NBER WORKING PAPER SERIES

WHAT DRIVES (GAPS IN) SCIENTIFIC STUDY PARTICIPATION? EVIDENCE FROM A COVID-19 ANTIBODY SURVEY

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

Working Paper 30880
http://www.nber.org/papers/w30880

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 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. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w30880

What Drives (Gaps in) Scientific Study Participation? Evidence from a COVID-19 Antibody Survey

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

NBER Working Paper No. 30880
January 2023
JEL No. C40,C42,C83,I1,I14,I30,O31

ABSTRACT

Underrepresentation of minority and poor households in scientific studies undermines policy decisions and public health. We study data from a serological study that randomized participation incentives. Participation is low (6% at $0, 17% at $100, 29% at $500) and unequal: minority and poor households are underrepresented at low incentive levels. We develop a framework for disentangling non-contact and ``participation hesitancy'' in explaining non-participation. We find that underrepresentation occurs because poor and minority households are more hesitant, not because they are harder to contact. The $500 incentive appears to overcome differences in hesitancy and restore representativeness along observable dimensions.

Deniz Dutz
1126 East 59th Street
Chicago, IL 60637
ddutz@uchicago.edu

Michael Greenstone
University of Chicago
Department of Economics
1126 E. 59th Street
Chicago, IL 60637
and NBER
mgreenst@uchicago.edu

Ali Hortaçsu
Kenneth C. Griffin Department of Economics
University of Chicago
1126 East 59th Street
Chicago, IL 60637
and NBER
hortacsu@uchicago.edu

Santiago Lacouture
University of Chicago
924 E 57th St
Chicago, IL 60637
slacouture@uchicago.edu

Magne Mogstad
Department of Economics
University of Chicago
1126 East 59th Street
Chicago, IL 60637
and NBER
magne.mogstad@gmail.com

Azeem M. Shaikh
Department of Economics
University of Chicago
1126 E. 59th Street
Chicago IL 60637
amshaikh@uchicago.edu

Alexander Torgovitsky
University of Chicago
1126 E. 59th Street Chicago, IL
60637 atorgovitsky@gmail.com

Winnie van Dijk
Department of Economics
Harvard University
Littauer Center
1805 Cambridge Street
Cambridge, MA 02138
and NBER
winnie_vandijk@fas.harvard.edu

A data appendix is available at http://www.nber.org/data-appendix/w30880

1 Introduction

It is widely documented that scientific studies with human subjects suffer from low and unequal participation rates across socioeconomic and demographic groups. Less is known about the key barriers to participation and how they can be addressed. A recent report for the National Academies of Science Engineering and Medicine (NASEM, 2022, p. 107) summarizes the state of the literature:

> *"There is substantial quantitative data demonstrating the size and scope of the problem of underrepresentation and exclusion of populations in research; however, there is a dearth of critical qualitative data about facilitators of successful inclusion."*

The report argues that, as a consequence,

> *"...large swaths of the U.S. 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."*

Nonrepresentative studies can be misleading for policy, especially when considering distributional goals.

There are two broad reasons for non-participation: either researchers are unable to make contact with a sampled household (non-contact), or a contacted household does not participate because the perceived costs of doing so exceed the perceived benefits (hesitancy). In typical settings, it is difficult to disentangle and separately quantify these two reasons for non-participation, as participation rates are all that is observed in the data.

This paper provides a rare opportunity to learn about the causes of low and unequal participation in the context of a COVID-19 serological study. Unlike typical settings, the Representative Community Survey Project's (RECOVER) COVID-19 serological study experimentally varied financial incentives for participation. The study was conducted on households in Chicago. 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 our partner research lab in order 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 show how to use these randomized financial incentives to disentangle and separately quantify non-contact and hesitancy as drivers of non-participation. The random assignment of incentives creates ex-ante identical groups whose participation rates only differ if the financial compensation causes the benefits to exceed the costs. We develop a framework that uses this experimentally-induced variation in participation together with a simple model of participation behavior to separately identify and estimate the relative importance of non-contact and hesitancy for non-participation. When applied to the RECOVER study, our estimates show that both non-contact and hesitancy are important drivers of low participation, but

that inequality in participation across demographic groups is primarily driven by variation in hesitancy.

We describe the RECOVER serological study in Section 2. RECOVER was designed and implemented using best practices in collaboration with NORC, a leading national statistical agency, and the Wilson Antibody Biology Laboratory at the University of Chicago. The collected data consists of the randomly-assigned compensation offer, participation status, and addresses for each sampled household. These addresses are used to link households to a rich set of neighborhood-level (i.e., zipcode-level) characteristics, such as poverty, racial composition, and health, independently of whether the households participated in the study. This allows us to compare the representativeness of participants compared to the target invited population.

We estimate participation rates across the randomly assigned incentive levels in Section 3. In the control group, participation rates are just 6%, which is comparable to the participation rate in other serological studies (Bobrovitz et al., 2021, and our Appendix C). We find that financial compensation has a powerful effect on participation: the $100 incentive increases participation to 17%, and the $500 incentive increases it to 29%. We also find striking differences in participation rates by neighborhood characteristics. For example, in the unincentivized arm, only 2% of households in high poverty neighborhoods participate, compared to 10% in low poverty areas. The $100 incentive substantially increases participation among all groups, but widens differences in participation rates. The $500 incentive increases participation further and, more importantly, it entirely closes the gap in participation.

