

# An Empirical Justification for the Use of Racially Distinctive Names to Signal Race in Experiments

Daniel M. Butler and Jonathan Homola

Washington University in St. Louis, Department of Political Science, Campus Box 1063, One Brookings Drive, St. Louis, Missouri, MO 63130-4899, USA. Email: [daniel.butler@wustl.edu](mailto:daniel.butler@wustl.edu)

## Abstract

Researchers studying discrimination and bias frequently conduct experiments that use racially distinctive names to signal race. The ability of these experiments to speak to racial discrimination depends on the excludability assumption that subjects' responses to these names are driven by their reaction to the individual's putative race and not some other factor. We use results from an audit study with a large number of aliases and data from detailed public records to empirically test the excludability assumption undergirding the use of racially distinctive names. The detailed public records allow us to measure the signals about socioeconomic status and political resources that each name used in the study possibly *could* send. We then reanalyze the audit study to see whether these signals predict legislators' likelihood of responding. We find no evidence that politicians respond to this other information, thus providing empirical support for the excludability assumption.

## 1 Introduction

Racial discrimination affects numerous aspects of social and political life (DeSante 2013; Hutchings and Piston 2011; Pager 2007). Social scientists interested in understanding and reducing this bias frequently use experiments because doing so allows them to hold other factors constant when measuring racial bias (e.g., Hutchings and Piston 2011). While these experiments take a variety of forms, and are conducted on surveys, in the lab, and in the field, researchers will often use racially distinctive names (i.e., names that are much more likely to be associated with a given racial group) as a key tool to measure racial bias (e.g., Bertrand and Mullainathan 2004). For example, surveys will often present respondents with a vignette about an individual, randomizing whether the individual has a name that is distinctively black or distinctively white, and then ask them to evaluate that individual (e.g., DeSante 2013). Similarly, field experiments will ask politicians for help, using the name of the individual to signal the putative race of the one sending the email (e.g., Broockman 2013; McClendon 2016; Mendez and Grose 2014; White, Fallor and Nathan 2015). The researchers then use the differences in respondents' reactions to the putatively white individuals versus the putative minorities as a measure of racial bias.

The ability of these experiments to inform us about racial discrimination depends in large part on the excludability assumption (Gerber and Green 2012). In the majority of studies using names to signal the race of a subject, the excludability assumption is that the research subjects' response to the name is driven *solely* by the signal that the name provides about race. In other words, these studies assume (at least implicitly) that the research subjects are *not* responding to some other relevant information that might be signaled with racially distinctive names. In economics,

*Political Analysis* (2017)  
vol. 25:122–130  
DOI: 10.1017/pan.2016.15

Published  
21 February 2017

Corresponding author  
Daniel M. Butler

Edited by  
Jens Hainmueller

© The Author(s) 2017. Published  
by Cambridge University Press  
on behalf of the Society for  
Political Methodology.

*Authors' note:* Dan Butler thanks the Weidenbaum Center at Washington University in St. Louis for financial support. The authors also thank Tim Teehan and Paul Graham at L2 for their help with the voter data. Finally, the authors thank Bernard Fraga, Hans Hassell, Hakeem Jefferson, Michelle Torres, Jens Hainmueller, and the anonymous reviewers for comments on this paper. The replication data (Butler and Homola 2016) is available on the PA Dataverse at: <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/LUGBL1>.

the validity of using names to measure the effect of ethnic/racial bias has been a matter of active debate because of concerns about excludability. Most prominently, Fryer and Levitt (2004) argue that racially distinctive names are also strong signals about socioeconomics. The issue is not just that black and Latino individuals have lower incomes than white individuals. But that even among black and Latino individuals, those with distinctively black and Latino names tend to be different than others of their race. Fryer and Levitt argue that once you control for the initial differences at birth, names do not have an impact on a wide range of outcomes (see also Heckman and Siegelman 1992; Heckman 1998). However, several recent studies have taken extensive steps to control for such background characteristics and find that names can have very sizable effects on outcomes (Abramitzky, Boustan and Eriksson 2016; Biavaschi, Giulietti and Siddique 2013; Goldstein and Stecklov 2016).

