sociological science

Supplement to:

Lundberg, Ian. 2025. "The Causal Impact of Segregation on a Disparity: A Gap-Closing Approach" Sociological Science 12: 871-890.

Online Supplement

The causal impact of segregation on a disparity:

A gap-closing approach

Ian Lundberg*

October 13, 2025

A Weighting of the CPS-ASEC

The CPS-ASEC samples people with unequal probabilities. All analyses in the main text weight by the ASECWT weighting variable designed for the ASEC, with inference calculated from variances estimated with replicate weights (Appendix D). Because the U.S. population grows over time, I allow the weight variable to place greater weight on the more recent years proportional to the relatively larger size of the population (see Fig 14).

In addition to weighting population averages, the analyses also weight the generalized additive models (Wood, 2017) used to estimate conditional outcomes given predictors. The reason to weight the GAM is because it is only an approximation to a conditional mean function that is likely to be more complex; in fact, there are likely to be omitted interactions that the additive aspect of the GAM will miss. The aim is for the GAM to provide a useful and parsimonious approximation to the more complex conditional mean function. In order

*Department of Sociology, 264 Haines Hall, 375 Portola Plaza, UCLA, Los Angeles, CA 90095, ianlundberg.org, ianlundberg@ucla.edu.

to prioritize the quality of that approximation in regions of the covariate space that are more heavily weighted by the sampling weights (where the predictions will matter most), the GAM estimation is weighted.

B Methods for descriptive scatterplots (Fig 1)

Because of the large number of occupations, some occupations are sparsely populated. The descriptive occupation-level analyses in Fig 1 use a partial-pooling strategy to improve precision. I first score each occupation on two dimensions: (1) the racial/ethnic composition of those employed in the occupation and (2) the proportion of those employed in the occupation this year to report a work-limiting disability next year. I estimate each score by a regression model with occupation random intercepts and no other predictors, thereby reducing the high-variance estimates that would otherwise be produced in small occupations. The scatterplots visualize occupation-level onset of work-limiting disability plotted against occupation-level racial composition. The best-fit line is a linear regression through the estimated points (occupations), with each occupation weighted proportional to size.

C Measurement issues with work-limiting disability in the CPS-ASEC

There are two important limitations to the operationalization of work-limiting disability in the CPS-ASEC: changes in the survey instrument over time and limitations in the definition of disability for a self report.

C.1 Changes in the CPS-ASEC survey instrument

Prior to 2014, the CPS-ASEC instrument was relatively constant over time. The household respondent answered a work-limiting disability question for each household member: “(Do you/Does anyone in the household) have a health problem or disability which prevents (you/them) from working or which limits the kind or amount of work (you/they) can do?" Two important changes occurred after that, one in 2014 and one in 2016.

In 2014, the wording of this question remained the same but other aspects of the CPS-ASEC underwent a major redesign.[1] All respondents answered new questions about health insurance, and 3/8 of the sample received income questions in a new format. After 2014, the CPS-ASEC followed the new format for all respondents. Although the work-limiting disability question did not change, Fig 8 Panel A shows that the reported prevalence of work-limiting disability rose substantially under the new questionnaire structure. One reason for this increase might be the change in the order of the questionnaire. In the traditional questionnaire, the module immediately preceding work-limiting disability was about Social Security survivor benefits. In the redesigned questionnaire, the module immediately preceding work-limiting disability was about Supplemental Security Income (SSI), a Social Security program that provides transfers to individuals with disabilities (Czajka et al. 2015:5). Asking someone to report their receipt of SSI may prime them to report a work-limiting disability at a higher rate in the immediately subsequent questionnaire module.

In 2016, the wording of the work-limiting disability question changed to focus explicitly on work-limiting disability in the past calendar year, with an added prompt about limitations that were in place "even for a short time."[2] As illustrated in Figure 8, this wording change produced no apparent change in the proportion reporting a work-limiting disability (Panel A) or in the Black-white disparity (Panel B).The analyses therefore pool over all years.

