

NBER WORKING PAPER SERIES

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

Krista J. Ruffini  
Aaron Sojourner  
Abigail K. Wozniak

Working Paper 27792  
<http://www.nber.org/papers/w27792>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
September 2020

We thank Kirsten Cornelson, Kevin Rinz, and Zachary Swaziek, as well as participants in the Williams College summer economics and the HeLP! seminars for helpful comments. Errors are our own. The views expressed here are those of the author and do not necessarily represent those of the Federal Reserve Bank of Minneapolis, the Federal Reserve System, or the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2020 by Krista J. Ruffini, Aaron Sojourner, and Abigail K. Wozniak. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Who's In and Who's Out under Workplace COVID Symptom Screening?

Krista J. Ruffini, Aaron Sojourner, and Abigail K. Wozniak

NBER Working Paper No. 27792

September 2020

JEL No. I1,J5,J7,K3,M5

**ABSTRACT**

COVID symptom screening, a new workplace practice, is likely to affect many millions of American workers in the coming months. Eleven states already require and federal guidance recommends frequent screening of employees for infection symptoms. This paper provides some of the first empirical work exploring the tradeoffs employers face in using daily symptom screening. First, we find that common symptom checkers will likely 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.

Krista J. Ruffini  
Minneapolis Federal Reserve Bank  
Opportunity and Inclusive Growth Institute  
90 Hennepin Avenue  
Minneapolis, MN 55401  
k.j.ruffini@berkeley.edu

Abigail K. Wozniak  
Federal Reserve Bank of Minneapolis  
90 Hennepin Ave.  
Minneapolis, MN 55401  
and IZA  
and also NBER  
abigailwozniak@gmail.com

Aaron Sojourner  
University of Minnesota  
Carlson School of Management  
321 19th Ave S, 3-300  
Minneapolis, MN 55409  
asojour@umn.edu

## 1. Introduction

As employers seek to resume operations under the ongoing pandemic threat, many have begun screening their employees for signs of infection on a daily basis. Workplace screening measures commonly include some combination of temperature checks and symptom self-reports.<sup>2</sup> There is not yet a reliable measure of how common daily health screening is in the US, but the numbers are likely large. Tens of millions of US workers have seen their work disrupted by the coronavirus. In June 2020, more than 40.4 million workers were not working or working fewer hours because of the COVID pandemic, and one in three workers were working from home (BLS 2020). As these workers return to previous or new employment, workplace screening is likely to be part of their routines. Employers will also feel pressure to implement screens from many sources. Federal and state agencies are already recommending or requiring regular infection screening in the workplace (OSHA 2020). As of July 23, 2020, eleven states require employee infection screening of some or all businesses that reopen.<sup>3</sup> Two out of five U.S. workers in a recent survey chose temperature and symptom checker screens as a factor necessary to make them feel safe at work, making workforce screens among the five most important workplace measures from employees' perspectives (American Staffing Association 2020).

Employers are implementing these changes in the face of considerable uncertainty. Little evidence exists on key questions around widespread daily health screening, including how many workers screens detect; whether there are differences in detection across various screening options; whether screens are equitable in their detection; whether differences in underlying health or survey response behaviors lead screens to disproportionately flag certain groups; and whether workplace screens help reduce COVID-19 spread.

This paper offers the first evidence on many of these questions. Specifically, we draw on a novel, nationally-representative survey, the COVID Impact Survey (CIS), to provide the first evidence on how workplace screening programs based on temperature checks and self-reported symptoms might affect the U.S. workforce. 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 generate new evidence on what types of workers would be flagged by alternative screening practices. The symptom screens we explore fall into the category of “prudent questions” recently recommended by Anthony Fauci in place of temperature checks alone.<sup>4</sup>

Our findings suggest that employers will need to contend with several tradeoffs when using daily symptom screening. First, common symptom checkers, including temperature-taking, will likely identify many workers as high-risk on any given day. Our analysis finds that up to 7 percent of workers could be flagged as symptomatic for COVID-19 daily, depending on the measure used. These levels are broadly consistent with available comparable benchmarks. Second, the measures used will matter for three reasons: many respondents report any given symptom, so choosing to include more or fewer symptoms will identify

---

<sup>2</sup> Some high profile employers like major league sports teams and the White House are doing daily medical testing of employees, but these are likely exceptional.

<sup>3</sup> These states include AZ, CO, DE, KY, MI, MN, NH, NJ, UT, VT and WA. This list excludes health provider screening which may be more widely required. Source: <https://www.litler.com/publication-press/publication/wont-hurt-bit-employee-temperature-and-health-screenings-list>

<sup>4</sup> As reported in The Hill: <https://thehill.com/changing-america/well-being/prevention-cures/512769-fauci-explains-why-temperature-checks-to-fight>

different individuals; survey design matters; 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 particularly important from the standpoint of equity and discrimination.