We develop a framework for quantifying non-contact and hesitancy as drivers of non-participation in Section 4. In the RECOVER study—as in many other scientific studies and social surveys—a sampled household must first be successfully contacted. Because the randomly assigned incentive is only revealed after the household is contacted, non-contact rates do not depend directly on the incentive level. Conditional on being contacted, households decide whether to participate by comparing perceived costs and benefits, which are shifted by randomly assigned incentives. The key assumption is a bound on the proportion of contacted households who would decline to participate at $500. We show how to use the model to quantify non-contact and hesitancy as sources of non-participation. The analysis highlights the critical importance of randomly varying financial incentives, something which is rarely done in existing research.[1]

We apply these methods to the RECOVER study in Section 5. We find that both non-contact and hesitancy are important determinants of low participation. However, these determinants have different implications for who participates. Non-contact rates differ little by household demographics. But households from higher poverty and minority neighborhoods

[1]Bobrovitz et al. (2021) perform a systematic review of COVID-19 serological studies. In Appendix C, we use their metadata to identify serological studies that, like ours, randomly sampled subjects from a region within the United States. In all such studies, incentives are either not specified, are offered uniformly, or are varied in a non-random way. Dutz et al. (2021) reach a similar conclusion in their systematic review of social surveys used in empirical economics research.

have much higher hesitancy rates, implying that they have higher perceived costs of participation. For example, we estimate that 61% of *contacted* households in majority minority neighborhoods would not participate for $100, compared to only 14% in majority White neighborhoods. Hesitancy explains 89% of the participation gap at $0, and 93% at $100. We show that our findings remain qualitatively unchanged even if we assume that up to 25% of contacted households decline to participate at $500. Removing this assumption entirely actually *strengthens* our conclusions about hesitancy and its differential impact, although it weakens our conclusions about the importance of non-contact.

We estimate the extent to which financial incentives overcome the nonrepresentativeness caused by differential hesitancy in Section 6. We stratify average characteristics of participating households by incentive level, and compare these characteristics to those of the entire invited sample. We show that without financial incentives, the participating households are highly nonrepresentative of the invited sample along a range of socioeconomic, racial, and health dimensions, including the risk of COVID-19 infection. As one example, only 13% of the participants in the unincentivized arm are from higher poverty neighborhoods, compared to 46% in the target population. Although this gap largely persists at the $100 incentive, the $500 incentive closes it entirely. We find that the $500 arm is representative across a battery of socioeconomic, risk, and health measures.

Our paper contributes to ongoing discussions about the quality of COVID-19 serological studies. Serological studies were widely used to estimate epidemiological parameters that served as inputs to highly consequential health policy decisions, such as the Infection Fatality Rate (IFR, the likelihood of death conditional on infection), and the share of the population already infected. 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).[2] Due to the exponential nature of transmission models, even small biases can translate into large forecast errors (see, e.g., Ioannidis et al., 2022).

The problem of nonrepresentativity is not specific to studies that rely on convenience samples. It also occurs in studies with random sampling, because some types of households may be relatively difficult to reach or unwilling to participate. Our paper contributes to this discussion by developing tools to measure the causes of nonrepresentativeness. While we find non-contact rates that are high overall, we are also able to isolate the cause of nonrepresentativeness along racial and poverty lines to differences in hesitancy, rather than

[2]Concerns about nonrepresentativeness of study samples have also been raised in relation to other types of studies that aimed to inform pandemic-era health policy. For example, Bradley et al. (2021) argued that nonrepresentative surveys substantially overestimated US vaccine uptake due to overrepresentation of highly educated and white participants. Beyond pandemic health policy, concerns have recently been raised about under-counting of Hispanic, Black and Native American residents in the 2020 U.S. Census, which may lead to under-allocation of government resources to these groups (Wines and Cramer, 2022).

differences in non-contact rates. This finding implies that representativeness along these characteristics can be improved by providing higher incentives to participate.

Our paper is also related to a broader literature on unequal representation in scientific studies. A comprehensive review of this literature can be found in a recent report published by the National Academies of Science Engineering and Medicine (NASEM, 2022). As noted above, this report concludes that, to date, there is limited credible evidence on the *causes of* and *solutions to* low and unequal participation across demographic groups. Our study directly addresses these questions.

Our paper is complementary to recent work in economics that studies the *consequences* of unequal representation in clinical trials. Alsan et al. (2022) propose a model that predicts that unequal representation in clinical trials reduces the extent to which innovation benefits the underrepresented groups. Key parameters in the model are the perceived net benefits for White and Black patients, which drive differences in recruitment costs across these groups.[3] However, Alsan et al. (2022) note that there is no publicly available data on trial recruitment cost, let alone on how these costs vary across demographic groups.[4] While our setting is not a clinical trial, participation decisions in our setting are likely informed by similar considerations. Our evidence on the causes of non-participation and how they vary by race may be useful for informing clinical trial recruitment.

Finally, there is a literature in public economics on how to increase the take-up of social programs. The key barriers are lack of information about eligibility and transaction costs (including stigma) associated with enrollment (see e.g., Moffitt, 1983; Currie, 2006; Dahl et al., 2014; Deshpande and Li, 2018; Finkelstein and Notowidigdo, 2019). In contrast, all sampled households in the RECOVER study are eligible to participate. Instead, the barriers to participation are the researchers' inability to contact the sampled household and their hesitancy to participate conditional on being contacted. There is no private benefit to participation in our context, other than the randomized financial compensation, because participants would not be told the result of the test. Thus, participation decisions reflect a tradeoff between the perceived private costs and social benefits of participation, rather than eligibility or transaction costs.

[3]Based on national survey data collected by a nonprofit organization, Alsan et al. (2022) report that Black Americans are less likely to have confidence in research institutions, to believe science is beneficial for them, or to enroll in clinical trials. Responses to open-text questions in a survey conducted by Alsan et al. (2022) suggest similar conclusions, since Black patients are more likely to cite trust, privacy, and racism as reasons not to enroll, whereas White patients cite logistical barriers and co-morbidities. These explanations are consistent with our finding of racial gaps in hesitancy.

[4]On pg. 9 they write: "The cost—in terms of both money and time—of enrolling a new patient in a trial also varies across demographic groups. To the best of our knowledge, there are no publicly available estimates of trial recruitment costs."

2 Study design and implementation

2.1 Background

The RECOVER serological 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 meant to inform the sampling design of a larger study on seropositivity in Chicago. The larger study was never implemented, partly because of low overall participation rates in RECOVER (see Section 3), and partly because the advent of vaccines made seropositivity a lower public health priority.

2.2 Design and implementation

NORC randomly sampled 882 Chicago addresses from United States Postal Service data. Hence, the sampled households were representative of the population of households with a mailing address in the city. Sampled 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 pre-paid return package.

Households in the sample were randomly assigned one of three levels of compensation: \$0, \$100, or \$500. The latter is quite a high level of compensation for participation in a serological survey. The reason for offering such high compensation was to try to maximize participation in this arm of the experiment. In Section 4, we show how to use the \$500 incentive arm to disentangle non-contact and hesitancy rates for the lower incentive arms.

