Consider This Before Using the SARS-CoV-2 Pandemic as an Instrumental Variable in an Epidemiological Study

Michelle C. Dimitris and Robert W. Platt

Correspondence to Dr. Michelle C. Dimitris, Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, 1020 Avenue des Pins, Montreal, Quebec, Canada, H3A 1A2 (e-mail: michelle.dimitris@mail.mcgill.ca)

Author affiliations: Department of Epidemiology, Biostatistics, and Occupational Health, Faculty of Medicine, McGill University, Montreal, Quebec, Canada (Michelle C. Dimitris, Robert W. Platt); Research Institute, Children's Hospital of Eastern Ontario, Ottawa, Ontario, Canada (Michelle C. Dimitris).

There was no specific funding for this work.

There were no new data generated or analysed in support of this work.

Conflict of Interest: none declared

Running head: Consider Before Using Pandemic as Instrument

Editor's note: *The opinions expressed in this article are those of the authors and do not necessarily reflect the views of the* American Journal of Epidemiology. *A response to this article is forthcoming.*

Abstract

Epidemiologists sometimes use external sources of variation to explore highly-confounded exposure-outcome relationships or exposures that cannot be randomized. These exogenous sources of variation, or natural experiments, are sometimes proposed as instrumental variables to examine the effects of a given exposure(s) on a given outcome(s). Previous epidemiological studies have applied this technique using famines, earthquakes, weather events, and previous pandemics as exogenous sources of variation for other exposures; interest in applying this technique using the current severe acute respiratory system coronavirus 2 (SARS-CoV-2) pandemic is already documented. Yet, large-scale events like these likely have broad and complicated impacts on human health, which almost certainly violates the exclusion restriction assumption of instrumental variable analyses. We review the assumptions of instrumental variable analyses, highlight previous applications of this method with respect to natural experiments with broad impacts or "shocks", and discuss how these relate to our current observations of the SARS-CoV-2 pandemic. While we encourage thorough investigation of the broad impacts of the SARS-CoV-2 pandemic on human health, we caution against its widespread use as an instrumental variable to study other exposures of interest.

Keywords

Causal methods; COVID-19; instrumental variable; natural experiment; pandemic; quasi-experiment; SARS-CoV-2

Abbreviations

SARS-CoV-2: severe acute respiratory system coronavirus 2

COVID-19: coronavirus disease 2019

Causal inference relies on creating, assuming, or finding random variation. We *create* random variation when we randomly assign treatment, as in a clinical trial. We *assume* random variation when we propose exchangeability between two groups of people or two periods of time, conditional on/weighted for imbalances in measured characteristics, as in an observational study. Our only other option is to *find* a source of random variation in nature, namely an instrument as part of a natural experiment or quasi-experiment. Since most exposures are difficult and/or unethical to randomize, and exchangeability is often implausible in observational studies, use of exogenous sources of variation as instrumental variables have been popular among epidemiologists since their introduction to our literature.(1–3)

Briefly, a natural experiment or quasi-experiment is a source of variation that is thought to occur at random with respect to an outcome of interest, is thought to impact or be correlated with an exposure of interest, and is not thought to impact or be correlated with the outcome of interest through any other mechanisms. This is useful when the effect of said exposure on said outcome is often difficult to identify in observational settings due to concerns about unmeasured confounding. For the purposes of this commentary, we note that natural experiments, quasi-experiments, instrumental variables, and "shocks" are terms that are often used in similar contexts; however, not all natural experiments or quasi-experiments meet the assumptions required to be instrumental variables, and "shocks" often refer to exogenous events or special cases of natural experiments or quasi-experiments with broad impacts, such as pandemics or natural disasters. We additionally note that instrumental variable analyses have both conceptual and statistical similarities to causal mediation analyses (Figure 1) and that nuanced discussion of causality and the potential outcomes framework is available elsewhere.(4)

Instrumental variables have been illuminating in situations where the impact of an event is narrow in scope, such as fortuitous timing of a specific policy change. Smith and colleagues' use of Ontario's

2007 implementation of a targeted human papillomavirus vaccination program is often cited as an example of appropriate application of this method.(5) This study investigated the impact of human papillomavirus vaccination on sexual behavior among adolescent girls in Ontario, Canada. Girls who were eligible to receive human papillomavirus vaccination were otherwise similar with respect to causes of the outcome of interest, which was sexual behavior, before and after the program was implemented. Together with a large sample, the observation that the program increased human papillomavirus vaccination, and the plausible assumption that the program could only affect sexual behavior through human papillomavirus vaccination, Smith and colleagues demonstrated that human papillomavirus vaccination had no effect on sexual behavior among adolescent girls. In addition to this practical example of appropriate application, we recommend previous papers, reviews, and a recent chapter in *Modern Epidemiology* that detail instrumental variable analyses to epidemiologists considering its application.(3,6–9)