Concerns about excludability are equally important for political science because racially distinctive names can also signal politically relevant information. For example, Latinos and blacks in the United States, especially those with racially distinctive names, are more likely to have a lower SES (Fryer and Levitt 2004). Similarly, and very important for many political science studies, Latinos and blacks are much more likely than whites to vote for a Democratic candidate. The concern is that the signal about race is confounded with these other signals and that people are responding to these other pieces of information and not the race of the individual. Of course, in some projects, researchers' conceptualization of the race treatment will include these differences (see discussion in Sen and Wasow 2016). However, many projects using racially distinctive names are interested in identifying how race, separate from signals about partisanship and political resources, affects subjects' responses to the individual with the racially distinctive name.

One approach for mitigating concerns about confounding the effect of race with these other factors is to convey additional information to the research subjects about the individuals with racially distinctive names. If the experiment is a survey vignette, researchers can incorporate the relevant pieces of information into the vignette (e.g., Hainmueller, Hopkins and Yamamoto 2014). If one includes the relevant pieces of information, this can be effective at mitigating confounding concerns (Dafoe, Zhang and Caughey 2015). However, it is not always possible to convey additional information to the research subject in the study. There may be time and space constraints which make it difficult to convey the necessary information. For example, in audit studies it may be hard or impossible to also convey the additional information in the communication using the racially distinctive name. Thankfully, it may not be necessary to convey additional information. The key assumption is that race is the only signal associated with the racially distinctive name that the research subjects are *responding* to. Even if a racially distinctive name signals other information, research subjects may not respond to this information. In other words, if research subjects are only responding to the signal about race, then the excludability assumption still holds.

We evaluate the excludability assumption undergirding the use of racially distinctive names in experiments by reanalyzing data from an audit study of how public officials respond to email requests from putative Latino, black, and white constituents (Butler 2014). In the audit study, state legislative offices were sent emails requesting information with the alias used in the email being randomly chosen from a set of over 100 different names. Each of these names signals information not only about race, but also about other factors. For our study, we merged the results from the audit study with detailed information from public records on individuals who share the first name with the aliases in the email experiment (as a robustness check we also run the analysis using full names). We thus have information about the specific signals that each name sends about the political resources and the partisan preferences of the putative individual sending the request.

We focus on political resources (in the form of turnout and one's SES) and partisanship because these are likely to be particularly important for most political science studies. Elections are

decided by votes and the financial contributions used to seek after votes. Thus actors in the political arena are likely to care about individuals' partisan preferences (because it affects who they vote for) as well as their political resources (i.e., their likelihood of voting and their ability to donate financial resources to candidates). In the study we reanalyze, the state legislators are deciding whether to give help to a constituent who contacts them. The original study concludes that white state legislators are less responsive to minorities because they are minorities. The concern is that this conclusion may be misplaced. Instead, officials may have used the names to make inferences about the political resources of the individual contacting them and then responded according to those inferences. If they are doing so, then what looks like racial bias would in fact be driven by their concerns for these other relevant factors.

In this letter we conduct two tests. First, we calculate the bivariate correlations between legislative responsiveness and signals that each name sends about political resources and partisan preferences. If officials are actually responding to these factors when they decide whether to respond to an email, then these factors should also be correlated with the levels of responsiveness observed within race. However, we find that these factors are uncorrelated with responsiveness. Second, we replicate the original analysis and then add in the information about the political resource and partisan preferences signals as additional controls. Adding in these controls has no effect on how the aliases are treated. We thus find no evidence that the political resource and partisan preference signals drove the elected officials' responses. This in turn provides empirical support for the exclusion restriction with regard to the most important confounders for studies of politics.

