



No. 109

DISCUSSION PAPER SERIES

# **Response to “The Many Misspellings of Albuquerque: A Comment”**

Peter Christensen

Christopher Timmins

This paper responds to:

Chen, Shi, Areez Gangji, Sunny Karim, Anthony McCanny, Matthew D. Webb. 2024. The Many Misspellings of Albuquerque: A Comment on ‘Sorting or Steering: The Effects of Housing Discrimination on Neighborhood Choice’. I4R Discussion Paper #108. Institute for Replication.

**March 2024**

## I4R DISCUSSION PAPER SERIES

I4R DP No. 109

# Response to “The Many Misspellings of Albuquerque: A Comment”

**Peter Christensen<sup>1</sup>, Christopher Timmins<sup>2</sup>**

<sup>1</sup>*University of Illinois, Urbana/USA*

<sup>2</sup>*Duke University, Durham/USA*

MARCH 2024

Any opinions in this paper are those of the author(s) and not those of the Institute for Replication (I4R). Research published in this series may include views on policy, but I4R takes no institutional policy positions.

I4R Discussion Papers are research papers of the Institute for Replication which are widely circulated to promote replications and meta-scientific work in the social sciences. Provided in cooperation with EconStor, a service of the [ZBW – Leibniz Information Centre for Economics](#), and [RWI – Leibniz Institute for Economic Research](#), I4R Discussion Papers are among others listed in RePEc (see IDEAS, EconPapers). Complete list of all I4R DPs - downloadable for free at the I4R website.

I4R Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

### Editors

**Abel Brodeur**  
*University of Ottawa*

**Anna Dreber**  
*Stockholm School of Economics*

**Jörg Ankel-Peters**  
*RWI – Leibniz Institute for Economic Research*

# Response to “The Many Misspellings of Albuquerque: A Comment”

Peter Christensen and Christopher Timmins

March 17, 2024

## Summary

We thank the authors of this report for their careful re-analysis of our paper, “Sorting or Steering: The Effects of Housing Discrimination on Neighborhood Choice” and the Institute for Replication for supporting this work. In the process of replicating the results published in Christensen and Timmins (2022), Chen et al. raise two concerns and provide an independent set of analyses to address the concerns. In particular, the authors report findings from two changes to the original analysis: (1) regressions that make use of a variant of the ‘city’ variable that is used throughout the paper as a control, and (2) dropping observations for testers identified as ‘other’ (not white, LatinX, or Asian). The authors find that several of the coefficients and standard errors reported in the study are sensitive to these changes. They conclude that while significance and magnitude are affected in certain instances, their re-analysis affirms the paper’s primary finding of substantial and nuanced housing racial discrimination. This document provides responses to the concerns discussed by Chen et al., additional analysis that addresses the concerns, and a discussion of the implications for the interpretation of findings in Christensen and Timmins (2022).

The analysis in this response uses the ‘clean city’ variable from the Chen et al. analysis and addresses inconsistency in the treatment of ‘other’ identifying testers. An important difference is that our response addresses the latter concern by treating ‘other’ identifying testers as a separate category in regression specifications rather than dropping those observations from the sample. In the following sections, we discuss the rationale for this choice and the implications of both changes for interpreting specific findings in Christensen and Timmins (2022). Similar to Chen et al., we report differences in the magnitudes of a set of coefficients reported in Christensen and Timmins (2022). Consistent with the conclusion made by the authors of the replication report, this sensitivity does not affect paper’s primary claim that minority and especially African American testers are steered into neighborhoods that confer disadvantage.

It is important to know whether the magnitudes of the estimates found in re-analysis are statistically different from the estimates reported in Christensen and Timmins (2022). Across the 28 point estimates discussed in the Chen et al. replication report, we do find 1 instance of a point estimate that is statistically different from the estimate reported in the original paper. This is the estimate for Hispanic/LatinX in the analysis of impacts

on steering of mothers into neighborhoods with lower elementary school rankings (Table 10). The estimate reported below is -0.87 and is statistically different from the original estimate of -1.92.

Changes to variables that enter regression specifications could also affect the power of tests reported in Christensen and Timmins (2022), particularly given the fact that the experimental sample analyzed in Christensen and Timmins (2022) was generated by a national HUD audit rather than on the basis of power calculations and a design developed on the basis of hypotheses proposed by the authors. We find that 18% of the estimates analyzed by Chen et al. change from significant in the original paper to not significant with the updated city variable and our correction for the ‘other’ race category. 11% change from not significant to significant. The statistical significance or non-significance of the remaining 71% of estimates remains the same.