Less convincing are published examples that used instrumental variables with broad impacts on health outcomes. Often, these examples leverage "shocks" such as famines, earthquakes, hurricanes, and pandemics. Such studies have investigated the effects of paternal malnutrition on education and education on cognition using the Chinese famine of 1959-1961(10,11); the effect of food security on HIV using the Malawi famine of 2001-2003(12); the effect of early hunger on long-term outcomes using famines more broadly(13); the effects of socioeconomic status on tooth loss and housing damage/loss of friends and family on dementia using the Japanese earthquake of 2011(14,15); the effects of maternal stress on birth outcomes using the Chilean earthquake of 2010 (16); the effect of gentrification on health outcomes using Hurricane Katrina of 2005 (17); the effect of poverty on mental health using weather (18); and the effects of fetal development on long-term outcomes using the influenza pandemic of 1918 (19–21). For example, one study that investigated the impact of birth during the influenza pandemic of 1918 on health-, education-, and income-related outcomes in adulthood attributed much of the observed effects to fetal development, particularly in the context of maternal influenza infection. This study did not explore broad social and economic consequences of the pandemic, and their concurrent impact on these long-term outcomes.(21) We are already aware of two papers recommending use of the severe acute respiratory system coronavirus 2 (SARS-CoV-2) pandemic as an instrumental variable to examine causes of preterm birth (22) and the effects of socioeconomic exposures, such as social interaction, education, and physical activity. (23)

Notably, these studies leverage “shocks” to make inferences about “soft” or difficult to measure exposures; however, each exposure listed above is only one of the many potential health-adjacent consequences of each corresponding “shock”. Interpreting instrumental variable analyses in this manner is equivalent to presuming the effect of each “shock” is completely mediated by each study’s exposure of interest.

Epidemiologists are often far removed from these natural experiments in space and/or time. However, for the first time in modern history, virtually all epidemiologists are experiencing a “shock” at the same time – namely, the SARS-CoV-2 pandemic. We use the ongoing SARS-CoV-2 pandemic as a didactic example to unpack the identifiability assumptions of instrumental variable analyses. We specifically highlight its complex impacts on population and public health to demonstrate that using the SARS-CoV-2 pandemic as an instrumental variable likely violates a key assumption of these analyses, and to dissuade widespread use of the SARS-CoV-2 pandemic as an instrumental variable in future studies.

To estimate upper and lower thresholds of an average causal effect, assumptions of instrumental variable analyses are: (1) the instrumental variable is associated with the exposure of interest; (2) there are no common causes (i.e. confounders) of the instrumental variable and outcome of interest; and (3) the instrumental variable affects said outcome only through said exposure.(6) To further generate a point estimate, additional assumptions are absence of additive effect modification of the exposure/outcome relationship by the instrumental variable among both exposed and unexposed groups (to estimate the average treatment effect) or among the exposed group only (to estimate the average treatment effect on the treated/exposed), and monotonicity of the instrumental variable/exposure relationship (to estimate the local average treatment effect).(6–9)

When we use "shocks" as instruments, we generally consider the first two assumptions plausible. Specifically, authors often present compelling evidence that a "shock" affects an exposure of interest, and the natural experiments we've described herein are generally considered to occur at random with respect to the outcome(s) of interest. However, the third criterion, called the exclusion restriction assumption, is almost certainly implausible in the case of SARS-CoV-2 and other "shocks" In the case of the SARS-CoV-2 pandemic, the pathways by which the proposed instrument operates are several, and we discover new effects of the virus itself and of the "shock" almost daily. For this reason, we highlight the exclusion restriction assumption in particular, and demonstrate its likely violation when using SARS-CoV-2 as an instrumental variable.

