AER: Insights 2024, 6(3): 313–323
https://doi.org/10.1257/aeri.20230195

Nonrepresentativeness in Population Health Research: Evidence from a COVID-19 Antibody Study[†]

By Deniz Dutz, Michael Greenstone, Ali Hortaçsu, Santiago Lacouture, Magne Mogstad, Azeem M. Shaikh, Alexander Torgovitsky, and Winnie van Dijk*

We analyze representativeness in a COVID-19 serological study with randomized participation incentives. We find large participation gaps by race and income when incentives are lower. High incentives increase participation rates for all groups but increase them more among underrepresented groups. High incentives restore representativeness on race and income and also on health variables likely to be correlated with seropositivity, such as the uninsured rate, hospitalization rates, and an aggregate COVID-19 risk index. (*JEL* C83, I11, I14, J15)

Many study samples are not representative of their target populations because inclusion in the study sample requires participation by those invited to participate. A recent report for the National Academies of Sciences, Engineering, and Medicine (NASEM 2022) documents that clinical trials have lower participation rates among racial minorities, despite those minorities making up an equal or greater share of the relevant patient populations. Recent concerns about nonrepresentativeness in health studies have been raised by Einav et al. (2020), Oster (2020), and Bradley et al. (2021), among others. An important nonhealth example of a nonrepresentative study sample was exhibited in the 2020 US decennial census, which likely undercounted racial minorities (US Census Bureau 2022).

Nonrepresentative samples are a problem because they can lead to biased estimates for the target population. Such bias can be highly consequential. The census undercount biases demographic estimates for the overall population, a bias that

*Dutz: Department of Economics, University of Chicago (email: ddutz@uchicago.edu); Greenstone: Department of Economics, University of Chicago (email: mgreenst@uchicago.edu); Hortaçsu: Department of Economics, University of Chicago (email: hortacsu@uchicago.edu); Lacouture: Department of Economics, University of Chicago (email: slacouture@uchicago.edu); Mogstad: Department of Economics, University of Chicago (email: mogstad@uchicago.edu); Shaikh: Department of Economics, University of Chicago (email: amshaikh@uchicago.edu); Torgovitsky: Department of Economics, University of Chicago (email: torgovitsky@uchicago.edu); van Dijk: Department of Economics, Yale University and NBER (email: winnie.vandijk@yale.edu).
Amy Finkelstein was the coeditor for this article. The authors are grateful to Danae Roumis for her invaluable support implementing the RECOVER study. Sarah Armstrong, Jackson Reimer, Maya Roy, and Charlie Sheils provided excellent research assistance. This study is approved by the IRB at the University of Chicago (IRB20-0721). This RCT was registered as AEARCTR-0010630 (Greenstone et al. 2023). We gratefully acknowledge financial support from the University of Chicago Women's Board, the Sloan Foundation, the Smith Richardson Foundation, the Becker Friedman Institute, the Booz Allen Foundation, and the Flu Lab. This paper was previously circulated under the title "What Drives (Gaps in) Scientific Study Participation? Evidence from a COVID-19 Antibody Survey."

[†]Go to https://doi.org/10.1257/aeri.20230195 to visit the article page for additional materials and author disclosure statement(s).

could lead to disparities in the allocation of public funding (Wines and Cramer 2022). The allocation of public health resources during the COVID-19 pandemic similarly relied on presumed representative estimates of infection rates and vaccine uptake for the overall population. Even if nonrepresentation does not introduce bias, the results of nonrepresentative studies might be less trusted by the underrepresented groups in ways that could undermine their well-being (Alsan et al. 2023). Nonrepresentativeness could also make studies less applicable to the underrepresented groups if it causes researchers to miss important heterogeneity. The NASEM report argues that, as a consequence of underrepresentation of racial minorities in clinical trials, "large swaths of the US population, and those that often face the greatest challenges, are less able to benefit from [new] discoveries because they are not adequately represented in scientific studies" (NASEM 2022, 107).

