara@cmu.edu

Xingchen Xu
Michael G. Foster School of Business, University of Washington, Seattle, WA 98195, USA
Email: xcxu21@uw.edu

Maria Mitkina
Uber Technologies, Inc., San Francisco, CA 94158, USA
Email: mmitkina@uber.com

Yong Tan
Michael G. Foster School of Business, University of Washington, Seattle, WA 98195, USA
Email: ytan@uw.edu

Appendix A: Summary of Identification Challenges and Related Checks

Table A1: Summary of Concerns and Analyses

Concerns	Related Reasoning and Analyses	Related Literature	Related Section
Reverse Causality: An increase in visits may lead to a greater volume of reviews, thereby enhancing the likelihood of receiving a Black-owned review (earlier).	1. Non-Switchback Treatment: Once a restaurant receives its first review mentioning Black ownership, it remains treated, as the Yelp system continues to display it in search results for "Black-owned restaurants". This context exhibits a clear *non-switchback* treatment structure, thereby fundamentally mitigating concerns about reverse causality.	(Bojinov et al. 2023, Chan et al. 2019, Burtch et al. 2018)	Section 4.2.1
	2. Random Entry: We use basic statistics and graphs to demonstrate that consumers rarely mention the owner's race in online reviews, and the timing at which different restaurants receive their first Black-owned review is randomly distributed.	(Gu and Kannan 2021, Petrova et al. 2021, Deshpande and Li 2019, Burtch et al. 2018)	Section 4.2.1 and Appendix B
	3. Statistical Test of Random Entry: We examine whether an increased number of reviews or any other restaurant characteristics raises the likelihood of a restaurant receiving Black-owned review, and finds no evidence supporting this relationship.	(Gu and Kannan 2021, Petrova et al. 2021, Deshpande and Li 2019, Burtch et al. 2018)	Section 4.2.1 and Appendix B
	4. Debiasing with the Semi-Parametric Odds Ratio Model (SORM): We employ a semi-parametric approach to jointly model the treatment assignment and outcome generation processes, effectively mitigating potential endogeneity issues.	(Qian and Xie 2024, Feit and Bradlow 2021, Qian and Xie 2014)	Section 4.2.1 and Appendix C
Simultaneity: In the first treatment month, both the treatment variable and the post-treatment outcome change concurrently.	5. Excluding the first month of treatment: We exclude the first treatment period for all treated units in our analysis. Upon re-running the analyses without this period, we obtain consistent results.	(Wooldridge 2010, Angrist and Pischke 2009)	Section 4.2.1 (Table 2, Col 4)
	6. Debiasing with SORM: As discussed in point 4, the SORM is also effective in addressing simultaneity concerns.	(Qian and Xie 2024, Feit and Bradlow 2021)	Section 4.2.1 and Appendix C
Omitted Variable Bias: Some unobservables might affect both the treatment and outcome variables.	7. Black-owned Sample Only: To enhance sample comparability, our primary analyses restrict the sample to Black-owned restaurants, although including non-Black-owned establishments yields similar results.	(Angrist and Pischke 2009)	Entire analyses and Appendix N
	8. Fixed Effects: If unobservable factors are restaurant-invariant, time-specific, or if there are city-level unobservable time trends, our fixed effects in the model account for these influences.	(Chan et al. 2019, Burtch et al. 2018, Angrist and Pischke 2009, Autor 2003)	Section 4.1 (Table 2, Col 1)
	9. Only "not yet" treated as control: By focusing exclusively on the treated sample, we eliminate potential biases arising from unobservables that might differentially affect treated and never-treated units.	(Hoynes et al. 2016)	Section 4.1 (Table 2, Col 2)
	10. Relative Time Model: By verifying that treated and control groups exhibit similar trends prior to treatment, we ensure that unobservable time-varying factors existing before treatment do not confound our treatment effect estimation.	(Angrist and Pischke 2009, Autor 2003, Bertrand et al. 2004)	Section 4.1 and Appendix P

Continued on next page

Table A1 continued

Concerns	Related Reasoning and Analyses	Related Literature	Related Section
	11. Additional Restaurant-Month Fixed Effect: We include additional restaurant-month fixed effects to account for restaurant-specific time-varying trends, thereby providing a conservative estimation.	(Burtch et al. 2018, Hendricks and Sorensen 2009)	Section 4.1 (Table 2, Col 3)
	12. Placebo Test: After assigning a placebo treatment to restaurants that were never treated, we find no significant differences between these placebo-treated restaurants and the control group.	(Burtch et al. 2018)	Section 4.2.2 and Appendix H
	13. Coarsened Exact Matching (CEM): Employing CEM, we mitigate the potential differences that can be identified from covariates and reinforce the robustness.	(Zervas et al. 2017, Iacus et al. 2012)	Section 4.2.2 and Appendix E
	14. Debiasing with SORM: As discussed earlier in point 4, the SORM is also effective in addressing the omitted variable bias.	(Qian and Xie 2024, Feit and Bradlow 2021)	Section 4.2.1 and Appendix C
	15. Sensitivity Analyses: We perform two sensitivity analyses to demonstrate that our main estimates remain robust even in the presence of high levels of unobservable confounders, strengthening the validity of our results.	(Stechemesser et al. 2024, Sen et al. 2024, Petrova et al. 2023, Cinelli and Hazlett 2020, Altonji et al. 2005)	Section 4.2.1 and Appendix D
Validity of the Outcome Variable.	16. Alternative Measurement: To ensure the validity of our outcome variable, we reconstruct it by incorporating time spent and demonstrating the robustness of our results.	(Li and Wang 2020)	Appendix F
Restaurants may receive more “Black-owned” reviews and visits during periods of heightened national interest.	17. Excluding the restaurants treated during increased national attention: In our main analyses, we exclude the restaurants that received their first “Black-owned” review during these periods. With these restaurants included, all the analyses still hold.	(Agarwal and Sen 2022)	Entire Analyses and Appendix L
	18. Subsample Analyses by Excluding the Periods after the BLM protests: We exclude the periods after May 2020 and rerun the analyses, showing consistent results.	(Agarwal and Sen 2022)	Section 4.2.1 (Table 2, Col 5)
Restaurants that receive “Black-owned” reviews might have visual cues that directly attract visitors and confound the effect.	19. Subsample Analyses with Offline Visual Cues removed: We manually exclude samples displaying visual cues indicating the owner’s race and perform subsample analyses, which rule out this alternative explanation.	(Zhang et al. 2021)	Appendix G
The impact of other similar platforms is not considered in this analysis.	20. Google Maps Control: We collect review data from Google Maps and reconstruct related variables. Incorporating these variables into our regression analyses, we demonstrate that our conclusions remain valid.	(Burtch et al. 2018)	Appendix I

Appendix B: Randomness Timing of First Review

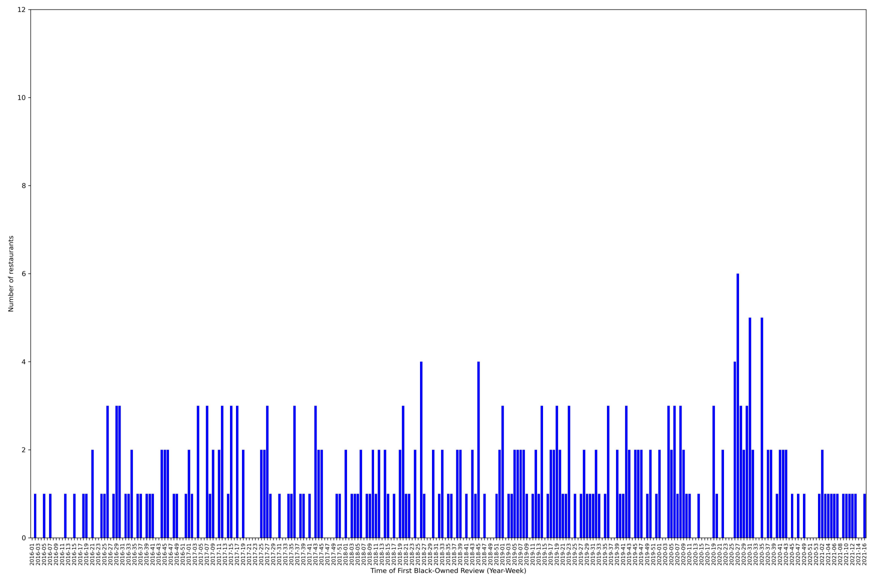


Figure B1 Distribution of Timing of First Black-owned Review

Note: The figure shows the number of restaurants with the first "Black-owned" review over time. Observations before 2016 are not shown in this graph.

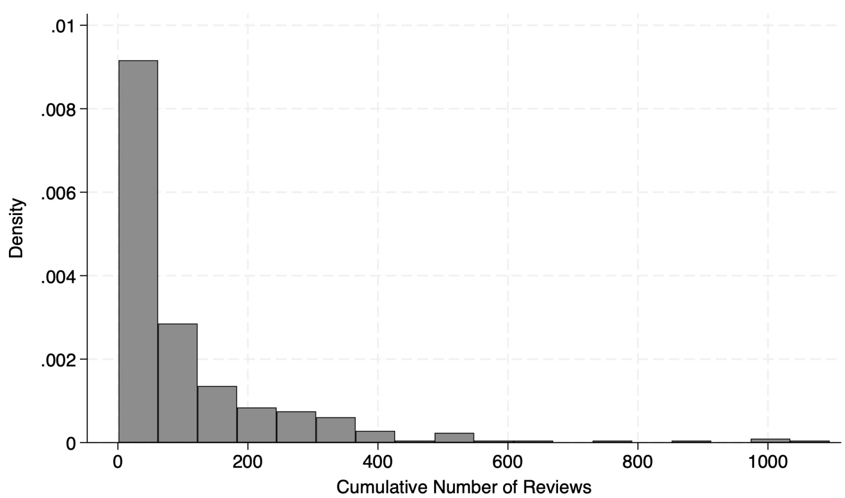


Figure B2 Distribution of Pre-Treatment Cumulative Number of Reviews

Note: The figure shows the distribution of cumulative number of reviews until the month when the first black-owned review appears.