The greatest sensitivity in the estimates from Christensen and Timmins (2022) is found in Table 7, which splits the sample to test for variation in steering into white neighborhoods by income level. We find similar patterns when we introduce the ‘clean city’ variable and assign ‘other’ identifying testers to a separate category, but we do not detect statistical differences from zero when these changes are made. This sensitivity deserves attention and is discussed in greater detail below.

## Specific Concerns Raised in the Chen et al. Report

### Concern 1: Use of ‘city’ variable

Chen et al. have identified an important shortcoming with respect to the variable indicating the housing market (i.e. city) where a particular audit experiment took place. The variable entered the original analysis in its raw form provided by HUD as one of more than 10 variables that were not related to the exclusion restriction in the model, but were used to increase the power of the experimental estimates. The ‘city’ variable was added during the referee process as part of a discussion about variation across cities. It did not yield statistical differences in magnitudes, but did increase statistical power in certain tests since the sampling strategy developed for the national HUD audit was not designed to provide powered tests of the hypotheses generated in Christensen and Timmins (2022). In their report, Chen et al. note correctly that there are many instances where a city name is mis-spelled (a notable example being Albuquerque). In Christensen and Timmins (2022), a different market indicator variable was assigned to each city name (multiple indicator variables were created for the multiple different spellings of Albuquerque). The Chen et al. report undertakes a detailed exercise to match mis-spelled markets and collapse them into a single indicator. In the discussion below, we describe how the use of this new variable, which is referred to in the comment as the ‘clean’ city indicator, affects our estimates with respect to a variety of different outcomes.

### Concern 2: Definition of ‘Racial Minority’ variable

Chen et al. also raise concern about the inconsistent categorization of testers who identify with racial/ethnic groups that are coded as ‘other.’ In several sections of Christensen and Timmins (2022), two sets of results are provided for each table: (1) estimates of the

differential treatment for racial minorities as an aggregate category (‘Racial Minority’) and (2) estimates that break out the differential treatment for each race/ethnicity group of interest. Effects on Black, LatinX and Asian testers are reported relative to the white reference group. Christensen and Timmins (2022) reports the two different versions of each model to provide maximum information to the reader. The ‘other’ group is not an object of interest in the paper because the group comprises a small percent of the study sample and is not racially/ethnically defined. In the racially disaggregated model, the ‘other’ group was coded as a separate category in order to estimate effects on all groups of interest using all experimental trials.

In Christensen and Timmins (2022), ‘other’ is defined as part of the ‘white’ category when constructing the combined ‘Racial Minority’ variable. This decision was made on the basis of evidence from the individual names that suggested that the group may be largely comprised of Hispanic-origin testers who do not uniquely identify as ‘LatinX’ but also did not identify as ‘white.’ The assumption is that the racial minority “signal” is not strong for these testers.

Chen et al. correctly point out that this choice compromises consistency in interpretation of estimated effects. This was part of the rationale for providing evidence using the disaggregated version of each model in addition to the aggregate version. A comparison of the first column to the second column of Tables 5, 7, and 10 below illustrates that estimates are never statistically different when regressions use the ‘Racial Minority’ category as defined in the original paper versus when the ‘other’ group is defined as a separate category to distinguish that from the two groups of interest (Minority vs. white).

Testing the sensitivity of results to changes in variable construction or model selection is a valuable undertaking and we appreciate the work done by the authors. In the Chen et al. analysis, observations of ‘other’ identifying testers are dropped. We do not view the choice of dropping observations of ‘other’ identified testers from the sample to be the optimal strategy for obtaining consistent/comparable results across regressions, as it results in inconsistency between the samples used to obtain the estimates for the aggregate (‘Racial Minority’) vs. the disaggregated (‘African American’, ‘Hispanic’, ‘LatinX’) racial/ethnic groups. While the sample of ‘other’ identifying testers constitutes a small fraction of the overall experimental sample, these ‘other’ identifying testers are participants in experimental trials that also include testers from racial/ethnic groups of interest. In the presence of trial fixed effects, dropping observations of ‘other’ identifying testers will therefore alter the estimates identified for the racial groups of interest. In the discussion that follows, we describe the impacts of estimating regressions while coding ‘other’ as a separate category on results reported in the comment.