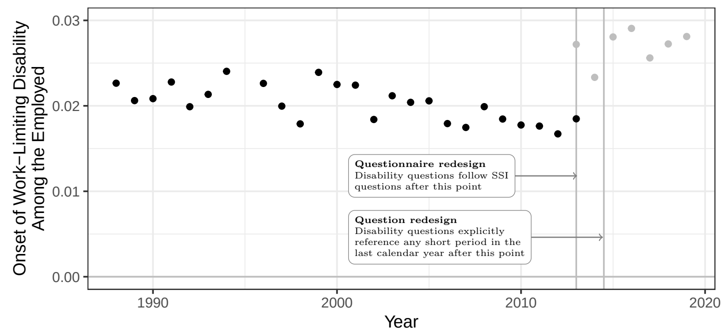
C.2 Validity in the meaning of the disability question

Self-reported work-limiting disability faces the same problems of survey measurement that affect any question: it captures disabilities that limit work only to the extent that respondents

[1]Details of the CPS-ASEC 2014 redesign are available in a BLS memo: `https://cps.ipums.org/cps/resources/other_docs/ASEC_Redesign.pdf`

[2]Full question wordings are available at `https://cps.ipums.org/cps-action/variables/DISABWRK#questionnaire_text_section`.

A) Reports of a work-limiting disability increased when the questionnaire changed



B) The key disparity of interest remained roughly constant

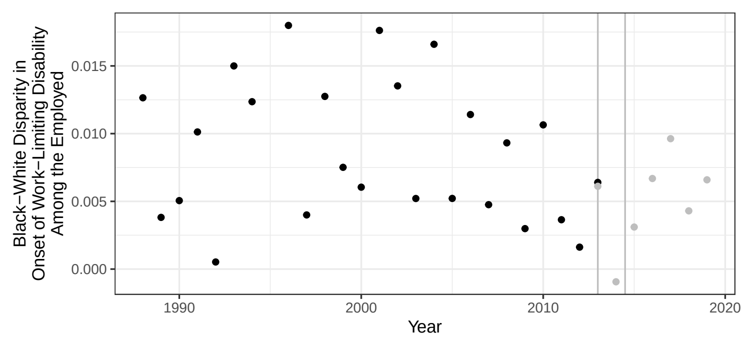


Fig. 8. Changes to the survey instrument affected levels but not disparities in reported work-limiting disability. For consistency with other plots, the x-axis here refers to the first year of observation, but the outcome is measured in the second year of observation. Therefore, CPS questionnaire changes in 2014 and before 2016 correspond to vertical lines placed at 2013 and before 2015 in this plot. In 2014, the CPS questionnaire was redesigned and modules were re-ordered so that the disability questions immediately followed the Supplemental Security Income questions instead of the Survivor Benefit questions. The redesigned questionnaire was given to 3/8 of the sample (gray dot on the vertical line), while 5/8 of the sample received the original questionnaire (black dot on the vertical line). Before 2014, everyone answered the traditional questionnaire (black dots) and after 2014 everyone answered the redesigned questionnaire (gray dots). Beginning in 2016, the question about "a health problem or a disability which prevents him/her from working or which limits the kind or amount of work" was changed to reference the past calendar year and to include the prompt for conditions lasting "even for a short time." The 2014 redesign increased reported disability, while the 2016 question change appears to have had little effect on reported disability. The Black-white disparity in disability onset is roughly constant over both changes to the survey instrument. The main analyses therefore pool over both formats of the question.

understand them as such. Hale (2001) critiques the CPS work-limiting disability question for defining disabilities in terms of work limitations rather than in terms of specific medical impairments. In an evaluation that compares the CPS work-limiting disability question to a broader definition of disability from a series of questions in the National Health Interview Survey, Burkhauser et al. (2002) show that the work-limiting disability question estimates a smaller population with a disability but that employment trends in both populations are similar. Maestas et al. (2019) collect new data to show that those who report a workplace accomodation for health reasons may still answer "no" to a question about whether they have a work-limiting disability. One should therefore understand a work-limiting disability as one particular operationalization of a broader concept of disability. It is possible that self-perceptions of disability deviate from how a medical professional would diagnose a disability, and those deviations may differ across individuals. Yet, one's own perception that one has a work-limiting disability may be a worthy object of study in its own right; it is real for the person involved even if it is distinct from the diagnosis a medical professional would provide. A further advantage of accepting these self-reports as worthy of study is the ability to use the very large samples of the CPS with detailed occupational information linked over a 12-month interval.

