Health Economics

WILEY

Health Economics

RESEARCH ARTICLE

Influenza Vaccination and Occupational Social Intensity: Evidence From U.S. Workers

Yancheng Xiao[1] | Matthew Harris[2] | Alissa O'Halloran[3] | Jamison Pike[3]

[1]Hunan University of Technology and Business, Changsha, China | [2]Department of Economics and Boyd Center for Business and Economic Research, Haslam College of Business, The University of Tennessee, Knoxville, Tennessee, USA | [3]National Center for Immunization and Respiratory Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia, USA

Correspondence: Matthew Harris (mharris@utk.edu)

Received: 29 April 2024 | **Revised:** 2 April 2025 | **Accepted:** 8 April 2025

Keywords: disease spread | externality | Immunization | influenza | social intensity | vaccination

ABSTRACT

We examine whether individuals with socially intensive occupations are empirically more likely to get vaccinated against flu. Socially active workers interact with a relatively large number of people and are therefore more likely to catch and spread infectious diseases. Vaccinating these workers against infectious disease therefore yields greater marginal social benefits than the average person. We construct multiple measures of social intensity from the Occupational Information Network. We find that workers in occupations with high social intensity are not significantly more likely to receive a flu shot. These results add to earlier findings that individuals' vaccination decisions are driven mainly by private valuations rather than social considerations. However, we provide the first direct evidence that individuals for whom vaccination would yield the greatest external benefits are not more likely to get vaccinated.

JEL Classification: I12, I18, D62, D85, H41

1 | Introduction

Vaccination protects not only vaccinated individuals but all those who interact with them. The Centers for Disease Control and Prevention (CDC) estimate that in the 2019-2020 season, 48.4% of adults and 63.7% of children were vaccinated for influenza. These vaccinations are estimated to have prevented 7.1 million influenza illnesses, 100,000 hospitalizations, and 7200 deaths in the United States (CDC 2020, 2023). Some vaccines provide long-term immunity, however, people must receive annual flu vaccination to increase their resistance. Although the annual seasonal flu vaccination coverage has risen steadily during the past decade, annual vaccination rates are still far below the level needed to achieve herd immunity.[1] Even when capacity constraints do not bind, herd immunity rates (let alone disease eradication) are still difficult to achieve.[2]

As is the case with all externality-producing goods, the marginal private benefits and marginal social benefits of individuals' vaccination decisions are not equal. Geoffard and Philipson (1997) show that as vaccine uptake increases, the private benefit of marginal vaccinations decreases, and the willingness to be vaccinated diminishes. Because motivating everyone to get vaccinated is difficult, policymakers are usually interested in vaccinating important subpopulations, such as those with health vulnerabilities and health care providers (Boulier et al. 2007).[3] While health care providers may present the highest propensity to spread the virus to vulnerable populations, any individual with high frequency of interpersonal interactions

We extend sincere gratitude to the editor and two anonymous referees for their constructive comments. We are also grateful to the following for helpful comments: participants of the 2019 Guangdong Economic Association Annual Meetings and the 2021 Sun Yat-Sen-Jinan-SCNU Economic Meetings, seminar participants at South China Normal University, Beijing Normal University at Zhuhai, Zhongnan University of Economics and Law, and Xiangtan Institute of Technology.

may disproportionately spread the virus (Stein 2011; Teicher 2023). Song et al. (2021) found that in the COVID-19 pandemic, dependents and cohabitants of essential frontline workers were over 50% more likely to catch the virus than those in households without an essential worker.[4] Therefore, relative to a random draw from the population, vaccinating individuals with high levels of social interaction may yield greater marginal social benefits. However, as Boulier et al. (2007) point out, these individuals are generally not prioritized in vaccination roll-outs.

In this paper, we investigate the empirical relationship between the social intensity of an individual's occupation and the probability of being vaccinated against influenza. We believe this is the first paper to do so. Our data are from the Occupational Information Network (O*NET) and restricted files from the Behavioral Risk Factor Surveillance System (BRFSS) that contain Standard Occupational Classification (SOC) codes at the individual level.[5] We construct measures of occupational social intensity from full and partial sets of social-related descriptors from the O*NET data, applying Principal Component Analysis to form single indices of social intensity for each SOC. We separate the O*NET measure of "exposure to disease" from the other social aspects of a job. "Exposure to disease" carries a clear private risk of infection, while higher measures of social intensity entail both greater private risk of infection, but also greater likelihood to spread the flu. Therefore, we interpret the associations between vaccination and social intensity as a measure of how individuals value that marginal private risk and risk of spreading the disease to others.

We find that workers in occupations with high social intensity measures—despite being more likely to catch and spread infectious disease—are not significantly more likely to report receiving a flu shot. Results from the preferred specification are a relatively precise zero. A one standard deviation greater social intensity score is associated with a 0.37 percentage points higher likelihood of vaccination. By comparison, a one standard deviation difference in the measure of direct exposure to diseases is associated with a 2.59 percentage points higher likelihood of vaccination.[6] Workers in clinical health care occupations may follow a different data generating process, due to the highly socially intense nature of their job, the consequences of spreading disease to vulnerable populations, and in some cases, vaccine mandates. However, when we restrict the sample to non-health care workers, the coefficient on "social intensity" is essentially unchanged. These results are consistent across all alternative measures of social intensity derived from O*NET data.[7]

From an econometric perspective, these results should be interpreted associatively rather than causally. From a policy perspective, we emphasize that the associative relationship may be equally - if not more important - than the causal estimate. In this context, a true causal interpretation would mean that if person i switched to a more socially active occupation, that increased social intensity makes them more likely to get vaccinated. Meanwhile, these results should be interpreted as: "aware that there is non-random selection into occupations, are the individuals found in socially intensive occupations more likely to get vaccinated?" From a social planning perspective, the associative interpretation is important in determining who should be targeted/prioritized for vaccination efforts. Nevertheless, we implement Oster (2018) bounds to quantify the potential implications of selection biases in the results.

Understanding the association between workplace-specific social intensity and vaccination probability is useful, even compared to some holistic measure of social activity (i.e., SafeGraph data). First, individuals with the same occupation face highly similar patterns of social interaction in the workplace, so an occupation-level social intensity measure can capture a representative worker's social activeness. Second, researchers find that workplaces are crucial channels spreading flu (Adda 2016; Pichler and Ziebarth 2017; Markowitz et al. 2019). Therefore, we focus on the vaccination decisions of employed individuals and utilize their occupational characteristics to construct measures of social intensity.

These estimates are consistent with prior work showing that individuals respond to the potential personal infection risk but do not internalize the external effects of their vaccination choices (Philipson 1996; de Janvry et al. 2010; Ward 2014; Oster 2018; White 2019; Schaller et al. 2019).[8] The estimates for all constructions of social intensity are either null or economically insignificant, even when statistically significant. Note that the social intensity of an occupation includes both potential external benefits from vaccination and some measure of private risk from interacting with others (beyond "exposure to disease"). To the extent our measures of social intensity do capture some amount of private risk, the small conditional association between social intensity and vaccine uptake further underscores the implication that individuals do not consider external benefits when making vaccination decisions.

This paper also contributes to the empirical literature on externalities from flu vaccination. Recent economics literature finds a large marginal externality when policies motivate individuals to get flu vaccines, either by mandating health care professionals (White 2019; Carrera et al. 2021) or by subsidizing working-age adults (Ward 2014). We find evidence consistent with workers not fully internalizing the social benefits of their vaccination choices, which is useful when considering optimal vaccination programs of any sort (Siciliani et al. 2020).

This paper also has immediate policy implications for any future vaccine roll outs. Prioritizing who gets vaccinated is complicated. Policymakers must evaluate various factors, such as vulnerability, mortality risk, and whether the individual's work is essential for social functioning. Buckner et al. (2021) find that depending on the policy objectives of a COVID-19 vaccine roll out, policymakers can prioritize either younger essential workers to control disease spread or seniors to reduce mortality after senior essential workers get vaccinated. Therefore, to the extent that vaccinating individuals with large *external* benefits is a priority, these results suggest that individuals do not place sufficient weight on the social benefits of their vaccination, or are unaware that their occupation may give them a disproportionate role in spreading infectious disease. To the extent that targeting high-intensity (and therefore high externality) individuals is a priority, these findings imply that some incentives may be necessary to raise uptake among high-intensity individuals.[9]

Finally, this paper contributes to the public health literature about factors related to workers' flu vaccination coverage. Researchers compare flu vaccination rates on the major occupation group level and find that many non-health care occupations with frequent public contact have lower coverage than the population average (Luckhaupt et al. 2014; O'Halloran et al. 2017; Hughes et al. 2019). We quantify the social intensity of detailed occupations and find vaccine uptake and social intensity are not correlated when we control for other variables.

2 | Data

Our primary data are protected individual-level cross-sectional data from the Behavioral Risk Factor Surveillance System (BRFSS) (CDC 2019) for the years 2013–2015, containing more detailed occupational specifics than the publicly available data set. BRFSS is the largest telephone health survey in the world, interviewing more than 400,000 adults every year. BRFSS collects information from U.S. residents on health-related risk behaviors, chronic health conditions, and utilization of preventive services among other health and demographic variables. BRFSS started an optional industry/occupation module in 2013. We obtained access to the 2013–2015 data, which include 31 states.[10] We restrict the sample to employed, civilian adults whose occupation and industry information were available, yielding an analytical sample of 229,144 individuals. We can observe their self-reported flu vaccine uptake, demographics, health conditions, insurance coverage, and restricted industry/occupation information. Table 1 presents the summary statistics. Although BRFSS is designed to be nationally representative, these restrictions make our analytical sample younger, more white, female, and college-educated than the U.S. on average.

The outcome variable in our study is whether the respondent reports receiving a flu shot with a needle or a nasal spray in the past 12 months. Our sample includes the month and year of a flu shot, which ranges from January 2012 to December 2015. Because the flu prevalence varies across seasons, we create a "season" variable to distinguish which flu season the respondent gets vaccinated. Because CDC's weekly flu activity report of a particular flu season starts in September, we define "season" as vaccine uptakes from August of a given year to July of the subsequent year. For example, the 2012-2013 season contains flu shots taken from August 2012 to July 2013.[11] The weighted sample mean vaccination coverage is 34%.[12]

Respondents report their current occupation if they are employed. We assume that respondents do not change their occupations between the time when they receive a flu shot and the interview time.[13] We categorize occupations based on the 2000 Standard Occupational Classification (SOC) codes.[14] Our sample contains individuals employed in 509 SOCs, including 371 detailed occupations, 126 broad occupations, and 12 minor groups. Individuals in our sample include employees in all non-military major and minor groups and over 90% of non-military broad occupations.[15] The coding of occupations to the SOC level in the restricted BRFSS allows us to match these individuals with a social intensity measures at a detailed occupational level.

We use the Occupational Information Network (O*NET) (National Center for O*NET Development 2019) data to construct measures of social intensity for each SOC occupation. O*NET surveys workers and provides data about specific job descriptions and requirements, which include numeric measures of work contexts and activities. We use information from O*NET to develop multiple measures of the social intensity of an occupation. In the most comprehensive approach, we use all available descriptors from the Work Activities and Work Context sections of O*NET that potentially reflect interactions with other people in the workplace. Twenty-three such descriptors pertained to social activities in some way. Six of these descriptors characterized work context and 16 descriptors were about work activities related to interpersonal interactions. Because we intend for these measures to capture the increased likelihood of catching *and* spreading flu, these measures should contain information about the external benefits of a worker in a given occupation being vaccinated. We partition out one variable, "Exposed to Disease or Infections," and consider this as a measure of "private risk" of infection alone.

Table 2 presents the summary statistics of the 23 occupational descriptors for 509 SOCs in Panel A, including "Exposed to Disease or Infections", which we partition out separately. We use Principal Component Analysis (PCA) to reduce these numerical ratings into a single index. Using weights (reported in Table 3) determined by the first principal component, we construct an index comprised of those 22 descriptors as the intensity score. We also construct indices using subsets of those 22 descriptors for robustness concerns and include the summary statistics of derived indices in Panel B.[16] Finally, we merge these measures of social intensity with the BRFSS data at the SOC level.[17]

Our measure of social intensity contains information on the external benefits of an individual getting vaccinated, but also contains information about private risk of exposure to flu.[18] In an effort to absorb some of the private risk of infections in socially intensive jobs, we include the descriptor "exposure to diseases" as a separate variable. Exposure describes the frequency of exposure to diseases, which captures the salient infection risk in the work environment. High-exposure workers are typically health care personnel (e.g., acute care nurses, dental hygienists, and general practitioners), but non-health care personnel can also have high exposure (e.g., fire fighters and police officers).[19] By explicitly controlling for "exposure to diseases" among workers in both health care and non-health care occupations, we focus our interpretation of "social intensity" as a measure of the external benefits of vaccination.

We view the index formed by combining all variables using weights from the first principal component as the preferred index. However, it is possible that including all variables in the index may mask relationships between key subsets of those occupational descriptors and vaccination uptake. To address that possibility, we construct three alternative indices for the social intensity of an individual's occupation, with the weights described in Table 3: one index that only contains information from the descriptors about work context, one index that only contains information on work activities, and a third alternative index that only contains the descriptors related specifically to physical contact.

TABLE 1 | Vaccination coverage for groups of selected characteristics—2013–2015, 31 states[a].

Characteristic	Observations	Weighted vaccination coverage	Std. Dev.
Total	229,144	0.34	0.47
Health care personnel[b]			
Yes	36,347	0.57	0.50
No	192,797	0.30	0.46
Age			
18–24	11,075	0.23	0.42
25–29	13,959	0.26	0.44
30–34	17,961	0.31	0.46
35–39	20,163	0.31	0.46
40–44	22,616	0.31	0.46
45–49	26,162	0.33	0.47
50–54	32,545	0.37	0.48
55–59	34,022	0.41	0.49
60–64	26,918	0.46	0.50
65–69	13,629	0.52	0.50
70–74	5989	0.57	0.50
75–79	2523	0.56	0.50
80+	1582	0.56	0.50
Sex			
Female	120,487	0.39	0.49
Male	108,657	0.29	0.46
Race/ethnicity			
White	186,699	0.36	0.48
Black	14,886	0.27	0.44
Hispanic	16,389	0.27	0.45
Other	11,170	0.37	0.48
Education			
Less than high school	9532	0.23	0.42
High school graduate or GED	52,287	0.27	0.44
Some college or technical school	63,463	0.32	0.47
College graduate	103,862	0.44	0.50
Income			
< 10,000	4426	0.23	0.42
10,000 to < 15,000	5587	0.23	0.42
15,000 to < 20,000	10,247	0.23	0.42
20,000 to < 25,000	14,801	0.26	0.44
25,000 to < 35,000	20,540	0.27	0.44
35,000 to < 50,000	32,218	0.30	0.46
50,000 to < 75,000	42,556	0.34	0.47
≥ 75,000	98,769	0.42	0.49
Marital status			
Married	139,613	0.39	0.49
Divorced	31,269	0.31	0.46

(Continues)

TABLE 1 | (Continued)

Characteristic	Observations	Weighted vaccination coverage	Std. Dev.
Widowed	8977	0.42	0.49
Separated	4710	0.28	0.45
Never married	36,570	0.25	0.43
Unmarried couple	8005	0.27	0.45
Number of children			
0	141,130	0.35	0.48
1	34,402	0.33	0.47
2	33,090	0.35	0.48
3	13,705	0.31	0.46
4	4746	0.31	0.46
5 or more	2071	0.26	0.44
High-risk conditions[c]			
Yes	52,375	0.43	0.49
No	176,769	0.32	0.47
Have personal health care provider			
Yes	185,183	0.39	0.49
No	43,961	0.18	0.38
Have medical insurance			
Yes	206,381	0.37	0.48
No	22,763	0.14	0.35

Note: Sample is restricted to employed, civilian adults whose occupation and industry information were available. All statistics are weighted using BRFSS weights.
[a]31 states are California, Colorado, Connecticut, Florida, Georgia, Idaho, Illinois, Iowa, Louisiana, Maryland, Missouri, Massachusetts, Michigan, Minnesota, Mississippi, Montana, Nebraska, New Hampshire, New Jersey, New Mexico, New York, North Carolina, North Dakota, Oregon, Tennessee, Utah, Vermont, Washington, West Virginia, Wisconsin, and Wyoming. Not all states collect occupational data in 2013–2015. See state-year crosswalk details in Appendix A.1.
[b]"Health care personnel" are clinical and nonclinical staff working in hospitals (NAICS 622), outpatient care and physician offices (NAICS 6214 and 6211), long-term care facilities (NAICS 6216, 6231, 6232, 6233, and 6239), and other clinical settings (NAICS 6212, 62131, 62132, 6213, 6215, and 6219).
[c]"High-risk conditions" are people who reported having at least one of the following: asthma, diabetes, myocardial infarction, angina or coronary heart disease, chronic obstructive pulmonary disease, emphysema or chronic bronchitis, or cancer (excluding skin cancer).
Source: BRFSS 2013-2015.

We also consider the effects of workers' industry, specifically health care. The vaccination decision for health care workers may be fundamentally different because these workplaces are more likely to require employees to obtain a flu vaccine.[20] Additionally, the mixture of private and external benefits from vaccination in socially intense occupations likely differs between health care and non health care occupations. We therefore follow O'Halloran et al. (2017) and define health care personnel as people working in a clinical setting based on the 2002 North American Industry Classification System (NAICS) industry codes.[21] Figure 1 contains a histogram of the number of workers by major occupation groups, which are ranked from the highest average intensity. Each group is classified by those who work in clinical settings and those who work in non-clinical settings. All major occupation groups except "farming, fishing, and forestry" include some health care personnel (HCP).

Table 4 provides some intuition for what kinds of occupations are more/less socially intensive by depicting the 20 occupations with the highest and lowest ratings for social intensity using the index comprised of all descriptors. For this comprehensive index, note that few of the highest rated occupations are in the health care sector. Appendix Tables A5–A7 provide similar lists for indices based on context, activities, and physical contact, respectively. Because health care occupations are so highly represented among the top occupations for context and physical contact, we separately delineate the top occupations for health care and non-health care sectors. This separation provides additional context for split-sample analysis where we divide the sample into health care and non-health care workers.

Figures 2–4 plot the flu vaccination coverage by the social intensity of 22 major occupation groups, of 93 minor occupation groups, and of 509 SOCs (mostly detailed-level occupations). The blue dots represent occupations or occupation groups whose proportion of health care personnel is over 50%, while the magenta dots in Figures 3 and 4 represent those under 50%. All figures include two upward-sloping fitted lines, with the solid one using all occupations and the dashed one excluding occupations with a majority share of health care workers. The upward-sloping lines show that workers' social intensity and flu vaccination coverage are positively correlated. Also, the dashed lines are flatter than the corresponding solid lines, which shows that the association between social intensity and vaccine uptake is smaller among non-health care personnel.