We include a set of control variables such as the cumulative number of reviews (serving as a proxy for prior traffic and popularity), distance from other Black-owned restaurants, price level, and the political

Table B1 Effect of Cumulative Number of Reviews and Other Characteristics on Getting Black-owned Reviews

	(1) AfterBlackOwnedReview=1	(2) AfterBlackOwnedReview=1
CumulativeNumReviewsPast	0.0000636	0.000909
	(0.000441)	(0.000648)
DistanceBlackOwned	0.0839	0.0980
	(0.294)	(0.221)
Democratic	-0.160	-0.165
	(0.151)	(0.120)
PriceLevel1	-0.261	-0.247
	(0.330)	(0.347)
PriceLevel2	-0.306	-0.282
	(0.225)	(0.274)
PriceLevel3	-0.134	-0.291
	(0.392)	(0.570)
Constant	-2.958***	-
	(0.290)	-
Method	Logistic Regression	Survival Analyses
Observations	9124	9124

Note: Robust standard errors are reported. * p<0.10, ** p<0.05, *** p<0.01.

orientation of the city (classified as Democratic or Republican based on the 2020 Presidential elections). Our dependent variable is a binary indicator that remains at 0 until treatment occurs and switches to 1 once the restaurant receives its first Black-owned review. We then employ both logistic regression and Cox proportional hazards models to examine whether any of these control variables significantly influence the probability that a restaurant receives a review mentioning Black ownership, thereby affecting its treatment status. Notably, in both models, the coefficients associated with the covariates are statistically insignificant, further supporting the randomness of treatment timing.

Appendix C: More Details on the Semiparametric Odds Ratio Model

Before constructing the model, we encounter an obstacle: unlike the cross-sectional setting, the DiD setting involves high-dimensional fixed effects. Directly including all these dummy variables is computationally infeasible, yet they are essential to the model specification. Therefore, we adopt a fixed-point approach to iteratively demean all three fixed effects (Gaure 2013), which is also the underlying philosophy of popular packages handling high-dimensional categorical variables, such as `reghdfe`. Specifically, as outlined in Algorithm C1, we construct adjusted treatment and outcome variables by simultaneously partialing out the fixed effects. In each iteration, for every type of fixed effect, we first compute the category mean for each possible value of that fixed effect and then demean each observation by its corresponding categorical mean. For instance, considering the restaurant fixed effect (α_i), we calculate the mean of both the treatment variable $\widehat{D_{it}}$ and the outcome variable $\widehat{Y_{it}}$ for each restaurant i, and subsequently subtract these mean values from the original variables. This demeaning process is applied iteratively to all three fixed effects until the values of both $\widehat{D_{it}}$ and $\widehat{Y_{it}}$ converge. Furthermore, if additional controls are employed for the outcome function and the nuisance function, these controls can be demeaned simultaneously, as demonstrated in Algorithm C1. Following this procedure, the influence of fixed effects is incorporated into the adjusted variables, allowing regressions to be performed directly on these variables without the need to include high-dimensional dummy variables. In fact, if we directly regress $\widehat{Y_{it}}$ on $\widehat{D_{it}}$, we will obtain the same estimate of β_1 as in Equation (1).

After the variable adjustment process, we consider two functions that must be estimated jointly: the outcome function (C.1), which quantifies the treatment effect of $\widehat{D_{it}}$ on $\widehat{Y_{it}}$ and represents our primary interest, and the nuisance function (C.2), which elucidates the relationship between the error term $\widehat{\epsilon_{it}}$ and the treatment variable $\widehat{D_{it}}$, thereby addressing potential endogeneity. It is worth noting that the numerical demeaning process disables the analytical traceability between the adjusted error term and the original ones. As a result, we must directly model these adjusted error terms, which may not be ideal but is worth documenting as a robustness check. Further, specifying both equations poses significant challenges, particularly when the treatment variable $\widehat{D_{it}}$ exhibits complex distributional characteristics—such as being bounded and derived from a binary variable in our case—and when the form of the regressor–error relationship is uncertain. To address these challenges and ensure robust estimation, the nuisance function is modeled semiparametrically using the SORM framework (Qian and Xie 2024, Feit and Bradlow 2021, Qian and Xie 2014, Chen 2004).

$$\widehat{Y_{it}} = \beta_0 + \beta_1 \widehat{D_{it}} + \widehat{\epsilon_{it}}. \tag{C.1}$$

$$f_\psi(\widehat{D_{it}} = d | \widehat{\epsilon_{it}} = e) = \frac{\eta_\gamma(d, d_0; e, e_0) f_\lambda(d|e_0)}{\int_d \eta_\gamma(d, d_0; e, e_0) f_\lambda(d|e_0)\, d(d)} \tag{C.2}$$

While the outcome function (C.1) is straightforward to interpret, analogous to a standard regression function (Qian and Xie 2024), debiasing it necessitates a comprehensive understanding of the nuisance function (C.2), which comprises two components: the baseline function $f_\lambda(d \mid e_0)$ and the odds-ratio function $\eta_\gamma(d, d_0; e, e_0)$.

The baseline function represents the marginal distribution of the treatment variable, conditional on the error term being fixed at a specific value e_0. To minimize the risk of model misspecification, we model this baseline function $f_\lambda(d \mid e_0)$ nonparametrically by adopting the empirical distribution of the treatment variable

$\widehat{D}$ derived from the data (Qian and Xie 2024). Specifically, conditional on e_0, we observe m unique values of $\widehat{D}$ with corresponding probabilities $\mathbf{p} = (p_1, \ldots, p_m)$. During the estimation process, these probabilities are reparameterized as $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m)$, where $\lambda_i = \ln\left(\frac{p_i}{p_m}\right)$ and $p_i = \frac{\exp(\lambda_i)}{\sum_{j=1}^{m}\exp(\lambda_j)}$, allowing for unrestricted optimization in the estimation procedure.

Furthermore, the odds-ratio function is more complex, as detailed below.

$$\eta_\gamma(d, d_0; e, e_0) = \frac{f_\psi(d|e) f_\lambda(d_0|e_0)}{f_\psi(d_0|e) f_\lambda(d|e_0)} \tag{C.3}$$

The odds-ratio function $\eta_\gamma(d, d_0; e, e_0)$ quantifies the ratio of the odds of observing the treatment variable $\widehat{D}$ at value d relative to d_0 as the error term ϵ varies from e_0 to e. This concept is analogous to the odds ratio used to interpret coefficients in logistic regression models. In principle, the odds-ratio function may assume any functional form; however, consistent with the approach of Qian and Xie (2024), we adopt a log-bilinear specification for the odds-ratio function due to its simple functional form and inherent flexibility.

$$\ln \eta_\gamma(d, d_0; e, e_0) = \gamma^E (e - e_0)(d - d_0) \tag{C.4}$$

Therefore, with both the baseline function $f_\lambda(d \mid e_0)$ and the odds-ratio function $\eta_\gamma(d, d_0; e, e_0)$ specified, we successfully derive the nuisance function using a semiparametric odds-ratio approach. Consequently, we obtain the corresponding likelihood contribution for each observation as $f_\psi(\widehat{D}_{it} \mid \widehat{\epsilon_{it}})$. For the outcome function, analogous to standard regression models, we assume that the error term ($\widehat{\epsilon_{it}} = \widehat{Y}_{it} - \beta_0 - \beta_1 \widehat{D}_{it}$) follows a normal distribution, $\widehat{\epsilon_{it}} \sim \mathcal{N}(0, \delta^2)$, thereby obtaining the corresponding likelihood. By combining both components across all units and time periods, we derive the total log-likelihood function as follows:

$$\log \mathcal{L}(\beta_0, \beta_1, \gamma^E, \delta^2) = \sum_{i=1}^{N}\sum_{t=1}^{T}\left[\ln f(\widehat{\epsilon_{it}} + \ln f_\psi(\widehat{D}_{it} \mid \widehat{\epsilon_{it}})\right] \tag{C.5}$$

Before conducting the formal estimation, a crucial identification assumption for the SORM is that the treatment variable should not follow a normal distribution. To assess this assumption, we perform both the Kolmogorov-Smirnov test and the Anderson-Darling, confirming that the distribution of $\widehat{D}_{it}$ significantly deviates from a normal distribution.

To construct uncertainty measurements while addressing potential model misspecification, we employ bootstrap sampling, repeating the estimation process 25 times to derive the standard errors. In each iteration, we randomly sample the same number of observations with replacement and perform maximum likelihood estimation (replacement at the restaurant level and re-demeaning yields consistent results).

Finally, to mitigate computational intensity, particularly for the nonparametric components, we adopt the profile likelihood approach and, following Qian and Xie (2024), retain only the first two decimal places for independent variables (treatment and control variables) during bootstrap sampling.[23]

In the absence of control variables, the debiased estimate of β_1 is 0.093 ($p < 0.01$). We further incorporate the cumulative number of reviews as a control variable in the outcome function and adjust the nuisance function to condition on both the error term and the (rounded) control variable. Upon re-running the analyses, the estimate of β_1 becomes 0.185 ($p < 0.01$). These results collectively reinforce the robustness of our findings.

[23] The estimates are robust to this approximation process.

Algorithm C1 Iterative Demeaning Procedure (Partial Out Fixed Effects)

Initialize adjusted variables:
$\widehat{Y}_{it} \leftarrow LogVisits_{it}$
$\widehat{D}_{it} \leftarrow AfterBlackOwnedReview_{it}$
$\widehat{X}_{it} \leftarrow Controls_{it}$
Set convergence parameters:
tolerance $\leftarrow 1 \times 10^{-6}$
Define fixed effects:
fixed_effects $\leftarrow \{\alpha_i, \gamma_t, \mu_{ct}\}$
Initialize variables for convergence tracking:
prev_$\widehat{Y}_{it} \leftarrow \widehat{Y}_{it}$
prev_$\widehat{D}_{it} \leftarrow \widehat{D}_{it}$
prev_$\widehat{X}_{it} \leftarrow \widehat{X}_{it}$
Initialize max_change $\leftarrow \infty$
while max_change > tolerance **do**
    for all fe **in** fixed_effects **do**
        Compute group means:
        group_mean_$\widehat{Y}_{it} \leftarrow$ GroupBy(fe)[$\widehat{Y}_{it}$].mean()
        group_mean_$\widehat{D}_{it} \leftarrow$ GroupBy(fe)[$\widehat{D}_{it}$].mean()
        group_mean_$\widehat{X}_{it} \leftarrow$ GroupBy(fe)[$\widehat{X}_{it}$].mean()
        Demean adjusted variables:
        $\widehat{Y}_{it} \leftarrow \widehat{Y}_{it} -$ group_mean_$\widehat{Y}_{it}$
        $\widehat{D}_{it} \leftarrow \widehat{D}_{it} -$ group_mean_$\widehat{D}_{it}$
        $\widehat{X}_{it} \leftarrow \widehat{X}_{it} -$ group_mean_$\widehat{X}_{it}$
    end for
    Check convergence:
    $\widehat{Y}_{\text{max\_change}} \leftarrow \max\left(\left|\widehat{Y}_{it} - \text{prev\_}\widehat{Y}_{it}\right|\right)$
    $\widehat{D}_{\text{max\_change}} \leftarrow \max\left(\left|\widehat{D}_{it} - \text{prev\_}\widehat{D}_{it}\right|\right)$
    $\widehat{X}_{\text{max\_change}} \leftarrow \max\left(\left|\widehat{X}_{it} - \text{prev\_}\widehat{X}_{it}\right|\right)$
    max_change $\leftarrow \max\left(\widehat{Y}_{\text{max\_change}}, \widehat{D}_{\text{max\_change}}, \widehat{X}_{\text{max\_change}}\right)$
    Update previous adjustments:
    prev_$\widehat{Y}_{it} \leftarrow \widehat{Y}_{it}$
    prev_$\widehat{D}_{it} \leftarrow \widehat{D}_{it}$
    prev_$\widehat{X}_{it} \leftarrow \widehat{X}_{it}$
end while

Appendix D: More Details on Omitted Variable Sensitivity Analyses

To evaluate the risk of omitted variable bias, we adopt the method proposed by Cinelli and Hazlett (2020) and assess the relative sensitivity of unobservables compared to a set of control variables derived from Yelp review content. To address the computational challenges posed by high-dimensional fixed effects - analogous to our approach for SORE - we apply the demeaning process described in Algorithm C1 to partial out these effects. Subsequently, we employ the method described by Cinelli and Hazlett (2020) and illustrate the sensitivity of the main estimates to unobserved confounders. In the following graphs, the horizontal axis represents the hypothetical residual share of the treatment variation explained by unobserved confounding, denoted as $R^2_{D\sim Z|X}$, which is derived from regressing the treatment variable D on the unobserved confounder Z while controlling for observed covariates X. The vertical axis denotes the hypothetical partial R^2 of the unobserved confounder with the outcome, $R^2_{Y\sim Z|X,D}$, obtained by regressing the outcome variable Y on Z while controlling for X and D, where Y is the outcome, D is the treatment, X represents the observed covariates, and Z is a single unobserved confounder.