D Statistical inference by replicate weights

Statistical inference (e.g. standard errors and confidence intervals) involves a claim about how an estimator would vary if applied to hypothetical new samples drawn from the sampling frame by the same procedure used to generate the observed sample. Standard procedures assume a simple random sample and apply only if that is how the data were generated. The CPS is not a simple random sample. Instead, geographic Primary Sampling Units (PSUs) are selected within geographic strata, and then geographic Ultimate Sampling Units (USUs, clusters of households) are selected within PSUs by systematic sampling after sorting

USUs systematically by labor market characteristics. Further, selection occurs with unequal probabilities in order to yield accurate estimates in subpopulations (e.g. state-specific unemployment estimates). This complex multistage cluster design renders standard inference procedures inappropriate. Further, as the analyst I cannot directly observe the strata, PSU, or USUs because these geographic identifiers are hidden in the data files for privacy reasons.

Inference therefore relies on a set of replicate weights constructed by the Bureau of Labor Statistics. Fig 9 summarizes the sampling process of the CPS and the procedure by which the BLS constructs the replicate weights (Bureau of Labor Statistics 2006 Ch. 14, Fay and Train 1995). Fig 10 illustrates the mechanical process by which the replicate weights yield a variance estimate: by conducting the estimation algorithm with each set of replicate weights in turn and estimating variation across the sets of replicate weights, one can simulate how the estimator would vary across repeated samples drawn from the CPS design. Because this entire procedure is computational, it is possible to estimate sampling variability by replicate weights even for potentially complicated estimation algorithms that involve regression, prediction, and aggregation to counterfactual summaries.

An advantage of the replicate weight procedure is that it applies equally well to any estimation algorithm, even one that is complex. Within each replicate, I estimate the predictive model for the outcome, impute unobserved counterfactuals, take averages, and report an estimate. The confidence intervals may not have proper coverage because the estimator is slightly biased due to regularization. However, they do capture how the estimator would vary across repeated samples drawn by the sampling design of the CPS.

Sampling Process	Replicate Process
Begin with the sampling frame: All residential addresses in the U.S.	Begin with the sample: Chosen households
Define geographic Primary Sampling Units (PSUs) within two types of geographic strata.	PSUs and strata are already defined.
Non-Self-Representing Strata (NSR): Each small-population PSUs is placed in a stratum with others.	For NSR, replicate the selection of PSUs. Combine pairs of NSRs into pseudostrata. Multiply sampling weight by a replicate factor of 1.5 and 0.5. More complex procedures are used for triplets and to adjust for unequal strata populations
Self-Representing Strata (SR): Each large-population PSU is placed in a stratum alone.	For SR, replicate the selection of USUs. Assign USUs into pairs by adjacency in the systematic sampling sort. Conduct an anlogous reweighting procedure.
Select Ultimate Sampling Units (USUs) (clusters of households) in each PSU	All USUs remain, but with new replicate weights per above.

For privacy reasons, these steps are conducted by the Bureau of Labor Statistics and are not visible to researchers.

Interview those households Estimate $\hat{\theta}$ with weights $\vec{w}$	Analyze those households Estimate $\hat{\theta}^*_r$ with each replicate weight $\vec{w}_r$
The variance of an estimator is how it would vary across repeated CPS samples drawn this way	The estimated variance is a function of variation across the 160 replicates: $\hat{\mathrm{V}}(\hat{\theta}) = \frac{4}{160}\sum_{r=1}^{160}\left(\hat{\theta}^*_r - \hat{\theta}\right)^2$ Note: The factor of 4 is needed because the replicates inflate and deflate weights without completely zeroing out any weights.