While we do not observe actual COVID infection rates, this analysis provides a foundation for potential future evaluations of the efficacy of various screens in a large-scale experimental setting. We also note that unlike actual workplace settings, the CIS is not a high stakes survey and respondents have no incentives to misreport information. Therefore, while the prevalence rates we observe may be better proxies of true underlying conditions, we are unable to estimate or approximate strategic reporting.

Our analysis raises some concerns about the precision of workplace screens, but these shortcomings may not disqualify workplace screens as an important public health tool. Public health experts have advocated prioritizing high-frequency testing over a particular test's sensitivity (Larremore et al. 2020). Also, daily health screens are likely to remain a widespread tool for containing the novel COVID-19 coronavirus, particularly in the US where other options are likely to remain limited for some time. Regular employer-based medical testing is currently prohibitive for most employers and for the overall US testing system as a whole, particularly tests with a rapid turnaround time. As of mid-July, about 750,000 COVID tests were administered each day, enough for less than 0.5 percent of the 160 million labor market participants (Johns Hopkins 2020). Due to bottlenecks, processing these tests was taking an average of 7-10 days, making them essentially uninformative for containment (Volz and Galewitz 2020). Requiring antibody tests, or "immunity passports," may be illegal under US disability law, and as of July 2020, only three states had adequate testing and tracing capacity to implement large-scale "test and trace" programs (Mohapatra, 2020; US EEOC 2020; TestAndTrace, 2020).<sup>5</sup> Therefore, as employers seek to manage the tension that human interactions power both the economy and virus spread, they will likely continue to rely on their employees refraining from coming to work if they may infect others. Regular symptom screenings offer a means to encourage this.

This paper relates to two literatures at the intersection of public health and economics. First, a large public health literature concludes non-pharmaceutical interventions mitigate infection spread when vaccination and antiviral medication are not available (see for example, Ferguson et al. 2006, Gostin et al. 2006, Kelso et al. 2009, Qualls et al. 2017, WHO 2019). Recent work explicitly considers tradeoffs 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 analyses complement this literature by quantifying how many and what types of workers are likely to be affected under various widespread symptom screens.

Second, a rich economics literature considers how employers may identify individuals by relying on imperfect signals of worker characteristics, including health risks (Autor and Scarborough 2008; Wozniak 2015). Our paper is the first study to empirically investigate the incidence of these 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.

---

<sup>5</sup> Phelan (2020) notes that immunity passports based on infection, rather than vaccination (currently unavailable), also create perverse public health incentives and may be banned for that reasons as well.

In our context, one consideration is whether self-reported symptom measures are informative about underlying infection. A public health literature finds that self-reported symptoms can accurately proxy objective measures (Buckley and Conine 1996, Chen et al. 2012, Jeong et al. 2010), but there is a lack of evidence on which types of workers would be flagged under various screens or how employers may select a screening device in a pandemic environment. We contribute the first evidence on how screens might work in a representative sample of workers facing no reporting constraints or consequences. We also provide evidence on considerations employers should account for in selecting a workplace symptom screen. While our data do not allow us to determine whether one screen is more accurate than another in detecting COVID infectiousness, we do provide evidence on whether self-reported symptom screens are likely to contain any information on infectiousness.

The remainder of this paper is organized as follows. Section 2 outlines a framework for how employers may weigh the tradeoffs associated with any screening device that relies heavily on self-reported information. Section 3 describes the COVID Impact Survey and how self-reported symptoms may be used to identify those with COVID-related symptoms. Section 4 presents our main results on the prevalence of fever- and COVID-related symptoms among CIS respondents in the labor force. Section 5 concludes.

## 2. 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. Employers may use workplace health screens to determine whether a worker is potentially infectious with coronavirus. Specifically, a screen takes the following form:

$$(1) d_{\tilde{s}}^i = \begin{cases} 1 & \text{if } f_{\tilde{s}}(S^i) \geq t_{\tilde{s}} \\ 0 & \text{otherwise} \end{cases}$$

where screen,  $d_{\tilde{s}}^i$ , is an indicator that takes the value 1 or 0 for whether worker  $i$  is indicated as infectious. Indicator screens can include a range of potential inputs. We denote the complete set of possible input measures as  $S$ , and the particular inputs used to generate the screen as  $\tilde{s}$ .  $f_{\tilde{s}}(\cdot)$  aggregates the particular inputs. A worker screens positive if the output of this function exceeds a given threshold  $t_{\tilde{s}}$ . For example, employers could ask about specific COVID symptoms or additional risky behaviors, such as attending large group gatherings. The potential inputs set includes all such factors, but 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. The screen would then identify anyone with any reported symptoms as potentially infectious.

Employers adopt health screens to prevent the spread of COVID-19 among employees and customers but they cannot observe each worker’s current potential infectiousness. Assume each day, an employee  $i$  has a true potential infectiousness  $g_i^* \in \{0,1\}$ .<sup>6</sup> Screens then provide the following information on infection probability:

---

<sup>6</sup> 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.

