Public Policy Implications of the COVID-19 Pandemic

WHO'S IN AND WHO'S OUT UNDER WORKPLACE COVID SYMPTOM SCREENING?

Krista Ruffini, Aaron Sojourner, and Abigail Wozniak

Abstract

COVID symptom screening, a new workplace practice, is already affecting many millions of American workers. As of this writing, 34 states already require, and federal guidance recommends, frequent screening of at least some employees for fever or other symptoms. This paper provides the first empirical work identifying major features of symptom screening in a broad population and exploring the trade-offs employers face in using daily symptom screening. First, we find that common symptom checkers could screen out up to 7 percent of workers each day, depending on the measure used. Second, we find that the measures used will matter for three reasons: Many respondents report any given symptom, survey design affects responses, and demographic groups report symptoms at different rates, even absent fluctuations in likely COVID exposure. This last pattern can potentially lead to disparate impacts and is important from an equity standpoint. © 2021 by the Association for Public Policy Analysis and Management

INTRODUCTION

Many employers have begun administering workplace screening to flag workers at higher risk of infectiousness so as to reduce the spread of COVID-19. Screening practices commonly include some combination of employee temperature checks and symptom self-reports, and such practices have spread rapidly.[1] Both Federal and state agencies already recommend or require regular infection screening in the workplace (OHSA, 2020), and 34 states require employee screening in some or all businesses as of December 8, 2020. Workers have also been generally receptive to these practices. Two out of five U.S. workers in a recent survey chose temperature

---

[1] Some high-profile employers such as major league sports teams and the White House are doing daily medical testing of employees, but, as of this writing, these remain exceptional.

and symptom checker screens as a factor necessary to make them feel safe at work, making these screens one of the most important workplace safety factors according to employees (American Staffing Association, 2020).

Symptom screening measures are also likely to become common in other settings as policymakers set guidelines for balancing productive, interpersonal activities with contagion risk. For example, some states recommend that elementary and secondary schools ask similar questions of students, and some businesses conduct similar screens of customers.[2] Workplace screens and related devices are likely to remain common in the medium term, as vaccine rollout is estimated to last well into 2021. Moreover, now that symptom screening has become widespread in the context of COVID, these practices are likely to remain a tool that organizations use to manage any future outbreaks of communicable diseases.

Despite the rapidly increasing prevalence of workplace screens, there is little evidence on basic questions about how these screens function, and this paper provides novel descriptive evidence on several key questions. First, how common are COVID-related symptoms in the U.S. workforce? Based on this, who are common symptom screens likely to detect and screen out? What are the differences in detection across various screening options? Are screens equitable in whom they flag? And do differences in underlying health or survey response behaviors lead certain screens to disproportionately flag certain groups?

To answer these questions, we draw on a novel survey, the COVID Impact Survey (CIS). The CIS was administered to approximately 6,500 adults in a nationally representative sample over three weeks in spring 2020, during the initial months of the pandemic. The CIS is unique in combining questions about employment and financial security with questions about symptoms, underlying health, and protective behavior. This combination allows us to approximate responses to a variety of self-reported symptom screens that employers may adopt and to generate new evidence on what types of workers would be flagged by alternative screening practices.

Our findings identify several important features of daily symptom screening. First, screens based on common symptom checkers, including temperature-taking, will likely classify many individuals as high-risk each day. Our analysis finds that up to 7 percent of workers could be flagged as high-risk for COVID-19 daily, depending on the measure used, and infrequent symptom reporting could indicate non-truthful reporting. Second, the screening mechanism used will matter for three reasons: Within a set of possible symptoms, choosing to flag workers as high-risk based on more or fewer reported symptoms will identify different individuals; survey design matters; and demographic groups report symptoms at different rates, even absent fluctuations in likely COVID exposure. This last feature can potentially lead to disparate impacts and is particularly important from the standpoint of equity and discrimination. Finally, although we cannot observe actual COVID infections, our results suggest that employers face a trade-off between screens that may have a higher false negative rate but fewer disparities in detection, versus lower false negative rates but more disparities.

These findings are relevant for both policymakers and employers. Most directly, this paper provides evidence for employers to consider in using a screening program. We identify factors that contribute to different levels of detection across screens and demographic groups, and our findings provide a benchmark against which an organization can check employee self-reported symptom rates. We encourage employers to consider these factors but also to monitor results and change course if needed.

---

[2] As of November 2020, the CDC does not recommend school-based daily symptom screening for students, but does encourage home-based symptom screening (CDC, 2020d). Customer screening may be required by law, as in the case of New York City indoor dining, or voluntarily adopted by employers.

For example, if case rates are surging in a firm's area, or if the firm has known cases, firms may wish to use a screen with a higher positivity rate at the risk of more and uneven false positives. Our findings are also important for policymakers seeking to develop supports that will enable symptomatic workers to stay home and obtain medical testing for COVID. Specifically, we find even during a period of relatively low COVID prevalence, large shares of workers experience symptoms that may merit medical follow-up to slow community spread. Our results also suggest that some populations that report symptoms less often may benefit from encouragement of symptom self-monitoring.

An important caveat is that our analysis cannot determine whether workplace screens ultimately reduce COVID-19 spread as we do not observe actual COVID infection rates nor do we have data on actual employer practices. However, our descriptive analysis provides a foundation for potential future evaluations of the efficacy of various screens in a large-scale experimental setting. A public health literature finds that self-reported symptoms can accurately proxy objective illness measures (Buckley & Conine, 1996; Chen et al., 2012; Jeong et al., 2010). Hence our analysis is informative about how these proxies function in a representative work-force sample. Related, since the CIS is not a high stakes survey and respondents have no incentives to misreport information, the prevalence rates we observe may be better proxies of true underlying conditions.

This paper relates to several literatures at the intersection of public health and economics. First, a large public health literature concludes non-pharmaceutical interventions (NPI) mitigate infection spread when vaccination and antiviral medications are not available (see, for example, Ferguson et al., 2006; Gostin, 2006; Kelso et al., 2009; Qualls et al., 2017; WHO, 2019). Recent work explicitly considers trade-offs between virus spread and economic activity, concluding that non-economic, non-pharmaceutical interventions—such as social distancing, increased testing, and self-quarantine measures, as well as restrictions targeted to the highest-risk groups—can limit the contagion effect without curtailing economic activity as severely as full lockdowns (Acemoglu et al., 2020; Baqaee et al., 2020). Our analysis adds to this literature by closely investigating how a representative labor force sample interacts with one particular NPI—workplace symptom screens—in order to quantify how many and what types of workers are likely to be affected.

Second, a rich economics literature considers how employers may identify individuals by relying on imperfect signals of worker characteristics, including health risks (Autor & Scarborough, 2008; Bertrand & Mullainathan, 2004; Craigie, 2020; Doleac & Hansen, 2020; Wozniak, 2015). Our paper is the first study to empirically investigate the incidence of similar signals in the COVID pandemic. Analyses of concrete testing and screening approaches in the workplace during a pandemic appear to be very few, with the notable exception of Augenblick et al. (2020) who investigate the properties of group testing using simulations. Although our paper is related to the literature on worker screening, there are some important differences. First, we focus on screening for a temporary characteristic that may have negligible impacts on an individual's long-term productivity but may substantially affect a firm's productivity via spillovers to other workers. In contrast, much of the screening literature focuses on permanent individual productivity. Second, levels—the share of workers an employer might expect to screen out on a daily or weekly basis—are a major quantity of interest in our analysis; by contrast, much of the screening literature examines relative impacts of screens across groups of workers, assuming that firms can choose among many "screened-in" workers but that screens will affect the prevalence of different workers in that group.

Third, this paper broadly relates to a literature examining how workplace policies affect whether workers stay home from work when experiencing symptoms of illness. Much of the existing work focuses on the availability of paid sick leave, and finds that employees with paid leave are more likely to take time off of work when

ill and that these policies can reduce the spread of communicable diseases (Abay, Rosa, & Pana-Cryan, 2017; DeRigne, Stoddard-Dare, & Quinn, 2016; Kumar et al., 2012; Pichler, Wen, & Ziebarth, 2020; Piper et al., 2017). We find that a large share of the workforce was experiencing symptoms of potentially contagious illness on a weekly basis, providing both an explanation for the effectiveness of paid leave policies, and suggesting how paid leave may complement workplace symptom screening for inducing potentially contagious individuals to stay home from work.

The remainder of this paper is organized as follows. The next section outlines a framework for how employers may weigh the trade-offs associated with any screening device that relies heavily on self-reported information. The third section describes the COVID Impact Survey and how self-reported symptoms may be used to identify those with COVID-related symptoms. The fourth section presents our main results on the prevalence of fever- and COVID-related symptoms among CIS respondents in the labor force. The fifth section concludes.