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

2.3 Data

Our data consists of the randomly assigned incentive level, participation status, and address for each sampled household. Although we observe additional data via the short questionnaire for participating households, we do not have such information for non-participating households. However, we are able to merge our data with address-based information from external sources. As a result, we can observe a rich set of address-based neighborhood characteristics

for each household, independently of whether they participated in the study. This allows us to analyze non-participation conditional on these characteristics and to compare how the composition of participants differs from that of the invited sample.

We focus our attention on two neighborhood characteristics that feature prominently in discussions of representativeness (NASEM, 2022): poverty status and racial composition. We classify a household as being from a higher poverty neighborhood if the percentage of households below 1.5 times the poverty line is above 30% (and classify it as lower poverty otherwise). We classify a household as being from a majority minority neighborhood if the share of adults identifying as non-Hispanic White is below 50% (and classify it as majority White otherwise). This data is measured at the zipcode level, and is obtained from the American Community Survey. In additional analyses, we also consider other dimensions of neighborhood characteristics, including labor market and health conditions. We obtain this information from the Chicago Health Atlas and the City Health Dashboard (see Appendix B for details).

3 Participation rates and responsiveness to incentives

Figure 1a reports the proportion of households who participated in RECOVER by incentive level. Only 6 percent of unincentivized households participated. This rate is comparable to other serological surveys that invited a random sample of households to be tested for COVID-19 antibodies.[5]

Participation rates increase substantially with the level of the incentive. Offering $100 for participation increases participation rates to 17%, almost triple the rate without incentives. Offering $500 for participation increases participation rates further to 29%.

Figures 1b and 1c report participation rates by neighborhood poverty status and racial composition. While 10% of households in lower poverty neighborhoods participate without financial incentives, only 2% in higher poverty neighborhoods do. While 9% of households in majority White neighborhoods participate without financial incentives, only 4% in majority minority neighborhoods do. The $100 incentive increases participation substantially in lower poverty and majority White neighborhoods, but only modestly in higher poverty and majority minority neighborhoods. However, the $500 incentive increases participation rates to almost parity, with no statistically significant differences by either poverty status or racial composition.

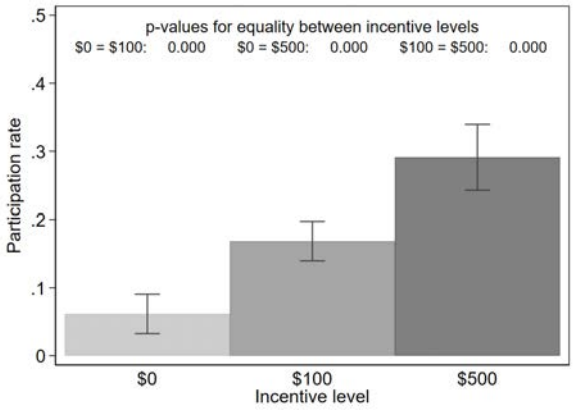
4 A model of study participation

There are two potential explanations for our finding of unequal participation rates. To participate, a household must first be successfully contacted and made aware of the study.

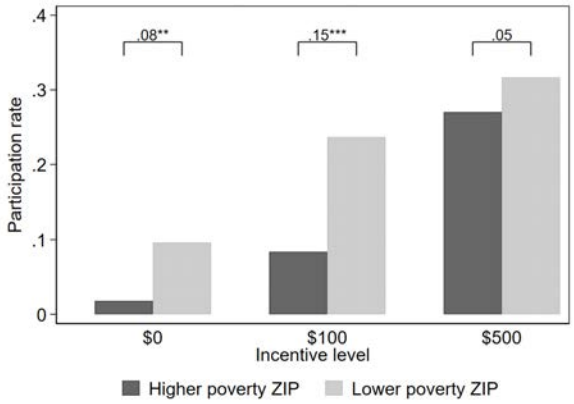
[5]Bobrovitz et al. (2021) perform a systematic review of serological studies. In Appendix C, we use their metadata to identify studies which, like ours, invited a random sample of subjects from a geographic region in the United States to be tested for COVID-19 antibodies. The average (median) participation rate for such studies that used mail outreach is 9.0% (8.3%).

Figure 1: *Participation rates*

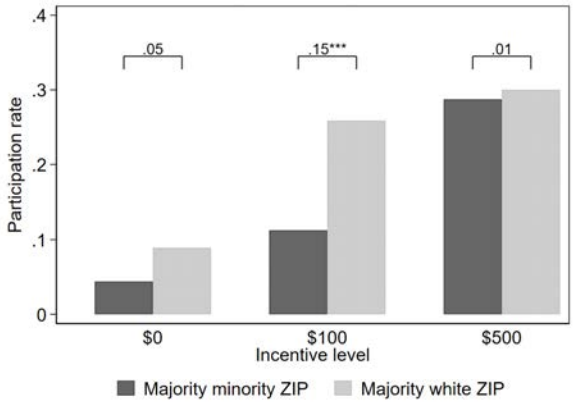
(a) *Overall*



(b) *By poverty status*



(c) *By racial composition*



Notes: This figure reports participation rates and 90% confidence intervals by incentive group for the overall sample (a), by neighborhood poverty status (b), and by neighborhood racial composition (c). In panel (a), p–values for testing the pairwise equality in participation rates across incentives are shown in the top. In panels (b) and (c), differences in participation levels across subgroups are depicted above the bars. Stars denote p-values smaller than .1 (*), .05 (**), and .01 (***).

Upon being contacted, the household must then decide to participate. Unequal participation rates are due to some combination of systematic differences in the difficulty of contact and hesitancy to participate. In this section, we model the participation process to quantify non-contact and hesitancy as sources of non-participation.

4.1 Model

Let $R_i(z) \in \{0, 1\}$ denote whether household i would participate if assigned incentive z. Participation is a two-step process in which the household is first *contacted*, and then *decides* to participate. Let $C_i(z) \in \{0, 1\}$ denote whether household i would be contacted under incentive level z, and let $D_i(z)$ denote whether they would decide to participate if contacted. Then household i's participation decision is $R_i(z) = C_i(z)D_i(z)$. We will estimate the model separately by demographic groups without any cross-group restrictions, so we suppress demographic conditioning in the notation.