## Implications for Interpretation

The authors of the report have re-analyzed the results from three of the tables in Christensen and Timmins (2022). Each table contains multiple results. In discussing their re-analysis, we break each table out into multiple panels for easier comparison of alternative choices of the ‘city’ variable and treatments of the ‘other’ race category. In the first column of each table, we reproduce the original estimates from Christensen and Timmins (2022), where ‘Racial Minority’ encompasses African American, Hispanic and

Asian testers. The omitted category in the top panel is white and ‘other’ identifying testers. The omitted category in the bottom panel, where race groups are disaggregated, is only white testers. The city control variable is used in its raw form from HUD. In the second column, ‘other’ identifying testers are treated as a separate category in the racial minority regressions (top panel). In the third column, we use the string-matched ‘clean’ city indicator variable provided by the Chen et al. In the fourth column, we treat ‘other’ testers as a separate category and use the cleaned city control. In the fifth column, we replicate the procedure used in the Chen et al. report by using the string-matched ‘clean’ city indicator and dropping testers that identify as ‘other’.

### **Table 5: Differences in Recommendations and Availability of Advertised Properties**

The authors of the report find that changes to the city variable do not change the interpretation of the original findings reported in the paper, which were statistical nulls. Their re-analysis yields the same results: “All point estimates in Column 2-5 still remain statistically insignificant as in the original case.” We find the same when replicating their results while treating ‘other’ identifying testers as a separate category rather than dropping those observations from the sample. The point estimates are a statistical zero in the original analysis and are also a statistical zero in the re-analysis.

### **Table 7: Discriminatory Steering by Income**

The greatest sensitivity in the estimates from the original paper is found in Table 7, which tests for variation in steering into white neighborhoods by neighborhood income level. This sensitivity deserves attention. The analysis is designed to explore results on steering into white neighborhoods from a prior table (Table 6). It relies upon variation in steering into white neighborhoods within three neighborhood income levels. We have re-produced the text that interprets the findings below:

*While steering into same-race neighborhoods certainly provides evidence of discriminatory behavior that could exacerbate segregation, the ultimate effects on the outcomes of buyer households are not obvious. Recent literature suggests that exposure to within group social models may have important (positive) effects on economic mobility and such recommendations could conform with homophily preferences of minority homebuyers (Chetty et al., 2018). As a result, while discriminatory steering itself is illegal, it is not clear whether increasing access to minority neighborhoods and restricting access to white neighborhoods will generate a welfare cost for minority households. Digging deeper, the estimates presented in Table 7 indicate that the results found in Table 6 are primarily driven by steering of African American buyers away from high income white neighborhoods. The steering effect is strong in high income white neighborhoods and is present for the minority group as a whole. It persists when we control for the listed price, the neighborhood racial composition, and the poverty rate of the advertised listing. These differences become much smaller for African American testers in medium-income white neighborhoods and disappear for the group of minority testers as a whole. The effect actually reverses for low-income white neighborhoods, such that Hispanic and Asian testers are more likely than their white counterparts (with the same income) to receive recommendations in a low-income white neighborhood.*

When we introduce the ‘clean city’ variable and assign ‘other’ identifying testers to a separate category, the estimated coefficients are not statistically different from zero after these changes are made. Coefficients signs are the same and magnitudes are not statistically different from those reported in Christensen and Timmins (2022). Patterns are qualitatively similar, though these estimates are not statistically significant and are therefore merely suggestive of a process where minority buyers are disproportionately steered away from high income white neighborhoods.

### **Table 10: Discriminatory Steering by Family Roles (Mothers)**

The information provided in the HUD HDS allows the researcher to determine whether a female tester presented themselves to a real estate agent as a mother. We consider the results discussed in the replication comment, focusing on the changes that result from using the ‘clean city’ variable along with treating the ‘other’ identifying testers as a separate category. Elementary school scores (Table 10A) remain strongly significant for both African American and Hispanic/LatinX testers. Coefficient magnitudes are statistically equivalent. For Elementary School Rankings (Table 10B), estimates for Racial Minority, ‘African American’ and ‘Asian’ testers are statistically significant. Point estimates are attenuated, but are not statistically different. Estimates for Hispanic/LatinX testers are not statistically significant. For the sample of mothers, switching to the ‘clean city’ variable and treating ‘other’ identifying testers as a separate category leads the the effect of steering of African American testers into lower skill neighborhoods (Table 10C) becoming non-significant (it remains significant for Hispanic/LatinX testers). For the estimates of the impact of steering into neighborhoods with more single-family households (Table 10D), estimates become statistically significant for African American and Hispanic/LatinX testers. For Asian testers, the estimates suggest that steering reduces the likelihood that Asian buyers are steered into neighborhoods with more single-family households.

