RESEARCH

Projected geographic disparities in healthcare worker absenteeism from COVID-19 school closures and the economic feasibility of child care subsidies: a simulation study (Supplementary Information)

Elizabeth T Chin^{*†}, Benjamin Q Huynh[†], Nathan C Lo, Trevor Hastie and Sanjay Basu

*Correspondence: etchin@stanford.edu Full list of author information is available at the end of the article †Equal contributor

Economic analysis

Here we describe how we obtained county-level estimates of childcare costs and wages. We use state-level child care costs from CCAoA and adjust them to countylevel by applying the ratio between state-level and county-level fair market rents from HUD. We calculate state-level rents from HUD by taking population-weighted averages of county rents. To estimate the number of healthcare workers with children at the county-level, we take the state-level proportion of healthcare workers with children from IPUMS and apply it to the county-level number of healthcare workers from ACS. We then calculate the county-level cost of providing child care to healthcare workers by multiplying child care costs by the proportion of healthcare workers with children.

For estimating county-level wages, some counties with low populations had redacted wages to preserve anonymity. We used multiple imputation by chained equations to impute these cases. To get all county-level wages, we multiplied the number of healthcare workers (by occupation group and sex) by their subgrouprespective county-level median wages.

Robustness checks

Here we provide sensitivity analyses and robustness checks of our estimates across various parameters. Table S2 shows different estimates of unmet child care needs based on different assumptions for determining child care needs. Table S3 Shows different estimates of school closure effectiveness based on varying the basic reproduction number from 2 to 6 when holding the reproduction number constant across all states and when varying the basic reproduction number using state-specific estimates across the mean and 95% confidence intervals.

Figure S1 displays the cut off points of rho and delta for 70% of counties to reach $\omega > 1$, where $\rho = \{0.5, 0.6, 0.7, 0.8, 0.9, 1\}$ is the proportion of those with unmet child care needs who go on to be absent from work and $\delta = \{1, 1.1, 1.2, 1.3, 1.4\}$ is the increased cost for emergency child care compared to normal costs. Even under pessimistic parameter assumptions ($\rho = 0.6, \delta = 1.4$), 70% of counties can still afford partial child care subsidies (70%) over bearing absenteeism costs.

Figure S2 shows that the rurality proportions across counties remains relatively constant across values of ρ , suggesting county characteristics do not change across parameter changes.

The list of occupation codes through the American Community Survey that were used to categorize essential workers is included as Additional file 3: Table S1: Occupation codes for essential worker classification.

Figure S1: Sensitivity analysis of parameter thresholds for 70% of counties to reach $\omega > 1$. Colored lines indicate different levels of subsidization rates.

Figure S2: Proportion of counties with higher rates of lost wages due to absenteeism than costs of child care $(\omega > 1)$ across $\rho = \{0.5, 0.6, 0.7, 0.8, 0.9, 1.\}$. Bars are shaded based on the level of rurality of counties.

Additional maps

Absenteeism, complication factors, and wages

Figure S3: County-level comparison of percent of healthcare worker households with unmet childcare needs and cardiovascular disease mortality (deaths per 100,000 people). Counties with confidence interval sizes in the 90th percentile or below are shown.

Figure S4: County-level comparison of percent unmet childcare needs and ω . Counties with confidence interval sizes in the 90th percentile or below are shown.

Figure S5: County-level comparison of percent of healthcare worker households with unmet child care needs and effectiveness of school closures using estimated reduction of peak ICU bed demand. Counties with confidence interval sizes in the 90th percentile or below are shown. No within-state normalization used.

Sensitivity analyses

Unmet childcare needs estimate

Table S2: Sensitivity analysis of unmet childcare need estimates using various population seeds. National Household Education Surveys Program (NHES) found that 50% of households had difficulty finding or could not find satisfactory child care. Integrated Public Use Microdata Series (IPUMS) are state specific seeds derived from the household structure of healthcare workers. Survey data from both the Pew Research Center and the US Census Bureau indicating that 89% of working couples rely on the mother for primary child care. To test sensitivity, we calculated unmet childcare need by assuming that 60% of working couples rely on the mother for primary child care. The NE, GP, OC, and OA model assumptions denote non-essential workers, grandparents, older children, and other adults (respectively) in the household are excluded as potential caregivers.

Transmission models

Table S3: Sensitivity analysis of transmission models under varying R_0 values and contact conditions. Estimates of the mean and 95% confidence intervals for initial reproduction number by states were retrieved on 2020-05-29. School closures (SC) reduce the risk of child-child interactions by 90%. Household (HH) interactions increase child-other age group interactions by 10%.

Model output

```
The glm() calls for our models and model output are below.
modelDiabetes <- glm(stateEstMeans~Diabetes.prevalence.raw.value+
X..65.and.older.raw.value+femalePct+pctMarried+
                X..below.18.years.of.age.raw.value+
                X..Non.Hispanic.African.American.raw.value+
                factor(state)+
                X..Hispanic.raw.value+
                X..American.Indian.and.Alaskan.Native.raw.value
                +Population.raw.value+X..Rural.raw.value,
              weights=numHCW,family=quasipoisson,
              data=regressionData)
modelCVD <- glm(stateEstMeans~cvdMortality+
X..65.and.older.raw.value+
femalePct+fmr+pctMarried+
                X..below.18.years.of.age.raw.value+
                X..Non.Hispanic.African.American.raw.value+
                factor(state)+
                X..Hispanic.raw.value+
                X..American.Indian.and.Alaskan.Native.raw.value+
                Population.raw.value+X..Rural.raw.value,
              weights=numHCW,family=quasipoisson,
              data=regressionData)
modelControls <- glm(stateEstMeans~X..Rural.raw.value+
X..65.and.older.raw.value+
femalePct+fmr+pctMarried+
                X..below.18.years.of.age.raw.value+
                X..Non.Hispanic.African.American.raw.value+
                factor(state)+
                X..Hispanic.raw.value+
                X..American.Indian.and.Alaskan.Native.raw.value+
                Population.raw.value,
              weights=numHCW,family=quasipoisson,
              data=regressionData)
```

```
summary(modelDiabetes)
summary(modelCVD)
summary(modelControls)
```


 $\frac{***p}{x}$ < 0.001, $*^{*}p$ < 0.01, $*p$ < 0.05

Table S4: Regression output for models on diabetes, cardiovascular disease, percent rural, and controls.