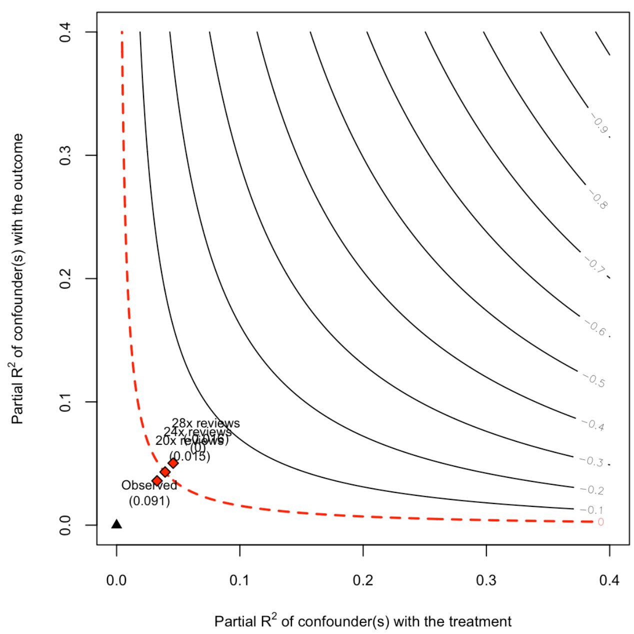


Figure D1 Sensitivity Analyses (Cumulative Reviews as the Benchmark)

Since the unobserved variable Z is not available, we approximate its effect by leveraging the observed covariates. The contour lines in the plot represent the potential treatment effect estimates that would have been obtained from a full regression model incorporating hypothetical unobserved confounders with specified strengths. The dashed line indicates a hypothetical treatment effect of zero, whereas the diamond markers denote the estimated treatment effects under scenarios in which the unobserved confounder is assumed to be a scalar multiple of the observed covariates. Each marker is accompanied by a numeric label that indicates the estimated treatment effect when the confounding is as strong as the corresponding level of the observed

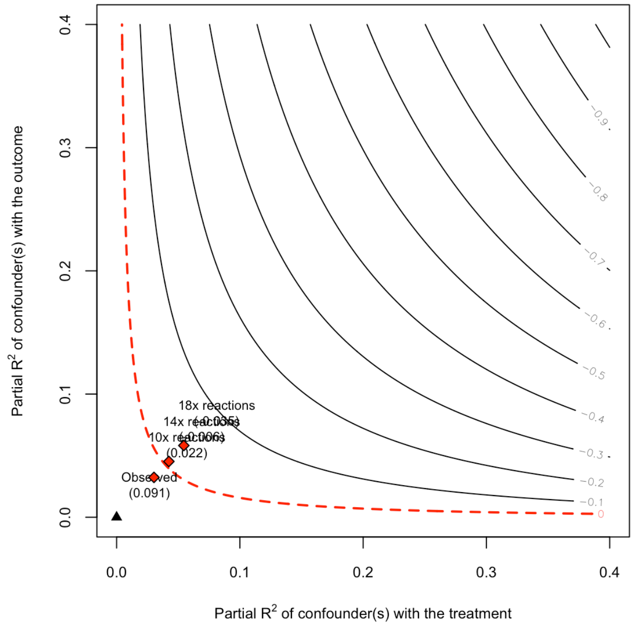


Figure D2 Sensitivity Analyses (Cumulative Reviews with Reactions as the Benchmark)

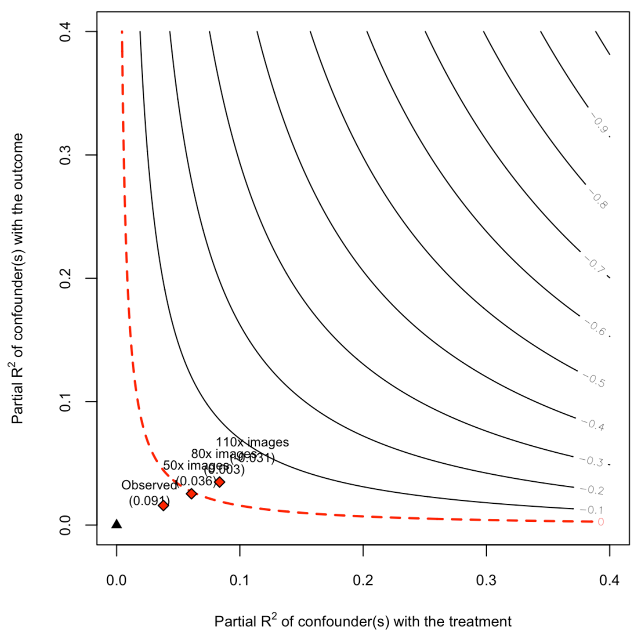


Figure D3 Sensitivity Analyses (Cumulative Reviews with Images as the Benchmark)

covariates. For example, the left marker in Figure D1 illustrates that when the confounding is 20 times as strong as this covariate, the effect remains positive (0.015).

Collectively, these contour plots demonstrate that the direction of the treatment effect remains robustly negative even when confounding is more than 14 times as strong as that of the observed covariates, a threshold that substantially exceeds the sensitivity limits reported in previous literature (Stechemesser et al. 2024).

In addition to the Cinelli-Hazlett method, we also perform sensitivity analyses using the traditional approach proposed by Altonji et al. (2005), which has been widely employed in previous empirical studies (Sen et al. 2024, Petrova et al. 2023, Galor and Özak 2016, Nunn and Wantchekon 2011). Specifically, we re-estimate the model by incorporating additional control variables to obtain an alternative set of estimates. We then compare the new coefficient to that obtained from the regression without the additional controls, using the ratio defined as $\frac{\beta_{\text{original}}}{\beta_{\text{original}}-\beta_{\text{control}}}$. The underlying rationale for this ratio is straightforward: a smaller difference between the original and controlled estimates suggests that the parameter of interest is less affected by selection on observables, thereby requiring a stronger selection on unobservables (relative to observables) to fully account for the observed effect. Conversely, a larger numerator indicates that a greater proportion of the effect is attributable to selection on unobservables, resulting in a higher focal ratio.

$$\begin{aligned} LogVisits_{it} = \beta_0 + \beta_1 AfterBlackOwnedReview_{it} + \beta_2 \mathbf{Control_{it}} \\ + \alpha_i + \gamma_t + \mu_{ct} + \epsilon_{it} \end{aligned} \tag{D.1}$$

Table D1 **Effect of Yelp Reviews with Yelp Review Controls**

	(1) LogVisits
AfterBlackOwnedReview=1	0.0905**
	(0.0372)
Restaurant FE	Yes
Year-Month FE	Yes
City-Year-Month FE	Yes
Observations	18070
R^2	0.865

Note: Cluster-robust standard errors are reported at the restaurant level. * p<0.10, ** p<0.05, *** p<0.01.

After incorporating these observables, the coefficient β_1 changes from 0.0952 to 0.0905. The similarity of the coefficients, as reflected by a ratio of $\frac{0.0952}{0.0905-0.0952} = -20.3$, suggests that selection on unobservables would have to be more than 20 times stronger than selection on observables to fully explain the observed effect, thereby further mitigating concerns regarding omitted variable bias.

Appendix E: Coarsened Exact Matching Results

Although our fixed effects account for unobservable differences among restaurants in the DiD framework and prior analyses indicate minimal endogeneity concerns, we aim to enhance the comparability of observations before estimation as another robustness check.

To achieve this, we apply coarsened exact matching (CEM) to derive sample weights based on observable restaurant characteristics (Pattabhiramaiah et al. 2022, Mayya et al. 2021, Blackwell et al. 2009). Continuous covariates ($Dissimilarity$ and $Isolation$) are coarsened into four equally spaced bins, and $AverageVisits$ is divided into three; all remaining categorical covariates are matched exactly. Covariate balance is then assessed via weighted t-tests on the original variables using the CEM-derived weights. As shown in Table E1, post-weighting differences between treatment and control groups in all the variables are statistically insignificant.

Table E1 Covariate Balance

Variable	Treated Mean	Control Mean	% Bias	t-stat	p-value
AverageVisits	183	167.31	12.4	1.55	0.123
Dissimilarity	32.71	33.241	-4.7	-0.68	0.495
Isolation	54.333	54.681	-6.1	-0.80	0.423
PriceLevel	1.8338	1.8073	3.7	0.58	0.559
Democratic	0.74212	0.74924	-1.6	-0.21	0.837
BlackNeighborhood	0.56447	0.56575	-0.3	-0.03	0.974

Applying the CEM-derived weights, we re-estimate the regression models with the full set of fixed effects to further mitigate bias from unobservables. The weighted estimates in Table E2 remain consistent with our original findings.

Table E2 Effect of Yelp Reviews after CEM

	(1)
	LogVisits
AfterBlackOwnedReview=1	0.1318***
	(0.0455)
Restaurant FE	Yes
Year-Month FE	Yes
City-Year-Month FE	Yes
Restaurant-Month FE	Yes
Observations	16184
R^2	0.901

Note: Cluster-robust standard errors are reported at the restaurant level. * p<0.10, ** p<0.05, *** p<0.01.