We impose three baseline assumptions on this model. First, since the assigned incentive is only revealed after the household is contacted and opens the package, we assume that contact does not depend on z, so that $C_i(z) \equiv C_i$. Second, we assume that $D_i(z)$ is non-decreasing in z for all i, so that households are more likely to participate under higher incentives. This is the Imbens and Angrist (1994) monotonicity assumption, which Vytlacil (2002) showed is equivalent to assuming that $D_i(z) = \mathbb{1}[H_i \leq z]$ for some latent variable H_i. Together, these two assumptions imply that

$$R_i(z) = C_i \mathbb{1}[H_i \leq z]. \tag{1}$$

We interpret $z - H_i$ as household i's net benefit from participating, and call H_i their *hesitancy* to participate. Third, we assume that the assigned incentive, Z_i, is independent of (C_i, H_i), which is justified by random assignment of incentives.

4.2 Contact and hesitancy rates

We define the *contact rate* as $\gamma \equiv \mathbb{P}[C_i = 1]$ and the *non-contact rate* as $1 - \gamma$. We define the *hesitancy rate* as $\eta(z) \equiv \mathbb{P}[H_i > z | C_i = 1]$, which is the probability that a household would not participate under incentive z if they were contacted. We measure the hesitancy rate conditional on being contacted in order to hold fixed the implementation protocol of the scientific study. Variation in $\eta(z)$ describes how incentives affect the participation rates for households who are able to be contacted using the current outreach method.

4.3 Identification and estimation

The researcher does not observe (C_i, H_i), but only the incentive level, Z_i, and the participation decision $R_i \equiv R_i(Z_i)$ under this incentive level. From these observables, they can estimate the *participation rate*

$$\rho(z) \equiv \mathbb{P}[R_i = 1 | Z_i = z] = \mathbb{P}[C_i = 1, H_i \leq z], \tag{2}$$

where the equality follows from the model (1) and random assignment of the incentive, Z_i. Measuring the contact and hesitancy rates requires determining the relative contribution of the unobservables C_i and H_i to ρ, while allowing these unobservables to be dependent.

To make progress, we consider what can be said about the contact and hesitancy rates under assumptions on the magnitude of the hesitancy rate at the highest incentive, $\bar{z}$. In the RECOVER survey, $\bar{z} = \$500$ is large, suggesting that $\eta(\bar{z})$ is small, and that non-participation in the \$500 treatment arm is primarily or solely due to non-contact. Since contact is not affected by the incentive level, the participation model then allows us to infer the hesitancy rates at lower incentives as well.

To see how this works, suppose that we know $\eta(\bar{z})$ exactly and decompose it as

$$\eta(\bar{z}) = \frac{\mathbb{P}[C_i = 1, H_i > \bar{z}]}{\mathbb{P}[C_i = 1]} = \frac{\mathbb{P}[C_i = 1] - \overbrace{\mathbb{P}[C_i = 1, H_i \leq \bar{z}]}^{= \rho(\bar{z}) \text{ by (2)}}}{\underbrace{\mathbb{P}[C_i = 1]}_{\equiv \gamma}} = 1 - \frac{\rho(\bar{z})}{\gamma}.$$

Rearranging shows that the contact rate γ (and non-contact rate $1 - \gamma$) is identified:

$$\gamma = \frac{\rho(\bar{z})}{1 - \eta(\bar{z})}. \tag{3}$$

Hesitancy rates at other incentive levels can then be identified by the following argument:

$$\begin{aligned} \eta(z) &= \mathbb{P}[z < H_i \leq \bar{z} | C_i = 1] + \mathbb{P}[H_i > \bar{z} | C_i = 1] \\ &= \frac{\rho(\bar{z}) - \rho(z)}{\gamma} + \eta(\bar{z}) = \left(\frac{\rho(\bar{z}) - \rho(z)}{\rho(\bar{z})} \right) (1 - \eta(\bar{z})) + \eta(\bar{z}), \end{aligned} \tag{4}$$

where the second equality used (2), and the third equality substituted in the identified contact rate from (3). We estimate (3) and (4) through their sample analogs by substituting the estimated participation rates $\rho(z)$ and $\rho(\bar{z})$.

Our baseline estimates set $\eta(\bar{z}) = 0$, which corresponds to the assumption that any household would have participated at \$500 incentive had they been aware of it (had they been contacted). Given the generosity of the incentive, we view this as a reasonable assumption. However, we also report estimates that allow $\eta(\bar{z})$ to vary in the set $[0, \alpha]$, where α is a number smaller than $1 - \rho(\bar{z})$, the largest value that keeps γ a proper probability via (3). Under this assumption, bounds on γ and $\eta(z)$ are given by

$$\rho(\bar{z}) \leq \gamma \leq \frac{\rho(\bar{z})}{1 - \alpha} \quad \text{and} \quad \frac{\rho(\bar{z}) - \rho(z)}{\rho(\bar{z})} \leq \eta(z) \leq \frac{\rho(\bar{z}) - \rho(z)(1 - \alpha)}{\rho(\bar{z})}. \tag{5}$$

The widest "worst-case" bounds are obtained at $\alpha = 1 - \rho(\bar{z})$. In Appendix D, we prove that these bounds are sharp (best possible, given the assumptions) for any choice of α, as long as observed participation rates $\rho(z)$ are increasing in z.

5 The causes of low and unequal participation rates in RECOVER

We now use the method in the previous section to separately estimate non-contact and hesitancy in the RECOVER study.

5.1 Baseline estimates

Table 1 reports our baseline estimates, which are constructed under the assumption that all households would choose to participate at \$500 if they were aware of the study ($\eta(\bar{z}) = 0$).