FRAMEWORK: SCREENING OUTCOMES AND INFECTION INFORMATION

A screen translates a set of input measures into a "high-risk" flag that triggers a consequence, such as barring an employee from entering the workplace that day or requiring additional, more-expensive screening. Specifically, a screen takes the following form:

$$d^i = \begin{cases} 1 \; if \; f(S^i) \geq t \\ 0 \; otherwise \end{cases}, \tag{1}$$

where screen, d^i, is an indicator that takes the value one or zero for whether worker i is indicated as high risk of being infectious. Indicator screens can include a range of potential inputs. We denote the complete set of possible input measures as S. $f(.)$ aggregates the particular inputs, possibly giving zero weight to some. A worker screens positive as high-risk if the output of this function exceeds a given threshold t. While S includes all COVID symptoms and additional risky behaviors, such as attending large group gatherings, a symptom-only screen would only use self-reported COVID symptoms as its inputs. An employer could then set the threshold at "1" for the number of affirmatively reported symptoms in order to flag anyone with any reported symptom as high risk.

Workplaces adopt health screens to prevent the spread of COVID-19, but they cannot observe each worker's current potential infectiousness.[3] Assume each day, an employee i has a true potential infectiousness $g_i^* \in \{0, 1\}$.[4] Screens then provide the following information on infection probability:

$g_i^* = \ldots$		1	0
$d^i = \ldots$	1	tp	fp
	0	fn	tn

[3] This framework is readily adapted to other situations where screens may be used, including businesses screening customers and schools screening students.
[4] Although individuals may vary in their contagiousness, both over the course of their infection and across infected people, this simplification aligns with medical diagnosis of COVID infection as either present or absent.

Accuracy is the rate at which a screen correctly identifies infectious and non-infectious workers, $[TP + TN]/[TP + TN + FP + FN]$, where uppercase indicates totals in each category for a tested population.[5] Errors can arise because of fundamental issues with the screen or human behavior. Workers may misreport their own health intentionally if they are able and face incentives to do so, but even absent manipulation, each symptom is a noisy signal of true infectiousness.

In a more complete model, one might specify a decisionmaker's objective function and seek an optimal screen. The screening technology can be conceptualized as an investment, with costs incurred in the short term when potentially contagious employees are not at work and, therefore, not producing goods and services, but with benefits realized in the longer term through lower workplace and community virus spread. In such a model, testing errors in either direction present costs to workers, organizations, and society. First, a screen with a high false negative rate may appear beneficial in the short term for the worker and the organization because an asymptomatic, infectious worker may be productive at work. However, in the medium term, these workers accelerate viral spread, resulting in lower economic activity and higher infection rates later. Secondly, a screen with a high false positive rate results in too many healthy people sent home and lost short-term productivity for workers, organizations, and society.

In practice, employers are proceeding with screens that they have chosen from a limited set of feasible options, taking two broad forms—medical and non-medical testing. Medical screening uses either diagnostic tests for current infection or, less frequently, blood tests for a past infection, which may indicate longer-run immunity to COVID-19.[6] While medical testing is often described as the ideal measure of an individual's infection potential, widespread testing is currently prohibitively expensive for most employers. Currently, rapid medical testing also varies widely in quality and may not be obviously superior to symptom screening (Pradhan, 2020).[7]

Our empirical analysis focuses on non-medical screens. Non-medical screening can include at least three strategies. First, employers may conduct regular or daily temperature checks before employees start work. If an employee has a temperature above a threshold, he or she is not able to work on site that day. Second, employers may require employees to regularly report their own body temperatures, fever symptoms, or COVID symptoms before arriving at work, and may condition an employee's ability to work on these self-reported responses. Third, employers may require employees to report high-risk behaviors, such as travel or attending large gatherings. These general approaches can be used in isolation or combination, and they may vary in how often and under what conditions employees report this information to their employers. Our analysis focuses on temperature checks and mandated self-reported health screens as employers frequently contemplate these strategies and the CIS data are well-suited to approximate these measures.

MEASURING WORKER HEALTH USING THE COVID IMPACT SURVEY

The COVID Impact Survey (Wozniak et al., 2020) was developed as a prototype of a survey tool that could assess many well-being measures at high frequency. Importantly for this paper, the CIS tracks individuals' physical symptoms, COVID exposure, mitigation behaviors, and labor market involvement. The survey instrument,

[5] "Accuracy" is the term in the diagnostic medical literature. Other relevant concepts are "sensitivity," the share of true positives among those testing positive, and "specificity," the share of true negatives among those testing negative (Baratloo et al., 2015).
[6] Requiring immunity to infection as a condition of work is currently illegal in the U.S. (EEOC, 2020).
[7] This is particularly true for the rapid tests most likely to be used to screen workers for attendance.

administered by the National Opinion Research Center (NORC), is described in greater detail in Wozniak (2020) and the online documentation.[8]

During each of three weeklong waves from mid-April through the first week of June 2020, CIS respondents were recruited through two channels. About 2,000 respondents were recruited through NORC's AmeriSpeak Panel, a nationally representative, address-based sample frame. An additional 6,500 respondents were recruited from address-based oversamples of 18 subnational areas ("places") by postcard invitation for a total of about 7,500 to 8,500 respondents for each cross-section.[9]

The CIS includes a battery of health screening questions chosen as the most promising instruments for wide-scale use. These inquire about symptoms commonly checked through health screening websites and apps, and they parallel questions employers may use as screening devices. Two separate checklists elicit self-reported fever- and COVID-related symptoms (CDC, 2020a). In addition, about half of all respondents provide a thermometer-based temperature reading.

We develop seven screens for COVID infection risk from this broad set of questions. The screens differ according to whether they use self-reported symptoms or a thermometer-measured temperature, whether they only ask about fever symptoms or general COVID symptoms (including COVID-related fever symptoms), and whether they derive from a short or long symptom checklist.[10] The short symptom list asks about three symptoms of fever, derived from emergency department fever screening approaches, while the long checklist asks about 17 temporary poor health items from various symptom checklists circulating in late spring 2020. Table A1 in the Appendix (which follows this article) defines the seven screens we examine, each defined as an indicator variable with 1 indicating a positive, higher-risk result.

The CIS also contains a battery of complementary questions about health, employment, protective and social behavior, and demographics. Questions on employment and reasons for non-work identify active labor force participants. Health-related questions in the CIS allow us to identify individuals with at least two factors that put them at risk for a severe COVID infection, such as asthma, diabetes, or hypertension. We also use information on whether there had ever been COVID diagnosis in the household, and whether the worker has been close to a person who died from COVID or respiratory illness since March 1, 2020—both of which are unavailable in other major surveys.[11] Our main analyses restrict the sample to labor force participants from the national panel and provide supplemental analysis for the subnational sample. Table A2 contains descriptive statistics from our primary analytic sample. All estimates are produced using the appropriate CIS sample weights.

---

[8] Full documentation is available at https://www.covid-impact.org/results.

[9] Geographic areas and the national AmeriSpeaks panel were re-sampled each CIS wave, creating repeated cross-sections. Places include 10 states (CA, CO, FL, LA, MN, MO, MT, NY, OR, and TX) and eight metropolitan statistical areas (Atlanta, Baltimore, Birmingham, Chicago, Cleveland, Columbus, Phoenix, and Pittsburgh). Wave 3 of the CIS was administered during the week of May 30 through June 8, 2020, coinciding with widespread protests against policy brutality in the U.S. NORC detected no compositional change in respondents, but the overall level of response to the mail recruitment for the subnational samples was lower than in Waves 1 and 2. Wave 3 contains a total of 7,505 responses, with 2,047 from the AmeriSpeaks panel and 5,458 in the subnational samples.

[10] There are many potential COVID symptoms, not all of which are present in all cases (Parker-Pope, 2020; Sudre et al., 2020). The symptoms included in our COVID checklist include some of the most common symptoms. However, the broad range of potential COVID symptoms makes it difficult to identify true "placebo" symptoms that should not be associated with our screens.

[11] The Delphi Research Group at Carnegie Mellon University and Facebook have partnered for another large survey effort (CMU-FB) that administers a COVID symptom checker to large numbers of Facebook users daily to estimate COVID spread. The CMU-FB survey is essentially one question from the physical health module of the CIS, and the representativeness of this sample is unknown, even after adjustments. By contrast, the CIS has smaller samples but richer data connected to the same surveillance questions.

RESULTS: INCIDENCE OF POSITIVE RESULTS ON COMMON SYMPTOM SCREENS

In this section, we estimate the incidence of positive screens across the alternative measures, assess each screen's informational value, and examine whether positive results are associated with other worker characteristics.

Prevalence of Self-Reported Symptoms in the Population

We first find that a substantial share of the workforce would screen positive under any of the screens described above, though the shares vary widely across screens. For example, 4 percent of respondents report currently having a temperature of 99 degrees F or higher based on the contemporaneous (current day only) thermometer reading (Table 1, row 1). All other symptom questions refer to a seven-day reference period, rather than only the current day, so levels on the other measures are not directly comparable to the temperature screen, and daily rates are likely lower than those reported here.[12] With this in mind, most respondents report experiencing at least one COVID-related symptom in the previous week, and 10 to 35 percent report experiencing at least one fever-related symptom over the same period on the short (fever symptom only) checklist.[13] Across measures, positive rates are considerably lower when screens require two or more symptoms for a positive result. For instance, whereas 51 percent report having at least one COVID symptom, only 26 percent report at least two. Also, rates of reported fever symptoms are much lower on the short fever-specific checklist than the long 17-item checklist. Finally, COVID symptoms are much more prevalent than fever-only symptoms. We examine what these patterns might mean for the informational content of different screens in the next subsection.