Appendix F: Alternative Outcome about Time Spent at a Restaurant

In this subsection, we verify that the outcome variable, defined as a logarithm of total foot traffic, is a reasonable approximation of sales. We utilize an alternative approach to rule out short visits unrelated to actual sales, similar to Li and Wang (2020). In particular, we use an alternative metric SafeGraph provides that breaks down the visits by the times spent at the location. For this analysis, we only include visitors who spent more extended periods (between 10 minutes and 2 hours) at the venue. We define the new outcome variable $LogEatInVisits$ as the logarithm of the number of more extended visits. We present our results in Table F1.

Table F1 Effect of Yelp Reviews on Dine-In Visits

	(1) LogEatInVisits
AfterBlackOwnedReview=1	0.122***
	(0.0378)
Restaurant FE	Yes
Year-Month FE	Yes
City-Year-Month FE	Yes
Observations	17339
R^2	0.859

Note: Cluster-robust standard errors are reported at the restaurant level. * p<0.10, ** p<0.05, *** p<0.01.

As the table demonstrates, our results still hold and are similar to the baseline results, where we include all foot traffic. This result demonstrates that the overall foot traffic is a good approximation for longer visits.

Appendix G: Alternative Mechanism Check about Visual Cues of Black Ownership

One of the potential factors that may impact our estimates of the impact of online reviews revealing Black ownership on consumption is offline visual cues. For example, a restaurant can add posters outside the venue that inform potential customers about minority ownership.

To address this concern, we manually checked more than five thousand photos on the Yelp page of all the restaurants in our sample. The photos include visuals of the exteriors and interiors and checked if the restaurants had any posters or signs identifying them as Black-owned. Such visual cues are rare, with only 7 restaurants out of 696 restaurants having such displays. We re-run our baseline analysis excluding these restaurants from the sample and present our results in Table G1. Our results remain consistent even after removing restaurants with posters and signs indicating Black ownership.

Table G1 Effect of Yelp Reviews after Removing Restaurants with Offline Labels

	(1) LogVisits
AfterBlackOwnedReview=1	0.0959**
	(0.0380)
Restaurant FE	Yes
Year-Month FE	Yes
City-Year-Month FE	Yes
Observations	17888
R^2	0.864

Note: Cluster-robust standard errors are reported at the restaurant level. * p<0.10, ** p<0.05, *** p<0.01.

Appendix H: Placebo Test

The Placebo Test is a standard additional method for verifying the parallel trends assumption. It compares the performance of the control group with the performance of the placebo group, which is usually a subsample of control units with randomly assigned "treatment" periods (Gertler et al. 2016). This test checks whether the placebo group deviates from the control group's performance. The two groups should demonstrate similar performance trends over time since none had an actual treatment.

For this test, we select only the restaurants that were never visible as Black-owned on Yelp. Next, we randomly choose half of them in the "treatment group." The restaurants from the treatment group also get assigned a random month at the start of the treatment. Next, we run the basic specification on this sample. We present our results in Table H1. We find that none of the regression coefficients attributed to treatment are significant, demonstrating the test's validity.

Table H1 Effect of Yelp Reviews - Placebo Test

	(1) LogVisits
FalseTreatment=1	-0.0494
	(0.0482)
Restaurant FE	Yes
Year-Month FE	Yes
City-Year-Month FE	Yes
Observations	8700
R^2	0.851

Note: Cluster-robust standard errors are reported at the restaurant level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Appendix I: Controlling for exposure on Google Maps

In our context, the impact of Yelp reviews on visits to Black-owned restaurants is primarily driven by online discoverability after the search. More specifically, Yelp places restaurants with at least one black-owned review in the search results when consumers search for "Black-owned" restaurants.

It is possible that some of the foot traffic came to restaurants through searches on other platforms like Google. Google Maps works differently in that businesses could opt for Black-owned business badges, which were adopted later than Yelp. The badges are not visible from a traditional Google search but only when the users search from within Google maps. Further, reviews within Google Maps are not searchable, and the businesses shown for Black-owned are not necessarily the ones with Black-owned reviews or Black-owned badges. Therefore, it is less likely that traffic from Google would be biasing our results.

However, it is possible that consumers might browse a large number of reviews to identify Black-owned restaurants, potentially introducing bias into our results. Fortunately, fewer than 20% of the restaurants in our sample with Black-owned reviews on Yelp also have Black-owned reviews on Google. We utilize this information to conduct two additional analyses.

First, we identify the first Black-owned review for each restaurant on Google Maps and construct a new treatment variable, $AfterBlackOwnedReviewBoth$, which equals 1 if the restaurant has received Black-owned reviews on either platform and 0 otherwise.[24] Since the number of restaurants that obtained the Black owned reviews on Google before Yelp was so few, we do not expect the results to change that much. Our results show that this is indeed true, as shown in Column (1) of Table I1. Additionally, we rerun the regression with the cumulative number of Google reviews as a proxy of traffic based on Google reviews, and as indicated in Column (2), the treatment effect remains consistent.

$$\begin{aligned} LogVisits_{it} = \beta_0 + \beta_1 AfterBlackOwnedReviewBoth_{it} + \beta_2 GoogleControl_{it} \\ + \alpha_i + \gamma_t + \mu_{ct} + \epsilon_{it} \end{aligned} \tag{I.1}$$

Table I1 Effect of Yelp Reviews Considering Google Map's Impact

	(1)	(2)
	LogVisits	LogVisits
AfterBlackOwnedReviewBoth=1	0.121***	0.121***
	(0.0354)	(0.0354)
Google Controls	No	Yes
Restaurant FE	Yes	Yes
Year-Month FE	Yes	Yes
City-Year-Month FE	Yes	Yes
Observations	18070	18070
R^2	0.865	0.865

Note: Cluster-robust standard errors are reported at the restaurant level. * p<0.10, ** p<0.05, *** p<0.01.

[24] In the month that receives either type of review, this treatment variable is also set to 1. Excluding this month from the treated sample and considering only the month following treatment leads to consistent results.

Appendix J: Calculating Exposure

Massenkoff and Wilmers (2023) use the same data from SafeGraph that indicates the number of users from a census block group. We closely follow Massenkoff and Wilmers (2023) to calculate the probability that a restaurant patron of one group would encounter another patron of any other pre-defined group. We can calculate the exposure metrics for any group. In our analysis, we focus on racial groups. We aggregate the visits from each census block group to a census tract (CT) level. Further, we categorize each census tract as White or non-White based on the dominant racial group. First, we calculate the exposure of members of each group G in each CT to each establishment k without counting the focal group. The metric is calculated using the equation below:

$$FirmExposure_{iGk} = \frac{Visitors_{k(G-i)}}{\sum_{m \in I: m \neq i} Visitors_{k(m)}} \tag{J.1}$$

We calculate the $FirmExposure_{iGk}$ metrics for each group G in each census tract i for each restaurant without including the focal CT i. Then, we average the $FirmExposure_{iGk}$ metrics weighted by total visits from each CT i. In other words, the exposure of the non-White group (NW) to the White group (W) has two components. The numerator is the product of Firm exposure of members of W CTs at a restaurant k and the total number of visitors to the restaurant k from CT i. We sum over all W CTs indexed by i. Then, we sum the total visits by residents of W CTs. More specifically, we define $Exposure_k(NW, W)$ as

$$Exposure_k(NW, W) = \frac{\sum_{i \in W} FirmExposure_{iGk} \times Visitors_{k(i)}}{\sum_{i \in W} Visitors_{k(i)}} \tag{J.2}$$

The above equation gives us the mean exposure of White CT residents to non-White CT residents, weighted by total visits to a restaurant k. The procedure is exactly similar to the one followed in Massenkoff and Wilmers (2023) except that we do not sum over k. $Exposure_A(NW, W)$ is 0.65, which indicates that a visitor from a non-White CT at restaurant A is likely to be exposed to 65% of White residents. Given that there are only two groups, the average exposure is 50% when no segregation exists.

Thus, we calculate the exposure for residents from non-White CTs to White CTs for each of the restaurant in our sample over all the months in our study period. We use this as a dependent variable to study how reviews mentioning Black ownership change the exposure experienced by different racial groups at different restaurants.

Appendix K: EatOkra

We leverage EatOkra data to identify Black-owned restaurants, which might not have been identified as such on Yelp. Other studies in this literature address the issue of ground truth by performing image analytics on business owners' images (Zhang et al. 2021). However, Yelp restaurants rarely have photos of business owners, making this issue even more difficult. Fortunately, we obtain this information from the website "EatOkra".[25] Starting in 2016, EatOkra initially used various online resources to list Black-owned restaurants.

EatOkra is the only resource to identify Black-owned restaurants in a central location. We contacted the chiefs of Black Business Bureaus in all the ten cities in our data through our university's Consulting and Business Development Center, which had a Black Future Co-op Fund and Black-Led Business program. We even interviewed one of the Black business leaders in Seattle. All our contacts replied that such a consolidated list does not exist, and even governments that contract out to minority-business owners don't have access to such a list and rely on user-generated content in the form of lists that were available on forums and social media.

We looked into how EatOkra obtained these lists, and according to an interview with Salon,[26] "Over the next six months, Edwards built the website whenever he wasn't working at his full-time job at a construction project management software company. The couple input all the data themselves—some 350 mostly New York-based restaurants they'd unearthed like so many diners before: through word of mouth, articles, and informally shared Google Doc lists. They debuted the app in May 2016, calling it EatOkra in a nod to their Southern heritage and the ingredient's literal and symbolic role as a binding agent and connector of Black Americans. " They also started having partnerships with other organizations like U.S. Black Chambers (USBC) which offers certification for Black businesses.[27] In summary, EatOkra is a combination of user-generated lists, with manual verification and certifications from USBC.

As of September 2022, "there are more than 11,000 Black-owned restaurants nationwide are listed on EatOkra." and they were awarded one of 15 annual App Store Awards in December 2021.

[25] https://www.eatokra.com/

[26] https://www.salon.com/2022/09/13/eatokra-aims-to-be-the-next-for-black-owned-businesses/

[27] see https://certification.byblack.us/?_gl=1*12rxe79*_ga*MTI5NzQyODEyMS4xNjYxNzg1MzAy*_ga_T23J87ML1W*MTY3MjgONDk3My4xLjEuMTY3MjgONTA2Mi4wLjAuMA.. for the certification process

Appendix L: Analyses including Restaurants Treated during BLM-period

In this appendix, we replicate all regression analyses using a sample of restaurants that were treated during the period of peak attention following the BLM protests. Our findings demonstrate that all results remain consistent within this less restricted sample.