We analyze representativeness in a unique COVID-19 serological study. Unlike most studies, the Representative Community Survey Project (RECOVER) COVID-19 serological study experimentally varied financial incentives for participation. The study was conducted on households in Chicago (the target population). Randomly sampled households were sent a package that contained a self-administered blood sample collection kit and were asked to return the sample by mail to be tested for the presence of COVID-19 antibodies ("seropositivity"). Households in the sample were randomly assigned one of three levels of financial compensation for participating in the study: $0, $100, or $500.

We find that households in neighborhoods with high shares of minority and poor households are grossly underrepresented at lower incentive levels. High incentives increase participation rates for all groups but increase them more among underrepresented groups. A $500 incentive restores representativeness in terms of neighborhood-level race and poverty status. Representativeness is also restored in health variables likely to be correlated with seropositivity, such as the uninsured rate, hospitalization rates, and an aggregate COVID-19 risk index. Since incentives were randomly assigned and only revealed after the household was contacted, the noncontact rates at $0 and $100 should be the same as at $500, implying that differential hesitancy to participate is responsible for much of the nonrepresentativeness that we find at lower incentives.

We are not aware of studies that randomize financial incentives in population health studies. It is well appreciated that racial minorities and lower-income households participate in health research at lower rates.[1] The impact of incentives on survey participation rates conditional on demographic characteristics has been studied in the survey methodology literature (see Groves et al. 2009; Singer and Ye 2013, and references therein). The incentives used in this literature are typically an order of magnitude smaller than the incentives in the RECOVER study.

More importantly, these studies typically don't start by inviting a sample that is randomly selected from the general population. For example, the Nonresponse Study described in Juster and Suzman (1995) considered a sample of individuals who previously refused to participate in the Health and Retirement Survey. The failure to start by inviting representative samples makes it impossible to answer

[1] See, for example, Yancey, Ortega, and Kumanyika (2006). This review also surveys strategies to recruit minority research participants, including providing incentives.

questions about the effect of incentives on representativeness in population surveys. The only two exceptions to this rule that we are aware of are from a nonhealth context. Mack et al. (1998) study randomly assigned incentives of $0, $10, and $20 in the first wave of the Survey of Income and Program Participation. Berlin et al. (1992) randomize incentives in the National Adult Literacy Survey. Both of these studies consider participation rates by race and poverty status and find that incentives have a larger effect on participation rates among minorities and low-income groups.

More recently, Dutz et al. (2021) use Norwegian data to study the effect of incentives on selection on unobservables and nonresponse bias in labor market statistics. In a companion paper to this one, Dutz et al. (2023) apply the methods developed in Dutz et al. (2021) to the data from the RECOVER study to bound the population average COVID-19 risk score. In contrast, the focus of this paper is on the impact of varying financial incentives on representativeness of participants, including on racial, socioeconomic, and health variables.

Substantively, our paper contributes to ongoing discussions about the quality of COVID-19 serological studies, in particular the challenges that stem from nonrepresentativeness.

Although serological studies were implemented in part to address bias due to the existence of asymptomatic and untested infections (Aspelund et al. 2020; Manski and Molinari 2021), systematic reviews and meta-analyses have emphasized that they often relied on nonrepresentative ("convenience") samples, exposing them to a different potential source of bias (see, e.g., Bobrovitz et al. 2021; Chen et al. 2021). Due to the exponential nature of transmission models, even small biases can translate into large forecast errors (see, e.g., Ioannidis, Cripps, and Tanner 2022).

I. Study Design and Implementation

A. *Background*

The RECOVER study was carried out in Chicago between December 2020 and March 2021. The study was designed and conducted in collaboration with two partners from the University of Chicago: NORC, a leading survey and research organization, and the Wilson Antibody Biology Laboratory. The RECOVER study was a pilot study intended to measure participation rates at different levels of compensation across neighborhoods. The results of the pilot were intended to inform the sampling design of a larger study on seropositivity in Chicago. The larger study was never implemented because the advent of vaccines made measuring seropositivity rates a lower public health priority.