The CIS responses come from a low-stakes setting where respondents face no adverse consequences for reporting symptoms. Moving to a high-stakes, employer-administered survey may affect these patterns. For example, workers may under-report symptoms in order to avoid being temporarily barred from work. On the other hand, workplace consequences may heighten attention to health conditions and lead to more accurate self-reporting. The first factor is likely more important than the second, but both are possibly relevant.[14]

The remaining rows of Table 1 show positive rates are fairly level across weeks, consistent with a relatively flat number of new COVID cases during the CIS reporting period. Across groups defined by race and ethnicity, income, and age, Table 1 shows meaningful differences in positive screens rates. For example, rates of multiple COVID symptoms are lower for Black and Hispanic workers relative to White workers, for those in higher earning households relative to lower earners, and for older workers relative to younger. Similar group patterns are present in some, but not all, other screens. We explore this point in detail below.

These results suggest that symptom screening will detect a substantial share of the workforce at any point in time. These shares are roughly consistent with rates found in two other surveys: The Census Household Pulse Survey (HHPS) and Carnegie Mellon University's Delphi Project (CMU-DP). First, about 1 percent of HHPS

---

[12] To get a sense of potential daily rates, if symptoms resolve within one day for all respondents, a lower bound on the daily incidence would be one-seventh of the rates displayed in Table 1, or 0.4 to 7.4 percent.
[13] These positive rates are generally lower than the 20 to 30 percent positive rate for influenza-type-illness at the peak of flu season (CDC, 2020b).
[14] NORC examined all data for response patterns indicating inattention (e.g., checking all affirmative responses on a symptom checker or many skipped questions) and found few respondents exhibiting this behavior.

Table 1. Share of labor force flagged by workplace screens.

	(1) Therm (99)	(2) Fever short	(3) Fever	(4) Fever short 2+	(5) Fever 2+	(6) COVID	(7) COVID 2+	(8) COVID in home	(9) Family, friend COVID death
Weeks 1–3	0.038	0.104	0.346	0.025	0.093	0.512	0.260	0.017	0.051
Survey week									
Week 1	0.040	0.130	0.363	0.030	0.092	0.522	0.274	0.010	0.038
Week 2	0.035	0.108	0.343	0.028	0.103	0.509	0.247	0.017	0.061
Week 3	0.039	0.072	0.331	0.017	0.086	0.504	0.259	0.024	0.054
Race/ethnicity									
White	0.037	0.095	0.362	0.026	0.096	0.524	0.278	0.011	0.035
Black	0.037	0.111	0.282	0.026	0.076	0.450	0.209	0.017	0.124
Hispanic	0.050	0.119	0.328	0.024	0.072	0.511	0.238	0.030	0.072
Other	0.031	0.118	0.357	0.024	0.142	0.508	0.255	0.027	0.033
Household income									
<$40,000	0.030	0.121	0.379	0.033	0.111	0.552	0.295	0.020	0.062
$40,000-75,000	0.034	0.110	0.362	0.027	0.101	0.533	0.288	0.017	0.052
$75,000+	0.046	0.085	0.307	0.017	0.074	0.463	0.211	0.015	0.043
Age									
18-29	0.046	0.145	0.436	0.033	0.105	0.575	0.314	0.014	0.032
30-44	0.053	0.101	0.313	0.029	0.085	0.504	0.249	0.023	0.055
45-59	0.020	0.086	0.305	0.018	0.094	0.483	0.231	0.017	0.061
60+	0.026	0.065	0.338	0.015	0.092	0.466	0.244	0.007	0.063
N, Wk 1–3	2,442	4,377	4,301	4,377	4,372	4,375	4,375	4,309	4,274

Notes: Source: CIS national sample, all waves (Week 1 occurred April 20 to 26, 2020; Week 2 occurred May 4 to 10, 2020; Week 3 occurred June 1 to 8, 2020). Sample is restricted to respondents in the labor force.

Not at work due to coronavirus or possible infection
Percent of Labor Force
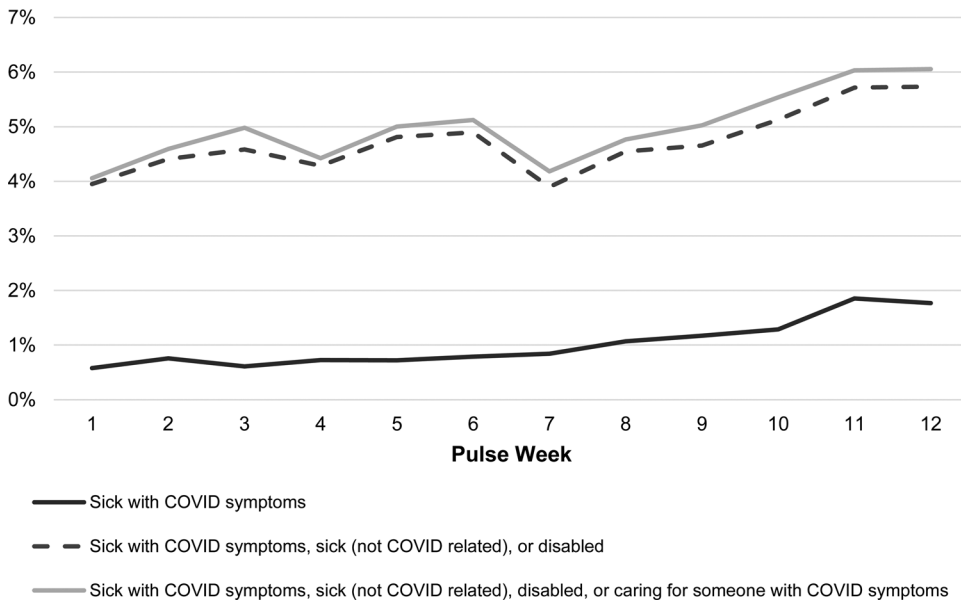
Notes: Census Household Pulse Survey data, weeks 1 to 12, spanning April 23 through July 21, 2020. Labor force participation is approximated from "reasons not at work last week."

Figure 1. Not at Work Due to Coronavirus or Possible Infection.

respondents were not at work in the past seven days because they were ill or symptomatic for COVID for the period overlapping with the CIS collection period, and this share rose to 2 percent later in summer 2020 as the pandemic spread, reasonably similar to our estimate that 0.4 to 7.4 percent of workers could screen positive on any given day (Figure 1).[15] Second, the CMU-DP daily symptom survey of Facebook users shows that between 0 and 5 percent of respondents report COVID-like symptoms in any 24-hour period; the mean is 0.52 percent (tabulations available upon request) with similar shares reporting flu-like symptoms. Although not limited to labor force participants, the CMU-DP data produce symptom rates similar to our estimates of daily shares of workers screening positive.

Consistency Across Screens

Given the variation in positive rates across screens in Table 1, it is reasonable to ask whether these screens are reliable. While we cannot correlate screens with medical testing for COVID-19, this subsection assesses the potential informational content of the screens in several ways.

---

[15] While the HHPS rates are lower than those reporting symptoms on the CIS screens, these patterns may predominantly include respondents staying home for an entire week or not capture individuals with symptoms who did not stay home at all or who worked from home.

First, we examine consistency across measures in Table 2. We are interested in whether different screens identify similar sets of workers. Each row of Table 2 reports the share of workers screening positive on each measure in columns, conditional on screening positive on that row's measure. The shares screening positive from the full (unconditional) labor force sample, from Table 1, are repeated at the top of the table for comparison. Scanning across the rows, the shares of workers screening positive on other measures, conditional on screening positive on one, varies widely—from 0.05 to 0.97. (Correlations denoted with a "†" are 1 by construction; we ignore those in this discussion.)

The conditional shares are a function of the overall unconditional level of workers reporting symptoms on different measures. Scanning down a column compares the unconditional share positive for that column's screen (the top number) to the share positive conditional on a positive result on a row's screen. For instance, in the first column, the first row indicates that 4 percent of workers report a temperature greater than or equal to 99 degrees F and the third row reports that conditional on reporting at least one symptom from the short fever screen, 13 percent of workers report a temperature of at least 99F. The conditional shares tend to be substantively larger than the unconditional shares across screens. The increased likelihood ranges from a 25 percent increase over the sample prevalence (respondents who report multiple COVID symptoms are 25 percent more likely to have a temperature of at least 99F) to a tenfold increase (respondents who report multiple fever symptoms on the short scale are 10 times more likely to have a temperature of at least 99F).

Taken together, Table 2 indicates that workers who screen positive under one measure are substantially more likely to screen positive on other measures, but the overlap is far from perfect. It is noteworthy that among those who screen positive on the single and multiple short fever symptom screens, the rate of at least one COVID symptom is 90 percent and 97 percent, respectively. This is higher than the rate of at least one COVID symptom (74 percent) among those with an elevated temperature. This suggests that self-reported symptoms may contain information beyond a single temperature screen. This is consistent with symptom screens falling into the category of "prudent questions" recently recommended by Anthony Fauci in place of temperature checks alone.[16]

We further assess consistency across our measures in Table A3, which shows that correlations are well below one across different screens, implying that screens identify different groups of workers. This is not surprising as each screen contains a noisy measure of true infection, and therefore overlap is unlikely to be perfect.