### **A Note on Clustering**

The authors of the report comment on our choice to cluster standard errors at the level of randomization (i.e. the trial). The choice made in the paper reflects careful consideration and decisions that were made throughout the writing and the referee process. Given that treatment assignment is random and housing listings (which define the departure point for each trial) were sampled at random from the population occurring within the sizable set of markets studied in the HUD audit, clustering at the level of randomization was considered the appropriate choice. While assignment of testers within a trial is random, clustering at the trial-level addresses the potential for correlated errors in outcomes arising from agent-specific or other factors that could affect the sample of properties drawn by a given agent during a trial.

## Table 5

Table 1. Differences in Recommendations and Availability of Advertised Properties

Table 5A. Number of Recommendations

	Original	Other as Separate	Clean City	Clean City + Other as Separate	Clean City + Drop Other
Racial Minority	-0.1419 (0.1987)	-0.0618 (0.1973)	-0.2828* (0.1660)	-0.2297 (0.1657)	-0.2140 (0.1662)
African American	-0.1690 (0.2707)	-0.1690 (0.2707)	-0.3598 (0.2203)	-0.3598 (0.2203)	-0.3493 (0.2207)
Hispanic	-0.1304 (0.2474)	-0.1304 (0.2474)	-0.1907 (0.2083)	-0.1907 (0.2083)	-0.1822 (0.2096)
Asian	0.0833 (0.2465)	0.0833 (0.2465)	-0.1371 (0.2116)	-0.1371 (0.2116)	-0.1122 (0.2116)
ln(Price) Advert Home	Y	Y	Y	Y	Y
Racial Comp Advert Home	Y	Y	Y	Y	Y
Observations	6,555	6,555	6,553	6,553	6,487
Adjusted R <sup>2</sup>	-0.2348	-0.2335	0.0274	0.0281	0.0202

Note: Coefficients report differences in the number of homes recommended to minority testers relative to a white tester (the omitted category). In the first column, we reproduce the original estimates from Christensen and Timmins (2022), where ‘Racial Minority’ encompasses African American, Hispanic and Asian testers. The omitted category in the top panel is white and ‘other’ identifying testers. The omitted category in the bottom panel, where race groups are disaggregated, is only white testers. The city control variable is used in its raw form from HUD. In the second column, ‘other’ identifying testers are treated as a separate category in the racial minority regressions (top panel). In the third column, we use the string-matched ‘clean’ city indicator variable provided by the Chen et al. In the fourth column, we treat ‘other’ testers as a separate category and use the cleaned city control. In the fifth column, we replicate the procedure used in the Chen et al report by using the string-matched ‘clean’ city indicator and dropping testers that identify as ‘other’. All regression specifications include trial fixed effects and control for the full set of actor characteristics, assigned characteristics, and search characteristics. Standard errors are clustered by trial, which is the level of randomization. Significance levels are: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01.

Table 5B. Home Availability

	Original	Other as Separate	Clean City	Clean City + Other as Separate	Clean City + Drop Other
Racial Minority	0.0057 (0.0183)	0.0017 (0.0186)	-0.0022 (0.0151)	-0.0058 (0.0154)	-0.0059 (0.0155)
African American	-0.0087 (0.0219)	-0.0087 (0.0219)	-0.0211 (0.0188)	-0.0211 (0.0188)	-0.0218 (0.0188)
Hispanic	-0.0077 (0.0258)	-0.0077 (0.0258)	0.0054 (0.0211)	0.0054 (0.0211)	0.0048 (0.0212)
Asian	0.0178 (0.0227)	0.0178 (0.0227)	0.0001 (0.0189)	0.0001 (0.0189)	0.0010 (0.0189)
ln(Price) Advert Home	Y	Y	Y	Y	Y
Racial Comp Advert Home	Y	Y	Y	Y	Y
Observations	6,562	6,562	6,560	6,560	6,493
Adjusted R <sup>2</sup>	-0.1774	-0.1778	0.1005	0.1005	0.0859