B. *Design and Implementation*

NORC randomly sampled 882 Chicago addresses from US Postal Service data. Hence, the sampled households were representative of the population of households with a mailing address in the city. Invited households were sent a package that contained a self-administered blood collection kit and were asked to return a blood sample to the Wilson Lab to be tested for seropositivity. The package

additionally contained a consent form with a short questionnaire, instructions on self-administering and returning a blood sample, a letter explaining the purpose of the study and providing information on financial compensation for participating (i.e., returning a blood sample), and a prepaid return package.

Households in the sample were randomly assigned one of three levels of compensation: $0, $100, or $500. By virtue of randomization, the incentive groups are probabilistically identical. Balance tests validating the random assignment of compensation levels are presented in online Appendix Table A.4.

Households were asked to select the adult with the earliest upcoming birthday for participation in the study. The letter informed the household members that the returned blood sample would be tested for seropositivity but that they would not be told the result of the test.[2] Hence, the desire to learn about seropositivity status likely does not contribute to the household's motivation for participating in the study. Online Appendix C contains copies of the written materials and additional details on sampling, randomization, and follow-up procedures.

C. *Data*

Our data on the invited sample consist of the randomly assigned compensation level, participation status, and address for each invited household. Using the address, we merge a set of neighborhood characteristics from external sources into our data. All of these characteristics are observed for each household, independently of whether they participated in the study.

Our analyses are centered on two neighborhood characteristics that feature prominently in discussions of representativeness in scientific studies (NASEM 2022): race and poverty status. In particular, we focus on the neighborhood-level share of non-White individuals and the share of poor households (below 200 percent of the poverty line).[3] We additionally consider other neighborhood characteristics, including the shares of working age (ages 25–60) and female individuals, and measures of labor market and health conditions. These characteristics are measured at the zip code level and are obtained from the American Community Survey, the Chicago Health Atlas, and the City Health Dashboard (see online Appendix D for details).

For participating households, we also observe responses to the short questionnaire included with the test kits sent to invited households. The questionnaire asked about the participant's race, gender, age, and household's interval-censored income (see online Appendix C for a copy of the questionnaire). We measure race, age, and gender in the individual data using the same definitions as in the neighborhood data. Individual responses for income were bracketed differently than the neighborhood data. For the individual data, we define a poor household as having yearly income below $50,000, which is close to 200 percent of the poverty line for a family of three ($41,222 in 2020; US Census Bureau 2021a). See online Appendix D for more detail.

[2] The consent form given to each household states, "You will not be notified of your test results." See online Appendix C for more details on outreach materials used in the study.

[3] Our results are not sensitive to how we define race and poverty status; see online Appendix B.

D. *Participants versus the Invited Sample*

Panel A of Table 1 shows that average neighborhood characteristics for participating households differ markedly from invited households. The average participating household lives in a neighborhood that is 53.2 percent non-White, compared to 63.5 percent for all invited households. Participating households reside in neighborhoods with fewer poor, unemployed, and uninsured residents, and fewer negative health outcomes. Participating households also reside in neighborhoods with lower COVID risk indices, a particularly important point given the goals of the RECOVER study. A joint test of equality of neighborhood characteristics between the invited and participating households overwhelmingly rejects (p-value < 0.01).

Panel B of Table 1 presents individual characteristics among participating households as reported in the responses to the study questionnaire. The probability that the participant reports to be non-White is even lower than the neighborhood-level non-White share among participants (37.9 percent versus 53.2 percent). Participants are also more likely to be of working age than the average resident in a participating household's neighborhood. These differences could arise for several reasons, including within-neighborhood selection of households, within-household selection of the participating member, and differences in measurement between the individual and public datasets due to, for example, differential wording of questions.