The first column of Table 1 shows estimates of the non-contact rate. Under our baseline

Table 1: *Estimated non-contact and hesitancy rates*

		Hesitancy rate	
	Non-contact rate	At $0	At $100
Overall	0.71	0.79	0.42
	(0.66, 0.76)	(0.68, 0.89)	(0.28, 0.56)
Higher poverty	0.73	0.93	0.69
	(0.68, 0.78)	(0.80, 1.00)	(0.55, 0.83)
Lower poverty	0.68	0.70	0.25
	(0.60, 0.76)	(0.54, 0.85)	(0.02, 0.49)
Difference	0.05	0.24	0.44
	(-0.05, 0.14)	(0.03, 0.45)	(0.17, 0.71)
Majority minority	0.71	0.85	0.61
	(0.66, 0.76)	(0.73, 0.97)	(0.47, 0.74)
Majority White	0.70	0.70	0.14
	(0.60, 0.80)	(0.50, 0.90)	(0.00, 0.47)
Difference	0.01	0.14	0.47
	(-0.09, 0.12)	(-0.08, 0.36)	(0.14, 0.80)

Notes: This reports estimates of non-contact and hesitancy rates under the baseline assumption that all contacted households would choose to participate if offered $500. 90% CIs are shown in parentheses.

assumption, all households who did not participate at $500 did so because they were not contacted, and so our estimates of the non-contact rate are the complement of the participation rates at $500 shown in Figure 1. Participation rates of 29% under the $500 incentive arm correspond to non-contact rates of 71%. Consistent with Figure 1, we find no large or statistically significant differences in non-contact rates across households by neighborhood poverty or racial composition.

The second column of Table 1 shows estimates of the hesitancy rate when no financial incentive is offered. With no financial incentive, 79% of contacted households would not participate. This figure increases to 93% for households in high poverty neighborhoods and to 85% for households in majority minority neighborhoods. These findings suggest that that the perceived costs of participation are empirically relevant barriers to participation, especially for poor and minority households.

The third column of Table 1 shows that a $100 incentive sharply decreases the overall hesitancy rate from 79% to 42%. However, the decrease is largely driven by households in lower poverty and majority White neighborhoods: in majority White neighborhoods, only 14% of contacted households decline to participate when offered the $100 incentive. Hesitancy rates remain substantial among households in higher poverty and majority minority neighborhoods. These findings suggest that the perceived costs of participation are high in general, and much higher for poor and minority households.

5.2 Decomposing the causes of unequal participation

A decomposition exercise helps clarify the relative importance of non-contact and hesitancy for explaining unequal participation by poverty status and racial composition. Suppose that majority minority households had the same hesitancy at $0 as majority White households. Then, instead of a participation rate of .043 at $0, majority minority households would have a $(1 - .71) \times (1 - .70) = .085$ participation rate, only slightly lower than the .090 participation rate for majority White households, and eliminating 89% of the participation gap. The same calculation for the $100 incentive brings participation for majority minority households from 11.3% to 24.8%, relative to 25.8% for majority White households, eliminating 93% of the participation gap. Similarly, if higher poverty households had the same hesitancy as lower poverty households, their participation would rise from 1.8% to 8.2% at $0 and 8.3% to 20.2% at $100, compared to 9.6% at $0 and 24.0% at $100 for lower poverty households. In all cases, setting hesitancy rates equal across households largely closes participation gaps across poverty status and racial composition. These results suggest that unequal participation rates are primarily driven by differences in hesitancy.

5.3 Sensitivity analysis

The estimates in Table 1 use the assumption that all contacted households would choose to participate if offered the $500 incentive. That is, the hesitancy rate at $500 is zero or, in the notation of Section 4, $\eta(\bar{z}) = 0$. In this section, we conduct a sensitivity analysis that estimates bounds on the same parameters under the weaker assumption that $\eta(\bar{z}) \leq \alpha$.

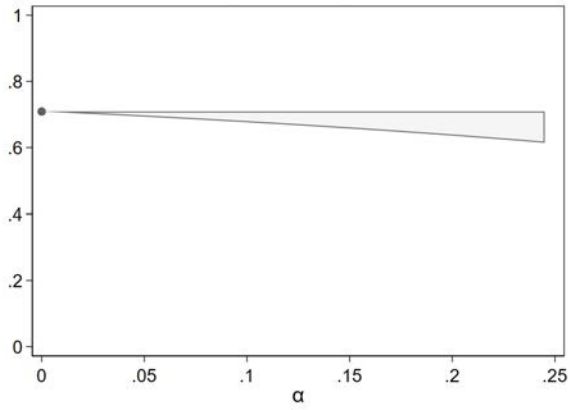
Figure 2 plots the estimated bounds on the overall non-contact and hesitancy rates for α up to .25. For example, allowing $\alpha = .20$ means assuming that up to 20% of contacted households decline to participate at $500 because they find the incentive not high enough to overcome their perceived costs. Even under this conservative assumption, Figure 2a shows that non-contact rates remain high at 64%. Higher hesitancy rates at $500 also rationalize higher hesitancy rates at lower incentive values (see (5)), reinforcing the conclusion that hesitancy is also an important source of non-participation. At $\alpha = .20$, between 79% and 83% of contacted households would not participate without an incentive.

Figure 3 plots estimated bounds by demographic group. Figures 3a and 3b show that the bounds on non-contact rates by demographic group largely overlap for all $\alpha \leq .25$, reinforcing the conclusion that non-contact rates do not vary systematically by demographics. Figures 3c–3f show that the opposite is true for hesitancy rates: even at $\alpha = .25$, hesitancy rates at both $0 and $100 differ markedly by both poverty status and racial composition. These results are consistent with the conclusions from the baseline case.

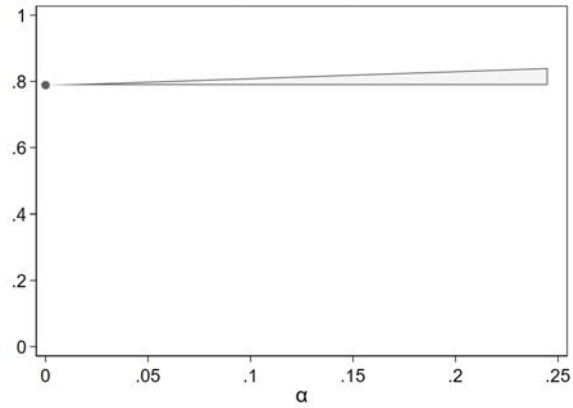
As discussed in Section 4, the largest value that we can set α to while still rationalizing the model is $1 - \rho(\bar{z})$, which we estimate to be 71% among the overall population. This value of α represents the "worst-case" assumption that everyone in the $500 incentive arm was contacted, but 71% declined to participate because $500 was not a sufficient incentive. If