Note: Coefficients report differences in the number of available homes recommended to minority testers relative to a white tester (the omitted category). In the first column, we reproduce the original estimates from Christensen and Timmins (2022), where ‘Racial Minority’ encompasses African American, Hispanic and Asian testers. The omitted category in the top panel is white and ‘other’ identifying testers. The omitted category in the bottom panel, where race groups are disaggregated, is only white testers. The city control variable is used in its raw form from HUD. In the second column, ‘other’ identifying testers are treated as a separate category in the racial minority regressions (top panel). In the third column, we use the string-matched ‘clean’ city indicator variable provided by the Chen et al. In the fourth column, we treat ‘other’ testers as a separate category and use the cleaned city control. In the fifth column, we replicate the procedure used in the Chen et al report by using the string-matched ‘clean’ city indicator and dropping testers that identify as ‘other’. Standard errors are clustered by trial, which is the level of randomization. Significance levels are: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01.



## Table 7

Table 7. White Household Share High Income

	Original	Other as Separate	Clean City	Clean City + Other as Separate	Clean City + Drop Other
Racial Minority	-0.0265*** (0.0080)	-0.0246*** (0.0080)	-0.0141** (0.0069)	-0.0110 (0.0070)	-0.0100 (0.0071)
African American	-0.0338*** (0.0104)	-0.0338*** (0.0104)	-0.0120 (0.0093)	-0.0120 (0.0093)	-0.0112 (0.0094)
Hispanic	-0.0147 (0.0108)	-0.0147 (0.0108)	-0.0138 (0.0091)	-0.0138 (0.0091)	-0.0131 (0.0092)
Asian	-0.0246*** (0.0095)	-0.0246*** (0.0095)	-0.0067 (0.0086)	-0.0067 (0.0086)	-0.0052 (0.0086)
Comparison Mean (White)	0.21	0.21	0.21	0.21	0.21
Share White Advert Home	Y	Y	Y	Y	Y
ln(Price) Advert Home	Y	Y	Y	Y	Y
Racial Comp Advert Home	Y	Y	Y	Y	Y
Poverty Share Advert Home	Y	Y	Y	Y	Y
Observations	21,442	21,442	21,442	21,442	21,213
Adjusted R <sup>2</sup>	0.7178	0.7178	0.7125	0.7127	0.7131

Note: Coefficients report differences in the racial composition of high income neighborhoods recommended to minority testers relative to a white tester (the omitted category). In the first column, we reproduce the original estimates from Christensen and Timmins (2022), where ‘Racial Minority’ encompasses African American, Hispanic and Asian testers. The omitted category in the top panel is white and ‘other’ identifying testers. The omitted category in the bottom panel, where race groups are disaggregated, is only white testers. The city control variable is used in its raw form from HUD. In the second column, ‘other’ identifying testers are treated as a separate category in the racial minority regressions (top panel). In the third column, we use the string-matched ‘clean’ city indicator variable provided by the Chen et al. In the fourth column, we treat ‘other’ testers as a separate category and use the cleaned city control. In the fifth column, we replicate the procedure used in the Chen et al report by using the string-matched ‘clean’ city indicator and dropping testers that identify as ‘other’. All regression specifications include trial fixed effects and control for the full set of actor characteristics, assigned characteristics, and search characteristics. Standard errors are clustered by trial, which is the level of randomization. Significance levels are: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01.

# Table 10

Table 10A. Mothers: Elementary School Scores

	Original	Other as Separate	Clean City	Clean City + Other as Separate	Clean City + Drop Other
Racial Minority	-0.1471*** (0.0483)	-0.1540*** (0.0511)	-0.1357** (0.0556)	-0.1516*** (0.0562)	-0.1009** (0.0500)
African American	-0.2231** (0.0954)	-0.2231** (0.0954)	-0.1796*** (0.0559)	-0.1796*** (0.0559)	-0.0042 (0.0561)
Hispanic	-0.2374*** (0.0729)	-0.2374*** (0.0729)	-0.2768*** (0.0970)	-0.2768*** (0.0970)	-0.2232*** (0.0855)
Asian	-0.0543 (0.1265)	-0.0543 (0.1265)	-0.0072 (0.0831)	-0.0072 (0.0831)	-0.0457 (0.0692)
Comparison Mean (White)	0.33	0.33	0.33	0.33	0.33
ln(Price) Advert Home	Y	Y	Y	Y	Y
Racial Comp Advert Home	Y	Y	Y	Y	Y
Outcome Advertised Home	Y	Y	Y	Y	Y
Observations	3,805	3,805	3,805	3,805	3,689
Adjusted R <sup>2</sup>	0.7625	0.7625	0.7564	0.7567	0.7590