E. *Participation Rates and Incentives*

Panel A of Table 2 reports the proportion of households that participated in the RECOVER study by incentive level. Only 6 percent of unincentivized households participated. This rate is similar to participation rates in comparable serological surveys that tested for COVID-19 antibodies.[4] Incentives have a large impact on participation rates, increasing them to 17 percent at $100 and to 29 percent at $500.

Panel B of Table 2 shows the impact of incentives by neighborhood racial composition. We classify a household as being from a majority non-White neighborhood if the neighborhood's share of adults identifying as non-Hispanic White is below 50 percent and classify it as majority White otherwise. Only 4.4 percent of unincentivized households in majority non-White neighborhoods participate, compared to 8.9 percent in majority White neighborhoods. The $100 incentive increases participation for both groups but increases it more for households in majority White neighborhoods (p-value 0.05). The $500 incentive further increases participation for both groups but increases it more for households in majority *non*-White neighborhoods (p-value 0.07). At $500, there is no meaningful difference in participation rates between households in majority White and non-White neighborhoods (p-value 0.84).

Panel C of Table 2 shows similar results for poverty status. We classify a household as being from a higher-poverty neighborhood if the neighborhood's share of households below 200 percent of the poverty line is above the median share of the

[4] In online Appendix E, we show that in serological studies that randomly sampled subjects in the United States, the average (median) participation rate for studies that used mail outreach is 9.0 percent (8.3 percent).

Table 1—Descriptive Statistics

	Participants	Invited
Panel A. Neighborhood characteristics (*external sources*)		
Share non-White (percent)	53.2	63.5
	[24.2]	[27.6]
Share poor (percent)	30.2	35.8
	[14.3]	[15.9]
Share working age (percent)	61.1	59.1
	[8.1]	[8.0]
Share female (percent)	51.4	51.6
	[2.2]	[2.6]
Share unemployed (percent)	6.7	8.5
	[4.2]	[5.3]
Share uninsured (percent)	7.8	8.6
	[4.3]	[4.2]
Drug-related hospitalization rate (per 10,000)	22.1	30.3
	[21.3]	[28.6]
Preventable hospitalization rate (per 10,000)	169.7	192.1
	[73.2]	[80.1]
COVID local risk index	4.3	5.3
	[2.6]	[3.0]
Panel B. Individual characteristics (*RECOVER questionnaire*)		
Participant is non-White (percent)	37.9	
	[48.7]	
Participant's household is poor (percent)	32.7	
	[47.1]	
Participant is working age (percent)	75.2	
	[43.3]	
Participant is female (percent)	58.5	
	[49.5]	
Observations	125	882

Notes: This table the presents average of participants (first column) and the average characteristics of the invited sample (second column). Panel A presents estimates for neighborhood characteristics obtained by merging addresses to external sources. Panel B presents estimates for individual characteristics as reported in the responses to the study questionnaire. Standard deviations are presented in square brackets below averages.

invited sample and classify it as from a lower-poverty neighborhood otherwise.[5] The results are similar to those in panel B, likely in part because neighborhood racial composition and poverty status are strongly correlated in Chicago (see online Appendix Table A.6).

II. How Financial Incentives Affect Representativeness

Heterogeneity in the effect of incentives on participation suggests that incentives might also affect the representativeness of participants relative to the invited sample. Panel A of Table 3 shows how the average neighborhood characteristics of participating households vary by incentive level. Larger incentives increase the neighborhood

[5] In online Appendix B, we show that the main conclusions are robust to different binary classifications. Small sample sizes prevent us from looking at finer partitions of the data.