$g_i^* = \dots$	1	0
$d_s^i = \dots$		
1	$tp_s$	$fp_s$
0	$fn_s$	$tn_s$

Accuracy is the rate at which a screen correctly identifies infectious and non-infectious workers,  $[TP + TN]/[TP + TN + FP + FN]_s$ , where uppercase indicates totals in each category for a tested population.<sup>7</sup> 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. Even absent manipulation, each symptom is a noisy signal of underlying true infection. The screens we consider likely vary in their error rates for both reasons.

In a more complete model, one might specify a decision-maker’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. Available screens take two broad forms – medical and non-medical testing – although in principle all such tests are potential elements of the set  $S$ . 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.<sup>8</sup> 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).<sup>9</sup>

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-

<sup>7</sup> “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).

<sup>8</sup> Requiring immunity to infection as a condition of work is currently illegal in the U.S. (EEOC 2020).

<sup>9</sup> This is particularly true for the rapid tests most likely to be used to screen workers for attendance.

reported responses. Third, employees 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 is well-suited to approximate these measures.

### 3. Measuring Worker Health using the COVID Impact Survey

The COVID Impact Survey (CIS; Wozniak et al. 2020) was developed as a prototype of a survey tool that could assess a broad spectrum of wellbeing 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, administered by the National Opinion Research Center (NORC), is described in greater detail in Wozniak (2020) and the online documentation.<sup>10</sup>

During each of three weeklong waves, 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 each cross-section.<sup>11</sup>

The CIS includes a battery of health screening questions chosen because they were 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 their temperature as measured by a thermometer during the survey.

We develop seven screens for COVID infection risk from this broad set of questions. These differ according to whether the screens use self-reported symptoms or a thermometer-measured temperature; whether they ask about fever symptoms only or COVID symptoms more generally (including all COVID-related fever symptoms); and whether they derive from a short or long symptom checklist.<sup>12</sup> In the CIS, the short symptom list asks about three symptoms of fever only. These are derived from emergency department fever screening approaches. The long symptom checklist asks about 17 temporary poor health items. The list combines items from various symptom checklists circulating in late spring 2020. Appendix Table 1

---

<sup>10</sup> Full documentation available at <https://www.covid-impact.org/results>.

<sup>11</sup> 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, TX) and 8 metropolitan statistical areas (Atlanta, Baltimore, Birmingham, Chicago, Cleveland, Columbus, Phoenix, 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 US. 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.

<sup>12</sup> 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.

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. As we are interested in workplace screening devices, we restrict our main sample to labor force participants from the national sample and provide supplemental analysis from the 10 states and 8 metropolitan areas in the subnational sample. 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 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 from this period.<sup>13</sup> Appendix Table 2 contains descriptive statistics from our primary analytic sample. Unless otherwise noted, all estimates are produced using the appropriate CIS sample weights.

#### **4. 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.

##### **4.1. Prevalence of Self-Reported Symptoms in the Population**

Our first finding is that a substantial share of the workforce would screen positive under any of the screens described above and in Appendix Table 1, though the shares vary widely. The first row of Table 1 shows the positive shares under each of the seven screens. Four percent of respondents report currently having a temperature of 99 degrees F or higher based on the contemporaneous (current day only) thermometer reading. All other CIS symptom questions refer to a seven-day reference period, rather than only the current day. Hence levels on the remaining measures are not directly comparable to the temperature screen, and daily rates of our screens are likely lower than those reported here. 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.<sup>14</sup> 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.

When considering how these rates may translate to an employment setting, we emphasize 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,

---

<sup>13</sup> 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.

<sup>14</sup> These positive rates are generally lower than the 20-30 percent positive rate for influenza-type-illness at the peak of flu season (CDC 2020b).



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.<sup>15</sup>

Table 1 also shows that 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 Section 4.2.

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/ethnicity, income, and age, Table 1 shows meaningful differences in positive screens rates. For example, rates of multiple COVID symptoms are lower for Blacks and Hispanics relative to whites; 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 in Section 4.3.

These results suggest that symptom screening will detect a substantial share of the workforce at any point in time. We assess whether these shares are reasonable by comparing them to two other sources: The Census Household Pulse Survey (HHPS) and Carnegie Mellon University's Delphi Project (CMU-DP).

The recent HHPS collected information on whether respondents were not at work due to COVID illness or symptoms. Figure 1 shows the share of labor force participants in the HHPS who were not at work in the past seven days because they were ill or symptomatic for COVID. The figure shows this was the case for about 1% of HHPS respondents early in the survey (overlapping with the CIS collection period), rising to 2% in more recent weeks as the pandemic has spread. This is lower than the rates of respondents reporting symptoms on the CIS screens, but due to the question wording in the HHPS, these numbers may predominantly reflect respondents staying home for an entire week. The HHPS share will also exclude individuals with symptoms who did not stay home at all as well as those who had symptoms but worked from home. The HHPS data suggest that 1-2% of the labor force may be absent for up to a week at a time. This seems not out of line with our ballpark of 0.4% to 7.4% of workers potentially screening positive on any given day.