## 2 Data

We use results from a previously conducted audit study (Butler 2014) and information on voters from a large-scale nationwide database to empirically test key parts of the excludability assumption in experiments with racially distinctive names.<sup>1</sup> In the original study, a total of 6,951 emails requesting information were sent to state legislators during the summer of 2010. Importantly, the experiment used a total of 123 different email aliases: 35 of them had a putatively black name, 42 had Latino names, and 46 had white names. Each alias was used to send between 51 and 66 different emails.<sup>2</sup> In line with other studies (Butler and Broockman 2011), the audit found that state legislators were more likely to respond to individuals from their own racial group. In order to analyze whether these effects could be driven by confounding factors such as socioeconomic status or political characteristics, we calculated the average response rate for each email alias' *first name* used in the audit study and then merged this information with socioeconomic and political data of individuals who actually have that name. We also examined the results when matching based on each email alias' *exact full name*. The results of this robustness check confirm the findings reported below and can be found in Tables SA.1 and SA.2 in the Appendix.

Our information on voters' socioeconomic status and political characteristics comes from the company L2.<sup>3</sup> L2's database contains over 166 million regularly updated voter records and includes data from state and county level registered voter files, current U.S. Census Data, telephone source files, election return data with results from every county in the U.S., and L2's own lifestyle and issue

- 1 One advantage of studying public officials' behavior is that they have the possibility to access the type of information about voters that we use in this study. Thus if anyone was in a position to respond to other signals associated with these racially distinctive names, it would be elected officials.
- 2 A full list of all aliases used in the study, the number of emails that was sent from each alias, and their respective response rates can be found in Table SA.15 in the Appendix.
- 3 <http://www.l2political.com>

data.<sup>4</sup> For our analysis, we received information on almost 9.3 million individuals in L2's database that have matching first names with the 123 aliases used in the audit study.<sup>5</sup>

For the analysis we used information from the L2 database to get estimates of each alias' partisanship and political resources. For partisanship we use the percent of individuals in the sample who belong to each respective party. For political resources, we use data from L2 on income, education, housing and turnout to compute a *Political Resources* factor score for each first name.<sup>6</sup> This score refers to the first dimension extracted from a factor analysis including these variables and reflects the idea that the concept of political resources is multidimensional. Using a factor score also increases the power of our test and reduces the danger of a Type II error.<sup>7</sup>

### 3 Analysis

We begin by confirming the original results in Butler (2014) regarding legislators' average response rates.<sup>8</sup> In particular, the first row of Table 1 looks at the correlation between the response rates and putative race of the alias for white legislators.<sup>9</sup> The correlation coefficient is positive and significant, confirming the main finding from Butler (2014): white legislators are significantly more likely to respond to emails from white constituents than from nonwhite constituents.

The original interpretation assumes that email aliases experience different response rates because of the respective aliases' putative race. Thus the excludability assumption would be violated if the legislators in the experiment inferred some kind of information from the email aliases that is distinct from race (i.e., socioeconomic or political characteristics) and based their behavior on that additional information. For example, this would be the case if legislators responded less to a Precious or a DeAndre in the experiment, *because* they assumed that they were less likely to donate money than an Emily or a Luke.

If socioeconomic and political information were the driving determinant of whether a legislator responds to constituents with racially distinctive names, we would expect there to also be large correlations between the variables capturing these characteristics and response rates in the audit study. Further, the correlations with political resources should be positive because individuals