Table 2—Participation Rates (in Percent) across Incentive Levels and Neighborhood Characteristics

	Incentive level			Incentive difference	
	$0	$100	$500	$100 − $0	$500 − $100
Panel A. Overall					
	6.1	16.8	29.1	10.7	12.3
	(1.8)	(1.8)	(2.9)	(2.5)	(3.4)
Panel B. Racial composition					
Majority White	8.9	25.9	30.0	17.0	4.1
	(2.8)	(2.8)	(5.3)	(4.0)	(6.0)
Majority non-White	4.4	11.3	28.7	6.9	17.5
	(2.2)	(2.2)	(3.5)	(3.1)	(4.1)
Panel C. Poverty status					
Lower poverty	9.7	23.7	31.6	14.0	7.8
	(2.4)	(2.4)	(4.5)	(3.4)	(5.1)
Higher poverty	2.2	9.1	27.3	6.8	18.2
	(2.5)	(2.5)	(3.8)	(3.6)	(4.6)
Observations	374	374	134		

Notes: This table presents participation rates by incentive group and by racial composition and income level (see text for variable definitions). Overall and subgroup-specific cell counts are presented in online Appendix Table A.2. Standard errors are presented in parentheses below the estimated rates. The number of invited households in each incentive level is presented in the bottom row.

non-White and poverty shares of participating households and have a similar monotonic impact on other socioeconomic and health measures.[6] Tests of equality across incentive levels reject at the 5 percent level for five out of the nine characteristics and at the 10 percent level for seven out of the nine characteristics. A joint test of equality rejects at the 10 percent level (p-value 0.06). Panel B of Table 3 shows that the individual characteristics of participants also differ substantially by incentive level. Only 21 percent of unincentivized participants self-identify as non-White, compared to almost 57 percent of participants in the $500 incentive arm.

Because we observe neighborhood characteristics for all invited households, we can evaluate representativeness by comparing these characteristics to participating households. For example, households that participate without incentives reside, on average, in neighborhoods where 48.6 percent of households are non-White, compared to an invited population average of 63.5 percent. This figure increases to 49.1 percent with a $100 incentive and to 62.3 percent with a $500 incentive. Formal tests strongly reject equality between the invited sample and the unincentivized and $100 participants ($p$-values less than 0.01) but not between the invited sample and $500 participants ($p$-value 0.80). Other socioeconomic characteristics show similar patterns. A joint test across all nine characteristics rejects equality between the invited sample and unincentivized participants (p-value 0.03) and between the invited sample and the $100 participants ($p$-value 0.02) but not between the invited sample and the $500 participants ($p$-value 0.35). These results imply that an unincentivized study and a study with a

[6] Some of the nine characteristics are strongly correlated, but many are not. A full correlation matrix is reported in online Appendix Table A.6.

Table 3—Representativeness and Composition of Participants across Incentive Groups

	Incentive level			Invited	*p*-value of selection	*p*-value of nonrepresentativeness		
	$0	$100	$500			$0	$100	$500
Panel A. Neighborhood characteristics (*external sources*)								
Share non-White (percent)	48.6	49.1	62.3	63.5	0.02	0.01	0.00	0.80
	(4.9)	(3.0)	(3.8)	(0.9)				
Share poor (percent)	25.8	27.7	36.9	35.8	0.00	0.00	0.00	0.66
	(2.9)	(1.7)	(2.2)	(0.5)				
Share working age (percent)	60.4	62.0	59.9	59.1	0.42	0.43	0.00	0.51
	(1.7)	(1.0)	(1.3)	(0.3)				
Share female (percent)	51.2	51.2	51.8	51.6	0.37	0.39	0.19	0.65
	(0.5)	(0.3)	(0.4)	(0.1)				
Share unemployed (percent)	6.1	6.2	8.0	8.5	0.08	0.03	0.00	0.53
	(0.9)	(0.5)	(0.7)	(0.2)				
Share uninsured (percent)	6.9	7.3	9.2	8.6	0.05	0.04	0.01	0.41
	(0.9)	(0.5)	(0.7)	(0.1)				
Drug-related hospitalization rate (per 10,000)	17.4	19.9	28.6	30.3	0.07	0.03	0.00	0.70
	(4.4)	(2.7)	(3.4)	(1.0)				
Preventable hospitalization rate (per 10,000)	158.0	158.8	193.8	192.1	0.04	0.04	0.00	0.89
	(15.0)	(9.1)	(11.5)	(2.7)				
COVID local risk index	4.1	3.9	5.3	5.3	0.03	0.05	0.00	0.89
	(0.5)	(0.3)	(0.4)	(0.1)				
Joint test					0.06	0.03	0.02	0.35
Panel B. Individual characteristics (*RECOVER questionnaire*)								
Participant is non-White (percent)	21.1	31.7	56.8		0.01			
	(10.8)	(6.1)	(7.8)					
Participant's household is poor (percent)	13.3	32.1	41.7		0.15			
	(12.1)	(6.2)	(7.8)					
Participant is working age (percent)	57.9	75.0	84.2		0.10			
	(9.8)	(5.5)	(7.0)					
Participant is female (percent)	68.4	47.5	71.1		0.04			
	(11.1)	(6.2)	(7.9)					
Joint test					0.01			