Main Results (Replication):

Table L1 Effect of Yelp Reviews Mentioning Black-Ownership on Restaurants' Visits

	(1) LogVisits	(2) LogVisits	(3) LogVisits	(4) LogVisits	(5) LogVisits
AfterBlackOwnedReview=1	0.117***	0.112***	0.132***	0.122***	0.0847**
	(0.0347)	(0.0352)	(0.0393)	(0.0365)	(0.0391)
Restaurant FE	Yes	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes	Yes	Yes
City-Year-Month FE	Yes	Yes	Yes	Yes	Yes
Never-Treated Excluded	No	Yes	No	No	No
Restaurant-Month FE	No	No	Yes	No	No
First Treatment Month Excluded	No	No	No	Yes	No
Obs after BLM Excluded	No	No	No	No	Yes
Observations	19162	10462	19144	18972	11066
R^2	0.866	0.881	0.907	0.865	0.923

Note: Cluster-robust standard errors are reported at the restaurant level. * p<0.10, ** p<0.05, *** p<0.01.

Event Study Analyses (Replication):

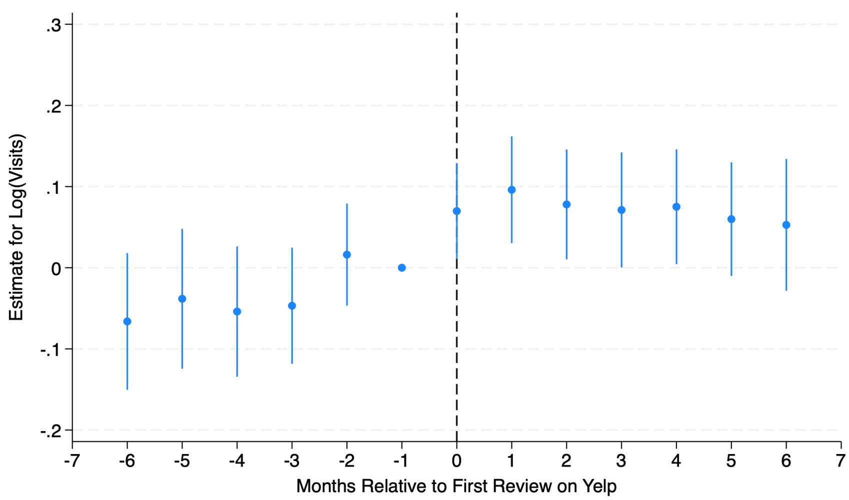


Figure L1 Event Study Analysis

Note: The figure shows relative time model estimates with log of visits as the dependent variable. The estimates are shown with a 95% confidence intervals.

Results of SORM (Replication):

We replicate the SORM analysis using the procedure detailed in Appendix C. The SORM estimate without controls is 0.113 ($p < 0.01$), while the estimate with controls is 0.253 ($p < 0.01$). Both estimates are positive and statistically significant.

Results of Sensitivity Analyses (Replication):

As shown in Table L1, the main estimate for the unrestricted sample is 0.117. After incorporating controls for this sample, the estimate decreases to 0.112, as presented in Table L2. Consequently, the ratio using Altonji's method is calculated as $\frac{0.117}{0.112-0.117} = -23.4$. Additionally, we present the sensitivity analysis graphs following the methodology of Cinelli and Hazlett (2020) in Figures L2, L3, and L4.

Table L2 Effect of Yelp Reviews with Yelp Review Controls

	(1)
	LogVisits
AfterBlackOwnedReview=1	0.112***
	(0.0345)
Restaurant FE	Yes
Year-Month FE	Yes
City-Year-Month FE	Yes
Observations	19162
R^2	0.866

Note: Cluster-robust standard errors are reported at the restaurant level. * p<0.10, ** p<0.05, *** p<0.01.

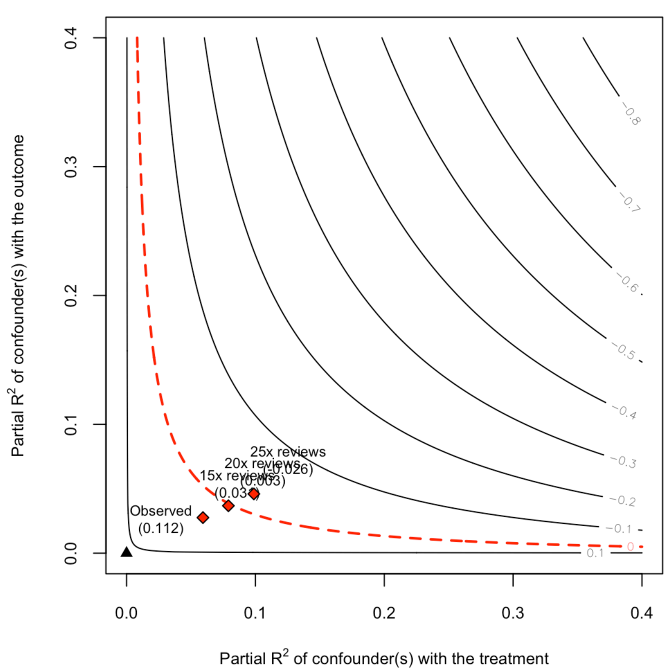


Figure L2 Sensitivity Analyses (Cumulative Reviews as the Benchmark)

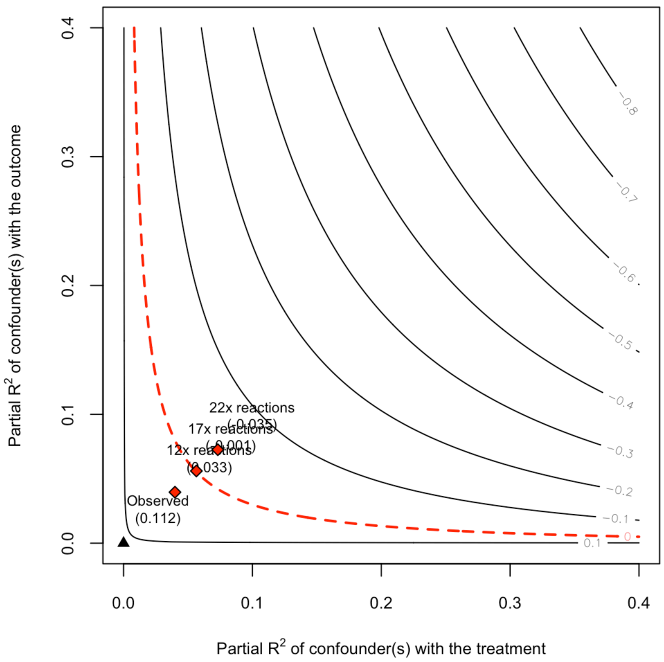


Figure L3 Sensitivity Analyses (Cumulative Reviews with Reactions as the Benchmark)

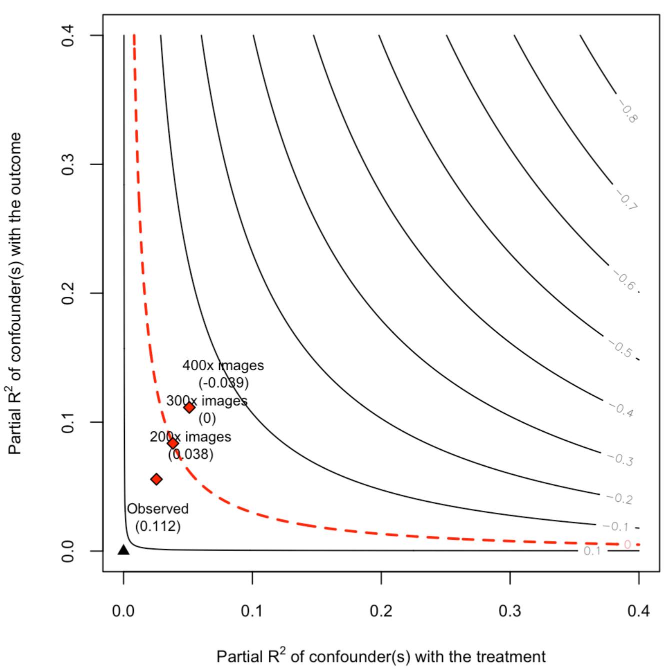


Figure L4 Sensitivity Analyses (Cumulative Reviews with Images as the Benchmark)

The contour plots indicate that the direction of the effect remains robustly negative even when confounding is more than 17 times as strong as that of the observed covariates, surpassing the threshold utilized in previous studies (Stechemesser et al. 2024).

Results of Coarsened Exact Matching (Replication):

We also perform coarsened exact matching on this sample and re-estimate the regression analyses, demonstrating that the results for the full sample remain robust to this matching procedure.

Table L3 Effect of Yelp Reviews after CEM

	(1) LogVisits
AfterBlackOwnedReview=1	0.1442***
	(0.0416)
Restaurant FE	Yes
Year-Month FE	Yes
City-Year-Month FE	Yes
Restaurant-Month FE	Yes
Observations	17484
R^2	0.902

Note: Cluster-robust standard errors are reported at the restaurant level. * p<0.10, ** p<0.05, *** p<0.01.

Results of Time Spent as the Alternative Outcome (Replication):

Furthermore, for this sample, we also define use an alternative outcome variable, *LogEatInVisits*, representing the logarithm of the number of more extended visits, and present the corresponding results in Table L4.

Table L4 Effect of Yelp Reviews on Dine-In Visits

	(1) LogEatInVisits
AfterBlackOwnedReview=1	0.143***
	(0.0352)
Restaurant FE	Yes
Year-Month FE	Yes
City-Year-Month FE	Yes
Observations	18376
R^2	0.859

Note: Cluster-robust standard errors are reported at the restaurant level. * p<0.10, ** p<0.05, *** p<0.01.

Results Considering Visual Cues of Black Ownership (Replication):

Table L5 Effect of Yelp Reviews after Removing Restaurants with Offline Labels

	(1) LogVisits
AfterBlackOwnedReview=1	0.121***
	(0.0352)
Restaurant FE	Yes
Year-Month FE	Yes
City-Year-Month FE	Yes
Observations	18928
R^2	0.865

Note: Cluster-robust standard errors are reported at the restaurant level. * p<0.10, ** p<0.05, *** p<0.01.

Results of Placebo Test (Replication):

Table L6 Effect of Yelp Reviews - Placebo Test

	(1) LogVisits
FalseTreatment=1	-0.0494
	(0.0482)
Restaurant FE	Yes
Year-Month FE	Yes
City-Year-Month FE	Yes
Observations	8700
R^2	0.851

Note: Cluster-robust standard errors are reported at the restaurant level. * p<0.10, ** p<0.05, *** p<0.01.

Results with Google Map's impact Considered (Replication):

Table L7 Effect of Yelp Reviews Considering Google Map's Impact

	(1) LogVisits	(2) LogVisits
AfterBlackOwnedReviewBoth=1	0.137***	0.137***
	(0.0328)	(0.0329)
Google Controls	No	Yes
Restaurant FE	Yes	Yes
Year-Month FE	Yes	Yes
City-Year-Month FE	Yes	Yes
Observations	19162	19162
R^2	0.866	0.866

Note: Cluster-robust standard errors are reported at the restaurant level. * p<0.10, ** p<0.05, *** p<0.01.