- 
- 4 Because the dataset only includes registered voters, it is likely that our measures of income and education might be higher for this sample than for the population as a whole. This is unlikely to be a problem for our analysis, however, because if public officials were to access such data when deciding how to treat constituents, this would be the very information they would have access to.
  - 5 Naturally, there is quite some variation regarding the number of matches for each alias. The median number of matches is 37,580; the mean is 91,060. The name with the highest number of matches is Matthew (718,729). The main analysis below will be based on all names that had any matches. However, all analyses are also robust to limiting them to those aliases that had more than 2,000 matches in the database. These results are presented in Tables SA.3 and SA.4 in the Appendix.
  - 6 Income is measured on a 12-point scale that ranges from \$1,000–14,999 to \$250,000+. Education captures an individual's number of education years and ranges from 5 to 18 in our sample. Housing indicates the estimated home value and is a continuous variable ranging from \$5,000 to \$9,999,999+ in our sample. The median estimated home value is \$212,000, the mean \$280,600. Finally, turnout is a 9-point scale ranging from 0 to 8 and reflects in how many general elections between 2000 and 2014 each voter turned out to cast their ballot. Correlations between the factor score and its constituting indicators are reported in Table SA.14.
  - 7 L2's consumer data uses high-quality third party sources such as Experian to provide both known and modeled income, education, and housing (as well as other data). Because some of the data is modeled, the variables relating to the SES portion of our political resources measure include some noise (as is true of all databases on the market). Despite that caveat, we think that the L2 database works very well for our purposes because we care about the signal that these names send about an individual's political resources. One of the advantages of using the L2 database is that politicians have the opportunity to use very similar databases. In other words, if politicians are using such data, they are getting the same information we are using. Thus our tests are using the same (or at least a very similar) signal that the politicians are potentially receiving. As a result, the uncertainty associated with modeled data is less of a concern for us than it might be for other projects. However, to further mitigate concerns about uncertainty we use these variables to create a political resources factor, thus decreasing the noise of this measure. In the Appendix, we also present an additional analysis, in which we create weights based on the number of matches for each alias with the idea that there is more certainty about the information/signal each name is sending if L2's database contains more people with that name (see Table SA.5).
  - 8 We focus on whether legislators responded to emails, but our results are also robust to using the audit study's other outcome variables: responding in a timely manner, and answering the original question. These results are presented in Tables SA.6 to SA.9 in the Appendix.
  - 9 This analysis, and all those in Table 1, use a dataset where each alias represents 1 observation.

**Table 1.** Correlation results for political resources and partisanship.

| Correlation                         | All aliases         |               |                |
|-------------------------------------|---------------------|---------------|----------------|
| <u>All White Legislators</u>        |                     |               |                |
| Responded/White Alias               | <b>0.33</b>         |               |                |
|                                     | <b>[0.14, 0.49]</b> |               |                |
| N (aliases)                         | 101                 |               |                |
| Correlation                         | White aliases       | Black aliases | Latino aliases |
| <u>All White Legislators</u>        |                     |               |                |
| Responded/Political Resources       | -0.05               | -0.31         | -0.01          |
|                                     | [-0.33, 0.25]       | [-0.61, 0.06] | [-0.39, 0.38]  |
| <u>White Democratic Legislators</u> |                     |               |                |
| Responded/Democrat                  | 0.15                | -0.27         | 0.12           |
|                                     | [-0.15, 0.42]       | [-0.58, 0.10] | [-0.28, 0.48]  |
| Responded/Republican                | -0.02               | -0.12         | -0.04          |
|                                     | [-0.31, 0.27]       | [-0.47, 0.25] | [-0.42, 0.35]  |
| <u>White Republican Legislators</u> |                     |               |                |
| Responded/Democrat                  | -0.10               | -0.11         | -0.14          |
|                                     | [-0.38, 0.20]       | [-0.46, 0.27] | [-0.50, 0.26]  |
| Responded/Republican                | 0.01                | 0.01          | 0.02           |
|                                     | [-0.28, 0.30]       | [-0.36, 0.37] | [-0.37, 0.41]  |
| N (aliases)                         | 46                  | 29            | 26             |
| N (matches)                         | 6,165,604           | 404,943       | 2,717,874      |

Note: The table reports Pearson’s correlation coefficients for a two-sided test. 95% confidence intervals in square brackets. Significant correlations are highlighted in **bold**. *Political Resources* refers to the first dimension extracted from a factor analysis including the variables for income, education, housing, and turnout. The eigenvalue of the first factor is 2.73.