Fig. 9. Replicate weights mimic repetitions of the CPS sampling process. Illustration is based on the BLS description (Bureau of Labor Statistics, 2006; Fay and Train, 1995). Because the CPS is not a simple random sample, standard approaches to statistical inference are inappropriate. Replicate weights simulate variation that comes from selection of geographic Primary Sampling Units (PSUs) within strata. For PSUs selected with probability one, replicate weights mimic variability within strata. Although the weight construction procedure is described by the BLS, as the researcher I only have access to the generated weights $\vec{w}$ and replicate weights $\{\vec{w}_r\}_{r=1}^{160}$. Inference in this paper proceeds by conducting the entire estimation algorithm repeatedly with each set of replicate weights, calculating the variance of the point estimate, and constructing a confidence interval by a normal approximation.

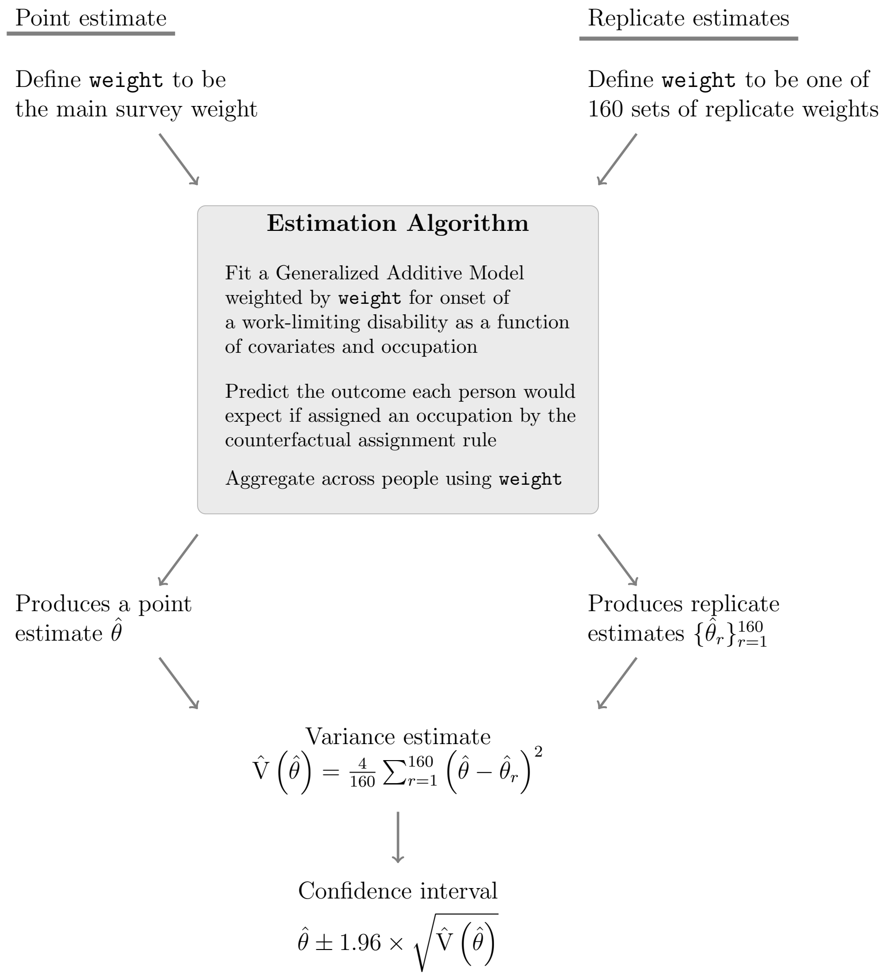


Fig. 10. Procedure to construct confidence intervals with replicate weights. The entire estimation procedure—regression, imputation, and aggregation—can be placed within an algorithm that accepts a weight as the input and returns an estimate. Passing the main survey weight through that algorithm yields a point estimate. Passing the replicate weights through the algorithm yields a set of 160 replicate estimates. Pooling across these estimates yields an estimated variance. By a normal approximation, the point estimate and estimated variance produce a 95% confidence interval.