Notes: This table presents the average characteristics of participants across incentive groups (first three columns), the average characteristics of the invited sample (fourth column), the *p*-value for equality of participant averages across incentive groups (fifth column), and the *p*-value for equality of the invited and participant averages for each incentive group (last three columns). Panel A presents estimates for neighborhood characteristics obtained by merging addresses to external sources. Panel B presents estimates for individual characteristics as reported in the responses to the study questionnaire. Standard errors are presented in parentheses.

$100 incentive would both be highly nonrepresentative, while a study using a $500 incentive would be representative, at least along the dimensions we observe.

Given the goal of the RECOVER study, a particularly important dimension of nonrepresentativeness is the COVID-19 local risk index. A study with no incentive or a $100 incentive would understate the average COVID-19 risk index in the target population by at least 1.2 points on a ten-point scale, but a study with a $500 incentive would be almost exactly representative. Taken together, our findings suggest that one should be cautious in using results from a study with lower incentives to draw inference about seropositivity rates in the population. In online Appendix Table A.7, we show that this conclusion holds true even if we reweight the unincentivized or $100 participants by neighborhood race and poverty status to achieve

representativeness on these dimensions. This suggests that study participation at lower incentives is correlated with these health variables, even conditional on neighborhood racial composition and poverty, casting doubt on reweighting as a solution to nonrepresentativeness.

Nonrepresentativeness relative to the invited sample is caused by differential nonparticipation. Nonparticipation occurs for one of two reasons: a sampled household is unable to be contacted (noncontact), or a contacted household does not participate because the perceived costs of doing so exceed the perceived benefits (hesitancy). Suppose that any contacted household would participate at $500. If this is the case, then all nonparticipation at $500 must be due to noncontact. Since incentives were randomly assigned and only revealed after the household was contacted, the noncontact rates at $0 and $100 should be the same as at $500, implying that the observed differences in the participation rates by neighborhood race and poverty status must be due to hesitancy. Given that response rates at $500 did not vary much by these neighborhood characteristics, this line of reasoning suggests that differential hesitancy is the main driver of nonparticipation and thus of nonrepresentativeness. In online Appendix F, we develop this argument formally.

III. Conclusion

Using data from a serological study with randomized participation incentives, we found large participation gaps by race and income with lower incentives but not with high incentives. High incentives also help restore representativeness in health variables likely to be correlated with seropositivity, such as the uninsured rate, hospitalization rates, and an aggregate COVID-19 risk index. While large financial incentives appear to be effective in achieving representative samples, they are costly to use and come with potential ethical concerns. There is a large medical literature that grapples with the tension between encouraging research participation and trying to avoid coercing participation from vulnerable populations (see, e.g., Halpern et al. 2004; Groth 2010; Largent and Lynch 2017). Our findings suggest that representativeness and its benefits could be considered when balancing these objectives.

REFERENCES

Alsan, Marcella, Maya Durvasula, Harsh Gupta, Joshua Schwartzstein, and Heidi L. Williams. 2023. "Representation and Extrapolation: Evidence from Clinical Trials." NBER Working Paper 30575.