Note: Coefficients report differences in elementary school scores for properties recommended to minority testers that are assigned the role of a mother (female gender with children in household) relative to a white tester with the same assigned role (the omitted category). In the first column, we reproduce the original estimates from Christensen and Timmins (2022), where ‘Racial Minority’ encompasses African American, Hispanic and Asian testers. The omitted category in the top panel is white and ‘other’ identifying testers. The omitted category in the bottom panel, where race groups are disaggregated, is only white testers. The city control variable is used in its raw form from HUD. In the second column, ‘other’ identifying testers are treated as a separate category in the racial minority regressions (top panel). In the third column, we use the string-matched ‘clean’ city indicator variable provided by the Chen et al. In the fourth column, we treat ‘other’ testers as a separate category and use the cleaned city control. In the fifth column, we replicate the procedure used in the Chen et al report by using the string-matched ‘clean’ city indicator and dropping testers that identify as ‘other’. All regression specifications include trial fixed effects and control for the full set of actor characteristics, assigned characteristics, and search characteristics. Standard errors are clustered by trial, which is the level of randomization. Significance levels are: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01.

Table 10B. Mothers: Elementary School Rankings

	Original	Other as Separate	Clean City	Clean City + Other as Separate	Clean City + Drop Other
Racial Minority	-3.0893*** (0.9501)	-2.3918** (0.9421)	-1.6457*** (0.4610)	-1.3552*** (0.4880)	-1.1849** (0.5781)
African American	-3.0262*** (0.7800)	-3.0262*** (0.7800)	-1.8995*** (0.5196)	-1.8995*** (0.5196)	-1.4308** (0.6757)
Hispanic	-1.9212** (0.9598)	-1.9212** (0.9598)	-0.8708 (0.6436)	-0.8708 (0.6436)	-0.8854 (0.7382)
Asian	-3.8178*** (0.9020)	-3.8178*** (0.9020)	-1.0773* (0.5705)	-1.0773* (0.5705)	-1.2051* (0.6442)
Comparison Mean (White)	6.5	6.5	6.5	6.5	6.5
ln(Price) Advert Home	Y	Y	Y	Y	Y
Racial Comp Advert Home	Y	Y	Y	Y	Y
Outcome Advertised Home	Y	Y	Y	Y	Y
Observations	4,520	4,520	4,520	4,520	4,398
Adjusted R <sup>2</sup>	0.6984	0.6988	0.6878	0.6879	0.6914

Note: Coefficients report differences in elementary school rankings for properties recommended to minority testers that are assigned the role of a mother (female gender with children in household) relative to a white tester with the same assigned role (the omitted category). In the first column, we reproduce the original estimates from Christensen and Timmins (2022), where ‘Racial Minority’ encompasses African American, Hispanic and Asian testers. The omitted category in the top panel is white and ‘other’ identifying testers. The omitted category in the bottom panel, where race groups are disaggregated, is only white testers. The city control variable is used in its raw form from HUD. In the second column, ‘other’ identifying testers are treated as a separate category in the racial minority regressions (top panel). In the third column, we use the string-matched ‘clean’ city indicator variable provided by the Chen et al. In the fourth column, we treat ‘other’ testers as a separate category and use the cleaned city control. In the fifth column, we replicate the procedure used in the Chen et al report by using the string-matched ‘clean’ city indicator and dropping testers that identify as ‘other’. All regression specifications include trial fixed effects and control for the full set of actor characteristics, assigned characteristics, and search characteristics. Standard errors are clustered by trial, which is the level of randomization. Significance levels are: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01.