TABLE 2 | Summary statistics of occupational descriptors and derived measures from O*NET.

Descriptors	Mean	Std. Dev.	Min	Max
Panel A. Occupational descriptors				
Exposed to disease or infections	24.73	26.09	0	100
Work context				
Physical proximity	60.55	15.41	17	100
Contact with others	83.51	10.73	50	100
Coordinate or lead others	61.85	13.24	15	91
Deal with external customers	59.62	21.68	2	99
Face-to-face discussions	88.21	8.435	39	100
Work with work group or team	78.10	11.17	25	100
Work activities				
Assisting and caring for others	45.46	18.56	10	97.25
Coaching and developing others	48.08	13.08	7	87
Communicating with persons outside organization	54.56	18.81	8	94
Communicating with supervisors, peers, or subordinates	73.13	10.16	34	96
Coordinating the work and activities of others	51.07	13.31	9	89
Developing and building teams	48.94	13.75	10	89
Establishing and maintaining interpersonal relationships	65.46	12.61	21	97
Guiding, directing, and motivating subordinates	44.97	15.03	6	90
Interpreting the meaning of information for others	53.95	14.82	19	93
Performing administrative activities	43.44	15.38	6	89
Performing for or working directly with the public	48.07	23.28	2	94
Provide consultation and advice to others	44.54	14.10	10	96
Resolving conflicts and negotiating with others	49.80	15.23	5	93
Selling or influencing others	37.36	17.13	5	99
Staffing organizational units	27.00	14.07	3	93
Training and teaching others	54.11	12.55	8	91.06
Panel B. Measures of disease exposure and social intensity				
Exposure	0	1	−0.95	2.89
Intensity	0.89	3.07	−7.91	8.41
Intensity (context)	0.56	2.23	−7.13	3.73
Intensity (activity)	0.72	2.89	−7.88	8.77
Intensity (contact)	0.58	1.48	−4.34	4.21

Note: Panel A presents the original occupational descriptors, and Panel B includes the derived measures. Exposure is the standardized variable of Exposed to Disease or Infections. Intensity measures are constructed by applying PCA to different subsets of the descriptors explained in Table 3. Column "mean" is average of the descriptor score; "Std. Dev." is the standard deviation; "Min" is the minimum; "Max" is the maximum For descriptors under Work Activities, the "importance" score is reported.
Source: O*NET.

3 | Methods

For our regression analysis, our main specification is a linear probability model. Specifically, we run the following regression:

$$\begin{aligned}\mathbf{1}\{y_{ikst}=1\} = \beta_0 + \beta_1 SI_k + \beta_2 Ex_k + \beta_3 \mathbf{1}\{hcp_i = 1\} \\ + \beta_4 SI_k \times \mathbf{1}\{hcp_i = 1\} + \boldsymbol{\gamma}\mathbf{X}_i + \mu_s \\ + \nu_t + s_i \cdot \eta_t + \epsilon_{ikst}\end{aligned} \quad (1)$$

where $\mathbf{1}\{y_{ikst} = 1\}$ is an indicator for whether person i working in occupation k in state s gets a flu vaccine in season t, SI_k is the occupational social intensity measure for the individual, Ex_k is the standardized measure of exposure to diseases of the occupation, $\mathbf{1}\{hcp_i = 1\}$ is the indicator for whether the individual works in a clinical setting, $\mathbf{X}_i$ are demographic, education, income and health-related control variables (listed in table notes), μ_s and ν_t are state and flu season fixed effects respectively, $s_i \cdot \eta_t$ are interactions of those fixed effects and ϵ_{ikst} is the error term. We report estimates with BRFSS sampling weights and errors clustered at the state level.

We recognize the associative relationship between occupational intensity and vaccine choices may not be linear. We therefore

TABLE 3 | Weights (%) used for principal component analysis of the occupational descriptors for the construction of measures of occupational social intensity.

Descriptors	Intensity	Context	Activity	Contact
Work context				
Physical proximity	1.12	11.17		20.97
Contact with others	3.85	19.27		22.82
Coordinate or lead others	4.80	18.95		
Deal with external customers	3.75	15.01		
Face-to-face discussions	3.57	16.71		12.67
Work with work group or team	4.14	18.90		
Work activities				
Assisting and caring for others	3.16		3.62	22.26
Coaching and developing others	5.57		7.33	
Communicating with persons outside organization	4.83		5.93	
Communicating with supervisors, peers, or subordinates	4.96		6.23	
Coordinating the work and activities of others	5.57		7.19	
Developing and building teams	5.89		7.62	
Establishing and maintaining interpersonal relationships	5.33		6.55	
Guiding, directing, and motivating subordinates	5.59		7.33	
Interpreting the meaning of information for others	4.69		6.05	
Performing administrative activities	4.75		5.88	
Performing for or working directly with the public	3.39		3.89	21.27
Provide consultation and advice to others	5.29		7.06	
Resolving conflicts and negotiating with others	5.64		7.00	
Selling or influencing others	4.02		5.08	
Staffing organizational units	5.60		7.30	
Training and teaching others	4.47		5.94	

Note: The table shows each descriptor's share of the occupational social intensity in percentage and may not sum up to 100 because of rounding. "Exposed to Disease or Infections" is an independent variable in the regression model to measure the salient infection risk so excluded from the construction of social intensity. For descriptors under Work Activities, the "importance" score is reported. All descriptors standardized before applying the principal component analysis.
Source: O*NET.

also employ a specification that includes indicators for whether the intensity measure falls in a particular range of values rather than treating intensity as a continuous variable. With intensity bins, we flexibly allow the coefficients to vary over the support of intensity. We use discrete changes on the integer level to categorize intensity into seventeen bins in $(-8, 9)$. Specifically, we run the regression with intensity bins:

$$\mathbf{1}\{y_{ikst} = 1\} = \beta_0 + \sum_j \beta_1^j \mathbf{1}\{bin_{kj} = 1\} + \beta_2 Ex_k + \beta_3 \mathbf{1}\{hcp_i = 1\} + \boldsymbol{\gamma}\mathbf{X}_i + \mu_s + \nu_t + s_i \cdot \eta_t + \epsilon_{ikst} \quad (2)$$

where $\mathbf{1}\{bin_{kj} = 1\}$ are indicators for whether the occupational social intensity falls into a particular bin j for every bin but the base level (i.e., [0,1]), and the same notations are used as in the above Equation (1). We also report estimates with BRFSS sampling weights and state-level cluster-robust standard errors.

For completeness, we also implemented a LASSO approach to determine if any of the 22 O*NET variables included by themselves would predict vaccination. None did. We do not report these results in any table as all of the variables of interest were dropped in the variable selection process.

We consider three factors that may affect estimates of the association between intensity and the probability of obtaining a vaccine. First, we control for heterogeneity in the vaccination cost by controlling for demographics and health-related characteristics. For instance, having health insurance significantly reduces the cost of obtaining a flu vaccine (typically for free). Second, demographics and health-related characteristics also control for heterogeneity in the cost of catching the flu. For example, if people aged 65+ or those with high-risk conditions contract flu, they are more likely to get hospitalized and thus encounter a greater cost than younger and healthier people. Last, we control for heterogeneity in the prevalence of influenza and vaccine effectiveness by including state and flu season fixed effects.

Finally, there may be unobservable factors that simultaneously influence vaccine uptake and career choice. For example,

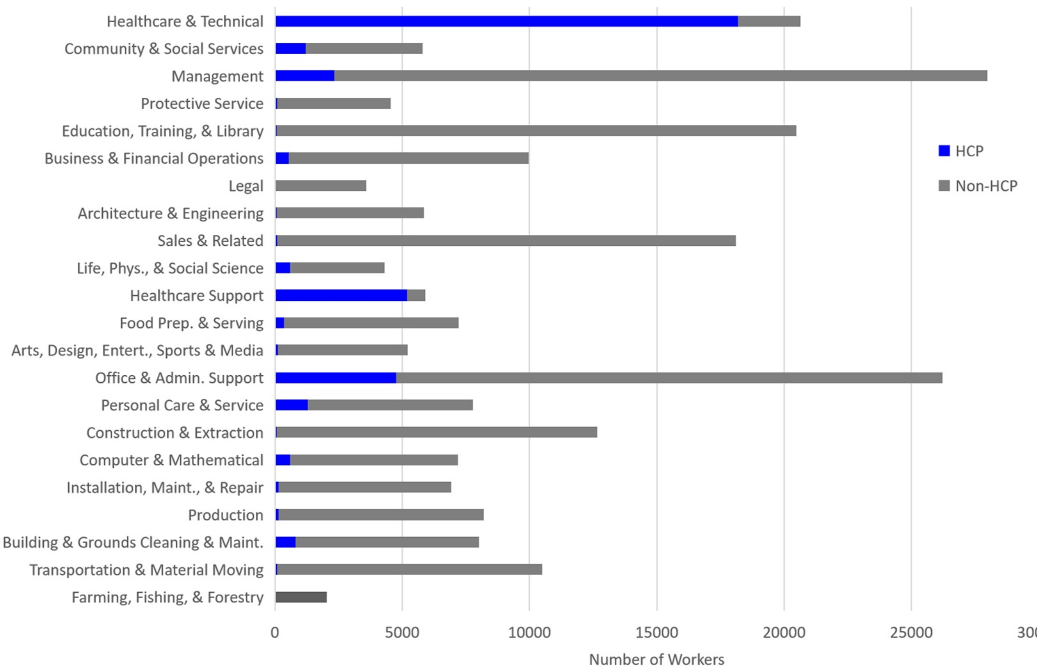

FIGURE 1 | The number of workers of 22 major occupation groups. Groups are ordered from the group with the highest weighted sample average social intensity. The blue pieces represent the number of health care personnel (HCP) who works in a clinical setting (unweighted). *Source:* BRFSS 2013–2015 and O*NET. [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 4 | Occupations with highest and lowest ratings of social intensity—Comprehensive index.

Highest rated		Lowest rated	
Occupation	**Score**	**Occupation**	**Score**
Supervisors of non-retail sales workers	8.41	Cut/Punch/Press machine operators	−5.16
Supervisors of police/detectives	8.16	Structural metal fabricators and fitters	−5.21
Human resource managers, all other	8.01	Automotive body repairers	−5.38
Supervisors of protective service workers	7.33	Motion picture projectionists	−5.43
Supervisors of fire fighting/prevention workers	7.30	Printing machine operators	−5.60
Directors of religious activities/education	7.26	Couriers and messengers	−5.77
Lodging managers	7.22	Machine feeders and offbearers	−5.98
Meeting/Convention planners	6.92	Cleaners of vehicles and equipment	−6.26
Postmasters and mail superintendents	6.84	Woodworking machine setters/operators	−6.26
Education administrators	6.67	Tire builders	−6.46
Supervisors of correctional officers	6.52	Sewing machine operators	−6.50
Chief executives	6.43	Animal breeders	−6.55
Dietitians and nutritionists	6.38	Textile knitting/weaving machine operators	−6.66
General/Operations managers	6.31	Shoe machine operators/tenders	−6.70
Urban/Regional planners	6.28	Shoe/Leather workers and repairers	−6.78
Medical/Health services managers	6.25	Postal service mail carriers	−6.95
Gaming managers	6.23	Pressers/Textile/Garment materials	−7.02
Dentists	6.21	Proofreaders and copy markers	−7.32
Human resources managers	6.15	Coin/Vending/Amusement machine servicers	−7.47
Religious workers, all other	6.00	Hunters and trappers	−7.91

Note: The left panel lists 20 occupations with the highest social intensity score, while the right panel lists the lowest 20 occupations. All "supervisors" refer to the first-line supervisors/managers of corresponding workers. The score is constructed by applying the principal component analysis to all 22 descriptors (excluding "Exposed to Disease or Infections"). Table 3 presents the weights of each descriptor.
Source: O*NET.

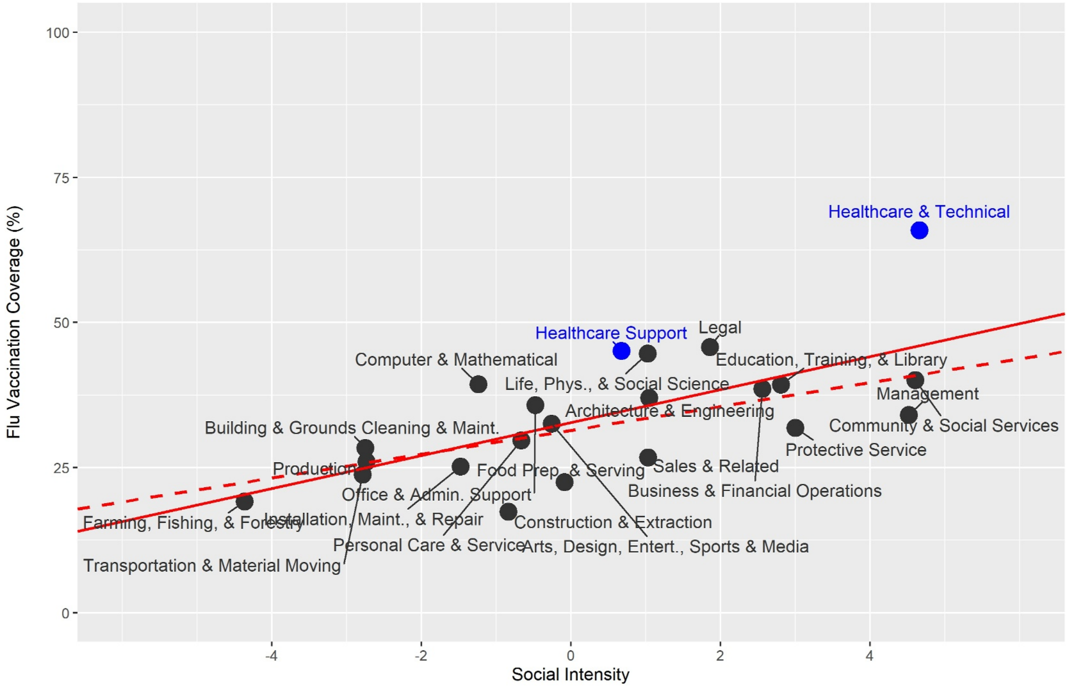

FIGURE 2 | The flu vaccination coverage by social intensity of 22 major occupation groups. The blue dots represent major occupation groups whose proportion of health care personnel is over 50%. The two red fitted lines include a solid one using all occupations and the dashed one excluding occupations with a majority share of health care workers. All graphs in the rest of the paper are created by tools from R Core Team (2018) and Wickham (2016). *Source:* BRFSS 2013–2015 and O*NET. [Colour figure can be viewed at wileyonlinelibrary.com]

certain occupations require workers to get vaccinated (e.g., medical professionals). Therefore, individuals who hold strong anti-vaccine beliefs may be less likely to choose these occupations. Moreover, if these anti-vaccine individuals randomly self-select into non-health care occupations, their average social intensity will be lower than the health care workers' because health care personnel's intensity is relatively higher. Therefore, if self-selection into occupation biases these estimates, that bias may lead these estimates to *overstate* the relationship between social intensity and vaccine uptake. That is, if vaccination and intensity are positively correlated, and "positive self-selection" is present, we overestimate a positive coefficient. To empirically evaluate the consequences of potential selection effects, we implement bounding procedures from Oster (2019).

4 | Results

4.1 | Main Results

Table 5 reports the main results about the association between occupational social intensity and flu vaccine uptake. Columns (1)–(4) use the full sample. Columns (5) and (6) split the sample into health care and non-health care personnel. We find that the coefficient of social intensity is economically insignificant, but being exposed to diseases and working at a clinic setting are strongly correlated with a worker's flu vaccine uptake.