E Supplemental Figures

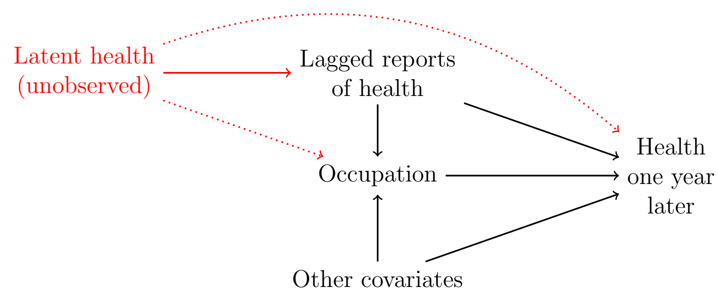


Fig. 11. Threat to causal identification: Coarse measurement of lagged health. The lagged health reports are likely to be coarse measures of latent health. To the degree that uncaptured components of latent health directly affect health outcomes and the occupation one holds (dotted edges), identification will be imperfect. This risk is mitigated to the degree that the lagged reports of health capture the relevant aspects of latent health that affect occupations and affect the health outcome, so that the dotted edges may be relatively unimportant.

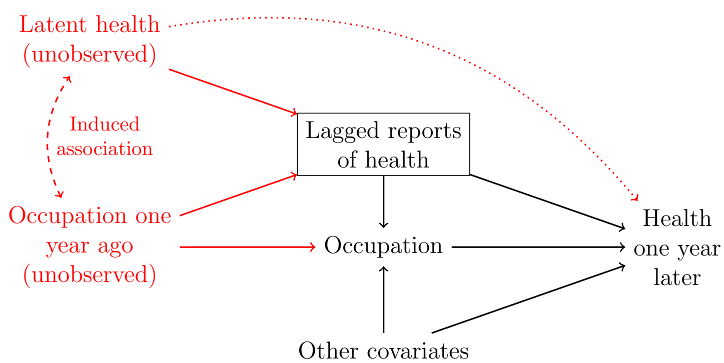


Fig. 12. Threat to causal identification: M-bias. Lagged health is a collider variable because it is likely to be a consequence of both latent health and the occupation one a year ago. Conditioning on lagged health induces an association between those causes. This opens a backdoor path (dashed) that would confound estimates of the causal effect of occupation on health one year later. This risk is mitigated to the degree that the lagged reports of health capture the relevant aspects of latent health that affect the health outcome, so that the dotted edge may be relatively unimportant.