Table 10C. Mothers: American Community Survey - High Skill

	Original	Other as Separate	Clean City	Clean City + Other as Separate	Clean City + Drop Other
Racial Minority	-0.0347** (0.0144)	-0.0430*** (0.0158)	-0.0186 (0.0148)	-0.0261* (0.0149)	-0.0263* (0.0153)
African American	-0.0622*** (0.0210)	-0.0622*** (0.0210)	-0.0296 (0.0194)	-0.0296 (0.0194)	-0.0423** (0.0198)
Hispanic	-0.0530*** (0.0171)	-0.0530*** (0.0171)	-0.0518*** (0.0172)	-0.0518*** (0.0172)	-0.0469*** (0.0176)
Asian	-0.0081 (0.0206)	-0.0081 (0.0206)	0.0206 (0.0181)	0.0206 (0.0181)	0.0219 (0.0185)
Comparison Mean (White)	0.48	0.48	0.48	0.48	0.48
ln(Price) Advert Home	Y	Y	Y	Y	Y
Racial Comp Advert Home	Y	Y	Y	Y	Y
Outcome Advertised Home	Y	Y	Y	Y	Y
Observations	7,997	7,997	7,997	7,997	7,827
Adjusted R <sup>2</sup>	0.6609	0.6609	0.6501	0.6502	0.6536

Note: Coefficients report differences in share of households with at least one member who is employed in a high skilled occupation for properties recommended to minority testers that are assigned the role of a mother (female gender with children in household) relative to a white tester with the same assigned role (the omitted category). In the first column, we reproduce the original estimates from Christensen and Timmins (2022), where ‘Racial Minority’ encompasses African American, Hispanic and Asian testers. The omitted category in the top panel is white and ‘other’ identifying testers. The omitted category in the bottom panel, where race groups are disaggregated, is only white testers. The city control variable is used in its raw form from HUD. In the second column, ‘other’ identifying testers are treated as a separate category in the racial minority regressions (top panel). In the third column, we use the string-matched ‘clean’ city indicator variable provided by the Chen et al. In the fourth column, we treat ‘other’ testers as a separate category and use the cleaned city control. In the fifth column, we replicate the procedure used in the Chen et al report by using the string-matched ‘clean’ city indicator and dropping testers that identify as ‘other’. All regression specifications include trial fixed effects and control for the full set of actor characteristics, assigned characteristics, and search characteristics. Standard errors are clustered by trial, which is the level of randomization. Significance levels are: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01.

Table 10D. Mothers: American Community Survey - Single-Parent HH

	Original	Other as Separate	Clean City	Clean City + Other as Separate	Clean City + Drop Other
Racial Minority	0.0001 (0.0125)	0.0115 (0.0127)	0.0066 (0.0104)	0.0127 (0.0103)	0.0128 (0.0106)
African American	0.0264 (0.0164)	0.0264 (0.0164)	0.0316** (0.0132)	0.0316** (0.0132)	0.0347** (0.0135)
Hispanic	0.0197 (0.0139)	0.0197 (0.0139)	0.0265** (0.0118)	0.0265** (0.0118)	0.0274** (0.0122)
Asian	-0.0163 (0.0164)	-0.0163 (0.0164)	-0.0262** (0.0123)	-0.0262** (0.0123)	-0.0316** (0.0124)
Comparison Mean (White)	0.14	0.14	0.14	0.14	0.14
ln(Price) Advert Home	Y	Y	Y	Y	Y
Racial Comp Advert Home	Y	Y	Y	Y	Y
Outcome Advertised Home	Y	Y	Y	Y	Y
Observations	7,997	7,997	7,997	7,997	7,827
Adjusted R <sup>2</sup>	0.4826	0.4828	0.4768	0.4769	0.4766

Note: Coefficients report differences in share of single parent households for properties recommended to minority testers that are assigned the role of a mother (female gender with children in household) relative to a white tester with the same assigned role (the omitted category). In the first column, we reproduce the original estimates from Christensen and Timmins (2022), where ‘Racial Minority’ encompasses African American, Hispanic and Asian testers. The omitted category in the top panel is white and ‘other’ identifying testers. The omitted category in the bottom panel, where race groups are disaggregated, is only white testers. The city control variable is used in its raw form from HUD. In the second column, ‘other’ identifying testers are treated as a separate category in the racial minority regressions (top panel). In the third column, we use the string-matched ‘clean’ city indicator variable provided by the Chen et al. In the fourth column, we treat ‘other’ testers as a separate category and use the cleaned city control. In the fifth column, we replicate the procedure used in the Chen et al report by using the string-matched ‘clean’ city indicator and dropping testers that identify as ‘other’. Standard errors are clustered by trial, which is the level of randomization. Significance levels are: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01.