The CMU-DP is conducting daily large scale symptom surveys over Facebook and making results available publicly at the county level. These data provide an additional source of information on symptom prevalence for a large population, although responses are not limited to labor force participants. The CMU data show that between 0 and 5% of county respondents report COVID-like symptoms in any 24-hour period; the mean is 0.52% (tabulations available upon request). Similar shares report flu-like symptoms, although we have not yet explored the overlap between these reports. Regardless, the CMU-DP data produce symptom reports very similar to our rough estimates of daily shares screening positive on our various measures.

---

<sup>15</sup> NORC examined all data for response patterns indicating inattention, e.g. checking all affirmative responses on a symptom checker, many skipped questions, and found few respondents exhibiting this behavior.

## 4.2 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.

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 Anthony Fauci's recent public comments on symptom screens versus temperature checks.<sup>16</sup>

We further assess consistency across our measures in Appendix Table 3, 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.<sup>17</sup> 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 and Conine 1996, Chen et al. 2012, Jeong et al. 2010).

---

<sup>16</sup> Reference in footnote 4.

<sup>17</sup> These results and those for 100+ temperatures, are available upon request.

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+) and using recent diagnosed COVID infection. Across the screens, roughly 1% to 10% of our sample reports elevated infection potential by one of our two proxies (tabulations available upon request). The median share is roughly 5%. This is close to the most recent rate of positive medical tests nationwide, as reported by the Harvard Global Health Initiative.<sup>18</sup> It is difficult to get data on reasons for obtaining a COVID test, but it is likely that most Americans getting a COVID test have reason to suspect they may be ill, either because of their own symptoms or suspected exposure.<sup>19</sup> Thus, on this dimension, respondents in the CIS reporting 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 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 Appendix Table 4, 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.

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 randomized across subjects to avoid order effects. The self-reported fever symptoms (short) screen is constructed from a separate, 3-item fever specific checklist. This is asked later in the survey, and items are not randomized.

Table 1 shows that positive screen rates are much lower for self-reported fever using the short checklist compared to the long checklist.<sup>20</sup> This pattern also applies to the one symptom that appears on both

---

<sup>18</sup> <https://globalepidemics.org/testing-targets/>. Accessed on 8/21/2020.

<sup>19</sup> Regular testing of asymptomatic individuals is still rare in the US.

<sup>20</sup> 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 don't 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 skips. 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.

checklists, the presence of chills.<sup>21</sup> 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.

### **4.3 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.<sup>22</sup> 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 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 January and June 2020, employment rates among Black Americans fell by 2.1 percentage points more and Hispanic Americans by 2.9 percentage points more than white Americans. This raises the question of whether these

---

<sup>21</sup> 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 responses rates are similar on the two checklists, so lower rates of symptom reporting on the later list is not due to item non-response.

<sup>22</sup> Health status includes disability and genetic information, which may include immunity status.

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.<sup>23</sup> 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.<sup>24</sup> 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/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).

The next four columns 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 and 3, respectively), but they do predict having at least two symptoms on both lists. Non-Hispanic Blacks and Hispanics are each less likely than non-Hispanic whites 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

---

<sup>23</sup> 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.

<sup>24</sup> In additional results, patterns are robust to including week dummies to assess time effects in reporting.

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 in. 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 and 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) reports coefficients on the health status covariates. The first of these is an indicator for a COVID diagnosis in the household, another measure of elevated infection risk in the absence of COVID infection information for individuals. Like a thermometer reading, a COVID diagnosis is a measure of infection risk that is independent of subjective symptom reporting. 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 screen 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/poor health represent an additional challenge for employers. In Appendix Table 1, 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.<sup>25</sup> 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 make 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/ethnic, gender, and age subsample separately and plots the resulting coefficients on having a COVID diagnosis in the household. The plots show that a COVID diagnosis in the home elevates the likelihood of a positive symptom screen across all screens and all demographic subsamples. 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

---

<sup>25</sup> [https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/people-with-medical-conditions.html?CDC\\_AA\\_refVal=https%3A%2F%2Fwww.cdc.gov%2Fcoronavirus%2F2019-ncov%2Fneed-extra-precautions%2Fgroups-at-higher-risk.html](https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/people-with-medical-conditions.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fcoronavirus%2F2019-ncov%2Fneed-extra-precautions%2Fgroups-at-higher-risk.html). Note: This page is frequently updated.

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, Appendix Table 5 shows the coefficients on race and ethnicity are stable when including information on per-capita caseloads. Turning to the coefficients on per-capita cases, Appendix Table 5 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 US 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.