Aspelund, Karl M., Michale C. Droste, James H. Stock, and Christopher D. Walker. 2020. "Identification and Estimation of Undetected COVID-19 Cases Using Testing Data from Iceland." NBER Working Paper 27528.

Berlin, M., L. Mohadjer, J. Waksberg, A. Kolstad, I. Kirsch, D. Rock, and K. Yamamoto. 1992. "An Experiment in Monetary Incentives." In *JSM Proceedings*, Survey Research Methods Section, 393–98. Alexandria, VA: American Statistical Association.

Bobrovitz, Niklas, Rrahul Krishan Arora, Christian Cao, Emily Boucher, Michael Liu, Claire Donnici, Mercedes Yanes-Lane, et al. 2021. "Global Seroprevalence of SARS-CoV-2 Antibodies: A Systematic Review and Meta-analysis." *PLOS ONE* 16 (6): e0252617.

Bradley, Valerie C., Shiro Kuriwaki, Michael Isakov, Dino Sejdinovic, Xiao-Li Meng, and Seth Flaxman. 2021. "Unrepresentative Big Surveys Significantly Overestimated US Vaccine Uptake." *Nature* 600: 695–700.

Chen, Xinhua, Zhiyuan Chen, Andrew S. Azman, Xiaowei Deng, Ruijia Sun, Zeyao Zhao, Nan Zheng, et al. 2021. "Serological Evidence of Human Infection with SARS-CoV-2: A Systematic Review and Meta-analysis." *Lancet Global Health* 9 (5): e598–609.

Chicago Health Atlas. 2022. *Morbidity: Hospitalization Rates*. Chicago: Chicago Department of Public Health. https://chicagohealthatlas.org/download (accessed September 24, 2022).

City Health Dashboard. 2021. *Health Outcomes and Social and Economic Factors. Chicago, IL, Census Tract Level*. New York: NYU Langone Health. https://www.cityhealthdashboard.com/ (accessed September 5, 2022).

Dutz, Deniz, Michael Greenstone, Ali Hortaçsu, Santiago Lacouture, Magne Mogstad, Azeem M. Shaikh, Alexander Torgovitsky, and Winnie van Dijk. 2023. "What Drives (Gaps in) Scientific Study Participation? Evidence from a COVID-19 Antibody Survey." NBER Working Paper 30880.

Dutz, Deniz, Michael Greenstone, Ali Hortaçsu, Santiago Lacouture, Magne Mogstad, Azeem M. Shaikh, Alexander Torgovitsky, and Winnie van Dijk. 2024. *Data and Code for: "Nonrepresentativeness in Population Health Research: Evidence from a COVID-19 Antibody Study."* Nashville, TN: American Economic Association; distributed by Inter-university Consortium for Political and Social Research, Ann Arbor, MI. https://doi.org/10.3886/E194205V1.

Dutz, Deniz, Ingrid Huitfeldt, Santiago Lacouture, Magne Mogstad, Alexander Torgovitsky, and Winnie van Dijk. 2021. "Selection in Surveys: Using Randomized Incentives to Detect and Account for Nonresponse Bias." NBER Working Paper 29549.

Einav, Liran, Amy Finkelstein, Tamar Oostrom, Abigail Ostriker, and Heidi Williams. 2020. "Screening and Selection: The Case of Mammograms." *American Economic Review* 110 (12): 3836–70.

Greenstone, Michael, Ali Hortacsu, Magne Mogstad, Alexander Torgovitsky, Winnie van Dijk, and Azeem Shaikh. 2023. *University of Chicago SARS-CoV-2 Serosurvey Pilot Study*. AEA RCT Registry. https://doi.org/10.1257/rct.10630-1.0.

Groth, Susan W. 2010. "Honorarium or Coercion: Use of Incentives for Participants in Clinical Research." *Journal of the New York State Nurses Association* 41 (1): 11–22.

Groves, Robert M., Floyd J. Fowler Jr., Mick P. Couper, James M. Lepkowski, Eleanor Singer, and Roger Tourangeau. 2009. *Survey Methodology.* 2nd ed. New York: Wiley.