While we cannot examine the relationship between our screens and actual COVID infection, we can investigate their relationship to one medical measure: elevated temperature on a thermometer reading. In a simple univariate regression, all four self-reported fever symptoms indicators significantly predict a high temperature with t-statistics ranging from 3.6 to 6.8.[17] This is consistent with the evidence in Table 2, as well as previous work documenting a strong correlation between self-reported fever symptoms and objective fever measures among emergency department and hospital patients (Buckley & Conine, 1996; Chen et al., 2012; Jeong et al., 2010).

As an additional check, we examine rates of elevated infection potential among those who are positive on our symptoms screens in the CIS. We proxy for elevated infection potential in our CIS data using an indicator for very high temperature (100+)

---

[16] As reported in The Hill: https://thehill.com/changing-america/well-being/prevention-cures/512769-fauci-explains-why-temperature-checks-to-fight.
[17] These results, and those for 100+ temperatures, are available upon request.

Table 2. Share of positive screens, conditional on other positive screen.

	(1) Pr(Temp (therm) = 1)	(2) Pr(Fever (short) = 1)	(3) Pr(Fever = 1)	(4) Pr(Fever (short 2+) = 1)	(5) Pr(Fever 2+) = 1)	(6) Pr(COVID = 1)	(7) Pr(COVID 2+ = 1)
CIS mean	0.04	0.10	0.35	0.03	0.09	0.51	0.26
Temp (therm) = 1	1.00	0.35	0.54	0.25	0.18	0.74	0.37
Fever (short) = 1	0.13	1.00	0.74	0.24	0.34	0.90	0.65
Fever = 1	0.06	0.22†	1.00	0.06	0.28	1.00†	0.65
Fever (short 2+) = 1	0.40	1.00†	0.90	1.00	0.51	0.97	0.83
Fever (2+) = 1	0.07	0.38	1.00†	0.14	1.00	1.00†	1.00†
COVID = 1	0.06	0.18	0.68	0.05	0.18	1.00	0.51
COVID (2+) = 1	0.05	0.26	0.86	0.08	0.36	1.00†	1.00

Notes: † Indicates "equals one" by construction. Sample includes CIS labor force participants from the national survey pooling all weeks. Cells show the fraction of respondents reporting a positive screen on the column header, conditional on having a positive screen on the row measure.

and using recent diagnosed COVID infection.[18] As each of these proxies indicates elevated risk for COVID or communicable illness, supplementing these questions or incorporating them into a workforce screen could be an effective way to minimize disease spread. Across the screens, roughly 1 to 10 percent of our sample reports elevated infection potential by one of these proxies (tabulations available upon request). The median share is roughly 5 percent. This is close to recent rates of positive medical tests nationwide, as reported by the Harvard Global Health Initiative.[19] Although reasons for obtaining a COVID test are not systematically collected, it is likely that most Americans getting a medical test have reason to suspect they may be ill, either because of their own symptoms or suspected exposure.[20] Thus, on this dimension, CIS respondents who report symptoms are roughly comparable to the population getting tested for COVID.

We can also use elevated temperature to explore false positives in other symptom screens. Screens with lower thresholds will tend to have higher false positive rates, flagging noninfectious people. For instance, consider deciding between thresholds of at least one or at least two reported symptoms using the short fever screen. The two-symptom threshold would pass a larger share of people as negative (97 percent) than the one-symptom threshold (90 percent). Of those with a positive flag with the one-symptom threshold, Table 2 indicates 13 percent have a high temperature (87 percent false positive rate). A positive flag on the two-symptom threshold is associated with a 40 percent chance of having a high temperature, or a 60 percent false positive rate. Hence, a larger share of the people flagged by the two-symptom threshold will have high temperature than those flagged by the one-symptom threshold, so the latter has a higher false positive rate; but we cannot say what these rates are with respect to true COVID infection.

In Table A4, we show that all symptoms that comprise each screen are fairly equally reported, and no particular symptom disproportionately accounts for positive rates across measures. Consequently, screens that inquire about more symptoms tend to have higher positive rates than those that depend on fewer symptoms, and screens that ask about different sets of symptoms will identify different workers, even if the number of questions is the same. If risk tolerances for false negatives increase when hospital and emergency resources fall, employers and policymakers may wish to use a stricter (any one-symptom indicator or a screen that asks about many symptoms) screen when local caseloads are increasing, and rely on a more lenient screen (asking about fewer symptoms or requiring more positive symptoms to flag individuals) when cases are not surging.

Design effects are another potential source of divergence across different screens. We define design effects to be differences in responses due to the way questions are worded or where they are asked in the survey. Recall that the long symptom checklist is used to construct both the self-reported fever (long) screen and the COVID symptom screens. This checklist appears early in the survey and asks about 17 possible temporary illness symptoms, some of which are not associated with COVID or fever, with the order of items being randomized across subjects to avoid order effects. The self-reported fever symptoms (short) screen is constructed from a separate, three-item fever-specific checklist. This is asked later in the survey, and items are not randomized.

---

[18] The CIS asked respondents if a medical provider had ever diagnosed them or someone in their household with COVID-19, but since the CIS was administered between April and early June of 2020, any infections were relatively recent.

[19] See https://globalepidemics.org/testing-targets/. Accessed on 8/21/2020.

[20] Regular testing of asymptomatic individuals is still rare in the U.S.

Table 1 shows that positive screen rates are much lower for self-reported fever using the short checklist compared to the long checklist.[21] This pattern also applies to the one symptom that appears on both checklists, the presence of chills.[22] These differences suggest that survey design may influence rates of symptom reporting. Differences in responses on the two sets of fever items could be due to differences across early and later survey questions, across similar questions worded differently, and across long, general checklists versus short, focused checklists.

We also explore whether the rate of self-reported symptoms changes over time. Time effects could be the result of both spurious and non-spurious factors. One possible spurious factor is salience: as individuals become more aware of symptoms associated with COVID and personal health, their symptom reporting may change. On the other hand, non-spurious time effects would arise if the screens are informative of underlying population infection rates, and if these rates are changing over time. The second set of rows in Table 1 show that week-by-week symptom rates are fairly stable for most screens, even as the share of respondents who have had a diagnosis of COVID in the household or who know someone who has died of the disease increase. This may be consistent with our data period, in which infection rates were relatively stable but cumulative cases were rising. However, the screens based on the short fever checklist are an exception; these show declining shares over time. We investigate the role of group and time effects in response rates in more detail in the next subsection.

Group Effects, Time Effects, and the Role of Underlying Health Conditions

Federal law prohibits employers from using workplace tools that have disparate impact (policies that appear to be group-neutral but disproportionately affect a protected group) across workers by race, ethnicity, gender, age, and health status.[23] The health and employment patterns seen over the course of the pandemic, in combination with the existing legal environment, make it crucial to examine whether screening measures disproportionately affect certain types of workers.

Disparate detection under symptom screening could lead to two sources of harm in the COVID environment. First, workers may be screened out of work for a period of time. If income supports and employment protection provide inadequate insurance, these workers will lose income and perhaps employment. Screens that systematically screen out particular groups at higher rates, regardless of underlying infection, will then generate disparate economic harm for these groups. On the other hand, screens that systematically miss symptomatic individuals in one group may expose workers in that group to a higher risk of infection. This health harm would be disproportionate if workers tend to be concentrated in particular occupations or

---

[21] Half our subjects take their own temperature before answering the short checklist, but we do not find substantively different responses to questions related to self-reported fever by whether respondents answered the thermometer-based question. It is therefore unlikely that lower positive rates on the short fever screen result from knowing one's actual temperature. Correlations across other screens are similar for respondents who do and do not take their own temperature. This suggests that thermometer screening may contain somewhat different information from self-reported symptoms. A related concern is attention effects. NORC did not find evidence consistent with low attention, such as checking "yes" to all items or many skipped questions. Our main sample is also from the standing survey panel, in which respondents are presumably accustomed to taking surveys. We therefore discount traditional inattention effects on later survey questions as an explanation for this difference.
[22] In additional analysis, we find a correlation of about 0.2 between the chills item responses on the long and short symptom checklists. The correlation is stable across demographic groups. Overall response rates are similar on the two checklists, so lower rates of symptom reporting on the later list is not due to item non-response.
[23] Health status includes disability and genetic information, which may include immunity status.

establishments with high levels of within-group interaction, such as meat-packing plants.

There are reasons to be concerned about disparate harms. Coronavirus case rates are higher among Black and Hispanic Americans than non-Hispanic Whites (CDC, 2020c). In addition, between November 2019 and 2020, employment rates among Black and Hispanic Americans fell by 1.4 percentage points more than White Americans. This raises the question of whether these groups would face higher rates of positive symptom screens under widespread workplace screening, potentially exacerbating the unequal employment consequences of COVID-19.

In addition to differences in underlying infection rates, groups may differ in how they perceive or report symptoms, leading to group effects in the level of positive screens that may be independent of variation in infection rates. The medical literature on cross-race differences in symptom reporting finds generally mixed evidence on this, with some studies documenting substantial cross-race differences in reporting and others finding no differences.[24] Hence, it is difficult to say whether one should expect group differences in symptom reporting a priori.