Heterogeneity by City Level Segregation (Replication):

Table L8 Effect of Yelp Reviews by City Level Segregation

	(1) LogVisits	(2) LogVisits	(3) LogVisits	(4) LogVisits
AfterBlackOwnedReview=1	0.0606	0.144***	0.0507	0.181***
	(0.0755)	(0.0368)	(0.0536)	(0.0439)
Restaurant FE	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes	Yes
City-Year-Month FE	Yes	Yes	Yes	Yes
High Dissimilarity	Yes	No	-	-
High Isolation	-	-	Yes	No
Observations	6811	12351	10264	8898
R^2	0.867	0.863	0.865	0.866

Note: Cluster-robust standard errors are reported at the restaurant level. * p<0.10, ** p<0.05, *** p<0.01.

Effects with Moderators Considered (Replication):

Table L9 Effect of Yelp Reviews with Moderators

	(1) LogVisits	(2) LogVisits	(3) LogVisits	(4) LogVisits	(5) Exposure
AfterBlackOwnedReview=1	0.0963*	0.120***	0.0337	0.140***	0.0374***
	(0.0569)	(0.0407)	(0.0680)	(0.0391)	(0.0103)
AfterBlackOwnedReview=1 × BlackInterest=1	0.0595***	0.0199	0.0365	0.0485***	
	(0.0197)	(0.0170)	(0.0242)	(0.0155)	
AfterBlackOwnedReview=1 × BlackNeighborhood=1					-0.0378***
					(0.0131)
Restaurant FE	Yes	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes	Yes	Yes
City-Year-Month FE	Yes	Yes	Yes	Yes	Yes
Black Neighborhood	No	Yes	-	-	-
w/ CuisineSignal	-	-	Yes	No	-
Observations	9879	9283	5485	13677	17095
R^2	0.870	0.867	0.888	0.859	0.811

Note: Cluster-robust standard errors are reported at the restaurant level. * p<0.10, ** p<0.05, *** p<0.01.

Appendix M: Survey for Cuisine Type

To bolster our subsample analyses concerning the inference of Black ownership based on cuisine type, we administered a survey via Prolific. The survey consisted of a single question in which participants were asked to select all cuisines from a provided list that they believed were most frequently associated with Black-owned restaurants in the United States. One hundred USA-based, English-speaking participants were recruited from Prolific's subject pool to complete a choice-based questionnaire. Based on the demographic information provided by Prolific, 34 participants identified as male and 66 as female, with an average age of 38.5 years. All survey responses were included in the subsequent analysis. The results, depicted in the bar chart below, indicate the number of participants (out of 100) who associated each cuisine with Black ownership. Notably, 11 cuisines received votes from over 60% of the respondents, aligning with our expectations. These cuisines include: Soul food, Jamaican restaurant, African restaurant, East African restaurant, West African restaurant, Creole/Haitian restaurant, Caribbean restaurant, Ethiopian restaurant, Barbecue, Cajun restaurant, Southern restaurant (US). For our subsample analyses, a restaurant is considered to exhibit a cuisine signal (i.e., w/ CuisineSignal = True) if it is categorized under one of these 11 cuisines.

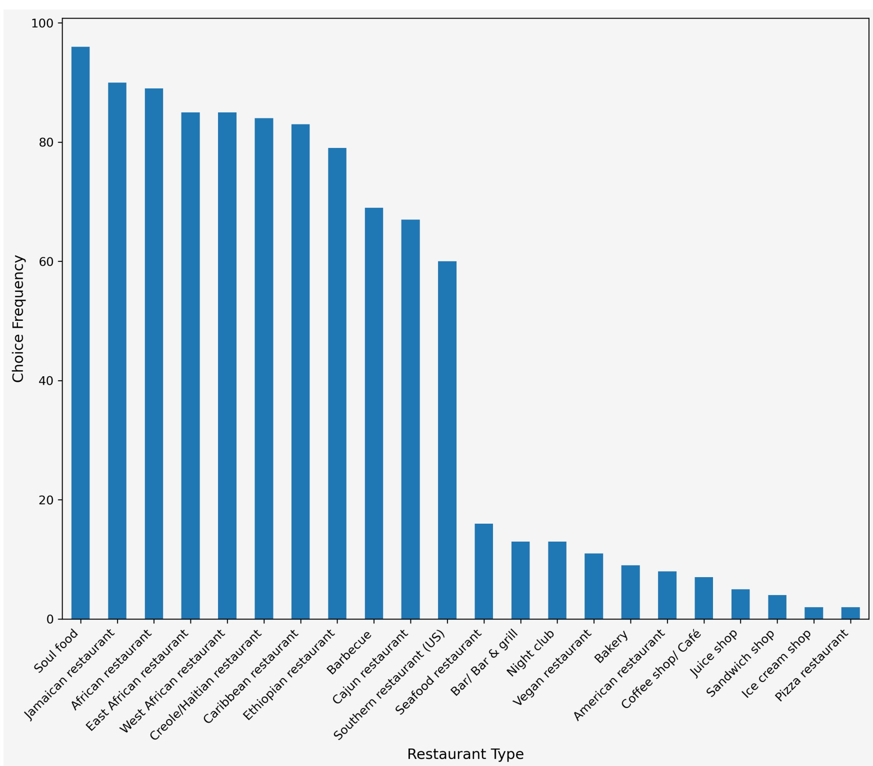


Figure M1 Survey Results (Inference from Cuisine)

Note: The figure shows the number of respondents who chose each cuisine type.

Appendix N: Analysis including non-Blacked-owned Restaurants

To further demonstrate the robustness of our findings, we expand the control group to include non-Black-owned restaurants. Restaurants that have received Black-owned badges remain excluded from our analyses. Consequently, the control group comprises Black-owned restaurants that have not yet (or never) received any review mentioning Black ownership, as well as non-Black-owned restaurants. Our results indicate that the main effect remains robust within this alternative sample.

Table N1 Estimation with Non-Black-Owned Restaurants Included

	(1)	(2)
	LogVisits	LogVisits
AfterBlackOwnedReview=1	0.143***	0.174***
	(0.0316)	(0.0350)
Restaurant FE	Yes	Yes
Year-Month FE	Yes	Yes
City-Year-Month FE	Yes	Yes
Restaurant-Month FE	No	Yes
Observations	715453	715069
R^2	0.864	0.901

Note: Cluster-robust standard errors are reported at the restaurant level. * p<0.10, ** p<0.05, *** p<0.01.

Appendix O: Time To First Review Analysis

We calculate the median interval (in months) between a Black-owned restaurant's opening and the receipt of its first review mentioning specific attributes. The attribute "staff" with service-related keywords is mentioned within the first month of a restaurant's opening. However, the median time for a restaurant to get a review mentioning vegan or vegetarian food as an attribute is over 10 months since opening. Rarer, non-obvious attributes about local ownership, locally sourced ingredients, or the one in our study, such as Black-ownership takes over 22 months since opening. This suggests that the attributes consumers mention in reviews are idiosyncratic, and we can leverage the variation in the time consumers mention attributes in reviews (and thus, make them part of search results for these attributes) to obtain the impact of these attributes. Figure B1 shows the distribution of the number of restaurants that get the first review mentioning Black ownership over time. This appears visually uniform, demonstrating that the timing of the first Black-owned review is more or less random.

Table O1 Time in Months since opening to get the first attribute related review

Attribute (Type)	**Median time in months of the first attribute related review since the time of restaurant opening**
Staff related keywords	1.37
Drinks or bar related keywords	3.77
Music related keywords	5.17
Decoration related keywords	6.6
Restaurant discovery related keywords	9.93
Vegetarian food related keywords	10.87
Parking related keywords	11.4
Local food keywords	22.67
Black-ownership related keywords	25.78

Appendix P: Event Study Leads and Lags

We extend Model (1) by incorporating a complete set of lead and lag dummy variables. Specifically, this approach captures the distance in months from the time of treatment, where leads represent periods after the treatment and lags represent periods before the treatment. This method is commonly used to alleviate concerns about the validity of control groups, and there is a plethora of studies in the economics, information systems, and marketing literature that employ the relative time model to address this issue (Ananthakrishnan et al. 2025, Chan et al. 2019, Burtch et al. 2018, Greenwood and Agarwal 2016, Angrist and Pischke 2009, Autor 2003).

$$LogVisits_{ict} = \sum_{j=-6}^{6} \beta_j Review_{ict}(t = k + j) + \alpha_i + \gamma_t + \mu_{ct} + \epsilon_{it}. \tag{P.1}$$

In this equation, each coefficient β_j ($j = -6, -5, \ldots, 5, 6$) captures the treatment effect at j months relative to the time when treated units receive their initial review mentioning Black ownership. In Figure P1 (same as the figure in Section 3, reproduced here for convenience), the point estimates to the left of the dashed line represent the foot traffic to restaurants prior to receiving their first review mentioning Black ownership, in comparison to other Black-owned restaurants that have not yet received such reviews. These estimates are statistically insignificant, thus supporting the parallel trends assumption and demonstrating the validity of the control group.

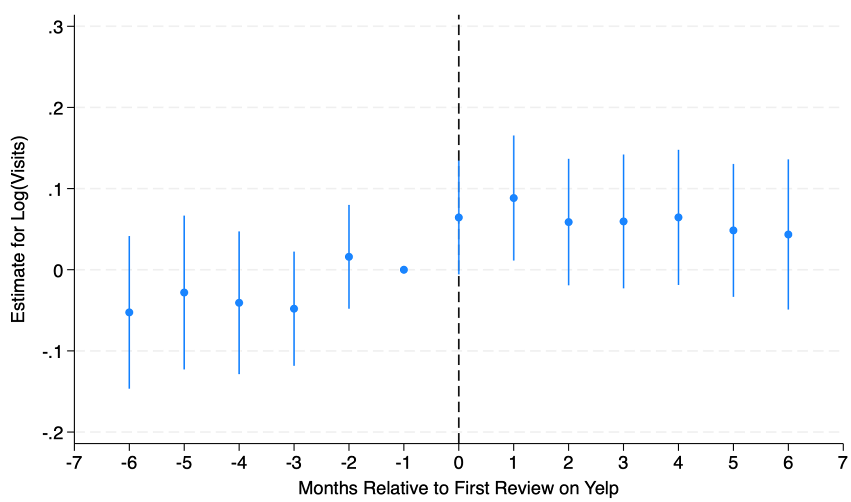


Figure P1 Event Study Analysis

Note: The figure shows relative time model estimates with log of visits as the dependent variable. The estimates are shown with a 95% confidence intervals.