Figure 2: *Bounds on non-contact and hesitancy rates*

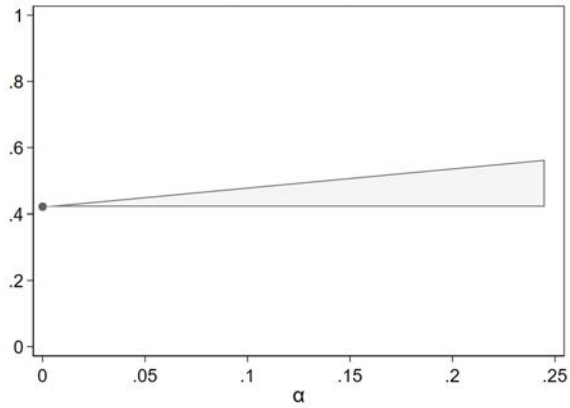
(a) *Non-contact rate*



(b) *Hesitancy rate at $0*



(c) *Hesitancy rate at $100*



Notes: These figures report estimates of the bounds in (5) for different levels of α.

this were true, then non-contact rates would be zero, and hesitancy rates at lower incentives would be even larger; for example between 79% and 94% at $0. Thus, even without taking a stand on α, we can conclude that hesitancy is an important barrier to participation. However, our view is that allowing for the possibility that 71% of contacted households would not trade $500 for a quick at-home blood sample is unreasonable. Under smaller—but still large—values of α, we find non-contact to also be an important cause of non-participation.

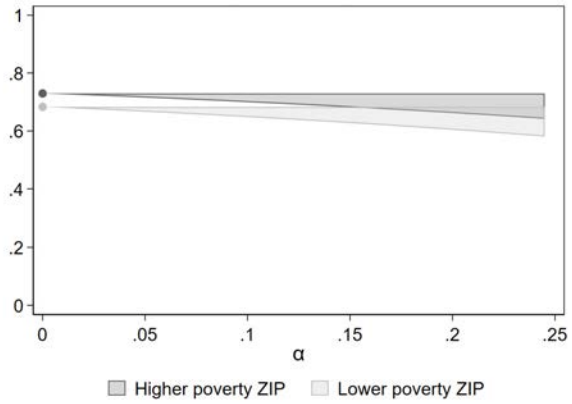
6 How financial incentives affect representativeness

The findings in the previous section show that both non-contact and hesitancy are important causes of low participation, but that only hesitancy differs by poverty status and racial composition. This result implies that different perceived costs of participation lead to non-representativeness in unincentivized studies, but that incentivized studies can improve representativeness. In this section, we quantify the extent to which this is the case.

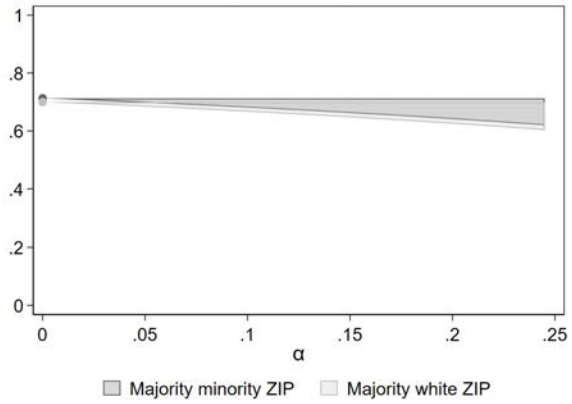
Figure 4 reports average neighborhood-level characteristics of participating households. The estimates are stratified by incentive level; intuitively, one can think of each incentive level

Figure 3: *Bounds on non-contact and hesitancy rates by demographics*

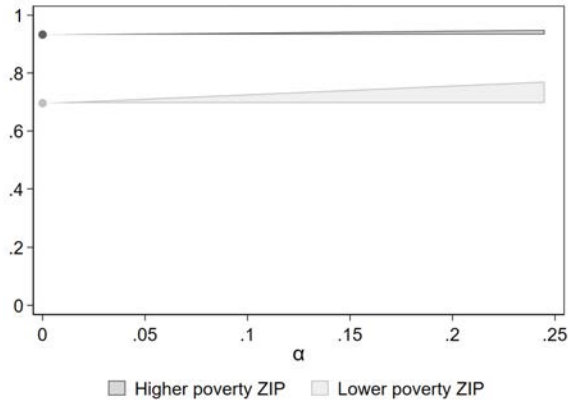
(a) *Non-contact rate by poverty status*



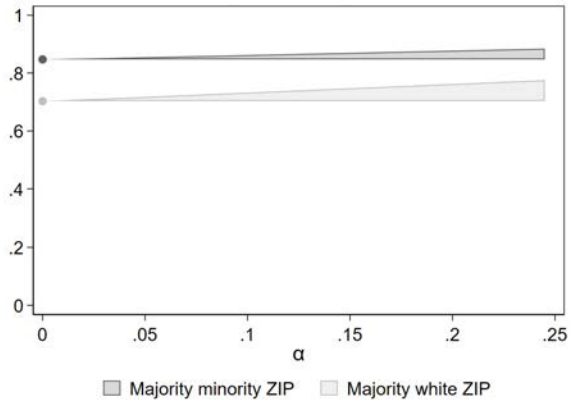
(b) *Non-contact rate by racial composition*



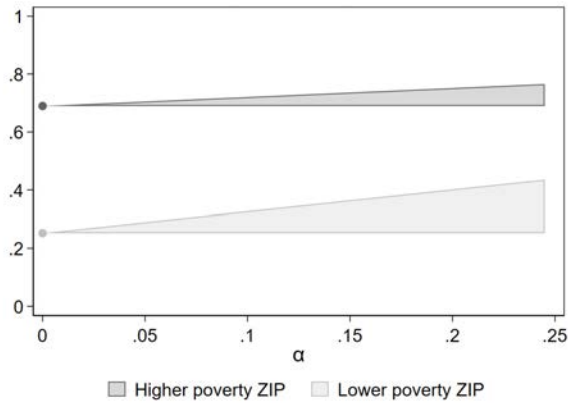
(c) *Hesitancy rate at $0 by poverty status*