In the CIS data, Table 1 suggests that self-reported symptoms do not always move consistently with self-reported experiences with coronavirus or caseload data. Consistent with caseload data, Black and Hispanic respondents are more likely than non-Hispanic Whites to report having a confirmed COVID case in their household, as are lower-income households. At odds with confirmed caseload patterns, however, Black and Hispanic respondents are less likely than White respondents to report experiencing any COVID-related symptom in the past week. Younger people are more likely to report experiencing a COVID symptom but are not more likely to have a confirmed case in their household and are less likely to know someone who has died of the disease. These patterns suggest there may be group-specific level effects in symptom reporting independent of a group's COVID infection intensity.

To further investigate potential group reporting effects, we regress our screen indicators on a broad set of individual characteristics. Specifically, we regress our indicators on race, ethnicity, gender, and age. We then add controls for marital status, presence of children, income, and population density in the county of residence. Finally, we add covariates reflecting current health status, including indicators for a COVID diagnosis in the household or experience with a COVID death, as well as measures of longer-term health, including our indicator for 2+ risk factors for severe COVID and an indicator for poor or fair overall health. We also include survey week dummy variables to allow for time effects in reporting.

Table 3 reports results from our most comprehensive specification, since results are robust to the set of covariates included.[25] Overall, we find evidence that groups differ in their rates of measured and reported symptoms. We begin by examining whether positive screens are associated with race and ethnicity, gender, or age. We find that race and ethnicity do not predict whether someone has a temperature of at least 99 degrees. Women and younger workers are more likely to report a high temperature, but only when income, household structure, and urbanity are excluded (column 1 of Table 3).

The next four columns of Table 3 show results for self-reported fever using our four versions of this screen. Protected demographics are not predictive of screens that require at least one symptom on either the short or long fever checklists (columns 2

---

[24] Some studies of self-reported symptom reporting find lower reported symptom burden among Black respondents, although this is not present in all studies of demographic differences in symptom reporting (Schnall et al., 2018). This seems to depend in part on the symptoms under study, e.g., HIV symptoms versus asthma symptoms, and possibly also samples.

[25] In additional results, patterns are robust to including week dummies to assess time effects in reporting.

Table 3. Characteristics of respondents with positive screens.

	(1) Therm (99)	(2) Fever (short)	(3) Fever	(4) Fever (short 2+)	(5) Fever 2+	(6) COVID	(7) COVID 2+
Panel a: Race/ethnicity							
Black, non-Hispanic	0.01 (0.02)	−0.01 (0.02)	−0.01 (0.01)	−0.14*** (0.03)	−0.05** (0.02)	−0.13*** (0.04)	−0.13*** (0.03)
Hispanic	0.01 (0.01)	−0.01 (0.02)	−0.01 (0.01)	−0.06* (0.03)	−0.04** (0.02)	−0.04 (0.03)	−0.07** (0.03)
Non-Black, non-Hispanic	−0.01 (0.02)	0.01 (0.03)	−0.01 (0.01)	−0.01 (0.04)	0.04 (0.03)	−0.01 (0.04)	−0.03 (0.04)
Panel b: Gender							
Female	0.00 (0.01)	0.04*** (0.01)	0.00 (0.01)	0.06*** (0.02)	0.03** (0.01)	0.08*** (0.02)	0.05** (0.02)
Panel c: Age							
Age 30–44	0.01 (0.01)	−0.04** (0.02)	−0.01 (0.01)	−0.12*** (0.03)	−0.02 (0.02)	−0.07** (0.03)	−0.06** (0.03)
Age 45–59	−0.02 (0.01)	−0.05*** (0.02)	−0.01 (0.01)	−0.15*** (0.03)	0.00 (0.02)	−0.11*** (0.03)	−0.09*** (0.03)
Age 60+	−0.02 (0.02)	−0.08*** (0.02)	−0.02 (0.01)	−0.14*** (0.04)	−0.01 (0.02)	−0.16*** (0.04)	−0.12*** (0.03)

Table 3. (Continued).

	(1) Therm (99)	(2) Fever (short)	(3) Fever	(4) Fever (short 2+)	(5) Fever 2+	(6) COVID	(7) COVID 2+
Panel d: Health							
HH with COVID	0.03	0.18***	0.04***	0.16**	0.14***	0.15**	0.24***
	(0.02)	(0.04)	(0.01)	(0.07)	(0.04)	(0.08)	(0.06)
Known death	−0.01	0.02	−0.01	0.03	0.01	0.03	0.07*
	(0.02)	(0.02)	(0.01)	(0.05)	(0.02)	(0.05)	(0.04)
2+ risk factors	0.01	0.04***	0.01**	0.14***	0.02	0.15***	0.12***
	(0.01)	(0.01)	(0.01)	(0.03)	(0.01)	(0.03)	(0.02)
Poor or fair health	0.00	0.06***	0.03***	0.15***	0.09***	0.18***	0.18***
	(0.01)	(0.02)	(0.01)	(0.03)	(0.02)	(0.04)	(0.03)
Panel e: HH income							
Income $40k-75k (2019)	0.01	0.01	0.00	0.00	0.00	0.00	0.01
	(0.01)	(0.02)	(0.01)	(0.03)	(0.02)	(0.03)	(0.03)
Income $75k+ (2019)	0.02*	−0.01	−0.01	−0.04	−0.04**	−0.06*	−0.06**
	(0.01)	(0.02)	(0.01)	(0.03)	(0.02)	(0.03)	(0.03)
Panel f: HH type							
Childless HH	0.01	−0.02	−0.01	−0.02	0.00	−0.02	0.00
	(0.01)	(0.02)	(0.01)	(0.03)	(0.02)	(0.03)	(0.03)
HH with children, all under age 5	0.02	−0.01	0.00	−0.20***	−0.03	−0.16***	−0.14***
	(0.02)	(0.03)	(0.01)	(0.05)	(0.03)	(0.03)	(0.03)
HH with children, some older than 5	0.02	0.02	0.00	−0.04	0.01	−0.03	−0.02
	(0.02)	(0.02)	(0.01)	(0.05)	(0.03)	(0.05)	(0.04)
Other HH composition	0.04***	0.00	0.00	−0.04	−0.01	−0.03	−0.05
	(0.01)	(0.02)	(0.01)	(0.03)	(0.02)	(0.04)	(0.03)
Panel g: Geographic location							
Suburban	−0.02	−0.01	−0.03***	0.04	0.00	0.00	0.05
	(0.02)	(0.02)	(0.01)	(0.04)	(0.02)	(0.04)	(0.04)
Urban	−0.01	0.01	−0.01	0.01	0.00	−0.03	0.02
	(0.01)	(0.02)	(0.01)	(0.04)	(0.02)	(0.04)	(0.03)
N	2,346	4,180	4,180	4,115	4,179	4,180	4,180

Notes: Table shows the marginal change in the likelihood of screening positive under each screen. All outcomes are from probit regressions, evaluated at the sample means. Sample is restricted to CIS participants from the national study, pooling all weeks. Robust standard errors are in parentheses. *** p < 0.01; ** p < 0.05; * p < 0.10.

and 3, respectively), but they do predict having at least two symptoms on both lists. Non-Hispanic Black and Hispanic workers are less likely than non-Hispanic White workers to report having at least two fever symptoms while members of other racial and ethnic groups are equally likely as Whites to flag positive on the short list but more likely on the long list (columns 4 and 5). Women are more likely to report two or more fever symptoms than men, especially on the short list. Similarly, young workers are more likely to report two or more fever symptoms than older workers on the short list.

Screens based on COVID-symptom lists, rather than fever symptoms, show similar results. Again, Black and Hispanic workers are less likely and women and younger workers more likely to screen positive. Higher-income households are less likely to screen positive, but these differences are only significant for fever and multiple short fever or COVID symptoms. This pattern could reflect reduced illness among those with a greater ability to work from home (Dingel & Nieman, 2020).

We next turn to the question of how underlying health status—temporary or permanent—relates to the likelihood of screening positive. Panel (d) of Table 3 reports coefficients on the health status covariates. The first of these is an indicator for ever having a COVID diagnosis in the household, another measure of elevated infection risk in the absence of COVID infection information for individuals.[26] Like a thermometer reading, a COVID diagnosis is a measure of infection risk that is independent of subjective symptom reporting, and employers could incorporate similar measures in workplace screens to flag individuals who might have high risk of being contagious. We also include a measure for a COVID death among friends and family, which may indicate the spread of COVID in one's close circle. Finally, we include measures of self-reported general health and for risk of severe COVID infection. While each of these measures may vary across the demographic groups in the preceding panels, results in those panels are not sensitive to their inclusion.

We find that workers with a COVID diagnosis in the home are significantly more likely to have positive screens across all measures. Point estimates, while large, are far below one, as expected since COVID does not always infect all members in a household or appear as a symptomatic case. In contrast, COVID mortality in one's close circle is not predictive of positive screens.