## 5. 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.

These tools present tradeoffs 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. Our analysis shows that both the number of symptom inputs and the threshold matters: all else equal, screens that enquire 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, we find that 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. We find that some groups – women, younger workers, and non-Hispanic whites – report multiple symptoms at higher rates than men, older workers, 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. 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. First, other research finds that when workers can stay home when ill, overall flu infection rates fall (Pichler et al. 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. Also, regular workplace COVID symptom screening might encourage individuals to seek out medical testing if such testing is not readily available from their employer. Despite spiking cases this summer, testing rates began to fall in many states. Health officials have urged the public to continue getting medically tested if they have reason to be concerned they have COVID. Workplace symptom screening might reinforce this message.

Applying our findings to the workplace setting raises additional considerations. 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. 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 a worker's economic security does not depend on reported symptoms. Employers could also consider group-level precautions that escalate if large shares of workers report symptoms, or tying 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).

Our analysis offers guidance for employers, although data limitations and symptom patterns that vary across cases preclude us from developing an optimal symptom screening procedure that is feasible at a large scale. 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 tradeoff between screens that may have a higher false negative rate (e.g. temperature checks alone) but fewer disparities in detection, versus a lower false negative rates (temperature checks plus COVID symptom screens) but more disparities. 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. As Jones et al. (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.

In the absence of universal, near-daily COVID testing, employers instead will need 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.



## References

- Acemoglu, Daron, Victor Chernozhukov, Iván Werning, and Michael D. Whinston. "Optimal targeted lockdowns in a multi-group SIR model." NBER Working Paper 27102 (2020).
- American Staffing Association, "Most Employees Satisfied with Covid-19 Return-to-Work Plans." July 16, 2020. <https://americanstaffing.net/posts/2020/07/16/covid-19-return-to-work-plans/>. Accessed August 2, 2020.
- Augenblick, Ned, Jonathan Kolstad, Ziad Obermeyer, and Ao Wang. 2020. "Group Testing in a Pandemic: The Role of Frequent Testing, Correlated Risk, and Machine Learning." *NBER Working Paper #27457*.
- Autor, David and David Scarborough. 2008. "Does Job Testing Harm Minority Workers? Evidence from Retail Establishments" *Quarterly Journal of Economics*. 123(1): 219-277.
- Baqae, David, Emmanuel Farhi, Michael Mina, and James H. Stock. "Policies for a second wave." (2020). Brookings Papers on Economic Activity.
- Baratloo, Alireza, Mostafa Hosseini, Ahmed Negida, Gehad El Ashal. 2015. "Part 1: Simple Definition and Calculation of Accuracy, Sensitivity, and Specificity." *Emergency (Tehran)*. 3(2): 48-49.
- Buckley, Robert G., and Melanie Conine. "Reliability of subjective fever in triage of adult patients." *Annals of emergency medicine* 27, no. 6 (1996): 693-695.
- Center for Disease Control and Prevention. "Symptoms of Coronavirus." May 2020. <https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html>. Accessed 1 July 2020.
- Center for Disease Control and Prevention. "Influenza Positive Tests Reported to CDC by U.S. Clinical Laboratories, 2019-2020 Season." August 2020. [https://www.cdc.gov/flu/weekly/weeklyarchives2019-2020/data/whoAllregt\\_cl32.html](https://www.cdc.gov/flu/weekly/weeklyarchives2019-2020/data/whoAllregt_cl32.html). Accessed 18 August 2020.
- Centers for Disease Control and Prevention. 2020. "COVID-19 Hospitalization and Death by Race/Ethnicity." Updated August 18, 2020. Available at <https://www.cdc.gov/coronavirus/2019-ncov/covid-data/investigations-discovery/hospitalization-death-by-race-ethnicity.html>
- Chen, Shey-Ying, Yee-Chun Chen, Wen-Chu Chiang, Hsiang-Chi Kung, Chwan-Chuen King, Mei-Shu Lai, Wei-Chu Chie, Shyr-Chyr Chen, Wen-Jone Chen, and Shan-Chwen Chang. "Field performance of clinical case definitions for influenza screening during the 2009 pandemic." *The American journal of emergency medicine* 30, no. 9 (2012): 1796-1803.
- Dingel, Jonathan I., and Brent Neiman. *How many jobs can be done at home?*. No. w26948. National Bureau of Economic Research, 2020.
- Ferguson, Neil M., Derek AT Cummings, Christophe Fraser, James C. Cajka, Philip C. Cooley, and Donald S. Burke. "Strategies for mitigating an influenza pandemic." *Nature* 442, no. 7101 (2006): 448-452.
- Gostin, Lawrence, et al. "Public health strategies for pandemic influenza: ethics and the law." *Jama* 295, no. 14 (2006): 1700-1704.
- Jeong, Ina, Chang-hoon Lee, Deog Kyeom Kim, Hee Soon Chung, and Sang Won Park. "Mild form of 2009 H1N1 influenza infection detected by active surveillance: implications for infection control." *American journal of infection control* 38, no. 6 (2010): 482-485.
- Johns Hopkins University of Medicine. "Testing Trends Toolkit." Coronavirus Resource Center, Testing Hub. (2020). <https://coronavirus.jhu.edu/testing/tracker/overview>. Accessed 25 July 2020.
- Jones, Damon, David Molitor and Julien Reif. 2019. "What Do Workplace Wellness Programs Do? Evidence from the Illinois Workplace Wellness Study." *Quarterly Journal of Economics*. 134(4): 1747-1791.
- Kaiser Family Foundation. "Adults at Higher Risk of Serious Illness if Infected with Coronavirus." (2020). <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](#). Accessed 25 July 2020.