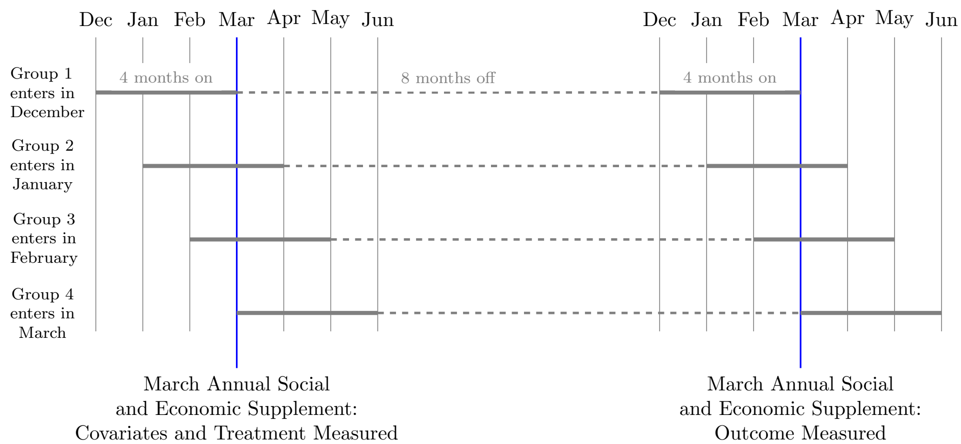


Fig. 13. Panel structure of the Current Population Survey. Households are sampled and assigned to rotation groups, which enter the panel on different months throughout the year. Each rotation group is included in the monthly survey for four months, excluded for eight months, and then included again for four more months. A basic survey is administered each month with the primary goal of tracking the unemployment rate. The Annual Social and Economic Supplement is a large survey administered each March. Because of the panel structure, households that enter the CPS panel in December to March are included in the ASEC survey for two adjacent years. The ASEC also includes oversamples which follow a different panel structure and are not linkable across years. For example, since 2001 the ASEC has included oversamples of low-income households drawn from those who enter the CPS in other months, with the goal of providing estimates relevant to evaluations of the State Children's Health Insurance Program. Because the oversamples are not linkable across years, they are excluded from the linked data analyses in this paper. See Flood and Pacas (2017) for more information.

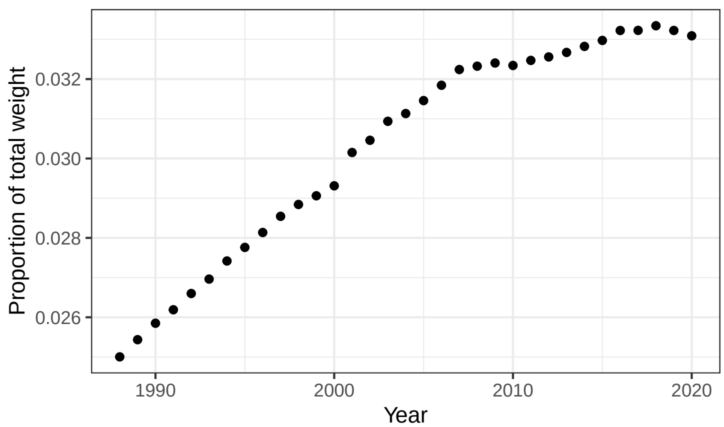


Fig. 14. The total weight on the sample increases over time. This is because the U.S. population ages 25–60 increases in size over time; when drawing inference about this population, the more recent years are weighed more heavily because they contain a larger share of the total population.

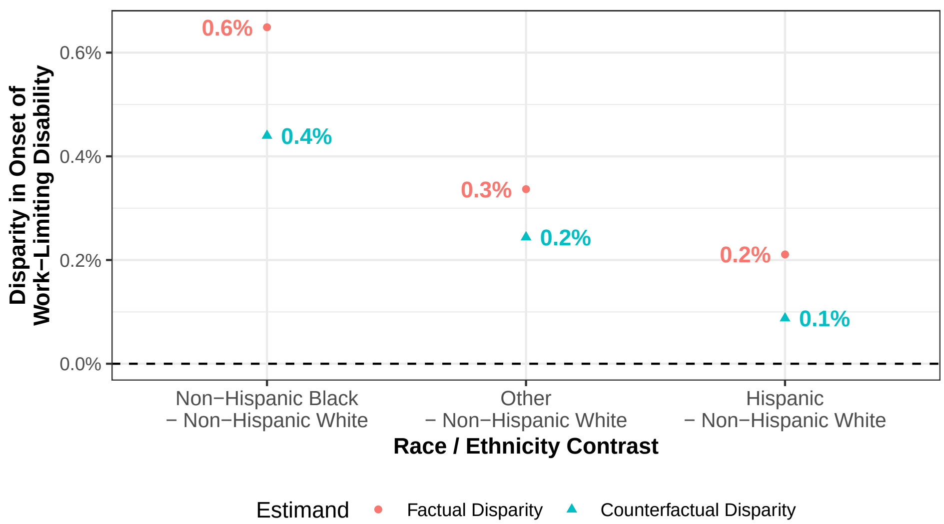


Fig. 15. Alternative specification: Analysis excluding immigrants. Analysis is the same as in the main specification (Figure 7 Panel B) but restricts to native-born respondents. While the main text showed that Hispanic respondents had lower rates of work-limiting disability than non-Hispanic white respondents, this figure shows that the pattern reverses once we restrict to native-born respondents. Thus, the main text finding that Hispanic respondents have better health than non-Hispanic white respondents is driven by the comparably good health of immigrants in the sample.