Appendix Q: Black-Owned Badges and Changes to Search Algorithm

Q.1. Black-owned Restaurants

We first consider the different mechanisms through which Black-owned restaurants get increased visibility on Yelp. First, some restaurants have user-generated reviews mentioning Black ownership. Next, we have voluntary adoption of Black-owned badges by some businesses after the introduction of these badges in June 2020. These two mechanisms are distinctly different — the reviewers drive the former in the form of UGC. The platform (as a feature introduction) and the restaurants on the platform (adoption of these badges) drive the latter. Each of these mechanisms confers different avenues of increased digital visibility, as described below. In our analysis, we want to study the impact of UGC in enabling political consumerism. More specifically, we want to measure the impact of online reviews in enabling the discovery process when consumers want to support Black-owned restaurants.

Q.2. Black-Owned Restaurants on Yelp Search

Users on Yelp can encounter Black-owned restaurants in two ways: a generic search and a "Black-owned" search. The differences in how Black-owned restaurants appear in each of these search results are critical to our identification strategy.

Generic search. We refer to searches that most customers would perform on Yelp as generic searches. For example, we categorize searches such as "Restaurants", "Takeout", "Italian Restaurants" under generic search. Such a search would display all restaurants based on Yelp's search algorithm (this is a function of the user, location, restaurant type, etc.).

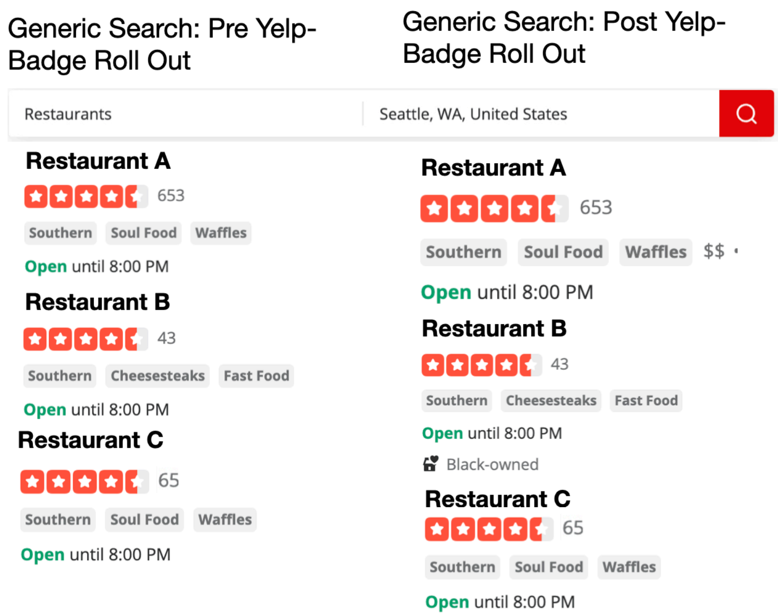
After June 2020, Black-owned restaurants that have voluntarily adopted a Black-owned badge displayed on the search results with a gem icon indicating Black ownership, as shown in the first restaurant in the top half of Figure Q3. In other words, if a Black-owned restaurant that has adopted a badge were going to be featured in a generic search result, this restaurant would be shown with increased visibility with an icon denoting Black ownership.[28] Black-owned restaurants that have not voluntarily adopted a badge would be displayed according to Yelp's search algorithm but featured without a badge. Thus, the increased visibility afforded by the Black-owned badge in a generic search depends on whether a Black-owned restaurant has adopted a badge.

Importantly, in our context, Black-owned restaurants that have not adopted a Black-owned badge but have a review mentioning Black ownership are *not treated* in any unique way in a generic search result (they don't have any icons to display Black-owned reviews).

Figure Q1 is a stylized example of changes to the search algorithm when the user performs a generic search (such as "restaurants"). In this example, Restaurant A has reviews mentioning Black ownership, Restaurant B has no reviews mentioning Black ownership but has voluntarily adopted a badge indicating Black ownership, and Restaurant C is not Black-owned. Before the introduction of badges and the changes on Yelp, Restaurants A, B, and C would look similar on the search result page. After the introduction of

[28] Yelp's official blog details how Black-ownership (voluntary adoption of badges) became a searchable attribute - *"If you choose to opt in, [...] when a user searches for Black-owned businesses, a gem will appear in the search results of your business page highlighting the attribute."*. Yelp does not verify Black ownership.

badges, Restaurant B would have a gem icon indicating Black-owned since it has voluntarily adopted a badge indicating Black ownership. There is no difference between Restaurants A and C since the user did not specifically search for a particular term.



Restaurant A has reviews mentioning Black ownership, **Restaurant B** has no reviews mentioning Black ownership but has voluntarily adopted a badge indicating Black ownership, **Restaurant C** is not Black-owned

Figure Q1 **Results For Generic Search**

Increase in political consumerism – “Black-owned” search. Next, we consider the searches that are more relevant to our analysis. Consumers conduct specific searches due to increased political consumerism. In other words, we have consumers explicitly looking to support Black-owned businesses due to offline events or an increase in national attention to Black causes. It is reasonable to assume that a consumer who searches for a “Black-owned” business does so to discover these businesses. It is crucial to note that such searches are not a rarity. Customers searched for “Black-owned” restaurants in an unprecedented manner in the aftermath of BLM protests. In fact, according to Yelp’s official blog, this search increased by 3600%.

“Black-owned” searches are essential in our analysis because of two reasons: First, they allow us to delineate political consumerism clearly. Second, the presence or absence of a restaurant in the results of a Black-owned search depends on whether the restaurant has received at least one review mentioning Black ownership at the time of the search. In other words, “Black-owned” searches help us measure the value of UGC in enabling political consumerism.

Specifically, when customers searched for Black-owned restaurants, the search results would display all restaurants with at least one review mentioning “Black-owned” as a simple string match. A lot of restaurants obtained reviews mentioning Black ownership much before 2020. Before June 2020, search results for the term “Black-owned business” would only include restaurants with at least one review mentioning “Black-owned” in the review text. After June 2020, search results for the term “Black-owned business” would include both restaurants with at least one mention of "Black-owned" in the reviews and restaurants that have voluntarily adopted a badge *even* if users had not mentioned Black-ownership in the text of their review. Further, during our study period, Yelp's search results mentioned why a restaurant was displayed when users searched for "Black-owned". For example, the search would say - restaurant A has ten reviews mentioning "Black-owned" or the search would have a gem icon indicating that restaurant B has adopted a badge as shown in the restaurant in Figure Q3.

Therefore, Black-owned restaurants with at least one review mentioning Black-ownership would appear in Black-owned searches if they had at least one review mentioning Black ownership at the time of search, during the entire period of our analysis from January 2019 to June 2021. Black-owned restaurants that have voluntarily adopted a badge declaring Black-ownership and *not* have any Black-owned review would appear in Black-owned searches only after they adopt the badge. This feature was released after June 2020. Note that only about 27 percent of all the Black-owned restaurants in our sample had voluntarily adopted a badge even by September 2021, so the majority of the restaurants still became visible to users who searched for Black-owned restaurants only through reviews that mentioned minority ownership. In fact, these restaurants with reviews were the only ones displayed on the "Black-owned" search before June 2020. We isolate restaurants that only have user-generated reviews mentioning Black ownership in our analysis and discuss more in our identification section below.

Figure Q2 is a stylized example of changes to the search algorithm when the user performs a "Black-owned" search. As before, Restaurant A has reviews mentioning Black ownership, Restaurant B has no reviews mentioning Black ownership but has voluntarily adopted a badge indicating Black ownership, and Restaurant C is not Black-owned. Before the introduction of badges and the changes on Yelp, a search for "Black-owned" restaurant would only display Restaurant A with an explanation of why Restaurant A was being displayed. Restaurants B and C would not be displayed in the Black-owned search since they don't have any reviews mentioning Black ownership. After the introduction of badges, the voluntary Black-owned badges became a searchable attribute, placing them in the "Black-owned" searches along with Restaurant A. Restaurant C would not be displayed since they do not have a review mentioning Black ownership nor have they voluntarily adopted a badge.

Q.3. Identification Strategy

How do we measure the impact of coordinating offline movement with online reviews that promote digital visibility? In other words, how do we measure the impact of user-generated content in enabling political consumerism? In our analysis, we choose to study the digital visibility afforded by *reviews* when users search to support Black-owned businesses. When a consumer searches for a specific term on Yelp, Yelp's search algorithm displays restaurants with at least one review containing the term that the consumer searches for that term. Restaurants have no control over whether and when consumers mention "Black-owned" in their reviews. Different Black-owned restaurants obtain reviews mentioning Black ownership at different times, if at all, which determines their presence in the "Black-owned" search results. This provides us with a staggered DiD framework.

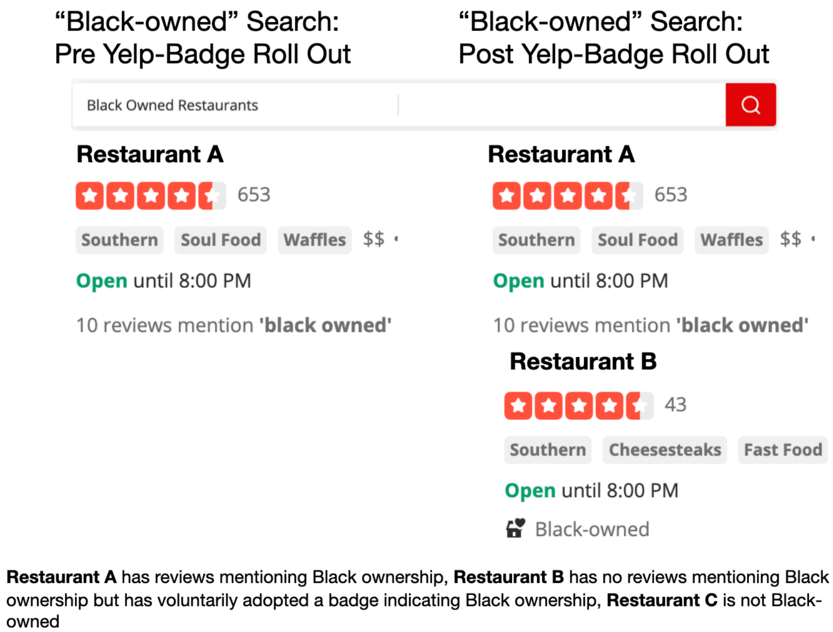


Figure Q2 **Results For Black-Owned Search**

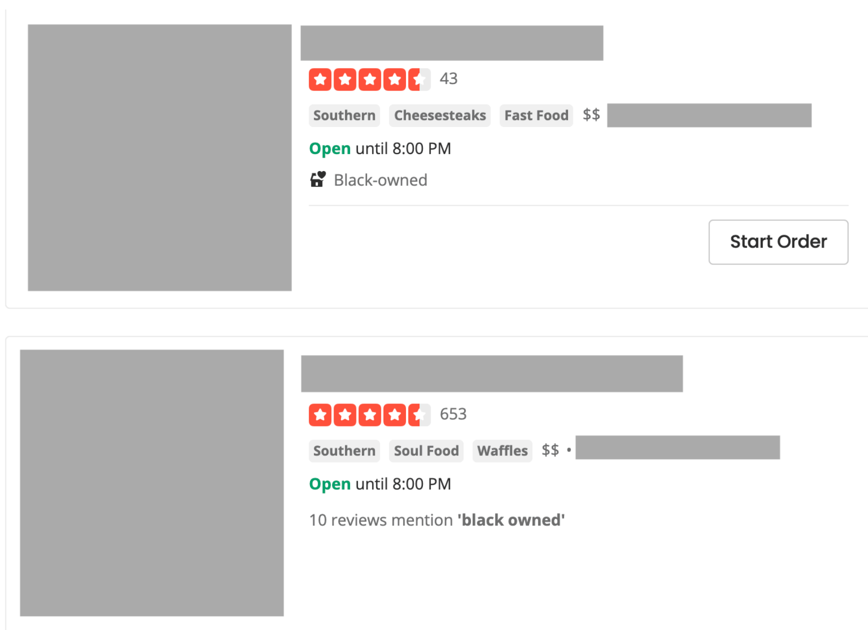


Figure Q3 **Difference between restaurants with reviews mentioning Black ownership and restaurants that have voluntarily adopted Black-owned badges in the search interface for Black-owned searches**

Black-owned restaurants only with reviews in the generic search result without any information about being Black-owned and would look the same way as any other restaurant on the search result but would be highlighted as Black-owned in the “Black-owned” search as shown in Figure Q4. Therefore, having at least one review mentioning Black-ownership would ensure a position in the “Black-owned” search results (searches proxying for political consumerism).