For emphasis, we anecdotally catalogue potential effects of the SARS-CoV-2 pandemic in our setting:

- SARS-CoV-2 infection and subsequent coronavirus disease 2019 or COVID-19 itself appears to cause a broad range of symptoms and severe complications, including death, particularly among older persons or those with preexisting health conditions.(24,25)
- School has been administered remotely/in online settings, which may impact education and socialization among children and young adults.(26)

- Some work has been conducted remotely/in online settings, while other work has become more onerous (i.e. with additional materials/responsibilities for infection control) and/or hazardous, which may impact physical and mental health among adults.(27)
- Unemployment and economic hardship may impact physical and mental health through a variety of mechanisms, including through food security and housing.(28)
- Government programs may mitigate effects of unemployment or precarious employment (29); it is possible that such programs may also increase economic stability among those experiencing pre-pandemic financial hardship.
- Social interaction has changed in both type and quality across age groups. Some report less face-to-face interaction with friends and family. Others report more meaningful interaction in the form of overdue telephone/online conversations.(30,31)
- Social isolation may also result in increased domestic/intimate partner violence.(32)
- Deleterious changes in diet and exercise patterns have been reported.(33)
- Changes in daily routine may have negative consequences for mental health (34), yet have also been credited in part with increased attention to the Black Lives Matter movement, facilitating important conversations about racism and social justice.(35,36)
- Environmental effects may be observed among lower commuting and air travel.(37)
- Infection control measures may result in less transmission of other infectious diseases, notably during the 2020/2021 influenza season.(38) Physicians have also cautioned against lower rates of childhood vaccinations due to missed routine visits.(26)
- Inequalities in effects of the SARS-CoV-2 pandemic have been observed, particularly by race and socioeconomic status.(28,39)

A study that leveraged the SARS-CoV-2 pandemic in an instrumental variable analysis to examine the health effects of one such exposure would be ignoring its broad effects. For example, a theoretical study of the impact of online schooling on child health outcomes that used the SARS-CoV-2 pandemic as an instrumental variable would effectively ignore concordant changes in the type and quality of children's social interactions; caregivers' employment and economic situations; domestic/intimate partner violence within children's households; children's diet and exercise; children's exposure to and experience of conversations about racism and social justice; environmental determinants of health; infection with SARS-CoV-2 itself, and their effects on these same child health outcomes. While there may be situations in which the effect of the SARS-CoV-2 pandemic on an outcome if interest is plausibly mediated by a single exposure of interest (for example, where SARS-CoV-2 infection is the exposure and COVID-19 hospitalization is the outcome of interest), we expect few exceptions to what we anticipate to be a limited application of this framework. The list above is not exhaustive; many current and future health-related impacts of the SARS-CoV-2 pandemic may emerge in the coming weeks, months, and years.

While we focus on exclusion-restriction as a necessary assumption of instrumental variable analyses, we also note that the monotonicity assumption is likely violated for many exposures that use the SARS-CoV-2 pandemic as an instrumental variable. Unlike the example we cited above, where a human papillomavirus vaccination program is expected to increase the likelihood of vaccination, it is plausible the pandemic may change many of the above exposures in the opposite direction than expected for a subgroup of individuals.

It is important to study the broad health effects of the SARS-CoV-2 pandemic, including its impact on several exposures and outcomes of public health interest, which will likely yield valuable and potentially hypothesis-generating information. Yet, it seems implausible to consider the SARS-CoV-2 pandemic as an instrument for only one of these exposures given the broad range of mechanisms with which it has affected society. Consider a natural experiment that assumed the SARS-CoV-2 pandemic impacted health only through one of the aforementioned pathways; would you be convinced by its conclusions? If the answer to this question seems obvious, consider the "shocks" cited above; were they likely to have occurred solely through the exposure of interest? We hope our collective lived experience amidst the SARS-CoV-2 pandemic will encourage epidemiologists to apply instrumental variable analyses more judiciously and interpret both past and future results with caution.

References

1. Angrist JD, Imbens GW, Rubin DB. Identification of Causal Effects Using Instrumental Variables. *J. Am. Stat. Assoc.* 1996;91(434):444.
2. Robins JM, Greenland S. Identification of Causal Effects Using Instrumental Variables: Comment. *J. Am. Stat. Assoc.* 1996;91(434):456.
3. Greenland S. An introduction to instrumental variables for epidemiologists. *Int. J. Epidemiol.* 2000;29(4):722–729.
4. Hernán MA, Robins J. Causal Inference. Boca Raton, FL: Chapman & Hall/CRC, forthcoming.; 2015.
5. Smith LM, Kaufman JS, Strumpf EC, et al. Effect of human papillomavirus (HPV) vaccination on clinical indicators of sexual behaviour among adolescent girls: the Ontario Grade 8 HPV Vaccine Cohort Study. *Can. Med. Assoc. J.* 2015;187(2):E74–E81.
6. Hernán MA, Robins JM. Instruments for Causal Inference. *Epidemiology*. 2006;17(4):360–372.
7. Davies NM, Smith GD, Windmeijer F, et al. Issues in the Reporting and Conduct of Instrumental Variable Studies. *Epidemiology*. 2013;24(3):363–369.
8. Swanson SA, Hernán MA. Commentary. *Epidemiology*. 2013;24(3):370–374.
9. Glymour MM, Swanson SA. Instrumental Variables and Quasi-Experimental Approaches. In:

Lash TL, VanderWeele TJ, Haneuse S, et al., eds. *Modern Epidemiology*. Pittsburgh, PA: Wolters Kluwer; 2018:677–709.

10. Kim S, Deng Q, Fleisher BM, et al. The Lasting Impact of Parental Early Life Malnutrition on Their Offspring: Evidence from the China Great Leap Forward Famine. *World Dev.* 2014;54:232–242.

11. Huang W, Zhou Y. Effects of education on cognition at older ages: Evidence from China's Great Famine. *Soc. Sci. Med.* 2013;98:54–62.

12. Loevinsohn M. The 2001-03 Famine and the Dynamics of HIV in Malawi: A Natural Experiment. *PLoS One*. 2015;10(9):e0135108.

13. van den Berg GJ, Pinger PR, Schoch J. Instrumental Variable Estimation of the Causal Effect of Hunger Early in Life on Health Later in Life. *Econ. J.* 2016;126(591):465–506.

14. Matsuyama Y, Aida J, Tsuboya T, et al. Are Lowered Socioeconomic Circumstances Causally Related to Tooth Loss? A Natural Experiment Involving the 2011 Great East Japan Earthquake. *Am. J. Epidemiol.* 2017;186(1):54–62.

15. Hikichi H, Aida J, Kondo K, et al. Increased risk of dementia in the aftermath of the 2011 Great East Japan Earthquake and Tsunami. *Proc. Natl. Acad. Sci.* 2016;113(45):E6911–E6918.

16. Torche F. The Effect of Maternal Stress on Birth Outcomes: Exploiting a Natural Experiment. *Demography*. 2011;48(4):1473–1491.

17. Schnake-Mahl A, Sommers BD, Subramanian S, et al. Effects of gentrification on health status after Hurricane Katrina. *Health Place*. 2020;61:102237.

18. Hanandita W, Tampubolon G. Does poverty reduce mental health? An instrumental variable analysis. *Soc. Sci. Med.* 2014;113:59–67.

19. Lin M-J, Liu EM. Does in utero exposure to Illness matter? The 1918 influenza epidemic in Taiwan as a natural experiment. *J. Health Econ.* 2014;37:152–163.

20. Nelson RE. Testing the Fetal Origins Hypothesis in a developing country: evidence from the

1918 Influenza Pandemic. *Health Econ.* 2010;19(10):1181–1192.

21. Almond D. Is the 1918 Influenza Pandemic Over? Long-Term Effects of In Utero Influenza Exposure in the Post-1940 U.S. Population. *J. Polit. Econ.* 2006;114(4):672–712.
22. Philip RK, Purtill H, Reidy E, et al. Unprecedented reduction in births of very low birthweight (VLBW) and extremely low birthweight (ELBW) infants during the COVID-19 lockdown in Ireland: a 'natural experiment' allowing analysis of data from the prior two decades. *BMJ Glob. Heal.* 2020;5(9):e003075.
23. Edwards JK, Lessler J. What Now? Epidemiology in the Wake of a Pandemic. *Am. J. Epidemiol.* 2021;190(1):17–20.
24. Rothan HA, Byrareddy SN. The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. *J. Autoimmun.* 2020;109:102433.
25. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet.* 2020;395(10229):1054–1062.
26. Chanchlani N, Buchanan F, Gill PJ. Addressing the indirect effects of COVID-19 on the health of children and young people. *Can. Med. Assoc. J.* 2020;192(32):E921–E927.
27. Béland L-P, Brodeur A, Mikola D, et al. The Short-Term Economic Consequences of COVID-19: Occupation Tasks and Mental Health in Canada. *IZA Discuss. Pap.* 2005;
28. Canadian Human Rights Commission. Statement - Inequality amplified by COVID-19 crisis. (https://www.chrc-ccdp.gc.ca/eng/content/statement-inequality-amplified-covid-19-crisis). (Accessed September 18, 2020)
29. Gentilini U, Almenfi M, Orton I, et al. Social Protection and Jobs Responses to COVID-19 : A Real-Time Review of Country Measures. (https://openknowledge.worldbank.org/handle/10986/33635). (Accessed September 18, 2020)
30. Folk D, Okabe-Miyamoto K, Dunn E, et al. Did Social Connection Decline During the First Wave of COVID-19?: The Role of Extraversion. *Collabra Psychol.* 2020;6(1).