At first glance, columns (1)–(2) indicate that there is a statistically significant relationship between social intensity and vaccine uptake. However, as columns (3)–(6) add controls for each occupation's exposure to diseases, that relationship weakens in both magnitude and statistical significance. The main specification is presented in column (3), and our variable of interest is intensity. We multiple the coefficient of intensity, 0.0012, by its weighted standard deviation, 3.07, and obtain 0.0037.[22] That is, when comparing two workers, if one's occupational social intensity exceeds the other's by one standard deviation, that worker is 0.37 percentage points more likely to be vaccinated. By contrast, for the same change in the magnitude of exposure to diseases, the rise in vaccine uptake likelihood is 2.59 percentage points on average.[23] These results indicate that the association between social intensity - individuals' propensity to catch and spread the flu at work - and vaccine uptake is close to a precise zero. The association between the personal risk of exposure to diseases and vaccine uptake is both strongly positive and persistently statistically significant across all specifications - both within and outside of health care contexts.[24]

To make sense of the magnitude of social intensity coefficient, we can examine an example of food preparation workers. A food preparation worker, whose social intensity score is 1.4, prepares ingredients for recipes and assists chefs and cooks in restaurants, canteens, and grocery stores. If we look for workers whose intensity is one standard deviation higher, we find public relations managers (4.54, as their social intensity score, and the same applies for the next three parentheses) and supervisors of landscaping workers (4.48). Then we predict that their vaccination coverage is roughly 0.37 percentage points higher than food preparation workers', ceteris paribus. Similarly, if we search for workers whose intensity is one standard deviation lower, we find database administrators (−1.64) and protective service workers

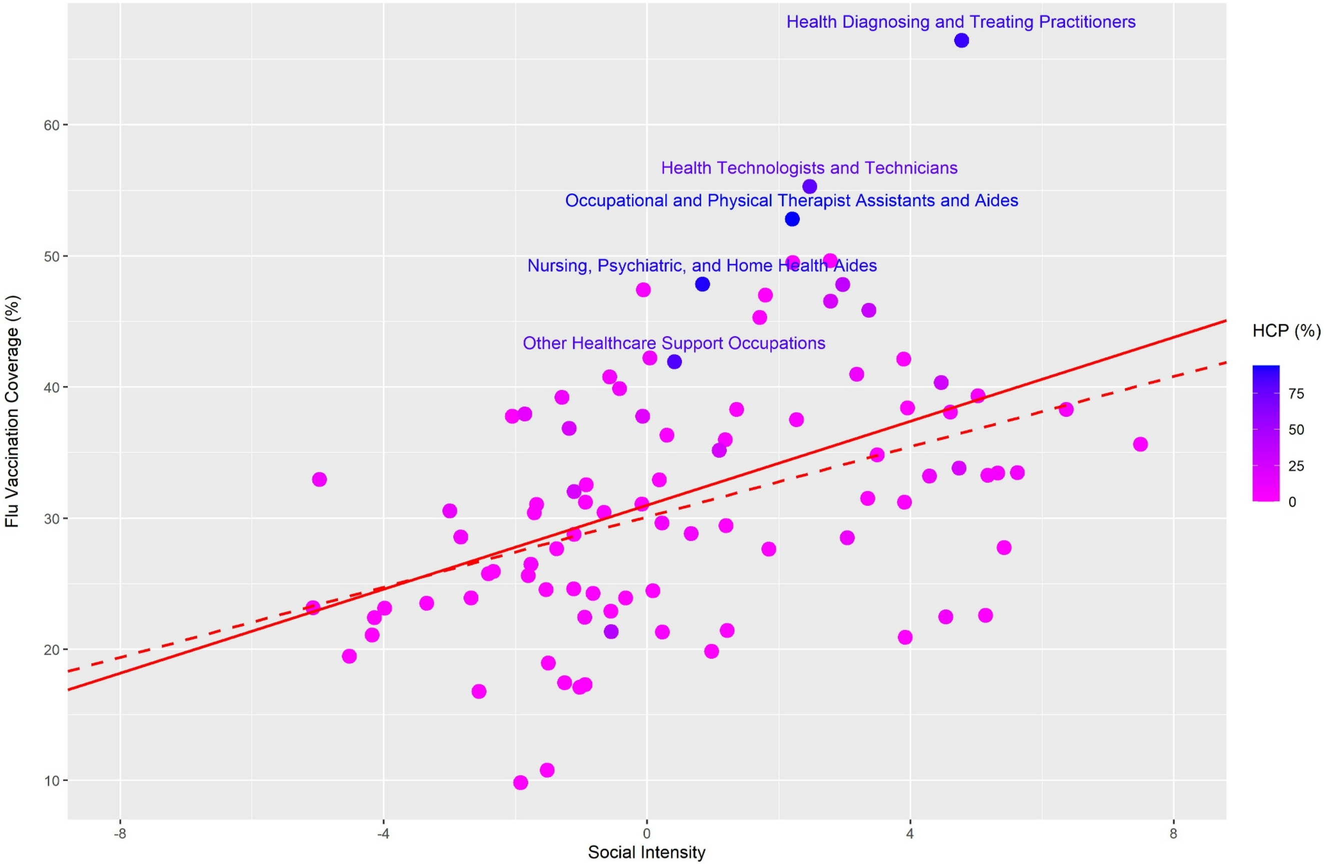

FIGURE 3 | The flu vaccination coverage by social intensity of 93 minor occupation groups. The blue dots with names displayed represent minor occupation groups whose proportion of health care personnel is over 50%, while the magenta dots represent those under 50%. The two red fitted lines include a solid one using all occupations and the dashed one excluding occupations with a majority share of health care workers. *Source:* BRFSS 2013–2015 and O*NET. [Colour figure can be viewed at wileyonlinelibrary.com]

(−1.67). Then we predict that their vaccination rate is 0.37 percentage points lower than food preparation workers', ceteris paribus.[25]

The association between being employed in health care and vaccine uptake is large in magnitude and statistically significant at 1% level in columns (1)–(4). Estimates in columns (3)–(4) show that workers in a clinical setting are 18 percentage points more likely to be vaccinated than workers in other settings. Vaccination requirements at health care facilities might be an important reason that promotes flu vaccine uptake. Therefore, our results suggest that workplace requirements for flu vaccination potentially provide a much stronger motivation than individuals' internalizing the external social benefits of their vaccination choices.

4.2 | Specifications With Greater Flexibility in Social Intensity and Vaccination

We also reanalyze our main results using discrete bins of social intensity rather than forcing social intensity to enter the model linearly. Appendix Table A12 presents estimates from the specification using indicators for whether the value of the comprehensive social intensity index falls in a discrete bin rather than the linear specification. All columns include seventeen bins for intensity while omitting the base level bin [0,1). Column (1) does not include exposure but columns (2)–(4) do. Each intensity bin coefficient is the point estimate of the difference in the probability of flu vaccine uptake between workers in that bin and workers at the base level. For ease of interpretation, we have included plots of our two preferred specifications here. Figure 5 plots column (2) of the binned estimates for social intensity, controlling for exposure to disease. Figure 6 plots column (4) of the same binned estimates, but for non-health care personnel only.[26]

These included figures and Table A12 provides three new insights. First, there is neither a clear monotonically increasing relationship nor a general increasing trend over the discrete bins of social intensity. However, workers with the absolute highest levels of social intensity are significantly more likely to get vaccinated than workers with merely above-average intensity. For instance, in Figure 5, the coefficient of bin [8, 9) is 8.31% greater than the base level and statistically significant. Third, the coefficients for the bins that comprise most of the distribution are relatively small and in many cases statistically insignificant. In column (2), the difference in the coefficients of bins from −5 to 5, which contain 86.8% of all workers (91.0% of non-health care workers) all fall within the [−0.005, 0.04] interval with no clear trend. These insights show that the linear

FIGURE 4 | The flu vaccination coverage by social intensity of 509 SOCs. The 509 Standard Occupational Classification codes (SOCs) includ 371 detailed occupations, 126 broad occupations, and 12 minor groups. The blue dots represent occupations whose proportion of health care personnel is over 50%, while the magenta dots represent those under 50%. The two red fitted lines include a solid one using all occupations and the dashed one excluding occupations with a majority share of health care workers. *Source:* BRFSS 2013–2015 and O*NET. [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 5 | Main results: Association between occupational social intensity and flu vaccine uptake.

	(1) Full	(2) Full	(3) Full	(4) Full	(5) HCP	(6) Non-HCP
Intensity	0.0023***	0.0013**	0.0012**	0.0008	0.0010	0.0011*
	(0.0006)	(0.0006)	(0.0005)	(0.0006)	(0.0017)	(0.0006)
Exposure			0.0259***	0.0249***	0.0306***	0.0218***
			(0.0023)	(0.0024)	(0.0047)	(0.0027)
Health care personnel	0.2212***	0.2046***	0.1808***	0.1740***		
	(0.0073)	(0.0078)	(0.0074)	(0.0078)		
Intensity × HCP		0.0075***		0.0038**		
		(0.0015)		(0.0015)		
Constant	0.9551***	0.9530***	0.9492***	0.9483***	0.2486***	0.9628***
	(0.0194)	(0.0190)	(0.0193)	(0.0191)	(0.0671)	(0.0189)
Observations	229,144	229,144	229,144	229,144	36,347	192,797

Note: Columns (1)–(4) use full sample. Column (5) uses health care personnel only, while column (6) excludes them. The full sample covers 22 major occupation groups, 93 minor occupation groups, and 509 detailed-level occupations in Standard Occupational Classification 2000. Intensity is constructed by applying the principal component analysis method to 22 descriptors related to social interactions; Table 2 presents the statistics of each descriptor, and Table 3 shows their weights. Exposure is the standardized index of disease exposure at work. We control for sex, age, race/ethnicity, education, income, marital status, number of children, and indicators for high-risk conditions, having personal health care provider, and having medical insurance. We also control for the interaction of state and flu season fixed effects. ***, ***, * mean that the coefficient is statistically different from zero at the 1%, 5%, and 10% level. All models weighted with BRFSS weights. Standard errors clustered at the state-level and reported in parentheses.
Source: BRFSS 2013–2015 and O*NET.

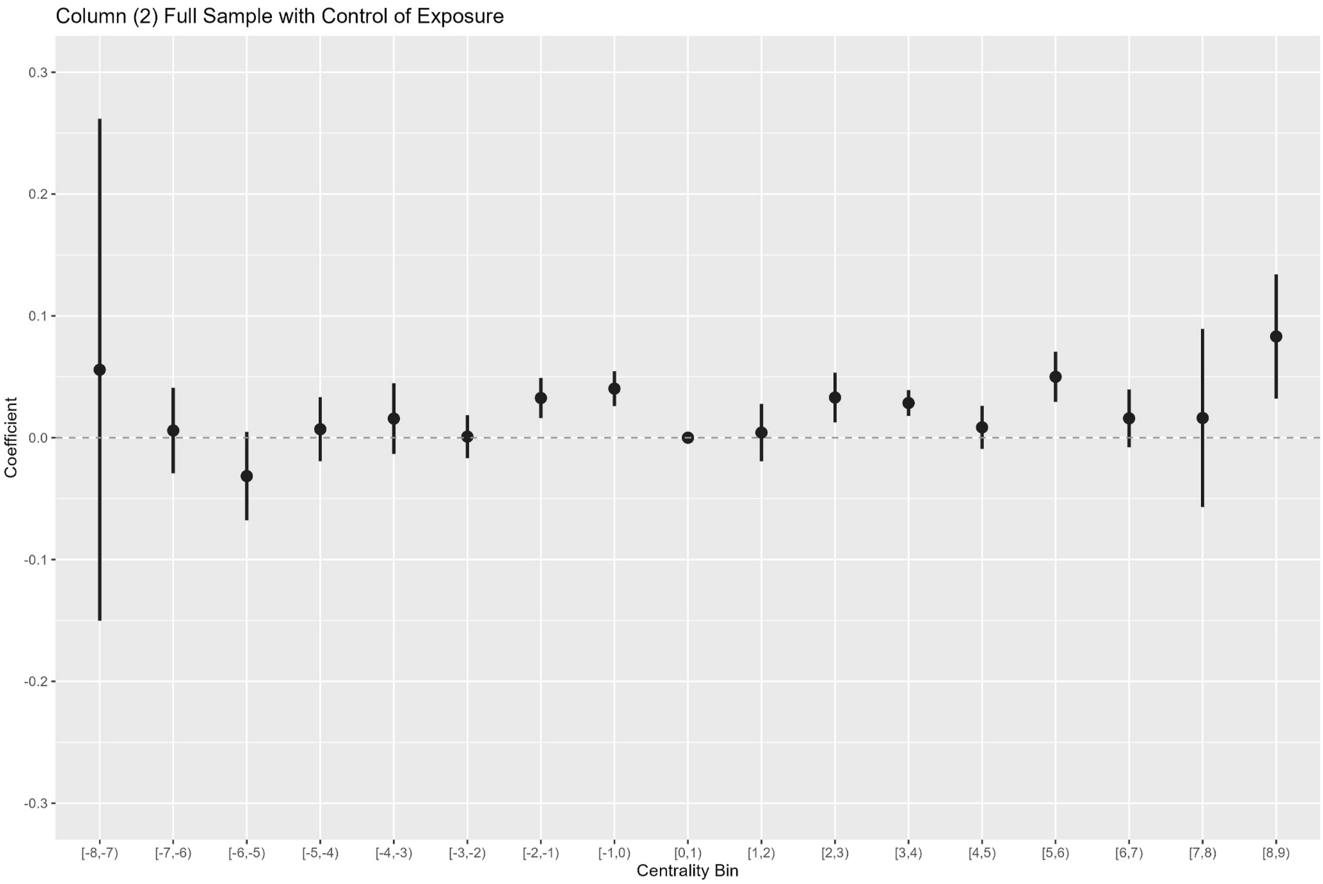

FIGURE 5 | Coefficient and 95% confidence interval estimates of intensity bins in Table A12. Column (2).

functional form for social intensity does not mask a more interesting non-parametric relationship. Across most workers, there is no relationship between social intensity and vaccination. This is consistent with workers not considering the external benefits of their vaccination choices unless those benefits are very salient. We again acknowledge that social intensity, even when controlling for exposure to disease, may capture some information regarding the marginal private risk of infection. However, because these results are generally null, this further underscores the implication that workers do not strongly consider potential external benefits in their vaccination choices.

4.3 | The Role of Education and Income

The results in Table 5 indicate that although there is an association between higher exposure to disease and the individual's vaccination decision, the relationship between the social intensity and vaccination is much weaker. This is consistent with previous work in the economics literature that individuals fail to internalize the external benefits of getting vaccinated.

However, there could be other mechanisms at work, including the possibility that individuals in socially intensive occupations are not well informed about the external benefits of their vaccination choices at a personal level. We conduct two sets of analysis to examine this possibility. First, in Table 6, we begin with a specification with no control variables in column (1) and gradually add controls to replicate our full sample estimates from our main specification in column (4) of Table 5. There are three main takeaways from these results. First, when there are no control variables, there is a much stronger association between social intensity and vaccination, consistent with the unconditional correlations shown in Figures 2 and 3. Second, when we control for education and income, the estimated coefficient decreases from 0.0093 in column (2) to 0.0035 in column (3) and 0.0016 in column (4). This indicates that the unconditional relationship between social intensity and vaccination is likely driven by individuals with more education are likely to select into socially intensive occupations and get vaccinated. Third, note that the estimated coefficients on exposure or working in health care do not substantively change when adding controls for education and income. While not definitive, these results suggest that education may play a role in individuals being aware of the external benefits of being vaccinated, but that all individuals may be relatively aware of the private benefits of their vaccination decisions.[27]

To further examine the role of education and income, we split the sample by education and household income level into four groups each. Results are shown in Table 7. In the samples split by education level, the coefficient on social intensity is only significant for those with some college, but not a bachelor's degree. However, the estimated coefficient does not increase monotonically with education as we would suspect if information limitations in and of themselves were responsible for the weak relationship between social intensity and vaccination

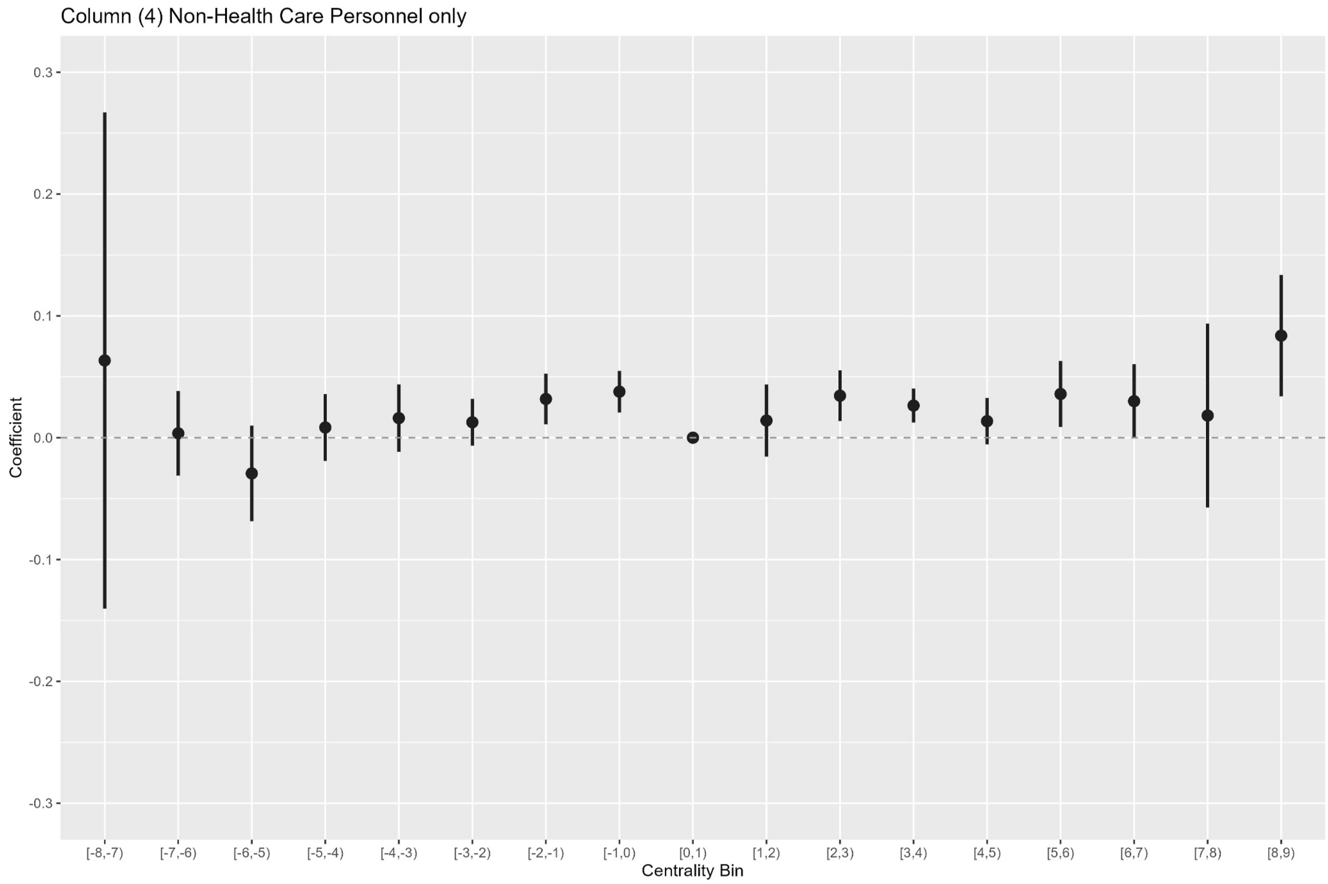

FIGURE 6 | Coefficient and 95% confidence interval estimates of intensity bins in Table A12. Column (4).

decisions. For the income subsamples in columns (5)–(8), note that the relationship between social intensity and vaccination probability *does* monotonically increase with income. Whether this is directly attributable to income per se, or other unobservable factors correlated with income is unknown, but is an area for future work.

We also conduct analysis on subsets of the sample, where we split the sample by key observable characteristics. Table A14 splits the sample by sex, age, and race/ethnicity. We find the coefficient for social intensity is positive for men, but not women. Among the age groups considered, the coefficient on intensity is positive only for "older working age" individuals aged 50–64. Splitting the sample by race, the coefficient on intensity is slightly larger for black workers, but only significant for white workers. Table A15 splits the sample by marital status and whether the respondent has children. Intensity is not significant for any of the subsamples delineated by marriage or cohabitation, and is only significant for childless individuals. Finally, Appendix Table A16 splits the sample by whether individuals have high-risk health conditions, have a primary care provider, and whether they have health insurance. The coefficient on intensity is only significant for individuals without high-risk conditions, with a health care provider, and with insurance. All of the coefficients that *are* significant are similar in magnitude to the main results. However, the results in Table A16 do indicate that if social intensity matters, it does only when vaccination is discretionary and at a low time-and out-of-pocket cost to the individual.