Kelso, Joel K., George J. Milne, and Heath Kelly. "Simulation suggests that rapid activation of social distancing can arrest epidemic development due to a novel strain of influenza." *BMC public health* 9, no. 1 (2009): 117.

Larremore, Daniel B., Bryan Wilder, Evan Lester, Soraya Shehata, James M. Burke, James A. Hay, Milind Tambe, Michael J. Mina, and Roy Parker. "Test sensitivity is secondary to frequency and turnaround time for COVID-19 surveillance." medRxiv (2020).

Mohapatra, Seema. "Why COVID-19 Immunity Passports May Violate US Law." *The Conversation*. May 27, 2020.

Parker-Pope, Tara. 2020. "The Many Symptoms of Covid-19." *New York Times* (5 August 2020).

Phelan, Alexandra. 2020. "COVID-19 immunity passports and vaccination certificates: scientific, equitable, and legal challenges." *The Lancet*. Comment. 395(May 23): 1595-1598.

Pichler, Stefan, Katherine Wen, Nicholas R. Ziebarth. 2020. "Positive Health Externalities of Mandating Paid Sick Leave." *Working paper, REPEC*.

Pradhan, Rachana. "As Problems Grow With Abbott's Fast COVID Test, FDA Standards Are Under Fire." Kaiser Health News. (22 June 2020).

Qualls N, Levitt A, Kanade N, et al. Community Mitigation Guidelines to Prevent Pandemic Influenza — United States, 2017. *MMWR Recomm Rep* 2017;66(No. RR-1):1–34.

Schnall, Rebecca, Karolynn Siegel, Haomiao Jia, Susan Olender, and Sabina Hirshfield. 2018. "Racial and Socioeconomic Disparities in the Symptom Profile of Persons Living with HIV." *AIDS Care*. 30(6): 774-783.

Sudre, Carole H. et al. Symptom clusters in Covid19: A potential clinical prediction tool from the COVID Symptom study app (2020) medRxiv

Test and Trace. "What U.S. States are Ready to Test and Trace Today?" <https://testandtrace.com/state-data/>. Accessed 27 July 2020.

Volz, Matt and Phil Galewitz. "As Long Waits for Results Render COVID Tests 'Useless,' States Seek Workarounds." Kaiser Health News. (23 July 2020). <https://khn.org/news/states-search-for-ways-to-deal-with-covid-19-testing-backlogs/>

World Health Organization (WHO). "Non-pharmaceutical public health measures for mitigating the risk and impact of epidemic and pandemic influenza." ISBN: 978-92-4-151683-9 (2019). [https://www.who.int/influenza/publications/public\\_health\\_measures/publication/en/](https://www.who.int/influenza/publications/public_health_measures/publication/en/). Accessed 1 July 2020.

Wozniak, Abigail. 2015. "Discrimination and the Effects of Drug Testing on Black Employment." *Review of Economics and Statistics*. 93(7): 548-566.

Wozniak, Abigail. Disparities and Mitigation Behavior during COVID-19. No. 32. Federal Reserve Bank of Minneapolis, 2020.

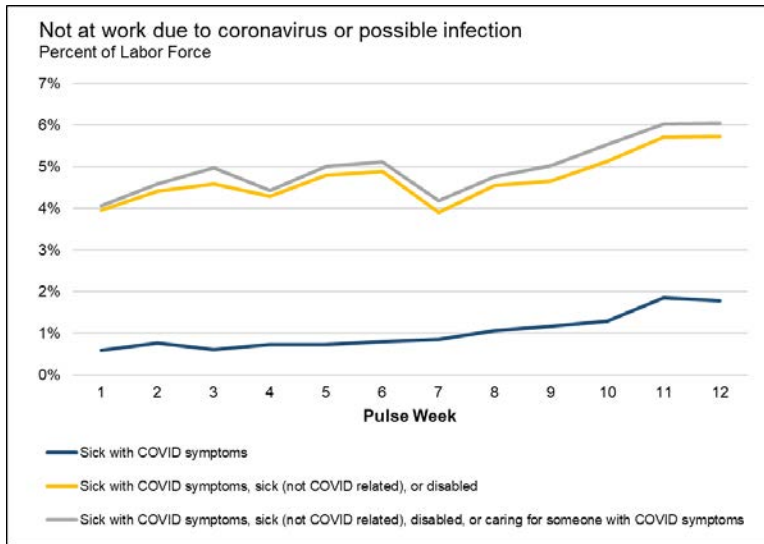
Wozniak, Abigail, Joe Willey, Jennifer Benz, and Nick Hart. 2020. COVID Impact Survey: Version 2[dataset]. Chicago, IL: National Opinion Research Center, 2020.