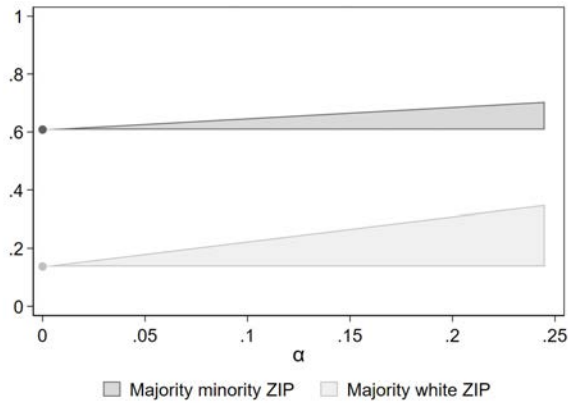
(d) *Hesitancy rate at $0 by racial composition*



(e) *Hesitancy rate at $100 by poverty status*



(f) *Hesitancy rate at $100 by racial composition*



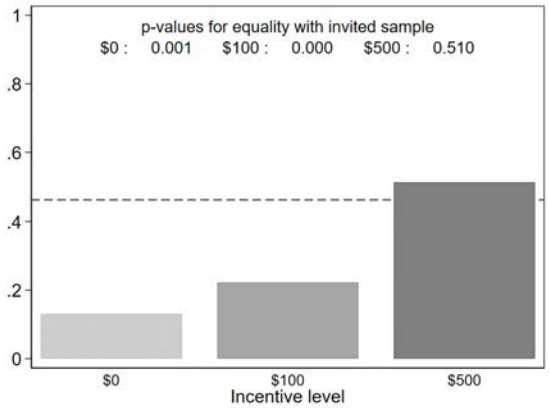
Notes: These figures report estimates of the bounds in (5) for different levels of α broken down by demographic group.

as representing different studies conducted among ex-ante identical populations. The dotted horizontal lines show the mean neighborhood-level characteristics of the target population. Estimates closer to the dotted line are more representative.

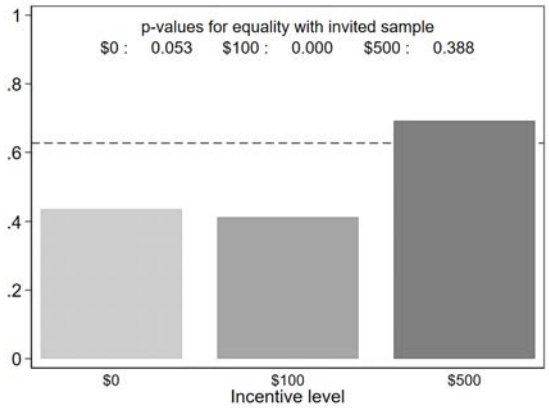
The results show that an unincentivized study would be highly nonrepresentative. A

Figure 4: *Effect of hesitancy on representativeness: main characteristics*

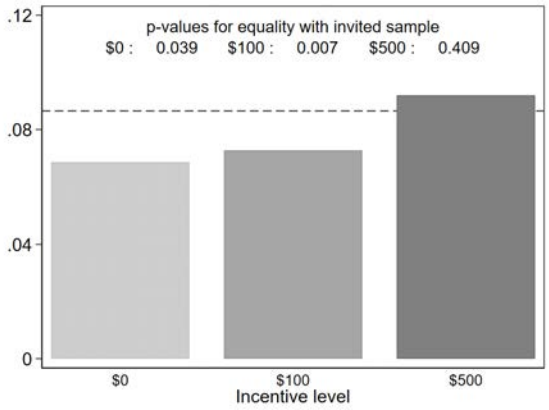
(a) *Indicator for higher poverty neighborhood*



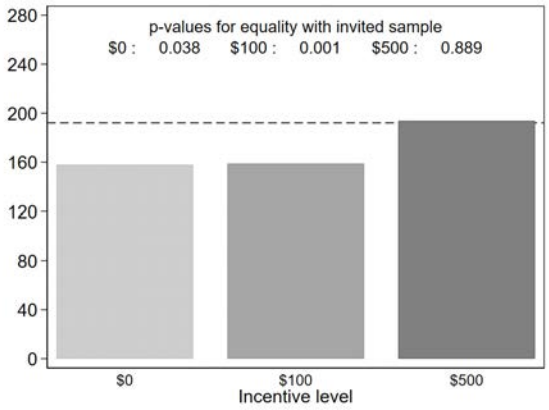
(b) *Indicator for minority neighborhood*



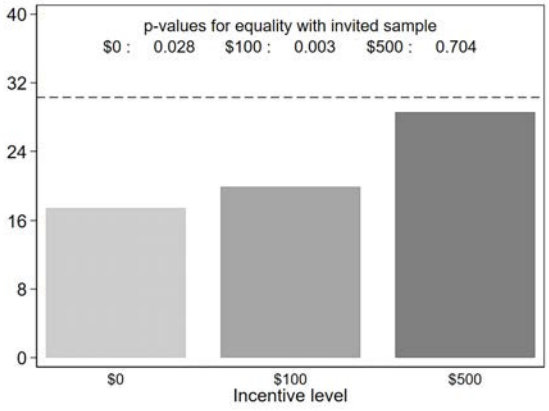
(c) *Uninsured rate*



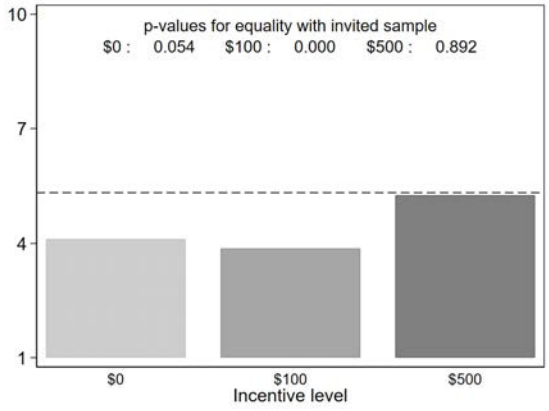
(d) *Preventable hospitalization rate (per 10,000)*



(e) *Drug-related hospitalization rate (per 10,000)*



(f) *COVID-19 local risk index (from 1-10)*



Notes: These figures show participant average neighborhood-level measures stratified by incentive group. The horizontal dashed line depicts the average among the invited sample. Variable definitions are given in Appendix B.

study with a $100 incentive would also be highly nonrepresentative on all six characteristics.[6] However, a study with a $500 incentive is representative on all six characteristics, with

[6]The six characteristics are positively correlated but typically not highly so; Appendix Table E.1 shows that the correlations lie within 0.11 and 0.87.

average participant characteristics that are statistically indistinguishable from those in the target population. For example, with no incentive only 13% of the participants are from higher poverty neighborhoods, compared to a target population average of 46%. This figure would increase to 22% with a $100 incentive, and to 51% with a $500 incentive.