Halpern, Scott D., Jason H. T. Karlawish, David Casarett, Jesse A. Berlin, and David A. Asch. 2004. "Empirical Assessment of Whether Moderate Payments Are Undue or Unjust Inducements for Participation in Clinical Trials." *Archives of Internal Medicine* 164 (7): 801–03.

Imbens, Guido W., and Joshua D. Angrist. 1994. "Notes and Comments: Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62 (2): 467–75.

Ioannidis, John P. A., Sally Cripps, and Martin A. Tanner. 2022. "Forecasting for COVID-19 Has Failed." *International Journal of Forecasting* 38 (2): 423–38.

Juster, F. Thomas, and Richard Suzman. 1995. "An Overview of the Health and Retirement Study." *Journal of Human Resources* 30 (S1): S7–56.

Largent, Emily A., and Holly Fernandez Lynch. 2017. "Paying Research Participants: Regulatory Uncertainty, Conceptual Confusion, and a Path Forward." *Yale Journal of Health Policy, Law, and Ethics* 17 (1): 61–141.

Mack, Stephen, Vicki Huggins, Donald Keathley, and Mahdi Sundukchi. 1998. "Do Monetary Incentives Improve Response Rates in the Survey of Income and Program Participation?" In *JSM Proceedings*, Survey Research Methods Section, 529–34. Alexandria, VA: American Statistical Association.

Manski, Charles F., and Francesca Molinari. 2021. "Estimating the COVID-19 Infection Rate: Anatomy of an Inference Problem." *Journal of Econometrics* 220 (1): 181–92.

NASEM (National Academies of Sciences, Engineering, and Medicine). 2022. *Improving Representation in Clinical Trials and Research: Building Research Equity for Women and Underrepresented Groups.* Washington, DC: National Academies Press.

Oster, Emily. 2020. "Health Recommendations and Selection in Health Behaviors." *American Economic Review: Insights* 2 (2): 143–60.

Singer, Eleanor, and Cong Ye. 2013. "The Use and Effects of Incentives in Surveys." *ANNALS of the American Academy of Political and Social Science* 645 (1): 112–41.

Torgovitsky, Alexander. 2019. "Partial Identification by Extending Subdistributions." *Quantitative Economics* 10 (1): 105–44.

US Census Bureau. 2020. *American Community Survey 5-Year Data* (*2015–2019*) *Subject Tables*. Washington, DC: US Department of Commerce. https://www.census.gov/programs-surveys/acs/data.html (accessed June 14, 2021).

US Census Bureau. 2021a. *Poverty Thresholds*. Washington, DC: US Department of Commerce. https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-poverty-thresholds.html (accessed August 18, 2021).

US Census Bureau. 2021b. *Relationship Files*. Washington, DC: US Department of Commerce. https://www.census.gov/geographies/reference-files/time-series/geo/relationship-files.html#zcta (accessed September 14, 2022).

US Census Bureau. 2022. "Census Bureau Releases Estimates of Undercount and Overcount in the 2020 Census." News Release CB22-CN.02, March 10, 2022. https://www.census.gov/newsroom/press-releases/2022/2020-census-estimates-of-undercount-and-overcount.html#:~:text=The%20Black%20or%20African%20American%20alone%20or%20in%20combination%20population,significant%20undercount%20rate%20of%204.99%25.

Vytlacil, Edward. 2002. "Independence, Monotonicity, and Latent Index Models: An Equivalence Result." *Econometrica* 70 (1): 331–41.

Wines, Michael, and Maria Cramer. 2022. "2020 Census Undercounted Hispanic, Black and Native American Residents." *New York Times*, March 10, 2022. https://www.nytimes.com/2022/03/10/us/census-undercounted-population.html.

Yancey, Antronette K., Alexander N. Ortega, and Shiriki K. Kumanyika. 2006. "Effective Recruitment and Retention of Minority Research Participants." *Annual Review of Public Health* 27: 1–28.