5 | Alternative Specifications and Robustness

Because these estimates are conditional associations rather than causal estimates, we may be concerned about the confounding effects of non-random selection into occupations. To address that concern, we implement the bounding approach developed by Altonji et al. (2005) and Oster (2019) to assess how important that non-random selection may be. Both approaches consider two regressions, a potentially biased regression without unobservable controls and a hypothetical one with them. The underlying assumption behind these bounding approaches is that the selection on unobservables is likely proportional in some way to the selection on observable variables. Therefore, the change in R-squared from a univariate regression to one including controls can be used to approximate the R-squared of a hypothetical regression in which *unobservable* selection variables were somehow included.

As an example of proportional selection, we cannot observe individuals' social intensity after work, but it is potentially correlated with their social intensity at work if individuals select into these jobs based on introversion/extroversion. If individuals consider their social interactions when making flu vaccination decisions, the assumption behind these bounding approaches implies that workers place equal weight (or in a proportion of the other) on the risk of social interaction in the workplace and elsewhere (if no workplace requirements). If this is valid, the intensities of the social interactions at work

TABLE 6 | Sensitivity analysis: Column (4) with state-by-season fixed effects.

	(1) Full	(2) Full	(3) Full	(4) Full	(5) Full	(6) Full
Intensity	0.0118***	0.0093***	0.0035***	0.0016***	0.0016***	0.0008
	(0.0007)	(0.0007)	(0.0006)	(0.0006)	(0.0006)	(0.0006)
Exposure	0.0241***	0.0235***	0.0248***	0.0285***	0.0274***	0.0249***
	(0.0024)	(0.0022)	(0.0022)	(0.0022)	(0.0022)	(0.0024)
Health care personnel	0.1904***	0.1832***	0.1813***	0.1792***	0.1806***	0.1740***
	(0.0072)	(0.0080)	(0.0081)	(0.0079)	(0.0080)	(0.0078)
Intensity × HCP	0.0043**	0.0042**	0.0032**	0.0015	0.0016	0.0038**
	(0.0016)	(0.0017)	(0.0015)	(0.0015)	(0.0015)	(0.0015)
Constant	0.9912***	0.9381***	0.9123***	0.9552***	1.0140***	0.9483***
	(0.0005)	(0.0083)	(0.0109)	(0.0168)	(0.0193)	(0.0191)
Sex, age, race		X	X	X	X	X
Education			X	X	X	X
Income				X	X	X
Marital, #children					X	X
Health-related						X
State × season FE	X	X	X	X	X	X
Observations	229,144	229,144	229,144	229,144	229,144	229,144
R^2	0.073	0.101	0.112	0.117	0.120	0.138

Note: While column (1) does not control for demographics, columns (2)–(6) control various sets of sex, age, race/ethnicity, education, income, marital status, number of children, and health-related indicators (tagged by X). All columns use full sample and control for the interaction of state and flu season fixed effects.
***, ***, * mean that the coefficient is statistically different from zero at the 1%, 5%, and 10% level. All models weighted with BRFSS weights. Standard errors clustered at the state-level and reported in parentheses.
Source: BRFSS 2013–2015 and O*NET.

TABLE 7 | Sub-sample analysis: Column (4) by education, income.

	(1) < High school	(2) High school	(3) Some college	(4) College +	(5) < 20K	(6) 20–50K	(7) 50–75K	(8) > 75K
Intensity	−0.0033	0.0003	0.0029**	0.0004	−0.0030	0.0003	0.0003	0.0024**
	(0.0020)	(0.0012)	(0.0013)	(0.0011)	(0.0020)	(0.0008)	(0.0011)	(0.0010)
Exposure	0.0304**	0.0158***	0.0212***	0.0296***	0.0188**	0.0275***	0.0241***	0.0262***
	(0.0116)	(0.0040)	(0.0051)	(0.0039)	(0.0085)	(0.0025)	(0.0047)	(0.0037)
Health care personnel	0.1020***	0.1717***	0.2054***	0.1523***	0.1044***	0.1830***	0.1698***	0.1905***
	(0.0292)	(0.0153)	(0.0094)	(0.0122)	(0.0261)	(0.0098)	(0.0186)	(0.0176)
Intensity × HCP	−0.0019	−0.0056	0.0018	0.0062**	−0.0017	−0.0041	0.0093**	0.0005
	(0.0115)	(0.0049)	(0.0033)	(0.0029)	(0.0078)	(0.0029)	(0.0041)	(0.0031)
Constant	0.9223***	0.9537***	0.9027***	0.8861***	0.5038***	0.9304***	0.6271***	0.5641***
	(0.0667)	(0.0369)	(0.0408)	(0.0317)	(0.0649)	(0.0324)	(0.0450)	(0.0308)
Observations	9532	52,287	63,463	103,862	20,260	67,559	42,556	98,769
R^2	0.147	0.113	0.133	0.120	0.109	0.131	0.128	0.126

Note: The sample is divided into subsamples by demographics: columns (1)–(4) by education; columns (5)–(8) by income. We control for sex, age, race/ethnicity, education, income, marital status, number of children, and indicators for high-risk conditions, having personal health care provider, and having medical insurance, unless the variable is used as a label for the subsample. We also control for the interaction of state and flu season fixed effects.
***, ***, * mean that the coefficient is statistically different from zero at the 1%, 5%, and 10% level. All models weighted with BRFSS weights. Standard errors clustered at the state-level and reported in parentheses.
Source: BRFSS 2013–2015 and O*NET.

and after work should affect her vaccination decision in similar (or proportional) magnitudes. Under these conditions, the relationship for an individual's social activeness at work and vaccination uptake is proportionally correlated with the relationship between her social activeness after work and vaccination uptake.

Table 8 presents the estimates using Oster's method. Column (3) shows the bounded interval of the effect of social intensity under the recommended parameters from Oster (2019), assuming $\delta = 1$ and $R_{Max} = 1.3R^2$. Column (4) shows the degree of proportional selection on unobservables versus observables (δ) needed to make the coefficient of intensity zero.[28] We find that all bounding intervals in column (3) contain zero, except the results for non-health care personnel. These results imply that non-random selection into occupations does not cover up a ""true" positive relationship between social intensity and vaccination decisions, and yields a false null result. For instance, in the first specification with the full sample but not controlling for HCP, the coefficient was statistically significant at the 5% level, but it would be insignificant if there was equal selection on observables and unobservables. Moreover, all estimates of δ in column (4) are less than one (specifically 0.02), which means that only a relatively small amount of selection on unobservables versus observables is needed to make the coefficient of intensity zero. However, this is unsurprising as the estimates for social intensity were already very close to zero. Most interestingly, results from the bounding procedure imply that the selection on unobservables *can* plausibly explain the association between disease exposure and vaccination.

Next, we check the sensitivity of our results to imputation for missing specific occupation codes. We restrict the sample to occupations exactly matched with O*NET. Section 2 mentioned that many occupations in BRFSS are not coded at the detailed level while most in O*NET are. Thus, we exclude those occupations with imputed intensity and obtain a sample of 359 SOCs.[29] We repeat the main specification used to generate Table 5 and summarize results in Table 9. We compare these two tables and find that the main results are qualitatively unchanged. The one exception is that for health care personnel (column (5)) the coefficient is considerably larger and statistically significant. A one standard deviation increase in intensity is associated with an increased probability of vaccination by 2.46% for health care personnel.[30] Generally, the results in this table demonstrate that fine-occupation imputation was not a consequential source of measurement error for most of our sample, except perhaps for health care personnel. For this particular group, our main results likely suffer from some

TABLE 8 | Robustness check: Oster (2019)'s bound estimates of the coefficients of occupational social intensity and disease exposure.

	Controlled effect		Oster (2019) bounds	δ^0 for
	$\tilde{\beta}$ (1)	R^2 (2)	$\delta = 1, R_{Max} = 1.3R^2$ (3)	$\beta = 0$ (4)
Panel A. Intensity				
(1) Full sample	0.0023***	0.136	[−0.0059, 0.0023]	0.0164
	(0.0006)			
(2) Full sample	0.0013**	0.137	[−0.0090, 0.0013]	0.0080
	(0.0006)			
(3) Full sample	0.0012**	0.138	[−0.0076, 0.0012]	0.0084
	(0.0005)			
(4) Full sample	0.0008***	0.138	[−0.0096, 0.0008]	0.0047
	(0.0006)			
(5) HCP	0.0010	0.108	[−0.0131, 0.0010]	0.0044
	(0.0017)			
(6) Non-HCP	0.0011*	0.107	[−0.0045, 0.0011]	0.0090
	(0.0006)			
Panel B. Exposure				
(1) Full sample	0.0259***	0.138	[−0.0209, 0.0259]	0.0376
	(0.0023)			
(2) Full sample	0.0249***	0.138	[−0.0254, 0.0249]	0.0342
	(0.0024)			
(3) HCP	0.0306***	0.108	[0.0258, 0.0306]	0.2456
	(0.0047)			
(4) Non-HCP	0.0218***	0.107	[0.0185, 0.0218]	0.2758
	(0.0027)			

Note: Estimates in Columns (1)–(2) are from Table 5. Estimates in Columns (3)–(4) are computed using Stata command *psacalc*. Each identified set in Column (3) includes the value of β in the controlled model and the value of β calculated for $R_{Max} = 1.3R^2$, under the assumption that selection on observables and unobservables is proportional. Panel A presents the bound estimates for the coefficients of Intensity, and Panel B presents the bound estimates for Exposure. State and flu season fixed effects are controlled.
***, ***, * means that the coefficient is statistically different from zero at the 1%, 5%, and 10% level. All models weighted with BRFSS weights. Standard errors clustered at the state-level and reported in parentheses.

TABLE 9 | Sub-sample test: Association between occupational social intensity and flu vaccine uptake for the matched 359 SOCs.

	(1) Full	(2) Full	(3) Full	(4) Full	(5) HCP	(6) Non-HCP
Intensity	0.0027***	0.0009	0.0016***	0.0006	0.0064**	0.0008
	(0.0006)	(0.0007)	(0.0005)	(0.0007)	(0.0024)	(0.0007)
Exposure			0.0290***	0.0260***	0.0250***	0.0262***
			(0.0032)	(0.0034)	(0.0053)	(0.0036)
Health care personnel	0.2158***	0.1790***	0.1697***	0.1520***		
	(0.0069)	(0.0093)	(0.0065)	(0.0089)		
Intensity × HCP		0.0131***		0.0080***		
		(0.0021)		(0.0022)		
Constant	0.9697***	0.9663***	0.9621***	0.9608***	0.2177**	0.9646***
	(0.0217)	(0.0216)	(0.0224)	(0.0223)	(0.0805)	(0.0247)
Observations	136,862	136,862	136,862	136,862	22,149	114,713
R^2	0.143	0.144	0.145	0.145	0.116	0.111

Note: SOCs mean occupation codes from Standard Occupational Classification 2000. The matched SOCs cover 22 major occupation groups, 77 minor occupation groups, and 359 detailed-level occupations. Columns (1)–(4) use full sample. Column (5) uses health care personnel only, while column (6) excludes them. We control for sex, age, race/ethnicity, education, income, marital status, number of children, and indicators for high-risk conditions, having personal health care provider, and having medical insurance. We also control for the interaction of state and flu season fixed effects.
***, ***, * mean that the coefficient is statistically different from zero at the 1%, 5%, and 10% level. All models weighted with BRFSS weights. Standard errors clustered at the state-level and reported in parentheses.
Source: BRFSS 2013–2015 and O*NET.

attenuation bias from classical measurement error. Most tellingly, however, these results are consistent with health care workers (a.) valuing the external benefits of their vaccination choices, and/or (b.) perceiving private risk in their social intensity because of their workplace context.

Finally, we implement three alternative social intensity measures constructed from the three different subsets of the 22 descriptors to check the robustness of our results to our definition of social intensity. As described earlier, we first construct two *mutually exclusive* intensity measures, "context" and "activity" because descriptors about work context and work activities may contain common information about social aspects of an occupation.[31] Therefore, we obtain a "context" index from 6 descriptors about work context and an "activity" index from 16 descriptors about work activities. Third, we select 5 descriptors that are mostly related to physical contact in interpersonal interactions to construct a "contact" to disentangle the ambiguity of physical contact versus non-touch social interactions.[32] Appendix Tables A3 and A4 show the summary statistics and correlations of "context", "activity", and "contact". They show that while these indices are correlated, as expected, they are not redundant.

Table 10 replicates Table 5 but substituting social intensity with the "context" measure. Tables 11 and 12 do the same with "activity" and "contact" measures. We find that estimates are generally consistent with our finding with the original intensity measure. Column (1) estimates in these tables show that all intensity measures are positively correlated with vaccine uptake. Columns (2)–(6) show that the coefficient of intensity may be negative after controlling for exposure to disease or infection, while the coefficients of exposure and health care personnel are always positive and statistically significant at 1% level.[33]

In terms of interpretation, we note that these results should be interpreted as the association between occupational average social intensity and the marginal probability of being vaccinated. Even so, these results are probably subject to some attenuation bias. Because these indices are constructed variables, that process in and of itself introduces some classical measurement error.

Second, if we interpret these results as the association between workplace social intensity at the of an individual level, concerns about measurement error should be a bit more pronounced. A complete picture of the social intensity of an individual's workplace includes characteristics of the firm, the number of individuals at their site, the built environment of their workplace, the number of customers, vendors, and contractors visiting their physical location in a given day, among other factors. We acknowledge that our results for all occupational measures may be biased toward zero, but do not believe that substantively affects the interpretation of our results, that individuals are more sensitive to private risk than the external beneifts of their vaccination choices.

From the literature on measurement error, we know that the magnitude of the bias of an estimator $\hat{\beta}$ is proportional to the share of total variation in the explanatory variable that goes unmeasured:

$$plim\hat{\beta} = \beta\left(1 - \frac{\sigma_u^2}{\sigma_u^2 + \sigma_x^2}\right) \tag{3}$$

In our context, we observe credible measures of "between occupation" variation in social intensity, but the "within occupation" variance is unknowable.

So, how concerned should we be about this attenuation bias? Because of the detailed level of occupation for which we

TABLE 10 | Robustness check: Association between occupational social intensity (context) and flu vaccine uptake.

	(1) Full	(2) Full	(3) Full	(4) Full	(5) HCP	(6) Non-HCP
Intensity (context)	0.0023	0.0002	−0.0020	−0.0025*	−0.0134**	−0.0016
	(0.0014)	(0.0014)	(0.0014)	(0.0014)	(0.0055)	(0.0014)
Exposure			0.0275***	0.0265***	0.0393***	0.0229***
			(0.0025)	(0.0026)	(0.0055)	(0.0028)
Health care personnel	0.2222***	0.1971***	0.1813***	0.1753***		
	(0.0075)	(0.0093)	(0.0074)	(0.0089)		
Intensity × HCP		0.0184***		0.0055		
		(0.0038)		(0.0040)		
Constant	0.9561***	0.9525***	0.9449***	0.9443***	0.2403***	0.9593***
	(0.0200)	(0.0194)	(0.0202)	(0.0198)	(0.0657)	(0.0197)
Observations	229,144	229,144	229,144	229,144	36,347	192,797
R^2	0.136	0.136	0.138	0.138	0.109	0.107

Note: Intensity (Context) is constructed by applying the principal component analysis method to 6 descriptors about work context in O*NET, and Table 3 presents the weight of each descriptor. Exposure is the standardized index of disease exposure at work. Columns (1)–(4) use full sample. Column (5) uses health care personnel only, while column (6) excludes them. We control for sex, age, race/ethnicity, education, income, marital status, number of children, and indicators for high-risk conditions, having personal health care provider, and having medical insurance. We also control for the interaction of state and flu season fixed effects.
***, ***, * mean that the coefficient is statistically different from zero at the 1%, 5%, and 10% level. All models weighted with BRFSS weights. Standard errors clustered at the state-level and reported in parentheses.
Source: BRFSS 2013–2015 and O*NET.

TABLE 11 | Robustness check: Association between occupational social intensity (activity) and flu vaccine uptake.

	(1) Full	(2) Full	(3) Full	(4) Full	(5) HCP	(6) Non-HCP
Intensity (activity)	0.0025***	0.0015**	0.0017***	0.0012*	0.0020	0.0015**
	(0.0006)	(0.0007)	(0.0006)	(0.0006)	(0.0016)	(0.0006)
Exposure			0.0259***	0.0250***	0.0301***	0.0219***
			(0.0023)	(0.0024)	(0.0046)	(0.0027)
Health care personnel	0.2214***	0.2080***	0.1806***	0.1749***		
	(0.0073)	(0.0076)	(0.0074)	(0.0077)		
Intensity × HCP		0.0073***		0.0038**		
		(0.0015)		(0.0016)		
Constant	0.9540***	0.9524***	0.9487***	0.9480***	0.2533***	0.9624***
	(0.0194)	(0.0191)	(0.0194)	(0.0192)	(0.0669)	(0.0190)
Observations	229,144	229,144	229,144	229,144	36,347	192,797
R^2	0.136	0.137	0.138	0.138	0.108	0.107

Note: Intensity (Activity) is constructed by applying the principal component analysis method to 16 descriptors about work activities in O*NET, and Table 3 presents the weight of each descriptor. Exposure is the standardized index of disease exposure at work. Columns (1)–(4) use full sample. Column (5) uses health care personnel only, while column (6) excludes them. We control for sex, age, race/ethnicity, education, income, marital status, number of children, and indicators for high-risk conditions, having personal health care provider, and having medical insurance. We also control for the interaction of state and flu season fixed effects.
***, ***, * mean that the coefficient is statistically different from zero at the 1%, 5%, and 10% level. All models weighted with BRFSS weights. Standard errors clustered at the state-level and reported in parentheses.
Source: BRFSS 2013–2015 and O*NET.

measure social intensity, the observed share of total variation in our context is considerable, ranging from teachers, hospitality workers, and sales managers at the high end, to mathematicians and auto repairers (those doing the work, not the counter clerks). While there will clearly be some unobserved within-occupation variation, it is not clear that the unobserved within-occupation share of total variation would even exceed that of the between-occupation variation. For attenuation bias to substantively affect the interpretation of our results, the unobserved share of workplace social intensity would need to be an order of magnitude larger than the observed between-occupation share of social intensity. In our preferred specification, the estimated coefficient on exposure is 20 times as large as that of social intensity. For measurement error to cause that difference, the unobserved within-occupation variation would need to be 19 times larger than the between-occupation share, and our estimates of exposure would need to be free from any attenuation bias themselves.

TABLE 12 | Robustness check: Association between occupational social intensity (contact) and flu vaccine uptake.