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

U.S. Equal Employment Opportunity Commission. 2020. "What You Should Know about COVID-19 and the ADA, the Rehabilitation Act, and Other EEO Laws." Technical Assistance Questions and Answers, June 17, 2020. [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=](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=)

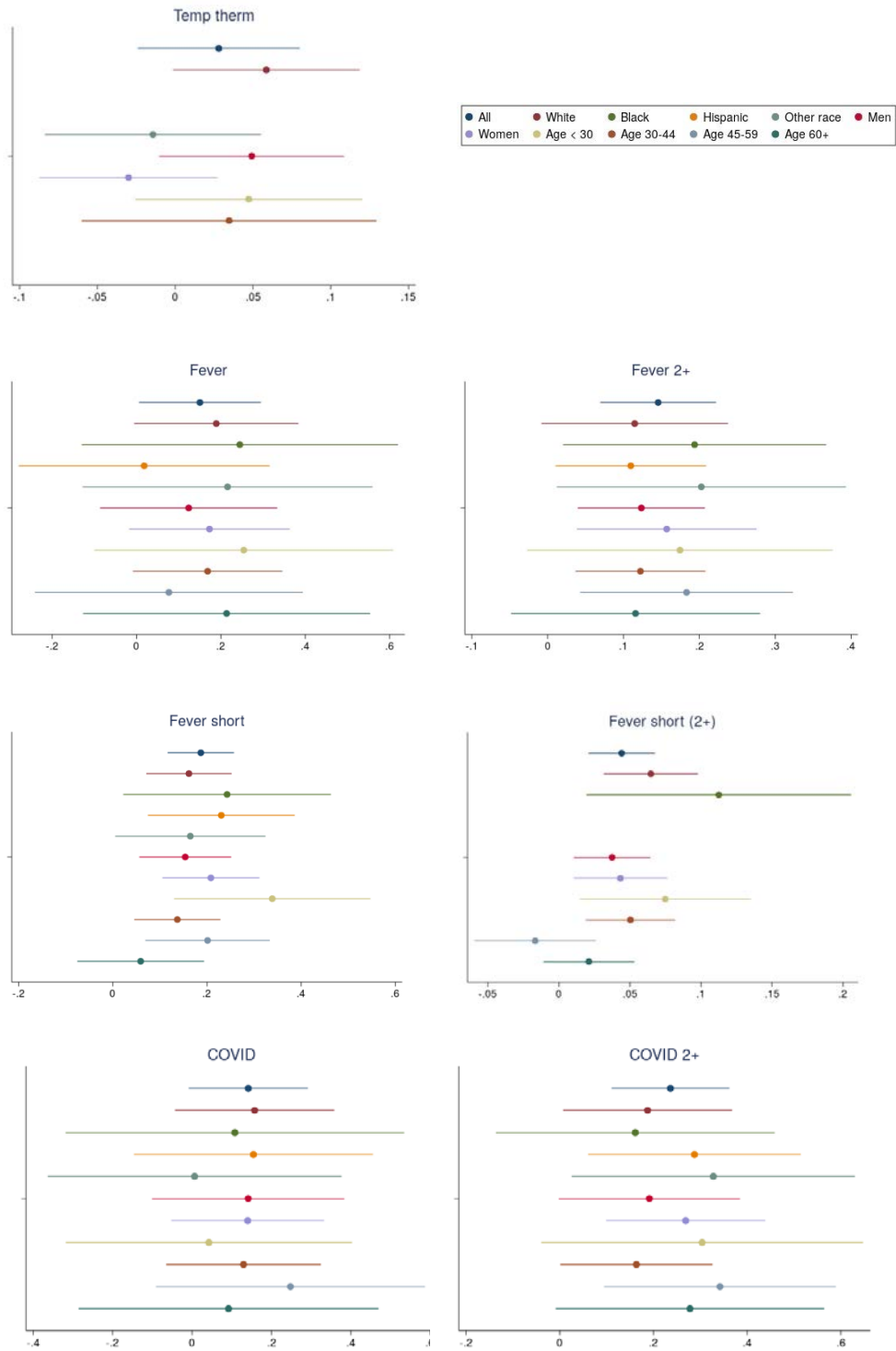
U.S. Occupational Health and Safety Administration. 2020. "Guidance on Returning to Work." OSHA 4045-06 2020.

**Figure 1: Not at Work Due to Coronavirus or Possible Infection**



Notes: Census Household Pulse Survey data, weeks 1 through 12, spanning April 23 through July 21, 2020. Labor force participation approximated from “reasons not at work last week.”

**Figure 2: Effect of household COVID diagnosis on positive screen rates by worker characteristics**



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. Source: CIS weeks 1-3, sample limited to labor market participants from the national sample. Missing point estimate indicates insufficient data to estimate for that subgroup.