with better financial resources would be able to donate more money and legislators have incentives to provide more help to individuals with higher turnout rates (Verba, Schlozman and Brady 1995). Similarly, individuals who share the legislators’ partisanship should be more likely to receive answers because they are more likely to vote and support the candidate at the polls. *If these confounding characteristics explain the differences in how racial groups are treated, then differences in these characteristics should also predict how different individuals within racial groups are treated.* To test whether this is the case, we computed a series of correlation coefficients to test for these potential effects. The tests are done within racial categories and the results are displayed in Table 1. We analyze correlations within racial categories because comparing across racial categories would confound these characteristics with race.<sup>10</sup>

The different columns in the bottom portion of Table 1 report the correlations between average response rates and political resources as well as partisanship respectively for different subsets of our data. The first results column reports the correlations for the white aliases in the study, the next column for black aliases, and the third results column for Latino aliases.<sup>11</sup>

The results show that all 15 of these correlations are statistically insignificant. In fact, the two correlation coefficients that come closest to being statistically significant and that have the highest absolute value are negative – a finding in the opposite direction of what a violation of the

<sup>10</sup> The estimation strategy of this first part of our analysis follows Bertrand and Mullainathan (2004).

<sup>11</sup> Table SA.10 in the Appendix also reports the bivariate correlations for each constituting variable of the political resources factor score separately.

excludability assumption would predict. And given that we are testing 15 different correlations, we would almost expect to observe at least one correlation to be significant at the 95% confidence level just by chance.

Of course this is a small dataset so we might not expect the results to be statistically significant; however, the results are also substantively small. To get an idea of the substantive meaning of these correlations, we can also interpret them as the square root of the  $R^2$  values of the respective bivariate regressions. In other words, if we were to regress the average response rate among white aliases on the political resources factor score, the  $R^2$  value for that regression would be  $-0.05^2 = 0.0025$ , indicating that the variance in political resources would only manage to explain 0.25% of the variation in response rates. It turns out that the mean (absolute) value for the 15 correlations is 0.099 (and the median is 0.1). Accordingly, the average amount of response rate variation explained by the socioeconomic and political variables is less than one percentage point. By contrast, the amount of variation explained by race is  $0.33^2 = 0.109$ , which is 10.9% or more than 10 times the amount explained by the mean correlation across the potential confounders. Also, of the 15 correlation coefficients, none is larger in magnitude than the correlation between response rate and race.

The results in Table 1 provide support for the notion that the potential socioeconomic and political confounders are not related to response rates among the legislators. However, an even more direct way to test whether these potential confounders would affect this finding is to directly control for them in the analysis. In Table 2 we therefore present a series of logit regressions that first replicate the original analysis by Butler (2014), and then also add in the new information about political resources and partisan signals of the aliases as additional controls.

More specifically, we used the original dataset of sent emails, and merged in the new information about the aliases' political resources and partisanship. For the latter, we created a new variable called *Copartisan* which indicates the probability that email sender and recipient share the same partisan identification. For example, 59% of all Alfonsos in our voter file data are Democrats. For an email received by a Democratic legislator and sent by the Alfonso alias, the *Copartisan* variable would therefore take on the value of 0.59.<sup>12</sup>

Model 1 of Table 2 only includes dummy variables to indicate emails sent using a black or Latino alias. Model 2 adds the *Political Resources* factor score for each name, Model 3 the *Copartisan* variable, and Model 4 both of these potential confounders. The results across all model specifications are very consistent. Both the black and Latino dummy variables have negative and statistically significant coefficient estimates in all four models, confirming the finding of the original study and suggesting that black and Latino aliases were less likely to receive an email response. At the same time, both the Political Resources and Copartisan variables fail to reach statistical significance and are relatively close to zero, regardless of whether they are included separately or jointly.

The results in Table 2 thus confirm the findings reported in Table 1. Together, these two analyses provide strong support for the excludability assumption as they suggest that the different response rates between white and nonwhite aliases are indeed driven by the putative race, and not by their putative political resources or partisanship.