Workers with multiple high-risk conditions and those in fair or poor health represent an additional challenge for employers. In Table A1, we show that these groups are large: one-tenth of the workforce reports being in fair or poor health and almost one-fifth (17 percent) has at least two characteristics that place them at high-risk for severe COVID as defined by the CDC.[27] This is close to an estimate using 2018 data, that found about one-fourth of U.S. workers have at least one underlying health condition that makes the risks and consequences of COVID particularly severe (Kaiser Family Foundation, 2020). Our data indicate that they are more likely to screen positive on workplace screens, although the thermometer readings are not significant for any of these groups. Designing workplace policies that protect such workers in the face of their greater vulnerability and higher symptom screening rates will be a challenge.

Figure 2 repeats this analysis for each racial and ethnic, gender, and age subsample separately and plots the resulting coefficients on ever having a COVID diagnosis in the household. The plots show that a COVID diagnosis in the home elevates the

---

[26] This measure includes recovered cases, but we note that our sample period occurred relatively early in the pandemic, when most cases were relatively recent.

[27] See https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/people-with-medical-conditions.html?CDC_AA_refVal and https://3A/2F/2Fwww.cdc.gov/2Fcoronavirus/2F2019-ncov/2Fneed-extra-precautions/2Fgroups-at-higher-risk.html. Note: This page is frequently updated.
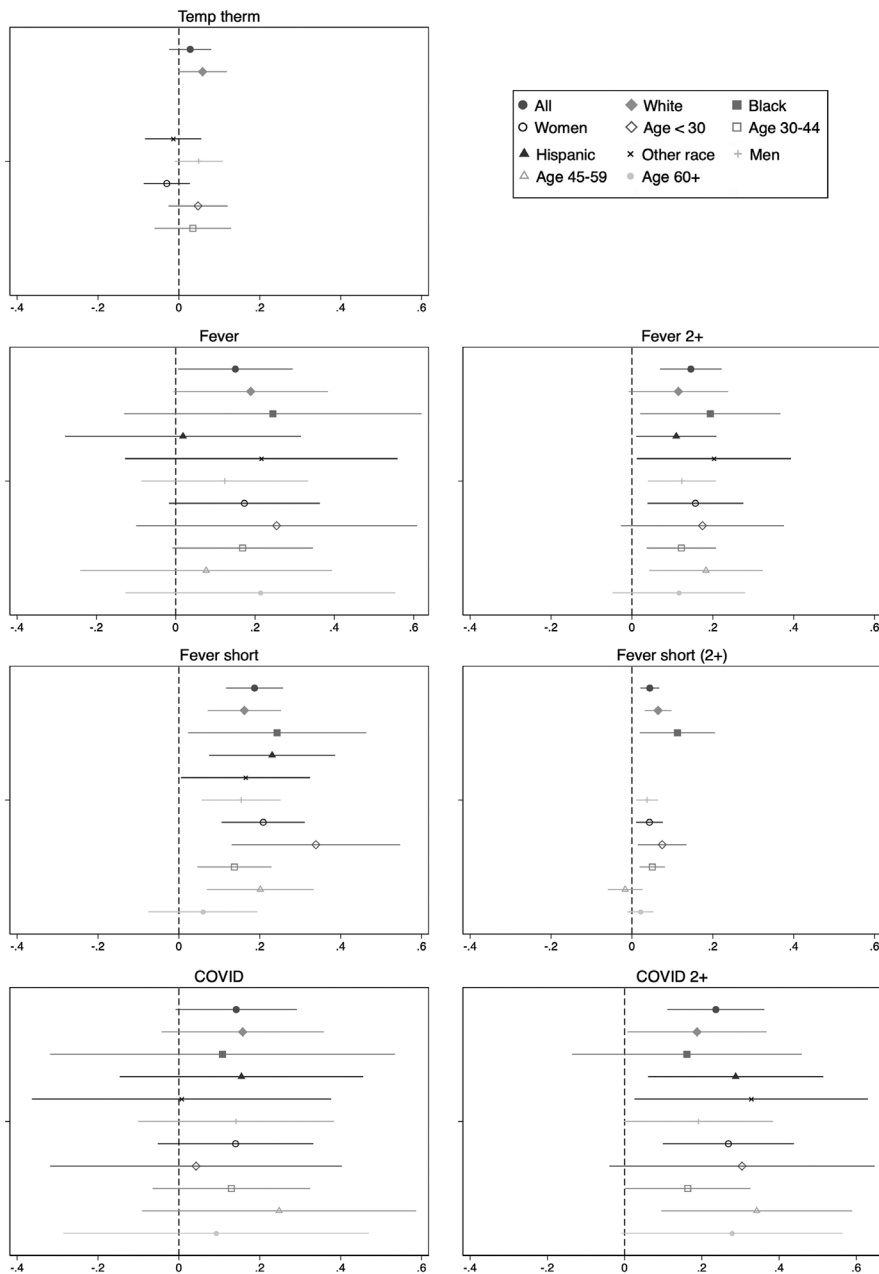
Source: CIS weeks 1 to 3. Sample is limited to labor market participants from the national sample. A missing point estimate indicates insufficient data to estimate for that subgroup.
Notes: Each figure plots the coefficients of having a COVID diagnosis in one's household on the probability of receiving a positive flag under that subfigure's screen by demographic group. All values are evaluated at sample means.

Figure 2. Effect of Household COVID Diagnosis on Positive Screen Rates by Worker Characteristics.

likelihood of a positive symptom screen across all screens and all demographic sub-samples. This suggests that the group effects identified in estimates using the pooled data (Table 3) are driven by underlying group differences in reporting rather than differential exposure to COVID.

Finally, we use the CIS geographic subsamples to explore whether local variation in COVID caseloads or death rates contributes to the group-level differences we observe. If there are group-specific-level effects in how screening questions are answered, so that Black and Hispanic respondents are less likely to report symptoms, controlling for cases may make the coefficients on racial and ethnic groups *more* negative as these groups have tended to have higher case rates. However, Table A5 shows the coefficients on race and ethnicity are stable when including information on per-capita caseloads. Turning to the coefficients on per-capita cases, Table A5 also shows a weak relationship between cases and screens. There are at least two plausible explanations for this pattern: first, testing capacity was still ramping up in the U.S. during the data collection period, or second, the relationship between symptom reports and infection rates is weak in small samples. Results using the current per capita death rate show similar patterns.

POLICY DISCUSSION AND CONCLUSION

The self-reported symptom and fever screens we examine are likely to be a central part of businesses' reopening plans for both legal and practical reasons. This paper provides the first empirical work exploring the incidence and properties of likely workplace screening devices. Our findings have implications for the use of health screens for other populations, including students and customers, as well as for potential future infectious disease outbreaks.

These tools present trade-offs and equity considerations. First, each screen we considered would likely identify a large share of the workforce as high-risk on any given day, limiting the number of people who might be allowed to attend work. Both the number of symptom inputs and the threshold matter: All else being equal, screens that inquire about more symptoms will screen out more workers than shorter screens, and screens that require multiple affirmative symptoms will screen out fewer workers than those requiring a single positive symptom.

Second, the survey design, the number of symptoms, and types of symptoms asked will screen out different individuals. Related to this, employers will likely have to contend with equity considerations, as demographic groups report symptoms at different rates, even after accounting for differential likely COVID exposure. Some groups—female, younger, and non-Hispanic White workers—report multiple symptoms at higher rates than male, older, Black, and Hispanic workers, even after accounting for differences in likely COVID exposure, income, and household structure.

Our evidence indirectly suggests that the value of these screens for identifying individuals with active COVID is likely to be low, particularly those with asymptomatic cases. Positive screen rates are much higher than actual infection rates, and local case rates are not predictive of positive screens. However, these findings do not necessarily imply that employers should do away with workplace health screens for at least three reasons. First, public health experts have advocated prioritizing high-frequency testing over a particular test's sensitivity (Larremore et al., 2020). Second, other research finds that when workers can stay home when ill, overall flu infection rates fall (Pichler, Wen, & Ziebarth, 2020). Hence, when symptomatic workers stay home, fewer other workers fall ill from COVID or other infectious diseases. Daily symptom screening may help encourage use of sick leave for any symptomatic workers and, thereby, contribute to reduced transmission of COVID. Finally, daily

health screens are likely to remain a widespread tool for containing COVID, particularly in the U.S. where other options are likely to remain limited for some time. While health officials have urged the public to get medically tested if they may have COVID, regular employer-based medical testing is currently prohibitive for most employers, particularly tests with a rapid turnaround time. As of mid-December, about two million COVID tests were administered each day, enough for less than 1.5 percent of the 160 million labor market participants, only five states had adequate testing and tracing capacity to implement large-scale "test and trace" programs (Johns Hopkins, 2020; Mohapatra, 2020; TestAndTrace, 2020). Therefore, as policymakers seek to manage the tension that human interactions power both the economy and virus spread, regular symptom screenings offer a means to isolate infectious individuals and encourage those with symptoms to get medically tested.