**Table 1: Share of Labor Force Flagged by Workplace Screens**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Therm (99)	Fever short	Fever	Fever short 2+	Fever 2+	COVID	COVID 2+	COVID in home	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
<b>Survey week</b>									
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
<b>Race/ethnicity</b>									
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
<b>Household income</b>									
<\$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
<b>Age</b>									
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. Restricted to respondents in labor force.

**Table 2: Share of Positive Screens, Conditional on Other Positive Screen**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Pr(Temp (therm)=1)	Pr(Fever (short)=1)	Pr(Fever =1)	Pr(Fever (short 2+)=1)	Pr(Fever 2+)=1)	Pr(COVID=1)	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.

**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+
<b>Panel a: Race/ethnicity</b>							
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)
<b>Panel b: Gender</b>							
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)
<b>Panel c: Age</b>							
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)
<b>Panel d: Health</b>							
HH with COVID	0.03 (0.02)	0.18*** (0.04)	0.04*** (0.01)	0.16** (0.07)	0.14*** (0.04)	0.15** (0.08)	0.24*** (0.06)
Known death	-0.01 (0.02)	0.02 (0.02)	-0.01 (0.01)	0.03 (0.05)	0.01 (0.02)	0.03 (0.05)	0.07* (0.04)
2+ risk factors	0.01 (0.01)	0.04*** (0.01)	0.01** (0.01)	0.14*** (0.03)	0.02 (0.01)	0.15*** (0.03)	0.12*** (0.02)
Poor/fair health	0.00 (0.01)	0.06*** (0.02)	0.03*** (0.01)	0.15*** (0.03)	0.09*** (0.02)	0.18*** (0.04)	0.18*** (0.03)
<b>Panel e: HH income</b>							
Income \$40k-75k (2019)	0.01 (0.01)	0.01 (0.02)	0.00 (0.01)	0.00 (0.03)	0.00 (0.02)	0.00 (0.03)	0.01 (0.03)
Income \$75k+ (2019)	0.02** (0.01)	-0.01 (0.02)	-0.01 (0.01)	-0.04 (0.03)	-0.04** (0.02)	-0.06* (0.03)	-0.06** (0.03)
<b>Panel f: HH type</b>							
Childless HH	0.01 (0.01)	-0.02 (0.02)	-0.01 (0.01)	-0.02 (0.03)	0.00 (0.02)	-0.02 (0.03)	0.00 (0.03)
HH with children, all under age 5	0.02 (0.02)	-0.01 (0.03)	0.00 (0.01)	-0.20*** (0.05)	-0.03 (0.03)	-0.16*** (0.05)	-0.14*** (0.04)

HH with children, some older than 5	0.02 (0.01)	0.02 (0.02)	0.00 (0.01)	-0.04 (0.03)	0.01 (0.02)	-0.03 (0.03)	-0.02 (0.03)
Other HH composition	0.04*** (0.01)	0.00 (0.02)	0.00 (0.01)	-0.04 (0.03)	-0.01 (0.02)	-0.03 (0.04)	-0.05 (0.03)
<b>Panel g: Geographic location</b>							
Suburban	-0.02 (0.02)	-0.01 (0.02)	-0.03*** (0.01)	0.04 (0.04)	0.00 (0.02)	0.00 (0.04)	0.05 (0.04)
Urban	-0.01 (0.01)	0.01 (0.02)	-0.01 (0.01)	0.01 (0.04)	0.00 (0.02)	-0.03 (0.04)	0.02 (0.03)
N	2346	4180	4180	4115	4179	4180	4180

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



## Appendix

**Appendix Table 1: 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 or more 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 or more 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 or more 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 or more 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. Appendix Table 4 lists COVID and fever symptoms from the long symptom checklist as well as all items from the short fever checklist.

**Appendix Table 2: 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 and 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.

**Appendix Table 3: Correlation across Screens**

	<b>Therm (99)</b>	<b>Fever (short)</b>	<b>Fever</b>	<b>Fever (short 2+)</b>	<b>Fever (2+)</b>	<b>COVID</b>	<b>COVID (2+)</b>	<b>Risk 2+</b>	<b>Fair/poor health</b>
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. Restricted to respondents in labor force. Table shows the raw correlation across measures.

**Appendix Table 4: 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

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 Appendix Tabel 1.

**Appendix Table 5: Implied Marginal Effects from Probit of Screen on Demographics, Regional Data**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Therm	Fever (short)	Fever	Fever (short 2+)	Fever (2+)	COVID (long)	COVID (long 2+)
Per capita cases	-0.12	0.16	-0.06	0.62**	0.17	0.00	0.16
<b>Race/ethnicity</b>							
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
<b>Gender</b>							
Female	-0.01	0.00	0.00	0.04**	0.02**	0.05**	0.05***
<b>Age</b>							
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 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 restricted to CIS participants from regionally-representative cross-sections, pooling all weeks. \*\*\*p < 0.01, \*\* p < 0.05, \*p < 0.10.