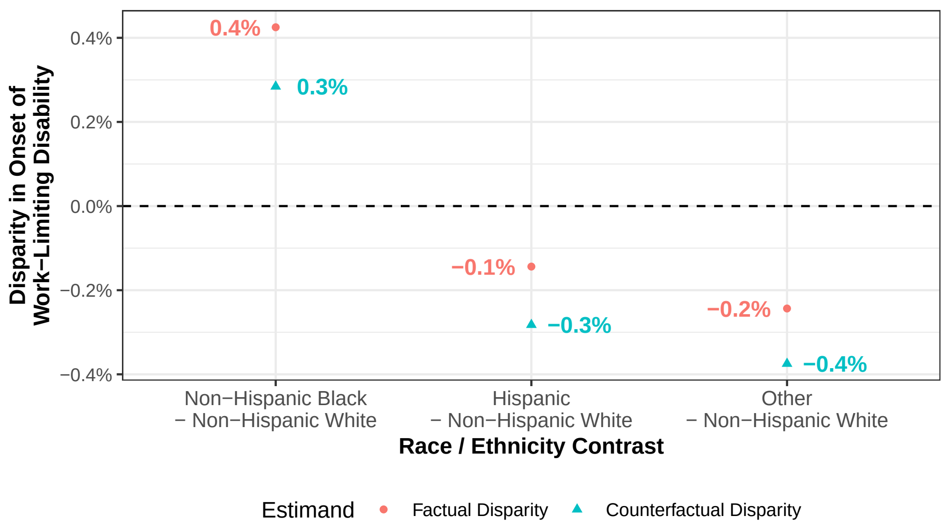


Fig. 16. Alternative specification: Further restriction on past disability. Analysis is the same as in the main specification (Figure 7 Panel B) but restricts to the years 2009–2020 and subsets to those who report never leaving a job for health reasons and who report no difficulty with hearing, vision, remembering, walking or climbing stairs, performing basic activities outside the home along, or taking care of personal needs ($N = 165,961$). This restriction improves the credibility of causal identification at the cost of narrowing the time period to this susbet of years.

References

Bureau of Labor Statistics 2006. *Design and Methodology: Current Population Survey Technical Paper 66*. U.S. Department of Labor.

Burkhauser, R. V., M. C. Daly, A. J. Houtenville, and N. Nargis 2002. Self-reported work-limitation data: What they can and cannot tell us. *Demography*, 39(3):541–555.

Czajka, J. L., R. Rosso, et al. 2015. Redesign of the income questions in the Current Population Survey Annual Social and Economic Supplement: Further analysis of the 2014 split-sample test. *Mathematica Policy Research*.

Fay, R. E. and G. F. Train 1995. Aspects of survey and model-based postcensal estimation of income and poverty characteristics for states and counties. In *Proceedings of the Section on Government Statistics, American Statistical Association, Alexandria, VA*, Pp. 154–159.

Flood, S. M. and J. D. Pacas 2017. Using the Annual Social and Economic Supplement as part of a Current Population Survey Panel. *Journal of Economic and Social Measurement*, 42(3-4):225–248.

Hale, T. W. 2001. The lack of a disability measure in today's Current Population Survey. *Monthly Labor Review*, 124:38.

Maestas, N., K. J. Mullen, and S. Rennane 2019. Unmet need for workplace accommodation. *Journal of Policy Analysis and Management*, 38(4):1004–1027.

Wood, S. N. 2017. *Generalized Additive Models: An Introduction with R*. Chapman and Hall/CRC.