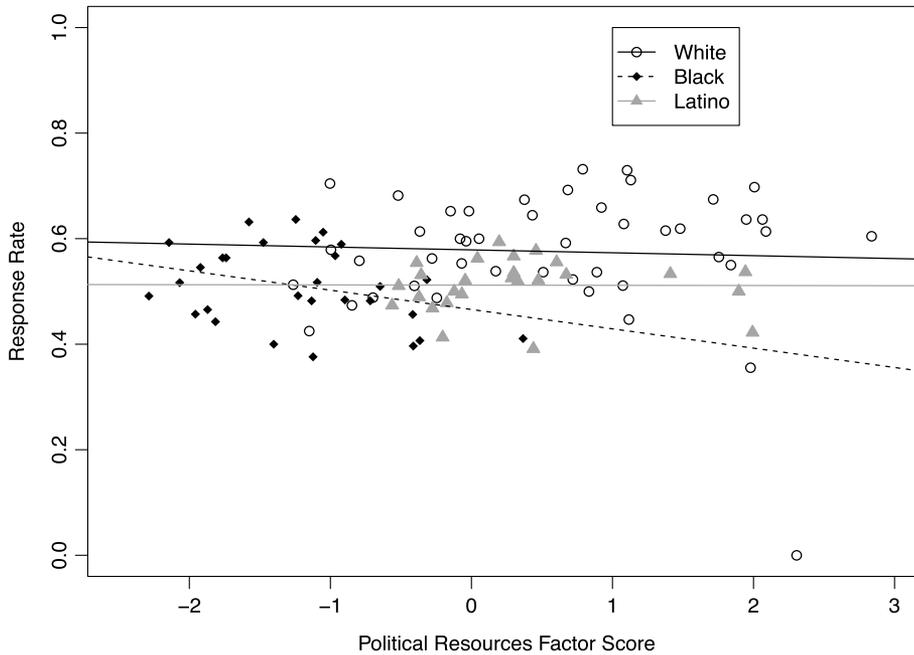
Figure 1 illustrates the findings graphically by showing the relationship between the political resources factor scores and the average response rates among white legislators for each group of aliases. Overall, there is not a strong relationship between political resources and the likelihood of receiving a response from the state legislators. The three respective regression lines are

<sup>12</sup> Only 15% of the Alfonsos in our dataset are Republicans. For emails received by Republican legislators and sent by the Alfonso alias, the *Copartisan* variable would therefore take on the value of 0.15.

**Table 2.** Pooled effects of race, political resources, and partisanship.

|                     | Outcome Variable: Responded Yes/No |                                 |                                 |                                 |
|---------------------|------------------------------------|---------------------------------|---------------------------------|---------------------------------|
|                     | (1)                                | (2)                             | (3)                             | (4)                             |
| Black Alias         | <b>-0.293</b><br><b>(0.064)</b>    | <b>-0.360</b><br><b>(0.086)</b> | <b>-0.292</b><br><b>(0.064)</b> | <b>-0.356</b><br><b>(0.087)</b> |
| Latino Alias        | <b>-0.252</b><br><b>(0.064)</b>    | <b>-0.263</b><br><b>(0.065)</b> | <b>-0.248</b><br><b>(0.064)</b> | <b>-0.260</b><br><b>(0.065)</b> |
| Political Resources |                                    | -0.038<br>(0.032)               |                                 | -0.036<br>(0.032)               |
| Copartisan          |                                    |                                 | -0.079<br>(0.128)               | -0.064<br>(0.129)               |
| Constant            | <b>0.307</b><br><b>(0.046)</b>     | <b>0.330</b><br><b>(0.049)</b>  | <b>0.332</b><br><b>(0.061)</b>  | <b>0.349</b><br><b>(0.063)</b>  |
| N (emails)          | 5,880                              | 5,880                           | 5,880                           | 5,880                           |
| N (aliases)         | 101                                | 101                             | 101                             | 101                             |
| N (matches)         | 9,288,421                          | 9,288,421                       | 9,288,421                       | 9,288,421                       |
| Log Likelihood      | -4,051.865                         | -4,051.174                      | -4,051.677                      | -4,051.052                      |
| Akaike Inf. Crit.   | 8,109.730                          | 8,110.349                       | 8,111.354                       | 8,112.104                       |