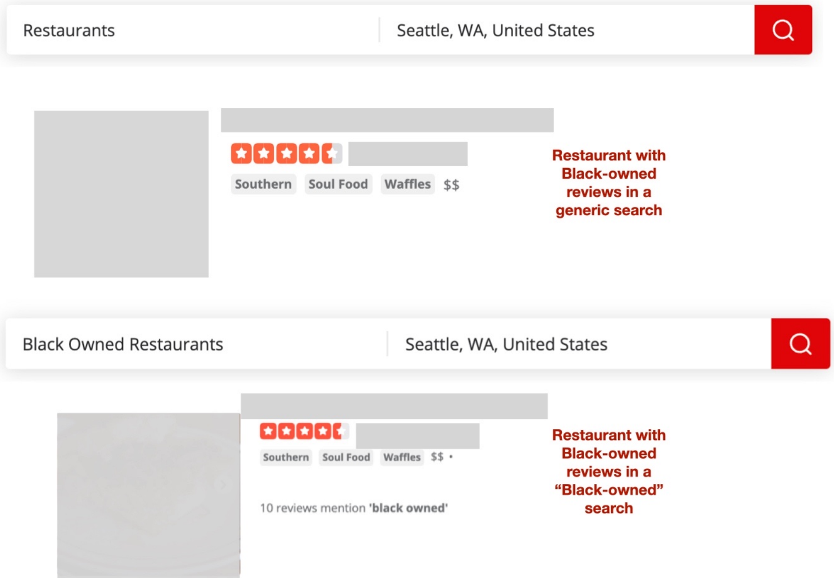


Figure Q4 Black-owned restaurants with reviews mentioning Black-ownership and no voluntary badges appearing in generic and Black-owned searches

Some nuances are worth clarifying. When users search for "Black-owned" restaurants, search results contain Black-owned restaurants that obtain at least one review mentioning Black ownership and those that have adopted a "Black-owned" badge (if the search happened after June 2020 after Yelp introduced these features and after the voluntary adoption of the badges by a restaurant). Our analysis only focuses on political consumerism and how consumers choosing to support Black-owned businesses discover these businesses through online reviews. We do not consider accidental encounters in generic search. Hence, we remove the restaurants that have voluntarily adopted badges for four reasons: First, such restaurants could get increased digital visibility on searches through the Black gem icon unrelated to offline political movements (generic search) and could confound our results. Second, these restaurants could be part of both generic and "Black-owned" searches, and since the time of adoption is unknown, we would not know when they became part of the "Black-owned" search result. Third, Black-owned restaurants adopting badges could self-select and be very different from other Black-owned restaurants. Fourth, badges require restaurant owners to adopt these badges. Technology adoption and increased visibility can increase revenue (Ghose et al. 2014, Agarwal et al. 2011, Ghose and Yang 2009). However, technology adoption among minority business owners is low, even if technology adoption reduces the revenue gap between Black and White business owners (Zhang et al. 2021). The number of such restaurants is less than 27% even during our data collection in September 2021. Other studies estimate this number as around 10% on all of Yelp one year after the introduction of these badges (Aneja et al. 2023). Fortunately, user-generated content does not require technical savvy among minority business owners or any other particular interventions from the platform.

This setting isolates political consumerism (with Black-owned searches as the proxy). Further, our modeling choice enables us to infer the precise time a restaurant would be featured in a Black-owned search based on the time of the first review mentioning Black ownership. Further, our decision to only consider Black-owned searches and the Black-owned restaurants with and without reviews mentioning Black-ownership would enable us to keep the treated and the control group among Black-owned restaurants.

Appendix References

Agarwal A, Hosanagar K, Smith MD (2011) Location, location, location: An analysis of profitability of position in online advertising markets. *Journal of marketing research* 48(6):1057–1073.

Agarwal S, Sen A (2022) Antiracist curriculum and digital platforms: Evidence from black lives matter. *Management Science* .

Altonji JG, Elder TE, Taber CR (2005) Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of political economy* 113(1):151–184.

Ananthakrishnan UM, Basavaraj N, Karmegam SR, Sen A, Smith MD (2025) Book bans in american libraries: Impact of politics on inclusive content consumption. *Marketing Science* .

Aneja A, Luca M, Reshef O (2023) Black ownership matters: Does revealing race increase demand for minority-owned businesses? Technical report, National Bureau of Economic Research.

Angrist JD, Pischke JS (2009) *Mostly harmless econometrics: An empiricist's companion* (Princeton university press).

Autor DH (2003) Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. *Journal of labor economics* 21(1):1–42.

Bertrand M, Duflo E, Mullainathan S (2004) How much should we trust differences-in-differences estimates? *The Quarterly journal of economics* 119(1):249–275.

Bojinov I, Simchi-Levi D, Zhao J (2023) Design and analysis of switchback experiments. *Management Science* 69(7):3759–3777.

Burtch G, Carnahan S, Greenwood BN (2018) Can you gig it? an empirical examination of the gig economy and entrepreneurial activity. *Management science* 64(12):5497–5520.

Chan J, Mojumder P, Ghose A (2019) The digital sin city: An empirical study of craigslist's impact on prostitution trends. *Information systems research* 30(1):219–238.

Chen HY (2004) Nonparametric and semiparametric models for missing covariates in parametric regression. *Journal of the American Statistical Association* 99(468):1176–1189.

Cinelli C, Hazlett C (2020) Making sense of sensitivity: Extending omitted variable bias. *Journal of the Royal Statistical Society Series B: Statistical Methodology* 82(1):39–67.

Deshpande M, Li Y (2019) Who is screened out? application costs and the targeting of disability programs. *American Economic Journal: Economic Policy* 11(4):213–248.

Feit EM, Bradlow ET (2021) Fusion modeling. *Handbook of Market Research*, 147–180 (Springer).

Galor O, Özak Ö (2016) The agricultural origins of time preference. *American economic review* 106(10):3064–3103.

Gaure S (2013) Ols with multiple high dimensional category variables. *Computational Statistics & Data Analysis* 66:8–18.

Gertler PJ, Martinez S, Premand P, Rawlings LB, Vermeersch CM (2016) *Impact evaluation in practice* (World Bank Publications).

Ghose A, Ipeirotis PG, Li B (2014) Examining the impact of ranking on consumer behavior and search engine revenue. *Management Science* 60(7):1632–1654.

Ghose A, Yang S (2009) An empirical analysis of search engine advertising: Sponsored search in electronic markets. *Management science* 55(10):1605–1622.

Greenwood BN, Agarwal R (2016) Matching platforms and hiv incidence: An empirical investigation of race, gender, and socioeconomic status. *Management Science* 62(8):2281–2303.

Gu X, Kannan P (2021) The dark side of mobile app adoption: Examining the impact on customers' multichannel purchase. *Journal of Marketing Research* 58(2):246–264.

Hendricks K, Sorensen A (2009) Information and the skewness of music sales. *Journal of Political Economy* 117(2):324–369.

Hoynes H, Schanzenbach DW, Almond D (2016) Long-run impacts of childhood access to the safety net. *American Economic Review* 106(4):903–934.

Iacus SM, King G, Porro G (2012) Causal inference without balance checking: Coarsened exact matching. *Political analysis* 20(1):1–24.

Li Z, Wang G (2020) The role of on-demand delivery platforms in restaurants. *Available at SSRN 3813891* .

Massenkoff M, Wilmers N (2023) Rubbing shoulders: Class segregation in daily activities. *Available at SSRN 4516850* .

Nunn N, Wantchekon L (2011) The slave trade and the origins of mistrust in africa. *American economic review* 101(7):3221–3252.

Petrova M, Sen A, Yildirim P (2021) Social media and political contributions: The impact of new technology on political competition. *Management Science* 67(5):2997–3021.

Petrova M, Sen A, Yildirim P (2023) New technologies and political competition: the impact of social media communication on political contributions. *The Political Economy of Social Media,'Campante, F, R Durante and A Tesei (eds)(2023). CEPR Press, Paris, London. https://cepr. org/publications/boo* .

Qian Y, Xie H (2014) Which brand purchasers are lost to counterfeiters? an application of new data fusion approaches. *Marketing Science* 33(3):437–448.

Qian Y, Xie H (2024) Correcting regressor-endogeneity bias via instrument-free joint estimation using semiparametric odds ratio models. *Journal of Marketing Research* 61(5):914–936.

Sen A, Grad T, Ferreira P, Claussen J (2024) (how) does user-generated content impact content generated by professionals? evidence from local news. *Management Science* 70(9):6045–6068.

Stechemesser A, Koch N, Mark E, Dilger E, Klösel P, Menicacci L, Nachtigall D, Pretis F, Ritter N, Schwarz M, et al. (2024) Climate policies that achieved major emission reductions: Global evidence from two decades. *Science* 385(6711):884–892.

Wooldridge JM (2010) *Econometric analysis of cross section and panel data* (MIT press).

Zervas G, Proserpio D, Byers JW (2017) The rise of the sharing economy: Estimating the impact of airbnb on the hotel industry. *Journal of marketing research* 54(5):687–705.

Zhang S, Mehta N, Singh PV, Srinivasan K (2021) Frontiers: Can an artificial intelligence algorithm mitigate racial economic inequality? an analysis in the context of airbnb. *Marketing Science* 40(5):813–820.