Applying our findings to the workplace setting raises additional considerations that employers and policymakers should consider. In the CIS survey setting, respondents have no incentives to misreport symptoms. In contrast, workplace screening policies may tie one's income, employment, and health insurance to one's responses, so respondents face incentives to answer strategically rather than truthfully. These concerns become particularly acute in a pandemic environment in which contagious individuals may be required to quarantine for several weeks. Weakening the link between individual responses and individual economic consequences reduces incentives to strategically misreport. This de-coupling could be achieved with complementary policy responses or screen design considerations. On the policy side, paid sick leave and employment protection for diagnostic absences promotes truthfulness by ensuring that a worker's economic security does not depend on reported symptoms. Evidence from this point comes from existing paid sick leave requirements, with studies finding paid leave increases the likelihood employees stay home from work when ill (DeRigne, Stoddard-Dare, & Quinn, 2016; Pichler, Wen, & Ziebarth, 2020). Employers could also consider group-level precautions that escalate if large shares of workers report symptoms, or could tie decisions to group average or pooled testing responses, rather than individual responses. While group-level strategies decouple consequences from individual reports, these types of screens are relatively blunt instruments that mis-assign more individuals to work (false negatives) or home (false positives).

Despite limitations that preclude us from recommending a particular screening approach as optimal, our analysis offers evidence to guide employers and policymakers. First, employers should expect substantial rates of affirmative symptom reporting. Infrequent symptom reporting could indicate non-truthful reporting. Second, some screens yield robust group differences (multiple symptoms, COVID symptoms, or targeted fever checklist), while others do not significantly differ by group (temperature checks or a single fever symptom from a long checklist). Finally, employers may use this information to inform a trade-off between screens that may have a higher false negative rate (e.g., temperature checks alone) but fewer disparities in detection, versus lower false negative rates (temperature checks plus COVID symptom screens) but more disparities depending on risk tolerance. If case rates are surging in a firm's area, or if the firm has known cases, it may make sense to screen more intensively at the risk of more and uneven false positives.

Regardless of the screens implemented, employers should monitor screening information to ensure reasonable and equitable treatment, particularly in the early stages of implementation. As Jones, Molitor, and Reif (2019) show, universal employee health programs may not always work as intended. Follow-up is particularly critical in the case of workplace COVID screening, since screens that do too much or too little are both damaging.

Absent universal, near-daily COVID testing or universal vaccination, employers will want to rely on measures that are predictive of likely COVID diagnosis. Regular symptom screening is a readily available technology for this. Our analysis shows that the design of these screens, including the number and types of questions, affects how many, and which, workers screen positive.

DISCLAIMER

The views expressed here are those of the authors and do not necessarily represent those of the authors' institutions or the Federal Reserve System.

KRISTA RUFFINI is a Visiting Scholar at the Federal Reserve Bank of Minneapolis and an Assistant Professor at the McCourt School of Public Policy at Georgetown University, 90 Hennepin Avenue, Minneapolis, MN 55401 (e-mail: k.j.ruffini@berkeley.edu).

AARON SOJOURNER is an Associate Professor at the Carlson School of Management at the University of Minnesota, 321 19th Avenue S., Minneapolis, MN 55409 (e-mail: asojourn@umn.edu).

ABIGAIL WOZNIAK is Senior Economist and Director of the Opportunity & Inclusive Growth Institute at the Federal Reserve Bank of Minneapolis, 90 Hennepin Avenue, Minneapolis, MN 55401 (e-mail: abigailwozniak@gmail.com).

ACKNOWLEDGMENTS

REFERENCES

Abay Asfaw, D., Rosa, R., & Pana-Cryan, R. (2017). Potential economic benefits of paid sick leave in reducing absenteeism related to the spread of influenza-like illness. Journal of Occupational and Environmental Medicine, 59, 822.

Acemoglu, D., Chernozhukov, V., Werning, I., & Whinston, M. (2020). Optimal targeted lockdowns in a multi-group SIR model. NBER Working Paper 27102. Cambridge, MA: National Bureau of Economic Research.

American Staffing Association. (2020). Most employees satisfied with Covid-19 Return-to-Work Plans. Retrieved August 2, 2020, from https://americanstaffing.net/posts/2020/07/16/covid-19-return-to-work-plans/.

Augenblick, N., Kolstad, J., Obermeyer, Z., & Wang, A. (2020). Group testing in a pandemic: The role of frequent testing, correlated risk, and machine learning. NBER Working Paper #27457. Cambridge, MA: National Bureau of Economic Research.

Autor, D., & Scarborough, D. (2008). Does job testing harm minority workers? Evidence from retail establishments. Quarterly Journal of Economics, 12, 219–277.

Baqaee, D., Farhi, E., Mina, M., & Stock, J. (2020). Policies for a second wave. Brookings Papers on Economic Activity. Washington, DC: Brookings Institution.

Baratloo, A., Hosseini, M., Negida, A., & El Ashal, G. (2015). Part 1: Simple definition and calculation of accuracy, sensitivity, and specificity. Emergency (Tehran), 3, 48–49.

Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. American Economic Review, 94, 991–1013.

Buckley, R., & Conine, M. (1996). Reliability of subjective fever in triage of adult patients. Annals of Emergency Medicine, 27, 693–695.

Center for Disease Control and Prevention (CDC). (2020). Symptoms of Coronavirus. Retrieved July 1, 2020, from https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html.

Center for Disease Control and Prevention (CDC). (2020). Influenza positive tests reported to CDC by U.S. clinical laboratories, 2019–2020 Season. Retrieved August 18, 2020, from https://www.cdc.gov/flu/weekly/weeklyarchives2019-2020/data/whoAllregt_cl32.html.

Centers for Disease Control and Prevention (CDC). (2020). COVID-19 hospitalization and death by race/ethnicity. Retrieved August 18, 2020, from https://www.cdc.gov/coronavirus/2019-ncov/covid-data/investigations-discovery/hospitalization-death-by-race-ethnicity.html.

Center for Disease Control and Prevention (CDC). (2020). Screening K-12 students for symptoms of COVID-19: Limitations and considerations. Retrieved December 9, 2020, from https://www.cdc.gov/coronavirus/2019-ncov/community/schools-childcare/symptom-screening.html.

Chen, S., Chen, Y., Chiang, W., Kung, H., King, C., Lai, M., … & Chang, S. (2012). Field performance of clinical case definitions for influenza screening during the 2009 pandemic. The American Journal of Emergency Medicine, 30, 1796–1803.

Craigie, T. (2020). Ban the Box, convictions, and public employment. Economic Inquiry, 58, 425–445.

DeRigne, L., Stoddard-Dare, P., & Quinn, L. (2016). Workers without paid sick leave less likely to take time off for illness or injury compared to those with paid sick leave. Health Affairs, 35, 520–527.

Dingel, J., & Neiman, B. (2020). How many jobs can be done at home? Journal of Public Economics, 189, (C).

Doleac, J., & Hansen, B. (2017). Moving to job opportunities? The effect of Ban the Box on the composition of cities. American Economic Review, 107, 556–559.

Ferguson, N., Cummings, D., Fraser, C., Cajka, J., Cooley, P., & Burke, D. (2006). Strategies for mitigating an influenza pandemic. Nature, 442, 448–452.

Gostin, L. (2006). Public health strategies for pandemic influenza: Ethics and the law. Journal of the American Medical Association, 295, 1700–1704.

Jeong, I., Lee, C., Kim, D., Chung, H., & Park, S. (2010). Mild form of 2009 H1N1 influenza infection detected by active surveillance: Implications for infection control. American Journal of Infection Control, 38, 482–485.

Johns Hopkins University of Medicine. (2020). Testing trends toolkit. Coronavirus Resource Center, Testing Hub. Retrieved July 25, 2020, from https://coronavirus.jhu.edu/testing/tracker/overview.

Jones, D., Molitor, D., & Reif, J. (2019). What do workplace wellness programs do? Evidence from the Illinois Workplace Wellness Study. Quarterly Journal of Economics, 13, 1747–1791.

Kaiser Family Foundation. (2020). Adults at higher risk of serious illness if infected with Coronavirus. Retrieved July 25, 2020, from https://www.kff.org/other/state-indicator/adults-at-higher-risk-of-serious-illness-if-infected-with-coronavirus/?currentTimeframe=0&sortModel=/7B/22colId/22/22Location/22/22sort/22/22asc/22/7D.

Kelso, J., Milne, G., & Kelly, H. (2009). Simulation suggests that rapid activation of social distancing can arrest epidemic development due to a novel strain of influenza. BMC Public Health, 9, 117.

Kumar, S., Quinn, S., Kim, K., Daniel, L., & Freimuth, V. (2012). The impact of workplace policies and other social factors on self-reported influenza-like illness incidence during the 2009 H1N1 pandemic. American Journal of Public Health, 102, 134140.

Larremore, D., Wilder, B., Lester, E., Shehata, S., Burke, J., Hay, J., … & Parker, R. (2020). Test sensitivity is secondary to frequency and turnaround time for COVID-19 surveillance. medRxiv.

Mohapatra, S. (2020). Why COVID-19 immunity passports may violate U.S. law. The Conversation, May 27, 2020.

Parker-Pope, T. (2020). The many symptoms of Covid-19. New York Times, August 5, 2020.

Phelan, A. (2020). COVID-19 immunity passports and vaccination certificates: Scientific, equitable, and legal challenges. Comment, The Lancet, 395, 1595–1598.