Note: Table entries are logit regression coefficients with standard errors in parentheses. Only white legislators are included in the analysis. *Political Resources* refers to the first dimension extracted from a factor analysis including the variables for income, education, housing, and turnout. The eigenvalue of the first factor is 2.73. Significance at the 95% level is highlighted in **bold**, one-tailed tests.



**Figure 1.** Relationship between average response rates (white legislators) and political resources factor score.

all relatively flat indicating no substantive associations.<sup>13</sup> And to the extent that there is a relationship, it is in the wrong direction.

<sup>13</sup> Figure SA.1 in the Appendix shows a similar plot including the 95% confidence intervals for each regression line—all three of which overlap substantially.

Finally, the patterns reported in Tables 1 and 2 are robust to limiting the analysis to voter information based on full name matches (see Tables SA.1 and SA.2 in the Appendix), or to aliases that had more than 2,000 matches in the voter database (see Tables SA.3 and SA.4 in the Appendix). The results are also robust when analyzing the audit study's other outcome variables: responding in a timely manner, and answering the original question (see Tables SA.6 to SA.9 in the Appendix). We also control for the possibility that examining the average levels of a characteristic for each name could mask important variation. For example, a given alias could have the same average income as another name, but also include a larger share of relatively poor and relatively rich people. In Table SA.11 in the Appendix, we therefore also include standard deviation measures of each indicator, and arrive at the same conclusions reported above.<sup>14</sup> All these results provide strong empirical evidence that the exclusion restriction (with regard to important political confounders) is not violated in experiments with racially distinct names.

#### 4 Conclusion

Excludability is a key assumption in all experimental work (Gerber and Green 2012). In the context of studies using racially distinctive names, excludability is the assumption that subjects are responding to the putative race of the individual and not to something else. We tested this assumption with data from an audit study of state legislators. While the names used in the audit study signal race, they could also signal information about the individuals' socioeconomic and political characteristics. We used the L2 company's voter database to learn what signals each name might send. We first looked within each of the three racial groups of aliases whether these other pieces of information predicted legislators' likelihood of responding in the original audit study. If the public officials were responding less to blacks because they were using race as a shortcut for other characteristics, then we should have observed that they were also treating aliases differentially based on those characteristics within each group. However, the evidence we present suggests that politicians are not responding to these characteristics when deciding how to respond to constituents. We also showed that the original results were similar when directly controlling for these potential confounders. Our findings are robust to a number of alternative model specifications and assumptions, and thus support the excludability assumption that politicians are not using racially distinctive names as a signal of potentially confounding factors. Most importantly, our results provide empirical support for the use of racially distinctive names as a way of signaling race. Knowing that racially distinctive names can be used as a tool to measure racial bias will allow researchers to continue to focus on other aspects of the experimental design in order to contribute to measuring, understanding and minimizing racial bias.

#### Supplementary material

For supplementary material accompanying this paper, please visit <https://doi.org/10.1017/pan.2016.15>.

#### References

- Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. 2016. Cultural assimilation during the age of mass migration. NBER Working Paper No. 22381. <http://www.nber.org/papers/w22381>.
- Bertrand, Marianne, and Sendhil Mullainathan. 2004. Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *The American Economic Review* 94:991–1013.
- Biavaschi, Costanza, Corrado Giulietti, and Zahra Siddique. 2013. The economic payoff of name Americanization. IZA Discussion Paper No. 7725, <http://ssrn.com/abstract=2363212>.