Given the RECOVER study's goal, a particularly important dimension of nonrepresentativeness is the COVID-19 local risk index (Figure 4f). This index directly relates to the goals of our partners at the Wilson Antibody Biology Laboratory.[7] A study with no incentive or a $100 incentive would understate the average COVID-19 risk index in the target population by more than 1.5 points on a 10-point scale, but a study with a $500 incentive would be almost exactly representative.

Appendix Figure E.1 shows similar patterns across a variety of additional risk and health measures. A study with a $500 incentive would be representative for the share of smoking adults, hospitalization rates related to alcohol, behavioral health, and mood or depressive orders, as well as for a credit insecurity index. In contrast, both an unincentivized study and one with a $100 incentive would not be representative for any of these measures, except the hospitalization rate related to alcohol use. Overall, these findings suggest that similar concerns will arise in studies or surveys designed to measure outcomes that are highly correlated with these measures.

7 Conclusion

Scientific studies with human subjects often report lower participation rates among Black, Hispanic, and low socioeconomic status households, but little is known about the causes of underrepresentation. Lack of representation poses a risk to public health and a challenge for policy discussions.

In this paper, we showed how randomized financial incentives can be used to measure the causes of non-participation and how they contribute to underrepresentation. In the context of the RECOVER serological study, we found that non-contact is a cause for low participation, but not of underrepresentation. Hesitancy among contacted households is a cause of both low participation and underrepresentation. High incentives for participation appear to restore representativeness.

[7] Although we observe antibody test results for participants, we do not observe them for non-participants. Our findings suggest one should be extremely cautious in using test results from only the self-selected group of participants.

References

Alsan, M., M. Durvasula, H. Gupta, J. Schwartzstein, and H. L. Williams (2022). Representation and Extrapolation:. *NBER working paper*. 5

Aspelund, K., D. M.C., J. Stock, and C. Walker (2020). Identification and estimation of undetected covid-19 cases using testing data from iceland. *working paper*. 4

Bobrovitz, N., R. K. Arora, C. Cao, E. Boucher, M. Liu, C. Donnici, M. Yanes-Lane, M. Whelan, S. Perlman-Arrow, J. Chen, H. Rahim, N. Ilincic, M. Segal, N. Duarte, J. V. Wyk, T. Yan, A. Atmaja, S. Rocco, A. Joseph, L. Penny, D. A. Clifton, T. Williamson, C. P. Yansouni, T. G. Evans, J. Chevrier, J. Papenburg, and M. P. Cheng (2021, June). Global seroprevalence of SARS-CoV-2 antibodies: A systematic review and meta-analysis. *PLOS ONE 16*(6), e0252617. 3, 4, 7, 29

Bradley, V. C., S. Kuriwaki, M. Isakov, D. Sejdinovic, X.-L. Meng, and S. Flaxman (2021, December). Unrepresentative big surveys significantly overestimated US vaccine uptake. *Nature 600*(7890), 695–700. 4

Chen, X., Z. Chen, A. S. Azman, X. Deng, R. Sun, Z. Zhao, N. Zheng, X. Chen, W. Lu, T. Zhuang, J. Yang, C. Viboud, M. Ajelli, D. T. Leung, and H. Yu (2021, May). Serological evidence of human infection with SARS-CoV-2: A systematic review and meta-analysis. *The Lancet. Global Health 9*(5), e598–e609. 4

Currie, J. (2006). The take up of social benefits. In A. Auerbach, D. Card, and J. Quigley (Eds.), *Public Policy and the Income Distribution*, Chapter 3. New York: Russell Sage Foundation. 5

Dahl, G. B., K. V. Loken, and M. Mogstad (2014). Peer effects in program participation. *The American Economic Review 104*(7), 2049–2074. 5

Deshpande, M. and Y. Li (2018). Who is screened out? application costs and the targeting of disability programs. *AEJ: Policy 11*(4), 213–248. 5

Dutz, D., I. Huitfeldt, S. Lacouture, M. Mogstad, A. Torgovitsky, and W. van Dijk (2021). Selection in surveys. Technical report, National Bureau of Economic Research. 3

Finkelstein, A. and M. J. Notowidigdo (2019). Take-up and targeting: Experimental evidence from snap. *The Quarterly Journal of Economics 134*(3), 1505–1556. 5

Imbens, G. W. and J. D. Angrist (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica 62*(2), pp. 467–475. 9

Ioannidis, J. P., S. Cripps, and M. A. Tanner (2022). Forecasting for COVID-19 has failed. *International Journal of Forecasting 38*(2), 423–438. 4

Manski, C. and F. Molinari (2021). Estimating the covid-19 infection rate: Anatomy of an inference problem. *Journal of Econometrics 220*(S2), 181–192. 4

Moffitt, R. (1983). An economic model of welfare stigma. *The American Economic Review 73*(5), 1023–1035. 5

NASEM (2022). *Improving Representation in Clinical Trials and Research: Building Research Equity for Women and Underrepresented Groups*. Washington, DC: The National Academies Press. 2, 5, 7

Torgovitsky, A. (2019, January). Partial identification by extending subdistributions. *Quantitative Economics 10*(1), 105–144. 31

Vytlacil, E. (2002). Independence, Monotonicity, and Latent Index Models: An Equivalence Result. *Econometrica 70*(1), 331–341. 9

Wines, M. and M. Cramer (2022, March). 2020 Census Undercounted Hispanic, Black and Native American Residents. *The New York Times*. 4