	(1) Full	(2) Full	(3) Full	(4) Full	(5) HCP	(6) Non-HCP
Intensity (contact)	0.0030*	−0.0006	−0.0107***	−0.0112***	−0.0219***	−0.0104***
	(0.0015)	(0.0016)	(0.0016)	(0.0017)	(0.0037)	(0.0016)
Exposure			0.0376***	0.0364***	0.0551***	0.0327***
			(0.0027)	(0.0028)	(0.0058)	(0.0029)
Health care personnel	0.2197***	0.1914***	0.1787***	0.1724***		
	(0.0075)	(0.0088)	(0.0073)	(0.0083)		
Intensity × HCP		0.0168***		0.0045		
		(0.0030)		(0.0031)		
Constant	0.9533***	0.9515***	0.9448***	0.9446***	0.2531***	0.9584***
	(0.0194)	(0.0190)	(0.0193)	(0.0192)	(0.0644)	(0.0190)
Observations	229,144	229,144	229,144	229,144	36,347	192,797
R^2	0.136	0.137	0.138	0.139	0.110	0.108

Note: Intensity (Contact) is constructed by applying the principal component analysis method to 5 descriptors directly related to physical contact: physical proximity, contact with others, face-to-face discussions, assisting and caring for others, and performing for or working directly with the public. Table 3 presents the weight of each descriptor. Exposure is the standardized index of disease exposure at work. Columns (1)–(4) use full sample. Column (5) uses health care personnel only, while column (6) excludes them. We control for sex, age, race/ethnicity, education, income, marital status, number of children, and indicators for high-risk conditions, having personal health care provider, and having medical insurance. We also control for the interaction of state and flu season fixed effects.
***, ***, * mean that the coefficient is statistically different from zero at the 1%, 5%, and 10% level. All models weighted with BRFSS weights. Standard errors clustered at the state-level and reported in parentheses.
Source: BRFSS 2013–2015 and O*NET.

6 | Conclusion

We use BRFSS data to investigate the relationship between individuals' flu vaccine uptake and social intensity of workers' occupations. We find that workers in occupations with high social intensity are not significantly more likely to report receiving a flu shot once we control for whether the individual works in a clinical setting. However, the level of potential direct exposure to diseases at work is strongly positively correlated with the worker's probability of obtaining a flu vaccine. In concert, these findings are consistent with the hypothesis that while all workers respond to the potential personal risk of infection, they do not place substantial weight on the external effects of their vaccination choices.

These results have immediate ramifications for any future vaccine roll outs, including flu and COVID-19. They are not intended to editorialize on policy prescriptions for rank-ordering priority access to vaccines. Vaccinating individuals who come into close contact with relatively high numbers of people may or may not take precedence over other considerations such as mortality risk. These results suggest individuals do not place sufficient weight on the external social benefits of vaccination. If high-social intensity individuals are deemed high-priority, early access and availability alone may not achieve optimal uptake. In this case, some incentives may help raise desired vaccination coverage among individuals with high external social benefits.

While the results yield important insight for maximizing the efficacy of vaccine roll outs, there are a few caveats that should be kept in mind when interpreting these results. First, occupational social intensity is not a perfect measure of the individual's overall social intensity. Other factors, such as personality type, age, life stage, social habits, engagement in community leadership, etc., will affect the total amount of interpersonal contact in which an individual engages. While it is true that individuals with preferences for social activity tend to select more social occupations, the occupational social intensity by itself is not a complete measure of individuals' "true social intensity". Second, the Occupational Information Network (O*NET) is not a designated data project to collect information about the risk of disease spreading, but a comprehensive database of career exploration and occupation analysis. Therefore, we should consider the derived intensity measure as a proxy to quantify social intensity at the workplace, but not a manual for rank-ordering vaccinations for specific occupations. Third, though we use "exposure" to measure the private benefits of vaccination, exposure may also be correlated with the external benefits. We are aware that high-exposure occupations include jobs that work with vulnerable people. For example, doctors and nurses work with patients who may suffer severe illnesses if they contract the flu. Finally, we do not measure unemployed individuals' social intensity because of data limitation, but vaccinating unemployed people is indispensable for a successful vaccination plan. Further research about unemployed individuals' social activities and their roles in disease spreading is necessary for effective vaccine roll outs.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from Centers for Disease Control and Prevention. Restrictions apply to the availability of these data, which were used under license for this study.

Data are available from the author(s) with the permission of Centers for Disease Control and Prevention.

Endnotes

[1] In 2020, the "Healthy People 2030" initiative set a goal of 70% flu vaccination coverage for non-institutionalized adults. We refer to non-institutionalized adults simply as adults hereafter. See details at https://health.gov/healthypeople.

[2] World Health Organization (WHO) launched the Expanded Programme on Immunization to eradicate vaccine-preventable diseases since 1974, but smallpox is the only infectious disease that has been eradicated (Hinman 1999). Although two of the three types of polio are eradicated, a few incidences of the third type still occur in some developing countries (Wilson et al. 2021; WHO 2021). In 2019, Global leaders pledged $2.6 billion to support continuing efforts to vaccinate every child against polio (WHO 2019).

[3] To the extent that costs of treatment are potentially cross-subsidized by healthier individuals, preventing vulnerable individuals from getting the flu yields substantial social benefits. For these reasons, medical professionals who frequently interact with vulnerable individuals may have vaccine mandates as one provider may spread the flu to many patients. A set of state vaccination laws can be found at https://www.cdc.gov/phlp/php/publications/vaccination-laws.html.

[4] Xie et al. (2023) also found that SMEs are more likely to shut down during a pandemic.

[5] We use O*NET (version 23.0, released August 2018) which contains occupation codes using the 2010 SOC. We transform these codes into the 2000 SOC to match those in BRFSS. The most recent O*NET (version 26.0) is using the 2019 SOCs. See details in the Appendix A.2.

[6] Results also show that individuals who work in a clinical setting are 18 percentage points more likely to get vaccinated than workers in non-clinical settings.

[7] For completeness, we also implemented a LASSO approach to determine if any of the 22 O*NET variables included by themselves would predict vaccination. None did.

[8] Considerable evidence from the economics literature shows that individuals generally make decisions on the basis of marginal private benefits rather than marginal social benefits. Oster (2018) and Schaller et al. (2019) show that individuals are more likely to get their children vaccinated for pertussis in response to an outbreak; but that these actions are consistent with myopia and irrationality. Field experimental work by de Janvry et al. (2010) shows that individuals can be outright non-altruistic: participants misrepresent their eligibility to be classified as "vulnerable", particularly when supplies are scarce. While there is evidence that individuals' vaccination choices are driven primarily by *private* benefits, it is not known whether the individuals whose vaccinations are likely to yield the *largest external benefits* are empirically more likely to do so.

[9] Clemens and Gottlieb (2021) have expressed support for the idea that incentives may be helpful in pushing out vaccines. The results in this paper support that notion as well, but with the additional emphasis that incentives for certain candidate recipients (not just providers) may be helpful in stewarding optimal outcomes. These incentives can be either monetary, such as gift cards, or non-monetary, such as on-site vaccination clinics at the workplace.

[10] BRFSS 2013-2015 data include 31 states, which are California, Colorado, Connecticut, Florida, Georgia, Idaho, Illinois, Iowa, Louisiana, Maryland, Missouri, Massachusetts, Michigan, Minnesota, Mississippi, Montana, Nebraska, New Hampshire, New Jersey, New Mexico, New York, North Carolina, North Dakota, Oregon, Tennessee, Utah, Vermont, Washington, West Virginia, Wisconsin, and Wyoming. See state-year crosswalk details in Appendix A.1. Data from Wyoming Department of Health, Public Health Division, Behavioral Risk Factor Surveillance System, were supported in part by Centers for Disease Control and Prevention Cooperative Agreement, U58/SO000016-1 through 3 [2011–2013]. Data from Washington State Department of Health, Center for Health Statistics, Behavioral Risk Factor Surveillance System were supported in part by Centers for Disease Control and Prevention Cooperative Agreement, U58/SO000047-1 through 3 [2011–2013]).

[11] BRFSS asks all respondents whether they received an influenza vaccine in the past 12 months. Therefore, our sample includes data from the 2011-2012 season (i.e. shots received in January 2012-July 2012) to the 2015–2016 flu season (i.e. shots received in August 2015-December 2015). About 37% respondents report their timing of flu shot in the raw data, and 38% respondents report their timing of flu shot in the sample. For respondents who do not report the timing of flu shot, we use the interview time as the flu shot time but lag 1 month. For instance, a vaccinated individual interviewed in September 2012 would be considered vaccinated in August 2012, if flu shot month and year are not available.

[12] For 2012–2015 flu seasons, the weighted sample mean vaccination coverages are 37.8% (2012–2013), 34.3% (2013–2014), and 34.0% (2014–2015). These estimates are generally lower than CDC's estimates for the population average because 90% of respondents are young adults aged 18–64. To compare, the CDC estimates for the vaccination coverage of adults aged 18–49 years during 2012–2015 seasons are 31.1%, 32.3%, 33.5% respectively (CDC 2023).

[13] We cannot verify this assumption because BRFSS does not ask about the duration of respondents' current occupations. However, the assumption is reasonable for two reasons. First, according to the U.S. Bureau of Labor Statistics (2014), the median duration of wage and salary workers with their current employee was 4.6 years in Jan 2014. Moreover, most people who change their jobs are likely to stay in the same occupation or occupations that require similar skills. For example, Parrado et al. (2007) estimate that about 90% changes are happened between these "similar" occupations during 1969–1993.

[14] BRFSS originally coded the industry/occupation module in the Census code. O'Halloran et al. (2017) converted the Census codes to the equivalent 2000 Standard Occupational Classification (SOC) codes and 2002 North American Industry Classification System (NAICS) industry codes.

[15] Our samfple covers 405 out of 446 non-military broad occupations. The 41 uncovered broad occupations include 10 miscellaneous occupations, 14 occupations of teachers in post-secondary and special areas (including 2 miscellaneous occupations).

[16] Because we standardize all descriptors before applying PCA, the intensity index value can be negative.

[17] O*NET (version 23.0) uses the 2010 SOC codes, so we follow the crosswalk guide provided by the U.S. Bureau of Labor Statistics (2012) to convert the codes to 2000 SOC. See detailed explanations about the construction of intensity in the Appendix. Also, for readers that are concerned that we have omitted a key descriptor of the social aspect of a job, Appendix Table A2 contains a list of all the occupational descriptors related to work activities and work contexts that we exclude in any social intensity indices.

[18] For health care workers in particular, the risk of exposure to flu may be correlated with the external benefits of vaccination. We therefore split the sample at times into health care and non-health care workers to examine whether there are distinct relationships between "social intensity" or "exposure to diseases" and vaccination between workers in health care and not in health care.

[19] For instance of "exposure", dental hygienists (SOC 29–2021, intensity 1.12) score 100 or 2.89 after standardized, and fire fighters (SOC 33–3051, intensity 3.96) score 61 or 1.37 after standardized.

[20] Non-health care industries rarely require workers to obtain flu vaccinations. One reason could be that almost all state laws allow religious or philosophical exemptions for vaccination requirements. See vaccination exemptions in state laws in National Conference of State Legislatures (2025).

[21] These are clinical and nonclinical staff working in hospitals (NAICS 622), outpatient care and physician offices (NAICS 6214 and 6211), long-term care facilities (NAICS 6216, 6231, 6232, 6233, and 6239), and other clinical settings (NAICS 6212, 62131, 62132, 6213, 6215, and 6219). Such a definition is consistent with the Advisory Committee on Immunization Practices (ACIP). ACIP defines health care personnel as "all paid and unpaid persons working in health-care settings who have the potential for exposure to patients or to infectious materials. These personnel might include (but are not limited to) physicians, nurses, nursing assistants, nurse practitioners, physician assistants, therapists, technicians, emergency medical service personnel, dental personnel, pharmacists, laboratory personnel, autopsy personnel, students and trainees, contractual staff, and other persons not directly involved in patient care but who can potentially be exposed to infectious agents (e.g., clerical, dietary, housekeeping, laundry, security, maintenance, administrative, and billing staff and volunteers)". See Grohskopf et al. (2020).

[22] See summary statistics of social intensity and other derived measures in Table 2, Panel B.

[23] We also investigate the 95% confidence interval of the coefficient of intensity in column (3), which is $(0.0002, 0.0022)$. To calculate the marginal association of a one standard deviation increase in social intensity at the upper bound of that confidence interval, we multiply the upper bound 0.0022 by the weighted standard deviation of intensity, 3.07, and get 0.0068. That is, when comparing two workers, if one's occupational social intensity exceeds the other's by one standard deviation, that worker is at most 0.68 percentage points more likely to be vaccinated.

[24] Such a magnitude of the effect of intensity is about a quarter of being a female (2.7%) or being married versus never married (3.08%), and less than one-tenth of having high-risk conditions (6.97%).

[25] The comparison between workers who are one standard deviation apart is for perceptual intuition only. For the record, in our sample, the vaccination rate is 22.9% for food preparation workers (SOC 35–2021); 33.9% for public relations managers (SOC 11–2031); 17.0% for supervisors of landscaping, lawn service, and groundskeeping workers (SOC 37–1012); 38.4% for database administrators (SOC 15–1061); and 23.5% for protective service workers (SOC 33–9090). As a potential outlier, economists (SOC 19–3011) have a relatively low intensity score −1.63 while 75.3% of them get vaccinated.

[26] Figures A2 and A3, which plot bin estimates for the full sample without controlling for exposure, and bin estimates for health care workers respectively, are available in the Appendix.

[27] A referee raised a concern about whether we were overfitting with fixed effects. Appendix Table A13 contains results from specifications like those used in Table 6, but without state-by-season fixed effects. Results in the two tables are statistically indistinguishable from one another. To address another concern about the lagged impact of the past flu season, we matched our sample with the state-level flu and pneumonia mortality data (CDC 2022). Table A17 shows that in our matched sample (without fixed effects), neither the coefficient of intensity nor that of flu mortality is significant. However, their interaction is positive and significant in all specifications except for health care personnel. This means that a non-health worker who lives in a state with a severe past flu season is more sensitive to her workplace infection risk than another worker in the same profession who did not experience such a severe flu season.

[28] Appendix Tables A8–A11 contain bounded estimates for the alternate measures of intensity and a sample restricted to non-imputed SOCs.

[29] The 359 SOCs include 358 detailed occupations and 1 broad occupation (biological scientists, SOC 19–1020).

[30] We multiply the coefficient 0.80% by 3.07, that is, the standard deviation of intensity, and get 2.46%. Appendix Table A8 shows the Oster's bound estimates of the coefficients of intensity and exposure.

[31] See survey questions of each descriptor in the Appendix Section A.3.

[32] Recall Table 3 in the Data section presents the weight of descriptors for each alternative intensity measure.

[33] We also bound estimates using alternative measures with Oster (2019)'s method and report the results in Tables A8–A11. We find that all intervals for intensity measures either contain zero or are less than zero, and intervals for exposure in Columns (3)–(4) also contain zero. But if we restrict the sample to non-health care personnel as in Column (6), intervals for exposure are always positive except for the intensity measure using descriptors of interpersonal physical contact only.

References

Adda, J. 2016. "Economic Activity and the Spread of Viral Diseases: Evidence From High Frequency Data." *Quarterly Journal of Economics* 131, no. 2: 891–941. https://doi.org/10.1093/qje/qjw005.

Altonji, J. G., T. E. Elder, and C. R. Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy* 113, no. 1: 151–184. https://doi.org/10.1086/426036.

Boulier, B. L., T. S. Datta, R. S. Goldfarb, V. Externalities, and B. E. The. 2007. "Vaccination Externalities." *Journal of Economic Analysis and Policy* 7, no. 1. https://doi.org/10.2202/1935-1682.1487.

Buckner, J. H., G. Chowell, and M. R. Springborn. 2021. "Dynamic Prioritization of COVID-19 Vaccines When Social Distancing Is Limited for Essential Workers." *Proceedings of the National Academy of Sciences* 118, no. 16. https://doi.org/10.1073/pnas.2025786118.

Carrera, M., E. C. Lawler, and C. White. 2021. "Population Mortality and Laws Encouraging Influenza Vaccination for Hospital Workers." *Annals of Internal Medicine* 174, no. 4: 444–452. https://doi.org/10.7326/m20-0413.

CDC. 2019. "Behavioral Risk Factor Surveillance System.". https://www.cdc.gov/brfss/.

CDC. 2020. "Flu Vaccination Coverage, United States, 2019–2020 Influenza Season." https://www.cdc.gov/fluvaxview/coverage-by-season/vaccination-2019-2020.html.

CDC. 2022. "Influenza/Pneumonia Mortality by State." https://www.cdc.gov/nchs/pressroom/sosmap/flu_pneumonia_mortality/flu_pneumonia.htm.

CDC. 2023. "Flu Burden Prevented From Vaccination 2019–2020 Flu Season." https://archive.cdc.gov/#/details?url=https://www.cdc.gov/flu-burden/php/data-vis-vac/2019-2020-prevented.html.

Clemens, J., and J. Gottlieb. 17 April 2021. *Cash Talks: The Need for Incentives in Vaccine Delivery*. Hill.: Retrieved from. https://thehill.com/opinion/healthcare/536703-cash-talks-the-need-for-incentives-in-vaccine-delivery?rnd=1612144901&rl=1.

de Janvry, A., E. Sodoulet, and S. Villas-Boas. 2010. "Short on Shots: Are Calls for Cooperative Restraint Effective in Managing a Flu Vaccines Shortage?" *Journal of Economic Behavior & Organization* 76: 209–224.

Geoffard, P.-Y., and T. Philipson. 1997. "Disease Eradication: Private versus Public Vaccination." *American Economic Review* 87, no. 1: 222–230.

Grohskopf, L. A., E. Alyanak, K. R. Broder, et al. 2020. "Prevention and Control of Seasonal Influenza With Vaccines: Recommendations of the Advisory Committee on Immunization Practices–United States, 2020-21 Season." *Morbidity and Mortality Weekly Report Recommendations and Reports* 69, no. 9: 1–24. https://doi.org/10.15585/mmwr.rr6908a1.

Hinman, A. 1999. "Eradication of Vaccine-Preventable Diseases." *Annual Review of Public Health* 20, no. 1: 211–229. https://doi.org/10.1146/annurev.publhealth.20.1.211.

Hughes, M. M., C. Reed, B. Flannery, et al. 2019. "Projected Population Benefit of Increased Effectiveness and Coverage of Influenza Vaccination on Influenza Burden in the United States." *Clinical Infectious Diseases* 70, no. 12: 2496–2502. https://doi.org/10.1093/cid/ciz676.