<sup>14</sup> A final robustness check regards an outlier observation in the data. The white alias "Jill Smith" received 0 replies. While we could not determine why this might be the case, Tables SA.12 and SA.13 in the Appendix replicate the main analysis for all white aliases except Jill Smith. The results confirm our findings discussed above and are not driven by this outlier.

- Broockman, David E. 2013. Black politicians are more intrinsically motivated to advance blacks' interests: A field experiment manipulating political incentives. *American Journal of Political Science* 57:521–536.
- Butler, Daniel M. 2014. *Representing the advantaged: how politicians reinforce inequality*. New York: Cambridge University Press.
- Butler, Daniel M., and David E. Broockman. 2011. Do politicians racially discriminate against constituents? A field experiment on state legislators. *American Journal of Political Science* 55:463–477.
- Butler, Daniel, and Jonathan Homola. 2016. Replication data for: an empirical justification for the use of racially distinctive names to signal race in experiments. doi:10.7910/DVN/LUGBL1, Harvard Dataverse, V1, [UNF:6:xVWZjghKCtaI26kPihDaRA==].
- Dafoe, Allan, Baobao Zhang, and Devin Caughey. 2015. Confounding in survey experiments: diagnostics and solutions. Manuscript. <http://www.allandafoe.com/confounding>.
- DeSante, Christopher D. 2013. Working twice as hard to get half as far: Race, work ethic, and America's deserving poor. *American Journal of Political Science* 57:342–356.
- Fryer, Roland G. Jr, and Steven D. Levitt. 2004. The causes and consequences of distinctively black names. *The Quarterly Journal of Economics* 119:767–805.
- Gerber, Alan S., and Donald P. Green. 2012. *Field experiments: design, analysis, and interpretation*. New York: W.W. Norton & Company.
- Goldstein, Joshua R., and Guy Stecklov. 2016. From Patrick to John F.: Ethnic names and occupational success in the last era of mass migration. *American Sociological Review* 81:85–106.
- Hainmueller, Jens, Daniel J. Hopkins, and Teppei Yamamoto. 2014. Causal inference in conjoint analysis: Understanding multidimensional choices via stated preference experiments. *Political Analysis* 22:1–30.
- Heckman, James J. 1998. Detecting discrimination. *Journal of Economic Perspectives* 12:101–116.
- Heckman, James J., and Peter Siegelman. 1992. The Urban institute audit studies: Their methods and findings, ed. Michael Fix and Raymond J. Struyk. *Clear and convincing evidence: measurement of discrimination in America*. Lanham, MD: Urban Institute Press, pp. 260–311.
- Hutchings, Vincent, and Spencer Piston. 2011. The determinant and political consequences of prejudice, ed. James N. Druckman, Donald P. Green, James H. Kuklinski, and Arthur Lupia. *Cambridge handbook of experimental political science*. New York: Cambridge University Press, pp. 306–319.
- McClendon, Gwyneth. 2016. Race and responsiveness: A field experiment with South African politicians. *The Journal of Experimental Political Science* 3:60–74.
- Mendez, Matthew S., and Christian R. Grose. 2014. Revealing discriminatory intent: Legislator preferences, voter identification, and responsiveness bias. USC CLASS Research Paper No. 14-17, <http://ssrn.com/abstract=2422596>.
- Pager, Devah. 2007. The use of field experiments for studies of employment discrimination: Contributions, critiques, and directions for the future. *The Annals of the American Academy of Political and Social Science* 609:104–133.
- Sen, Maya, and Omar Wasow. 2016. Race as a “Bundle of Sticks”: Designs that estimate effects of seemingly immutable characteristics. *Annual Review of Political Science* 19:499–522.
- Verba, Sidney, Kay Lehman Schlozman, and Henry E. Brady. 1995. *Voice and Equality: Civic Voluntarism in American Politics*. Cambridge: Harvard University Press.
- White, Ariel, Julie Faller, and Noah Nathan. 2015. What do I need to vote? Bias in information provision by local election officials. *American Political Science Review* 109:129–142.