Pichler, S., Wen, K., & Ziebarth, N. (2020). Positive health externalities of mandating paid sick leave. Working paper, REPEC.

Piper, K., Youk, A., James, A., & Kumar, S. (2017). Paid sick days and stay-at-home behavior for influenza. PLoS One, 12, e0170698.

Pradhan, R. (2020). As problems grow with Abbott's Fast COVID Test, FDA standards are under fire. Kaiser Health News, June 22, 2020.

Qualls, N., Levitt, A., Kanade, N., Wright-Jegede, N., Dopson, S., Biggerstaff, M., Reed, C., & Uzicanin, A. (2017). Community mitigation guidelines to prevent pandemic influenza—United States, 2017. Morbidity and Mortality Weekly Report (MMWR), Recommendations and Reports, 66, 1–34.

Schnall, R., Siegel, K., Jia, H., Olender, S., & Hirshfield, S. (2018). Racial and socioeconomic disparities in the symptom profile of persons living with HIV. AIDS Care, 30, 774–783.

Sudre, C., Lee, K., Lochlainn, M., Varsavsky, T., Murray, B., Graham, M., … & Drew, D. (2020). Symptom clusters in Covid-19: A potential clinical prediction tool from the COVID Symptom study app. medRxiv.

Test and Trace. (2020). What U.S. States are ready to Test and Trace today? Retrieved July 27, 2020, from https://testandtrace.com/state-data/.

U.S. Bureau of Labor Statistics. (2020). Supplemental data measuring the effects of the coronavirus (COVID-19) pandemic on the labor market. Retrieved July 30, 2020, from https://www.bls.gov/cps/effects-of-the-coronavirus-covid-19-pandemic.htm.

U.S. Equal Employment Opportunity Commission (EEOC). (2020). What you should know about COVID-19 and the ADA, the Rehabilitation Act, and other EEO Laws. Technical assistance questions and answers. Retrieved June 17, 2020, from https://www.eeoc.gov/wysk/what-you-should-know-about-covid-19-and-ada-rehabilitation-act-and-other-eeo-laws?utm_content=&utm_medium=email&utm_name=&utm_source=govdelivery&utm_term=.

U.S. Occupational Health and Safety Administration (OHSA). (2020). Guidance on returning to work. OSHA, 4045-06 2020.

Volz, M., & Galewitz, P. (2020). As long waits for results render COVID tests "Useless," States seek workarounds. Kaiser Health News. Retrieved July 25, 2020, from https://khn.org/news/states-search-for-ways-to-deal-with-covid-19-testing-backlogs/.

World Health Organization (WHO). (2019). Non-pharmaceutical public health measures for mitigating the risk and impact of epidemic and pandemic influenza. Retrieved July 1, 2020, from https://www.who.int/influenza/publications/public_health_measures/publication/en/.

Wozniak, A. (2015). Discrimination and the effects of drug testing on Black employment. Review of Economics and Statistics, 93, 548–566.

Wozniak, A. (2020). Disparities and mitigation behavior during COVID-19. No. 32. Minneapolis, MN: Federal Reserve Bank of Minneapolis.

Wozniak, A., Willey, J., Benz, J., & Hart, N. (2020). COVID Impact Survey: Version 2. Chicago, IL: National Opinion Research Center, 2020.

APPENDIX

Table A1. Workplace screening measures in the COVID Impact Survey.

Screen	Self-report or measured	Measures included	# of questions to receive a positive screen
Therm (99)	Measured	Temperature of 99.0 F or higher using a thermometer.	1/1
Fever (short)	Self-reported	Any self-reported fever symptoms from a 3-item fever checklist.	1/3
Fever (short 2+)	Self-reported	2+ self-reported fever symptoms from a 3-item fever checklist.	2/3
Fever	Self-reported	Any self-reported fever symptoms from a 17-item health symptom checklist. 3 possible fever items.	1/3
Fever (2+)	Self-reported	2+ self-reported fever symptoms from a 17-item health symptom checklist. 3 possible fever items.	2/3
COVID	Self-reported	Any self-reported COVID symptoms from a 17-item health symptom checklist. 7 possible COVID items; includes 3 possible items in fever (long).	1/7
COVID (2+)	Self-reported	2+ self-reported COVID symptoms from a 17-item health symptom checklist. 7 possible COVID items; includes 3 possible items in fever (long).	2/7
2+ risk factors	Self-reported	2+ risk factors for COVID complications from multiple questions about underlying health conditions.	2/9

Notes: Table defines each screening measure we consider. See text for details. Table A4 lists COVID and fever symptoms from the long symptom checklist as well as all items from the short fever checklist. The full list of symptoms is available at https://static1.squarespace.com/static/5e8769b34812765cff8111f7/t/5e99d902ca4a0277b8b5fb51/1587140880354/COVID-19+Tracking+Survey+Questionnaire+041720.pdf.

Table A2. Descriptive statistics, main sample.

	Mean	Obs
Black, non-Hispanic	0.113	4,349
Hispanic	0.194	4,349
Non-Black, non-Hispanic	0.094	4,349
Female	0.492	4,377
Age 30–44	0.331	4,377
Age 45–59	0.281	4,377
Age 60+	0.127	4,377
Income $40k-75k (2019)	0.275	4,377
Income $75k+ (2019)	0.394	4,377
Childless HH	0.209	4,377
HH with children, all under age 5	0.055	4,377
HH with children, some older than 5	0.258	4,377
Other HH composition	0.175	4,377
Suburban	0.178	4,375
Urban	0.742	4,375
Self or HH member has received COVID diagnosis	0.017	4,309
Close friend or family member died of COVID, respiratory illness	0.051	4,274
ONET work from home occ share (Dingel & Nieman)	0.417	2,340
High work from home indicator, occupation level (Aaronson et al.)	0.459	2,340
2+ risk factors for COVID complications	0.167	4,361
Self-described fair/poor general health	0.118	4,374

Notes: Table shows demographic and health characteristics for our main analysis sample, CIS respondents from the national sample, all weeks. See text for details.

Table A3. Correlation across screens.

	Therm (99)	Fever (short)	Fever	Fever (short 2+)	Fever (2+)	COVID	COVID (2+)	Risk 2+	Fair/poor health
Therm (99)	1.00								
Fever (short)	0.14	1.00							
Fever	0.08	0.27	1.00						
Fever (short 2+)	0.26	0.45	0.18	1.00					
Fever (2+)	0.06	0.25	0.44	0.21	1.00				
COVID	0.09	0.26	0.71	0.14	0.31	1.00			
COVID (2+)	0.06	0.29	0.65	0.22	0.54	0.58	1.00		
Risk 2+	0.00	0.03	0.11	0.04	0.04	0.12	0.11	1.00	
Poor/fair health	0.01	0.08	0.10	0.07	0.11	0.10	0.13	0.21	1.00

Source: CIS national sample, all waves.
Notes: Sample is restricted to respondents in the labor force. Table shows the raw correlation across measures.

Table A4. Share of respondents reporting COVID and fever symptoms, item-by-item.

Have you experienced [symptom] in the past 7 days?	Share of respondents
Panel (a): COVID and fever items from 17-symptom (long) checklist	
Cough	0.143
Shortness of breath	0.115
Fever[†]	0.164
Chills[†]	0.153
Muscle/body aches[†]	0.132
Sore throat	0.140
Changed/loss smell	0.137
Panel (b): Symptoms from fever-specific (short) checklist	
Felt hot/feverish	0.036
Felt cold/had chills	0.059
Sweating more than usual	0.042

Notes: Table shows the share of respondents affirmatively reporting each symptom. Sample includes all CIS workforce participants from the national sample.

[†] Denotes symptom on fever (long) screens in Table A1.

Table A5. Implied marginal effects from probit of screen on demographics, regional data.

	(1) Therm	(2) Fever (short)	(3) Fever	(4) Fever (short 2+)	(5) Fever (2+)	(6) COVID (long)	(7) COVID (long 2+)
Per capita cases	−0.12	0.16	−0.06	0.62**	0.17	0.00	0.16
Race/ethnicity							
Black, non-Hispanic	0.00	−0.03*	−0.01	−0.08***	−0.03**	−0.11***	−0.09***
Hispanic	−0.02	−0.01	−0.01	−0.01	−0.04***	−0.05*	−0.05*
Non-Black, non-Hispanic	0.00	0.00	−0.01	0.01	−0.02*	−0.02	−0.01
Gender							
Female	−0.01	0.00	0.00	0.04**	0.02**	0.05**	0.05***
Age							
Age 30–44	−0.01	−0.01	0.00	−0.02	0.01	−0.03	−0.01
Age 45–59	−0.03**	−0.03**	0.00	−0.07**	−0.02	−0.08***	−0.07***
Age 60+	−0.08***	−0.04**	−0.01	−0.09***	−0.02	−0.11***	−0.08***

Notes: Table shows the marginal change in the likelihood of screening positive under each screen. All outcomes are from probit regressions, evaluated at the sample means. Additional variables (coefficients not reported) include household income bracket, household type, and county suburban or urban location. Sample is restricted to CIS participants from regionally representative cross-sections, pooling all weeks. ***p < 0.01; **p < 0.05; *p < 0.10.