Luckhaupt, S. E., G. M. Calvert, J. Li, M. Sweeney, and T. A. Santibanez. 2014. "Prevalence of Influenza-like Illness and Seasonal and Pandemic H1N1 Influenza Vaccination Coverage Among Workers — United States, 2009–10 Influenza Season." *Morbidity and Mortality Weekly Report* 63, no. 10: 217–221.

Markowitz, S., E. Nesson, and J. J. Robinson. 2019. "The Effects of Employment on Influenza Rates." *Economics and Human Biology* 34: 286–295. https://doi.org/10.1016/j.ehb.2019.04.004.

National Center for O*NET Development. 2019. "Occupational Information Network.". https://www.onetonline.org/.

National Conference of State Legislatures. 2025. "State Non-Medical Exemptions From School Immunization Requirements." https://www.ncsl.org/health/state-non-medical-exemptions-from-school-immunization-requirements.

O'Halloran, A. C., P. Lu, W. W. Williams, et al. 2017. "Influenza Vaccination Among Workers - 21 U.S. States, 2013." *American Journal of Infection Control* 45, no. 4: 410–416. https://doi.org/10.1016/j.ajic.2017.01.005.

Oster, E. 2018. "Does Disease Cause Vaccination? Disease Outbreaks and Vaccination Response." *Journal of Health Economics* 57: 90–101. https://doi.org/10.1016/j.jhealeco.2017.10.003.

Oster, E. 2019. "Unobservable Selection and Coefficient Stability: Theory and Evidence." *Journal of Business & Economic Statistics* 37, no. 2: 187–204. https://doi.org/10.1080/07350015.2016.1227711.

Parrado, E., A. Caner, and E. N. Wolff. 2007. "Occupational and Industrial Mobility in the United States." *Labour Economics* 14, no. 3: 435–455. https://doi.org/10.1016/j.labeco.2006.01.005.

Philipson, T. 1996. "Private Vaccination and Public Health: An Empirical Examination for U.S. Measles." *Journal of Human Resources* 31, no. 3: 90–101. https://doi.org/10.2307/146268.

Pichler, S., and N. R. Ziebarth. 2017. "The Pros and Cons of Sick Pay Schemes: Testing for Contagious Presenteeism and Noncontagious Absenteeism Behavior." *Journal of Public Economics* 156: 14–33. https://doi.org/10.1016/j.jpubeco.2017.07.003.

R Core Team 2018. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.

Schaller, J., L. Schulkind, and T. Shapiro. 2019. "Disease Outbreaks, Healthcare Utilization, and On-Time Immunization in the First Year of Life." *Journal of Health Economics* 67: 1–12. https://doi.org/10.1016/j.jhealeco.2019.05.009.

Siciliani, L., C. Wild, M. McKee, et al. 2020. "And Members of the Expert Panel on Effective Ways of Investing in Health, Strengthening Vaccination Programmes and Health Systems in the European Union: A Framework for Action." *Health Policy* 124, no. 5: 511–518. https://doi.org/10.1016/j.healthpol.2020.02.015.

Song, H., R. M. McKenna, T. Chen Angela, G. David, and A. Smith-McLallen. 2021. "The Impact of the Non-Essential Business Closure Policy on Covid-19 Infection Rates." *International Journal of Health Economics and Management* 21, no. 4: 387–426. https://doi.org/10.1007/s10754-021-09302-9.

Stein, R. A. 2011. "Super-Spreaders in Infectious Diseases." *International Journal of Infectious Diseases* 15, no. 8: e501–e513. https://doi.org/10.1016/j.ijid.2010.06.020.

Teicher, A. 2023. "Super-Spreaders: A Historical Review." *Lancet* 23, no. 10: e409–e417. https://doi.org/10.1016/s1473-3099(23)00183-4.

U.S. Bureau of Labor Statistics. 2012. "Crosswalks Between the 2010 SOC and Systems Used by Other Federal and International Statistical Agencies." Retrieved from. https://www.bls.gov/soc/soccrosswalks.htm.

U.S. Bureau of Labor Statistics. 2014. "Employee Tenure in 2014." Retrieved from. www.bls.gov/news.release/archives/tenure_09182014.pdf.

Ward, C. J. 2014. "Influenza Vaccination Campaigns: Is an Ounce of Prevention Worth a Pound of Cure?" *American Economic Journal: Applied Economics* 6, no. 1: 38–72. https://doi.org/10.1257/app.6.1.38.

White, C. 2019. "Measuring the Social and Externality Benefits of Influenza Vaccination." *Journal of Human Resources*: 1118–9893R2.

WHO. February 2019. "Global Leaders Pledge US$2.6 Billion to Eradicate Polio at the Reaching the Last Mile Forum in Abu Dhabi." Retrieved from. https://www.who.int/news/item/19-11-2019-global-leaders-pledge-us-2.6-billion-to-eradicate-polio-at-the-reaching-the-last-mile-forum-in-abu-dhabi.

WHO. February 2021. "Statement of the Twenty-Ninth Polio IHR Emergency Committee." Retrieved from. https://www.who.int/news/item/20-08-2021-statement-of-the-twenty-ninth-polio-ihr-emergency-committee.

Wickham, H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag.

Wilson, N., O. D. Mansoor, M. J. Boyd, A. Kvalsvig, and M. G. Baker. 2021. "We Should Not Dismiss the Possibility of Eradicating COVID-19: Comparisons With Smallpox and Polio." *BMJ Global Health* 6, no. 8: e006810. https://doi.org/10.1136/bmjgh-2021-006810.

Xie, X., X. Jin, W. Guo, and C.-T. Chang. 2023. "Monitoring and Early Warning of SMEs' Shutdown Risk Under the Impact of Global Pandemic Shock." *Systems* 11, no. 5: 260. https://doi.org/10.3390/systems11050260.

Appendix A

A.1 | BRFSS Occupation and Industry Module: State Crosswalk

Appendix Table A1

TABLE A1 | The 2013–2015 state-year availability of the BRFSS industry and occupation module.

State	2013	2014	2015
California	X		
Colorado		X	X
Connecticut		X	X
Florida	X		X
Georgia		X	X
Idaho		X	X
Illinois	X	X	X
Iowa		X	X
Louisiana	X	X	X
Maryland		X	
Massachusetts	X	X	X
Michigan	X	X	X

(Continues)

TABLE A1 | (Continued)

State	2013	2014	2015
Minnesota	X	X	X
Mississippi	X	X	X
Missouri	X		
Montana	X	X	X
Nebraska	X	X	X
New Hampshire	X	X	X
New Jersey	X	X	X
New Mexico	X	X	X
New York	X	X	X
North Carolina		X	X
North Dakota	X	X	X
Oregon	X	X	
Tennessee		X	X
Utah	X	X	X
Vermont		X	
Washington	X	X	X
West Virginia			X
Wisconsin	X		X
Wyoming	X		

Source: BRFSS Industry and Occupation Module: 2013–2015.

A.2 | The Construction of Intensity

We apply the principal component analysis to the Occupational Information Network (O*NET) data to construct the occupation intensity. The O*NET data is publicly available at https://www.onetcenter.org/db_releases.html, and our data version is 23.0, released in August 2018. Principal component analysis is a classical linear dimensionality reduction technique in statistics. Principle component analysis converts a set of possibly correlated variables into a set of linearly uncorrelated variables. We call these linearly uncorrelated variables the "principal components". Specifically, we construct the occupation intensity in six steps:

1. We select 22 descriptors from Occupational Information Network (O*NET) that are related to interpersonal interactions. We exclude "exposed to diseases or infections" because it describes disease exposure but not interactions with other people. Table 2 presents summary statistics of the descriptors.

2. The O*NET dataset contains 967 8-digit specific occupation codes using the 2010 Standard Occupational Classification (SOC). We first collapse these 8-digit codes into 6-digit codes by taking a simple average and obtain descriptors for 770 occupations. Then we transform them into 2000 SOC using the crosswalk guide provided by the U.S. Bureau of Labor Statistics.

3. The 2013 Behavioral Risk Factor Surveillance System (BRFSS) dataset contains 494 unique occupations. We merge the BRFSS dataset with the transformed O*NET dataset and find 335 matched occupations, which are mostly specific occupations. The unmatched occupations are mostly coded as minor groups and broad occupations (SOC trailing in zeros) or miscellaneous occupations (SOC trailing in nines).

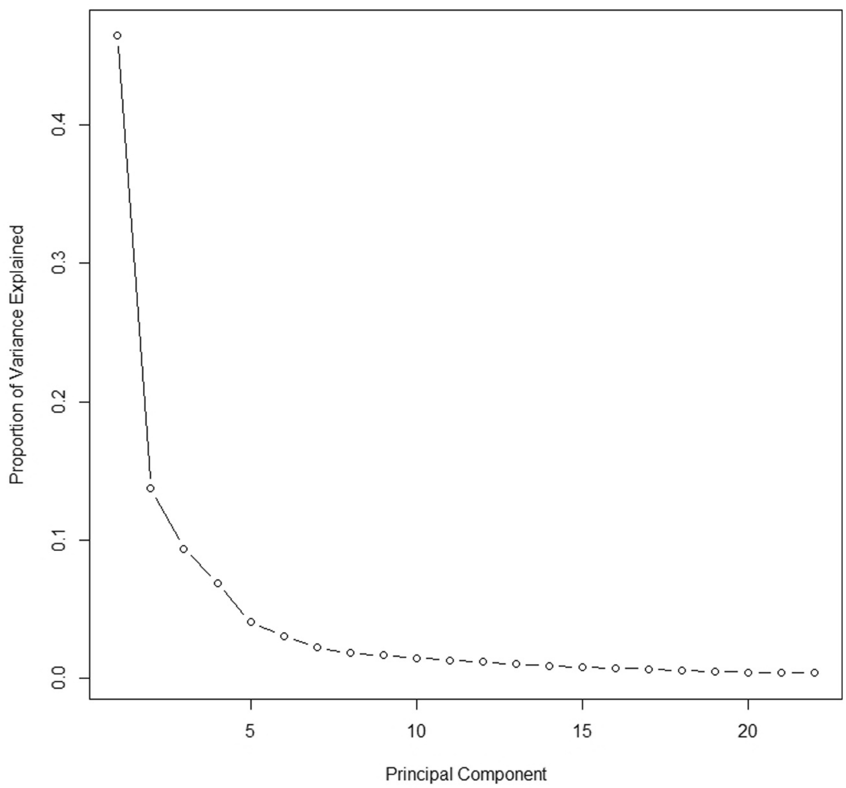

FIGURE A1 | The proportion of variance explained by every principal component with the first principal component as the social intensity.

4. We take the simple average of descriptors of specific occupations under the same general group to impute values for minor occupation groups, broad occupations, and miscellaneous occupations. Then we merge the BRFSS dataset with the imputed O*NET dataset and find 157 matches. We drop two occupations in BRFSS: Legislators (11–1031) because O*NET does not collect the information and Textile Bleaching and Dyeing Machine Operators and Tenders (51–6061) because BRFSS does not have all variables we need. Thus, we obtain descriptors for 492 occupations in BRFSS.
5. We apply principal component analysis to the 22 descriptors and retain the first principal component as the intensity score. The intensity score represents 47% of all variations of the 22 descriptors. Figure A1 displays the proportion of variance explained by each principal component.

A.3 | Descriptors From Occupational Information Network

We display the official definitions of the 23 descriptors from the Occupational Information Network (O*NET) here, including "exposed to diseases or infections" that is not used for the construction of intensity score. These descriptors are from three categories:

Work Context—Physical Work Conditions. This category describes the work context as it relates to the interactions between the worker and the physical job environment. Two descriptors are from this category:

1. Exposed to Disease or Infections—How often does this job require exposure to disease/infections?
2. Physical Proximity—To what extent does this job require the worker to perform job tasks in close physical proximity to other people?

Work context—Interpersonal Relationships. This Category describes the context of the job in terms of human interaction processes. Five descriptors are from this category:

3. Contact with Others—How much does this job require the worker to be in contact with others (face-to-face, by telephone, or otherwise) in order to perform it?
4. Coordinate or Lead Others—How important is it to coordinate or lead others in accomplishing work activities in this job?
5. Deal with External Customers—How important is it to work with external customers or the public in this job?
6. Face-to-Face Discussions—How often do you have to have face-to-face discussions with individuals or teams in this job?
7. Work with Work Group or Team—How important is it to work with others in a group or team in this job?

Work Activities—Interacting with Others. This category describes what interactions with other persons or supervisory activities occur while performing this job. Note that O*NET provides two scores for this category, "importance" and "level". We choose "importance" because "importance" indicates the intensity or frequency an activity is needed to the occupation, while "level" indicates the skill level needed to perform such activities. Sixteen descriptors are from this category:

8. Assisting and Caring for Others—Providing personal assistance, medical attention, emotional support, or other personal care to others such as coworkers, customers, or patients.
9. Coaching and Developing Others—Identifying the developmental needs of others and coaching, mentoring, or otherwise helping others to improve their knowledge or skills.
10. Communicating with Persons Outside Organization—Communicating with people outside the organization, representing the organization to customers, the public, government, and other external sources. This information can be exchanged in person, in writing, or by telephone or e-mail.
11. Communicating with Supervisors, Peers, or Subordinates—Providing information to supervisors, co-workers, and subordinates by telephone, in written form, e-mail, or in person.
12. Coordinating the Work and Activities of Others—Getting members of a group to work together to accomplish tasks.
13. Developing and Building Teams—Encouraging and building mutual trust, respect, and cooperation among team members.
14. Establishing and Maintaining Interpersonal Relationships—Developing constructive and cooperative working relationships with others, and maintaining them over time.
15. Guiding, Directing, and Motivating Subordinates—Providing guidance and direction to subordinates, including setting performance standards and monitoring performance.
16. Interpreting the Meaning of Information for Others—Translating or explaining what information means and how it can be used.
17. Performing Administrative Activities—Performing day-to-day administrative tasks such as maintaining information files and processing paperwork.
18. Performing for or Working Directly with the Public—Performing for people or dealing directly with the public. This includes serving customers in restaurants and stores, and receiving clients or guests.
19. Provide Consultation and Advice to Others—Providing guidance and expert advice to management or other groups on technical, systems-, or process-related topics.
20. Resolving Conflicts and Negotiating with Others—Handling complaints, settling disputes, and resolving grievances and conflicts, or otherwise negotiating with others.
21. Selling or Influencing Others—Convincing others to buy merchandize/goods or to otherwise change their minds or actions.
22. Staffing Organizational Units—Recruiting, interviewing, selecting, hiring, and promoting employees in an organization.
23. Training and Teaching Others—Identifying the educational needs of others, developing formal educational or training programs or classes, and teaching or instructing others.

A.4 Additional Tables and Figures

TABLE A2 | O*NET descriptors not included in measures of social intensity.

Work activities
Estimating the quantifiable characteristics of products, events, or information
Identifying objects, actions, and events
Inspecting equipment, structures, or materials
Getting information
Monitoring processes, materials, or surroundings
Monitoring and controlling resources
Analyzing data or information
Evaluating information to determine compliance with standards
Judging the qualities of objects, services, or people
Processing information
Developing objectives and strategies
Making decisions and solving problems

(Continues)

TABLE A2 | (Continued)

Work activities
Organizing, planning, and prioritizing work
Scheduling work and activities
Thinking creatively
Updating and using relevant knowledge
Documenting/Recording information
Drafting, laying out, and specifying technical devices, parts and equipment
Repairing and maintaining electronic equipment
Repairing and maintaining mechanical equipment
Working with computers
Controlling machines and processes
Handling and moving objects
Operating vehicles, mechanized devices, or equipment
Performing general physical activities
Work context
Electronic mail
Letters and memos
Telephone
Public speaking
Deal with physically aggressive people
Deal with angry or unpleasant people
Frequency of conflict situations
Responsibility for outcomes and results
Responsibility for others' health and safety
Spend time bending or twisting the body
Spend time climbing ladders, scaffolds, or poles
Spend time keeping or regaining balance
Spend time kneeling, crouching, stooping, or crawling
Spend time making repetitive motions
Spend time sitting
Spend time standing
Spend time using your hands to handle, control, or feel objects, tools, or controls
Spend time walking or running
Cramped work space, awkward positions
Exposed to contaminants
Exposed to whole body vibration
Exposed to bright or inadequate lighting
Sounds, noise levels are distracting or uncomfortable
Very hot or cold temperatures
Exposed to hazardous conditions
Exposed to hazardous equipment
Exposed to high places
Exposed to minor burns, cuts, bites, or stings
Exposed to radiation
Wear common protective or safety equipment
Wear specialized protective or safety equipment

(Continues)

TABLE A2 | (Continued)

Work context
In an enclosed vehicle or equipment
In an open vehicle or equipment
Indoors, environmentally controlled
Indoors, not environmentally controlled
Outdoors, exposed to weather
Outdoors, under cover
Level of competition
Consequence of error
Freedom to make decisions
Frequency of decision making
Impact of decisions on Co-workers or company results
Duration of typical work week
Pace determined by speed of equipment
Time pressure
Work schedules
Degree of automation
Importance of being exact or accurate
Importance of repeating same tasks
Structured versus unstructured work

TABLE A3 | Summary statistics of occupational disease exposure and social intensity.

Variable	Mean	Std. Dev.	Min	Max
Exposure	0	1	−0.95	2.89
Intensity	0.89	3.07	−7.91	8.41
Intensity (context)	0.56	2.23	−7.13	3.73
Intensity (activity)	0.72	2.89	−7.88	8.77
Intensity (contact)	0.58	1.48	−4.34	4.21

Note: BRFSS weights applied.
Source: O*NET.

TABLE A4 | Correlation table of occupational disease exposure and social intensity.

	Exposure	Intensity	Context	Activity	Contact
Exposure	1				
Intensity	0.27	1			
Intensity (context)	0.37	0.74	1		
Intensity (activity)	0.23	0.98	0.60	1	
Intensity (contact)	0.65	0.59	0.80	0.49	1

Note: 509 SOCs are used to calculate the correlation matrix (unweighted).
Source: O*NET.

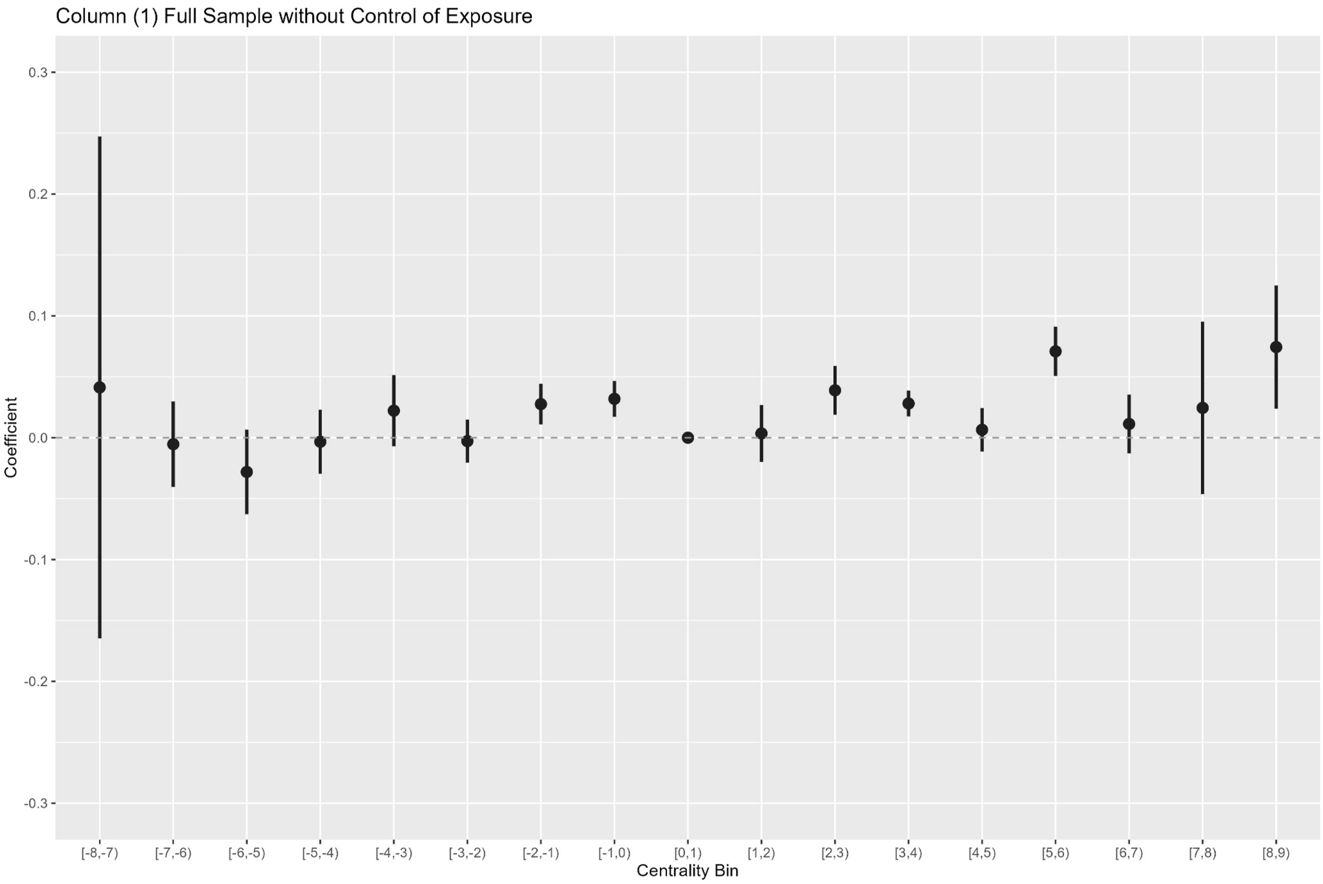

FIGURE A2 | Coefficient and 95% confidence interval estimates of intensity bins in Table A12. Column (1).

TABLE A5 | Occupations with highest and lowest ratings of social intensity—Work context.

Highest rated		Lowest rated	
Occupation (non-health care)	**Score**	**Occupation**	**Score**
First-line supervisors of law enforcement	3.70	Writers and authors	−3.47
Police and sheriff patrollers	3.31	Furniture finishers	−3.49
Supervisors of personal care workers	3.26	Lathe and turning machine tool setters, operators	−3.54
Special education teachers	3.24	Welding, soldering, and brazing workers	−3.55
Meeting and event planners	3.17	Metal furnace operators, tenders, porers, etc.	−3.63
First-line supervisors of food prep/Service	3.06	Tailors, dressmakers, and sewers	−3.66
Firefighers	3.06	Mathematicians	−3.66
First-line super. of mechanics repairers	2.97	Cutting, punching, and press machine operators	−3.88
Educational/Guidance/Career counselors	2.96	Refuse and recyclable material collectors	−3.95
Postmasters and mail superintendents	2.90	Extruding and drawing machine setters, operators	−4.02
First-line supervisors of law enf. Workers	2.87	Pressers, textile, garment, and related materials	−4.04
First-line supervisors of retail sales workers	2.84	Textile winding, twisting, machine setters	−4.07
Residential advisors/Activities coordinator	2.76	Sewing machine operators	−4.49
First-line supervisors of protective svc workers	2.75	Automotive technicians and repairers	−4.51
Preschool and kindergarten teacher	2.75	Audiovisual equipment installers and repairers	−4.56
Gambling managers	2.70	Knit/Weave machine setters, operators	−5.43
First-line supervisors of firefighters	2.66	Shoe machine operators and tenders	−5.52
Social workers	2.54	Animal breeders	−6.01
Clergy	2.60	Hunters and trappers	−6.21
Food service managers	2.59	Tire builders	−6.39
Occupation (health care)			**Score**
Dentists			4.52
Physical therapists			4.09
Paramedics			4.09
Veterinarians			3.98
Dental assistants			3.77
Respiratory therapists			3.74
Healthcare pracitioners			3.56
Occ therapists			3.52
Vocational nurses			3.50
Physicians and surgeons			3.46
Pharmacists			3.42
Chiropractors			3.39
Physican assistants			3.37
Podiatrists			3.29
Physical therapy aids and assistants			3.25
Radiation therapists			3.02
Diagnostic techs			2.87
Health practitioner support technicians			2.78
Miscellaneous health technicians			2.66
Optometrists			2.56

TABLE A6 | Occupations with highest and lowest ratings of social intensity—Workplace activities.

Highest rated		Lowest rated	
Occupation	**Score**	**Occupation**	**Score**
First-line supervisors of sales workers	8.77	Model makers and patternmakers, wood	−4.62
Human resource managers	8.45	Drywall and ceiling tile installers/Tapers	−4.68
First-line supervisors of law enforcement	7.44	Shoe machine operators and tenders	−4.73
First-line super. of protective service workers	7.03	Textile knitting and weaving machine operators	−4.80
First-line supervisors of firefighters	6.94	Fishers and fishing related	−4.88
Correctional officer supervisors	6.70	Riggers	−4.94
Directors of religious activities and educ.	6.62	Sewing machine operators	−4.97
Mfg. Building and mobile home installers	6.55	Jewelers and precious stone workers	−5.01
Hotel managers	6.54	Print binding and finishing workers	−5.02
Public relations managers	6.54	Automotive glass installers and repairers	−5.15
Principals and preschool administrators	6.51	Machine feeders and offbearers	−5.47
Medical/Health service managers	6.36	Structural metal fabricators and fitters	−5.56
Dieticians and nutritionalists	6.31	Hunters and trappers	−5.59
Regional planners	6.21	Pressers, textile, garment, and related materials	−5.93
CEOs	6.09	Cleaners of vehicles and equipment	−6.03
Postmasters and mail superintendents	6.05	Postal service mail carriers	−6.18
Management analysts	6.01	Woodworking machine setters, operators	−6.32
Meeting and event planners	5.94	Shoe and leather workers and repairers	−6.40
Purchasing managers	5.82	Couriers and messengers	−6.40
General managers/Operations managers	5.66	Coin, vending, and amusement machine service	−6.82
Marketing and sales managers	5.60	Proofreaders and copy markers	−7.88

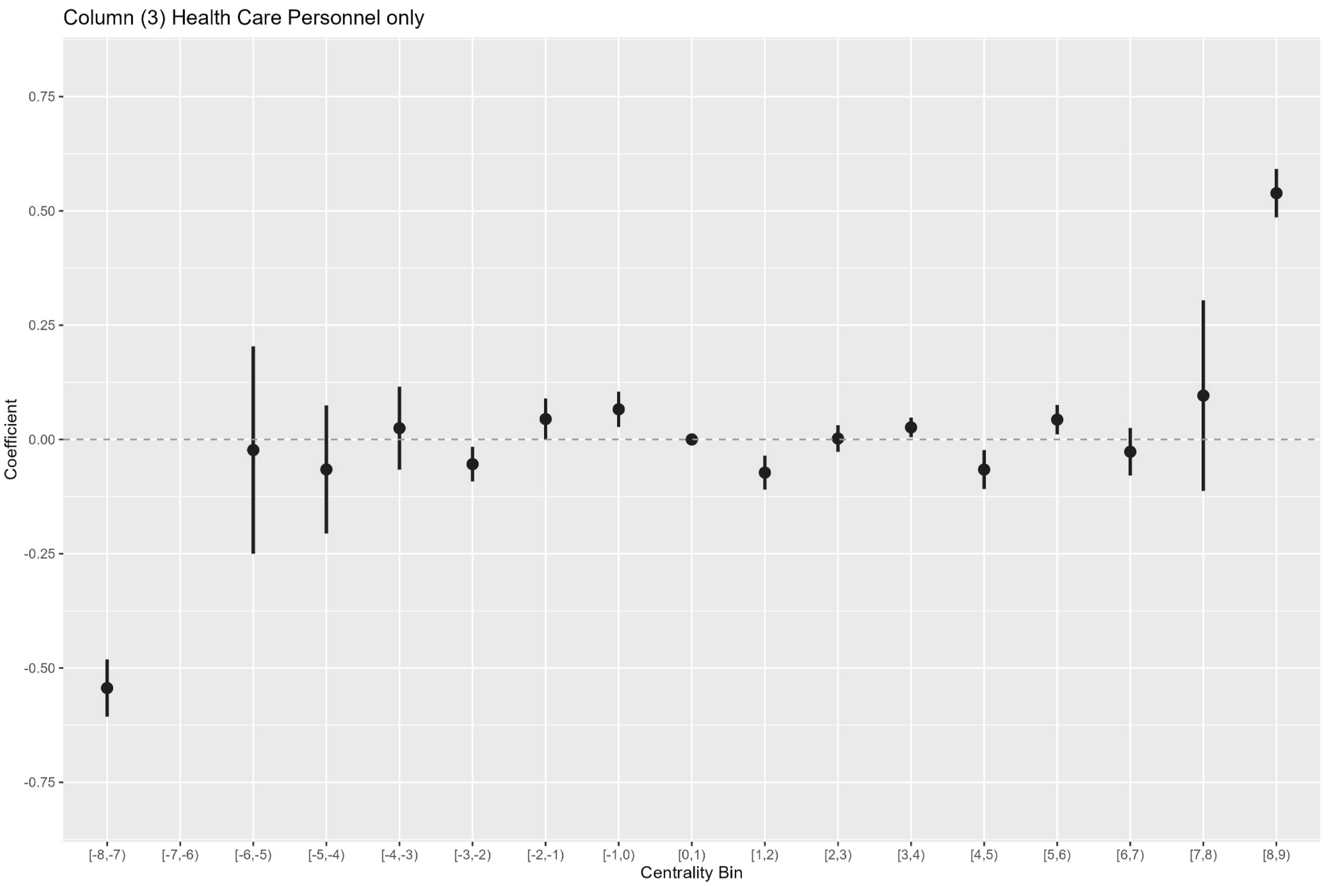

FIGURE A3 | Coefficient and 95% confidence interval estimates of intensity bins in Table A12. Column (3).

TABLE A7 | Occupations with highest and lowest ratings of social intensity—Interpersonal contact.

Highest rated		Lowest rated	
Occupation (non-health care)	**Score**	**Occupation**	**Score**
Hairdressers, hairstylists, and cosmetologists	3.19	Automotive technicians and repairers	−2.64
First-line supervisors of law enforcement	3.07	Sewing machine operators	−2.68
Supervisors of personal care workers	2.93	Cutting and press machine operators	−2.74
Police and sheriff patrollers	2.92	Computer hardware engineers	−2.77
Ticket agents and travel clerks	2.91	Metal furnace operators, tenders, and casters	−2.78
Educational, guidance, and career counselors	2.85	Conveyor operators and tenders	−2.81
Social workers	2.72	Extruding machine setters and operators	−2.85
Barbers	2.68	Refuse and recyclable material collectors	−2.86
Firefighers	2.68	Statisticians	−2.88
Residential advisors/Activities coordinator	2.62	Actuaries	−2.90
First-line supervisors of firefighters	2.59	Misc. Math/Science	−2.93
Special education teachers	2.54	Writers and authors	−2.97
Tellers	2.51	Economists	−3.11
Flight attendant	2.42	Shoe machine operators and tenders	−3.17
First-line super. of protective service workers	2.39	Mathematicians	−3.36
Meeting and event planners	2.26	Lathe and turning machine tool operators	−3.42
Preschool and kindergarten teacher	2.25	Animal breeders	−3.60
First-line super. of food service workers.	2.23	Textile knitting and weaving machine operators	−3.80
Retail salespersons	2.22	Tire builders	−3.91
Community and social service specialists	2.20	Hunters and trappers	−4.34
Occupation (health care)			**Score**
Physical therapists			4.21
Chiropractors			4.06
Dentists			3.99
Healthcare pracitioners			3.95
Radiation therapists			3.94
Vocational nurses			3.63
Paramedics			3.61
Dental assistants			3.59
Chiropractors			3.55
Podiatrists			3.55
Respiratory therapists			3.42
Physical therapy aids and assistants			3.41
Occ therapists			3.36
Optometrists			3.36
Misc. Health technologists and technician			3.31
Physicians and surgeons			3.26
Veterinarians			3.23
Diagnostic techs			3.22
Therapists, NEC			3.21
Physican assistants			2.87

TABLE A8 | Sub-sample test: Oster (2019)'s bound estimates of the coefficients of occupational social intensity and disease exposure for the matched 359 SOCs.

	Controlled effect $\tilde{\beta}$ (1)	Controlled effect R^2 (2)	Oster (2019) bounds $\delta = 1, R_{Max} = 1.3R^2$ (3)	δ^0 for $\beta = 0$ (4)
Panel A. Intensity				
(1) Full sample	0.0027***	0.143	[−0.0048, 0.0027]	0.0226
	(0.0006)			
(2) Full sample	0.0009	0.144	[−0.0090, 0.0009]	0.0059
	(0.0007)			
(3) Full sample	0.0016***	0.145	[−0.0063, 0.0016]	0.0130
	(0.0005)			
(4) Full sample	0.0006	0.145	[−0.0091, 0.0006]	0.0043
	(0.0007)			
(5) HCP	0.0064**	0.116	[−0.0100, 0.0064]	0.0265
	(0.0024)			
(6) Non-HCP	0.0008	0.111	[−0.0031, 0.0008]	0.0099
	(0.0007)			
Panel B. Exposure				
(1) Full sample	0.0290***	0.145	[−0.0209, 0.0290]	0.0421
	(0.0032)			
(2) Full sample	0.0260***	0.145	[−0.0321, 0.0260]	0.0335
	(0.0034)			
(3) HCP	0.0250***	0.116	[0.0091, 0.0250]	0.0917
	(0.0053)			
(4) Non-HCP	0.0262***	0.111	[0.0260, 0.0262]	2.1229
	(0.0036)			

Note: Estimates in Columns (1)–(2) are from Table 9. Estimates in Columns (3)–(4) are computed using Stata command *psacalc*. Each identified set in Column (3) includes the value of β in the controlled model and the value of β calculated for $R_{Max} = 1.3R^2$, under the assumption that selection on observables and unobservables is proportional. Panel A presents the bound estimates for the coefficients of Intensity, and Panel B presents the bound estimates for Exposure. State and flu season fixed effects are controlled. ***, ***, * means that the coefficient is statistically different from zero at the 1%, 5%, and 10% level. All models weighted with BRFSS weights. Standard errors clustered at the state-level and reported in parentheses.

TABLE A9 | Robustness check: Oster (2019)'s bound estimates of the coefficients of occupational social intensity (context) and disease exposure.

	Controlled effect $\tilde{\beta}$ (1)	Controlled effect R^2 (2)	Oster (2019) bounds $\delta = 1, R_{Max} = 1.3R^2$ (3)	δ^0 for $\beta = 0$ (4)
Panel A. Intensity (context)				
(1) Full sample	0.0023	0.136	[−0.0116, 0.0023]	0.0096
	(0.0014)			
(2) Full sample	0.0002	0.136	[−0.0168, 0.0002]	0.0009
	(0.0014)			
(3) Full sample	−0.0020	0.138	[−0.0198,−0.0020]	−0.0071
	(0.0014)			
(4) Full sample	−0.0025*	0.138	[−0.0221,−0.0025]	−0.0081
	(0.0014)			
(5) HCP	−0.0134**	0.109	[−0.0627,−0.0134]	−0.0185
	(0.0055)			
(6) Non-HCP	−0.0016	0.107	[−0.0098,−0.0016]	−0.0097
	(0.0014)			
Panel B. Exposure				
(1) Full sample	0.0275***	0.138	[−0.0254, 0.0275]	0.0367
	(0.0025)			
(2) Full sample	0.0265***	0.138	[−0.0369, 0.0265]	0.0313
	(0.0026)			
(3) HCP	0.0393***	0.109	[-4.4649, 0.0393]	0.1954
	(0.0055)			
(4) Non-HCP	0.0229***	0.107	[0.0199, 0.0229]	0.2713
	(0.0028)			

Note: Estimates in Columns (1)–(2) are from Table 10. Estimates in Columns (3)–(4) are computed using Stata command *psacalc*. Each identified set in Column (3) includes the value of β in the controlled model and the value of β calculated for $R_{Max} = 1.3R^2$, under the assumption that selection on observables and unobservables is proportional. Panel A presents the bound estimates for the coefficients of Intensity, and Panel B presents the bound estimates for Exposure. State and flu season fixed effects are controlled. ***, ***, * means that the coefficient is statistically different from zero at the 1%, 5%, and 10% level. All models weighted with BRFSS weights. Standard errors clustered at the state-level and reported in parentheses.

TABLE A10 | Robustness check: Oster (2019)'s bound estimates of the coefficients of occupational social intensity (activity) and disease exposure.

	Controlled effect		Oster (2019) bounds	δ^0 for
	$\tilde{\beta}$	R^2	$\delta = 1, R_{Max} = 1.3R^2$	$\beta = 0$
	(1)	(2)	(3)	(4)
Panel A. Intensity (activity)				
(1) Full sample	0.0025***	0.136	[−0.0057, 0.0025]	0.0179
	(0.0006)			
(2) Full sample	0.0015**	0.137	[−0.0089, 0.0015]	0.0092
	(0.0007)			
(3) Full sample	0.0017***	0.138	[−0.0070, 0.0017]	0.0116
	(0.0006)			
(4) Full sample	0.0012*	0.138	[−0.0092, 0.0012]	0.0072
	(0.0006)			
(5) HCP	0.0020	0.108	[−0.0106, 0.0020]	0.0093
	(0.0016)			
(6) Non-HCP	0.0015**	0.107	[−0.0042, 0.0015]	0.0120
	(0.0006)			
Panel B. Exposure				
(1) Full sample	0.0259***	0.138	[−0.0199, 0.0259]	0.0382
	(0.0023)			
(2) Full sample	0.0250***	0.138	[−0.0234, 0.0250]	0.0354
	(0.0024)			
(3) HCP	0.0301***	0.108	[0.0254, 0.0301]	0.2724
	(0.0046)			
(4) Non-HCP	0.0219***	0.107	[0.0187, 0.0219]	0.2846
	(0.0027)			

Note: Estimates in Columns (1)–(2) are from Table 11. Estimates in Columns (3)–(4) are computed using Stata command *psacalc*. Each identified set in Column (3) includes the value of β in the controlled model and the value of β calculated for $R_{Max} = 1.3R^2$, under the assumption that selection on observables and unobservables is proportional. Panel A presents the bound estimates for the coefficients of Intensity, and Panel B presents the bound estimates for Exposure. State and flu season fixed effects are controlled. ***, ***, * means that the coefficient is statistically different from zero at the 1%, 5%, and 10% level. All models weighted with BRFSS weights. Standard errors clustered at the state-level and reported in parentheses.

TABLE A11 | Robustness check: Oster (2019)'s bound estimates of the coefficients of occupational social intensity (contact) and disease exposure.

	Controlled effect		Oster (2019) bounds	δ^0 for
	$\tilde{\beta}$	R^2	$\delta = 1, R_{Max} = 1.3R^2$	$\beta = 0$
	(1)	(2)	(3)	(4)
Panel A. Intensity (contact)				
(1) Full sample	0.0030*	0.136	[−0.0114, 0.0030]	0.0129
	(0.0015)			
(2) Full sample	−0.0006	0.137	[−0.0247,−0.0006]	−0.0019
	(0.0016)			
(3) Full sample	−0.0107***	0.138	[−0.0549,−0.0107]	−0.0216
	(0.0016)			
(4) Full sample	−0.0112***	0.139	[−0.0621,−0.0112]	−0.0215
	(0.0017)			
(5) HCP	−0.0219***	0.110	[−0.1185,−0.0219]	−0.0240
	(0.0037)			
(6) Non-HCP	−0.0104***	0.108	[−0.0204,−0.0104]	−0.1115
	(0.0016)			
Panel B. Exposure				
(1) Full sample	0.0376***	0.138	[−0.0912, 0.0376]	0.0311
	(0.0027)			
(2) Full sample	0.0364***	0.139	[−0.1354, 0.0364]	0.0271
	(0.0028)			
(3) HCP	0.0551***	0.110	[−0.0346, 0.0551]	0.0900
	(0.0058)			
(4) Non-HCP	0.0327***	0.108	[-8.0767, 0.0327]	0.2709
	(0.0029)			

Note: Estimates in Columns (1)–(2) are from Table 12. Estimates in Columns (3)–(4) are computed using Stata command *psacalc*. Each identified set in Column (3) includes the value of β in the controlled model and the value of β calculated for $R_{Max} = 1.3R^2$, under the assumption that selection on observables and unobservables is proportional. Panel A presents the bound estimates for the coefficients of Intensity, and Panel B presents the bound estimates for Exposure. State and flu season fixed effects are controlled. ***, ***, * means that the coefficient is statistically different from zero at the 1%, 5%, and 10% level. All models weighted with BRFSS weights. Standard errors clustered at the state-level and reported in parentheses.

TABLE A12 | Association between occupational social intensity bins and flu vaccine uptake.

	(1) Full	(2) Full	(3) HCP	(4) Non-HCP
Intensity: [0,1) as base level				
[− 8, − 7)	0.0413	0.0558	−0.5438***	0.0634
	(0.1051)	(0.1051)	(0.0319)	(0.1039)
[− 7, − 6)	−0.0053	0.0059		0.0036
	(0.0179)	(0.0179)		(0.0177)
[− 6, − 5)	−0.0281	−0.0315*	−0.0231	−0.0293
	(0.0177)	(0.0185)	(0.1157)	(0.0200)
[− 5, − 4)	−0.0033	0.0070	−0.0656	0.0084
	(0.0134)	(0.0134)	(0.0715)	(0.0140)
[− 4, − 3)	0.0222	0.0157	0.0247	0.0161
	(0.0149)	(0.0148)	(0.0463)	(0.0141)
[− 3, − 2)	−0.0028	0.0009	−0.0540***	0.0066
	(0.0090)	(0.0090)	(0.0193)	(0.0098)
[− 2, − 1)	0.0276***	0.0326***	0.0447*	0.0318***
	(0.0085)	(0.0084)	(0.0229)	(0.0106)
[− 1, 0)	0.0319***	0.0403***	0.0660***	0.0378***
	(0.0075)	(0.0073)	(0.0197)	(0.0087)
[1, 2)	0.0035	0.0042	−0.0725***	0.0141
	(0.0119)	(0.0120)	(0.0189)	(0.0151)
[2, 3)	0.0389***	0.0330***	0.0021	0.0345***
	(0.0102)	(0.0104)	(0.0149)	(0.0106)
[3, 4)	0.0281***	0.0285***	0.0264**	0.0264***
	(0.0054)	(0.0054)	(0.0109)	(0.0071)
[4, 5)	0.0065	0.0085	−0.0658***	0.0136
	(0.0091)	(0.0090)	(0.0217)	(0.0097)
[5, 6)	0.0709***	0.0500***	0.0434**	0.0359**
	(0.0103)	(0.0105)	(0.0164)	(0.0138)
[6, 7)	0.0113	0.0159	−0.0269	0.0300*
	(0.0123)	(0.0121)	(0.0265)	(0.0155)
[7, 8)	0.0245	0.0162	0.0960	0.0182
	(0.0361)	(0.0373)	(0.1064)	(0.0385)
[8, 9)	0.0744***	0.0831***	0.5388***	0.0838***
	(0.0258)	(0.0260)	(0.0270)	(0.0254)
Exposure		0.0246***	0.0281***	0.0215***
		(0.0024)	(0.0056)	(0.0027)
Health care personnel	0.2153***	0.1809***		
	(0.0070)	(0.0077)		
Constant	0.9582***	0.9518***	0.2520***	0.9653***
	(0.0191)	(0.0188)	(0.0684)	(0.0181)
Observations	229,144	229,144	36,347	192,797

Note: Columns (1)–(2) use full sample. Column (3) uses health care personnel only, while column (4) excludes them. The base-level bin [0,1) is omitted. We control for sex, age, race/ethnicity, education, income, marital status, number of children, and indicators for high-risk conditions, having personal health care provider, and having medical insurance. We also controlfor the interaction of state and flu season fixed effects.
***, ***, * mean that the coefficient is statistically different from zero at the 1%, 5%, and 10% level. All models weighted with BRFSS weights. Standard errors clustered at the state-level and reported in parentheses.
Source: BRFSS 2013–2015 and O*NET.

TABLE A13 | Sensitivity analysis: Column (4) without state-by-season fixed effects.

	(1) Full	(2) Full	(3) Full	(4) Full	(5) Full	(6) Full	(7) Full	(8) Full
Intensity	0.0122***	0.0092***	0.0033***	0.0014**	0.0015**	0.0007	0.0007	0.0008
	(0.0007)	(0.0008)	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0006)
Exposure	0.0235***	0.0235***	0.0247***	0.0285***	0.0273***	0.0247***	0.0251***	0.0249***
	(0.0027)	(0.0025)	(0.0024)	(0.0024)	(0.0024)	(0.0026)	(0.0025)	(0.0024)
Health care personnel	0.1932***	0.1848***	0.1829***	0.1809***	0.1823***	0.1755***	0.1747***	0.1740***
	(0.0075)	(0.0084)	(0.0086)	(0.0083)	(0.0084)	(0.0083)	(0.0080)	(0.0078)
Intensity × HCP	0.0043**	0.0042**	0.0032*	0.0015	0.0016	0.0038**	0.0038**	0.0038**
	(0.0016)	(0.0017)	(0.0016)	(0.0016)	(0.0016)	(0.0016)	(0.0016)	(0.0015)
Constant	0.2992***	0.2072***	0.1625***	0.1554***	0.1858***	0.0806***	0.0672***	0.9483***
	(0.0112)	(0.0082)	(0.0081)	(0.0141)	(0.0149)	(0.0155)	(0.0187)	(0.0191)
Sex, age, race		X	X	X	X	X	X	X
Education			X	X	X	X	X	X
Income				X	X	X	X	X
Marital, #children					X	X	X	X
Health-related						X	X	X
State FE							X	—
State × season FE								X
Observations	229,144	229,144	229,144	229,144	229,144	229,144	229,144	229,144
R^2	0.049	0.079	0.090	0.095	0.098	0.117	0.125	0.138

Note: While column (1) does not control for demographics, columns (2)–(6) control various sets of sex, age, race/ethnicity, education, income, marital status, number of children, health-related indicators, and the interaction of state and flu season fixed effects (tagged by X). All columns use full sample.
***, ***, * mean that the coefficient is statistically different from zero at the 1%, 5%, and 10% level. All models weighted with BRFSS weights. Standard errors clustered at the state-level and reported in parentheses.
Source: BRFSS 2013–2015 and O*NET.

TABLE A14 | Sub-sample analysis: Column (4) by sex, age, race/ethnicity.

	(1) Female	(2) Male	(3) 18–49 years	(4) 50–64 years	(5) 65+ years	(6) White	(7) Black	(8) Hispanic	(9) Other
Intensity	−0.0010	0.0019**	−0.0002	0.0023**	0.0010	0.0014**	0.0026	−0.0011	−0.0000
	(0.0009)	(0.0007)	(0.0008)	(0.0009)	(0.0021)	(0.0007)	(0.0020)	(0.0022)	(0.0034)
Exposure	0.0206***	0.0309***	0.0279***	0.0210***	0.0289***	0.0259***	0.0092*	0.0196**	0.0413***
	(0.0032)	(0.0039)	(0.0028)	(0.0023)	(0.0072)	(0.0031)	(0.0052)	(0.0086)	(0.0113)
Health care personnel	0.1700***	0.1865***	0.1865***	0.1669***	0.0447	0.1893***	0.1594***	0.1250***	0.1860***
	(0.0112)	(0.0121)	(0.0118)	(0.0122)	(0.0282)	(0.0129)	(0.0202)	(0.0231)	(0.0315)
Intensity × HCP	0.0073***	−0.0029	0.0060**	0.0011	−0.0022	0.0015	0.0038	−0.0005	0.0061
	(0.0015)	(0.0022)	(0.0027)	(0.0021)	(0.0065)	(0.0027)	(0.0059)	(0.0073)	(0.0092)
Constant	0.7985***	0.9798***	0.9928***	0.8646***	0.5458***	0.8629***	0.8964***	0.9351***	0.8069***
	(0.0276)	(0.0189)	(0.0176)	(0.0294)	(0.0702)	(0.0292)	(0.0534)	(0.0553)	(0.0795)
Observations	120,487	108,657	111,936	93,485	23,723	186,699	14,886	16,389	11,170
R^2	0.129	0.134	0.133	0.115	0.082	0.144	0.111	0.131	0.180

Note: The sample is divided into subsamples by demographics: columns (1)–(2) by sex; columns (3)–(5) by age; columns (6)–(9) by race/ethnicity. We control for sex, age, race/ethnicity, education, income, marital status, number of children, and indicators for high-risk conditions, having personal health care provider, and having medical insurance, unless the variable is used as a label for the subsample. We also control for the interaction of state and flu season fixed effects.
***, ***, * mean that the coefficient is statistically different from zero at the 1%, 5%, and 10% level. All models weighted with BRFSS weights. Standard errors clustered at the state-level and reported in parentheses.
Source: BRFSS 2013–2015 and O*NET.

TABLE A15 | Sub-sample analysis: Column (4) by marital status, children.

	(1) Never married	(2) Unmarried couple	(3) Married	(4) Other	(5) No children	(6) Have children
Intensity	0.0010	0.0020	0.0009	0.0001	0.0015**	−0.0001
	(0.0013)	(0.0024)	(0.0009)	(0.0017)	(0.0006)	(0.0009)
Exposure	0.0208***	0.0324***	0.0256***	0.0257***	0.0217***	0.0286***
	(0.0060)	(0.0063)	(0.0029)	(0.0042)	(0.0024)	(0.0028)
Health care personnel	0.1589***	0.1724***	0.1840***	0.1698***	0.1709***	0.1768***
	(0.0169)	(0.0450)	(0.0151)	(0.0136)	(0.0104)	(0.0120)
Intensity × HCP	0.0075*	0.0063	0.0022	−0.0006	0.0033**	0.0041
	(0.0043)	(0.0094)	(0.0018)	(0.0041)	(0.0015)	(0.0028)
Constant	0.9653***	0.3206***	0.6692***	0.8642***	0.9278***	1.0483***
	(0.0217)	(0.0860)	(0.0631)	(0.0373)	(0.0224)	(0.0338)
Observations	36,570	8005	139,613	44,956	141,130	88,014
R^2	0.121	0.150	0.129	0.148	0.145	0.135

Note: The sample is divided into subsamples by demographics: columns (1)–(4) by marital status, and column (4) of other includes divorced, widowed, and separated individuals; columns (5)–(6) by number of children. We control for sex, age, race/ethnicity, education, income, marital status, number of children, and indicators for high-risk conditions, having personal health care provider, and having medical insurance, unless the variable is used as a label for the subsample. We also control for the interaction of state and flu season fixed effects.
***, ***, * mean that the coefficient is statistically different from zero at the 1%, 5%, and 10% level. All models weighted with BRFSS weights. Standard errors clustered at the state-level and reported in parentheses.
Source: BRFSS 2013–2015 and O*NET.

TABLE A16 | Sub-sample analysis: Column (4), health-related indicators.

	(1) High-risk Conditions	(2) No high-risk Conditions	(3) No healthcare Provider	(4) Have healthcare Provider	(5) No insurance	(6) Have insurance
Intensity	−0.0021	0.0016**	−0.0015	0.0015**	−0.0035**	0.0015**
	(0.0013)	(0.0006)	(0.0015)	(0.0006)	(0.0015)	(0.0006)
Exposure	0.0163***	0.0266***	0.0290***	0.0241***	0.0204***	0.0256***
	(0.0049)	(0.0027)	(0.0052)	(0.0024)	(0.0063)	(0.0025)
Health care personnel	0.1500***	0.1818***	0.1433***	0.1819***	0.1096***	0.1843***
	(0.0129)	(0.0093)	(0.0126)	(0.0084)	(0.0191)	(0.0081)
Intensity × HCP	0.0013	0.0041**	0.0066	0.0023	0.0084	0.0017
	(0.0034)	(0.0017)	(0.0051)	(0.0014)	(0.0055)	(0.0019)
Constant	1.0140***	0.9523***	0.9631***	0.7416***	0.8677***	1.0485***
	(0.0473)	(0.0126)	(0.0276)	(0.0201)	(0.0557)	(0.0234)
Observations	52,375	176,769	43,961	185,183	22,763	206,381
R^2	0.121	0.137	0.099	0.112	0.092	0.118

Note: The sample is divided into subsamples by health-related indicators: columns (1)–(2) by high-risk condition; columns (3)–(4) by healthcare provider; columns (5)–(6) by health insurance. We control for sex, age, race/ethnicity, education, income, marital status, number of children, and indicators for high-risk conditions, having personal health care provider, and having medical insurance, unless the variable is used as a label for the subsample. We also control for the interaction of state and flu season fixed effects.
***, ***, * mean that the coefficient is statistically different from zero at the 1%, 5%, and 10% level. All models weighted with BRFSS weights. Standard errors clustered at the state-level and reported in parentheses.
Source: BRFSS 2013–2015 and O*NET.

TABLE A17 | 2014 flu and pneumonia mortality and 2015 flu vaccine uptake (no FE).

	(1) Full	(2) Full	(3) Full	(4) Full	(5) HCP	(6) Non-HCP
Intensity	−0.0054	−0.0065	−0.0064	−0.0070	0.0038	−0.0071
	(0.0047)	(0.0047)	(0.0046)	(0.0047)	(0.0085)	(0.0043)
Mortality	0.0037	0.0037	0.0036	0.0036	0.0059	0.0033
	(0.0038)	(0.0038)	(0.0038)	(0.0038)	(0.0056)	(0.0036)
Intensity × mortality	0.0006*	0.0006*	0.0006*	0.0006*	−0.0001	0.0006**
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0006)	(0.0003)
Exposure			0.0266***	0.0253***	0.0271***	0.0235***
			(0.0035)	(0.0035)	(0.0066)	(0.0044)
Health care personnel	0.2182***	0.1989***	0.1765***	0.1678***		
	(0.0126)	(0.0146)	(0.0124)	(0.0145)		
Intensity × HCP		0.0087***		0.0049**		
		(0.0024)		(0.0023)		
Constant	0.0144	0.0133	0.0235	0.0224	−0.0387	0.0363
	(0.0550)	(0.0548)	(0.0547)	(0.0547)	(0.0629)	(0.0568)
Observations	72,084	72,084	72,084	72,084	11,570	60,514
R-squared	0.115	0.115	0.117	0.117	0.090	0.085

Note: Columns (1)–(4) use full sample. Column (5) uses health care personnel only, while column (6) excludes them. The 2014 Flu and Pneumonia Mortality is the number of deaths per 100,000 total population at the state level. The 2015 Flu Vaccine Uptake is whether the respondent interviewed in 2015 reports receiving a flu shot in the past 12 months. We control for sex, age, race/ethnicity, education, income, marital status, number of children, and indicators for high-risk conditions, having personal health care provider, and having medical insurance. We also control for the interaction of state and flu season fixed effects.
***, ***, * mean that the coefficient is statistically different from zero at the 1%, 5%, and 10% level. All models weighted with BRFSS weights. Standard errors clustered at the state-level and reported in parentheses.
Source: BRFSS 2013–2015, O*NET, and NCHS.