31. Wen T. How coronavirus has transformed the way we communicate. (https://www.bbc.com/worklife/article/20200408-coronavirus-how-lockdown-helps-those-who-fear-the-phone). (Accessed September 18, 2020)

32. Usher K, Bhullar N, Durkin J, et al. Family violence and COVID-19: Increased vulnerability and reduced options for support. *Int. J. Ment. Health Nurs.* 2020;29(4):549–552.

33. Mattioli A V., Sciomer S, Cocchi C, et al. Quarantine during COVID-19 outbreak: Changes in diet and physical activity increase the risk of cardiovascular disease. *Nutr. Metab. Cardiovasc. Dis.* 2020;30(9):1409–1417.

34. Galea S, Merchant RM, Lurie N. The Mental Health Consequences of COVID-19 and Physical Distancing. *JAMA Intern. Med.* 2020;180(6):817.

35. Chotiner I. A Black Lives Matter Co-Founder Explains Why This Time Is Different. *New Yorker*. 2020;

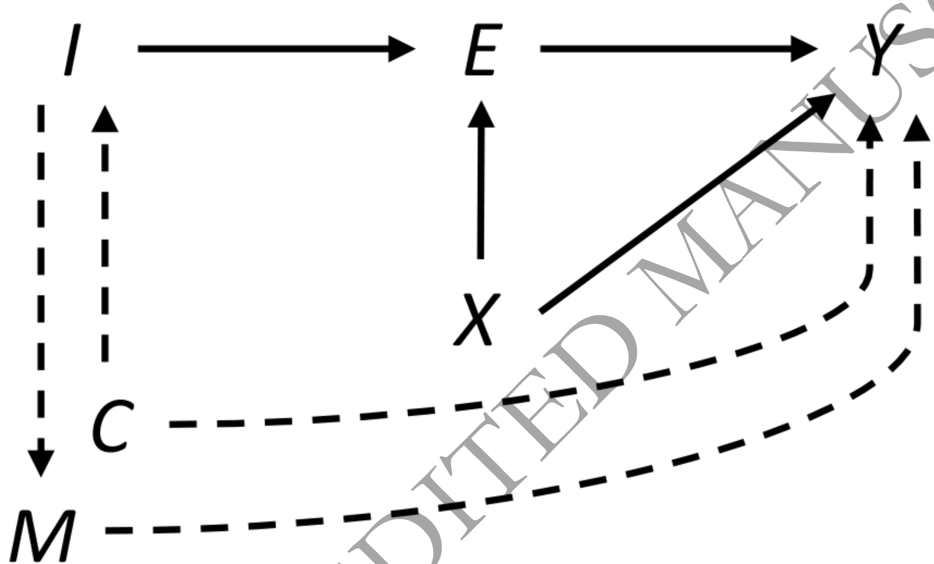
36. Galea S, Abdalla SM. COVID-19 Pandemic, Unemployment, and Civil Unrest. *JAMA.* 2020;324(3):227.

37. Muhammad S, Long X, Salman M. COVID-19 pandemic and environmental pollution: A blessing in disguise? *Sci. Total Environ.* 2020;728:138820.

38. Soo RJJ, Chiew CJ, Ma S, et al. Decreased Influenza Incidence under COVID-19 Control Measures, Singapore. *Emerg. Infect. Dis.* 2020;26(8):1933–1935.

39. Dorn A van, Cooney RE, Sabin ML. COVID-19 exacerbating inequalities in the US. *Lancet*. 2020;395(10232):1243–1244.

Figure 1. Directed acyclic graph for a typical instrumental variable analysis. *I* = instrumental variable, *E* = exposure, *Y* = outcome, *X* = exposure-outcome confounder, *C* = instrument-outcome confounder, *M* = mediator. Assumptions for instrumental variable analyses include: (1) instrument causes exposure (presence of arrow *I* to *E*) or is associated with exposure; (2) no instrument-outcome confounding (absence of *C*); and (3) no other causal pathways between exposure and outcome (absence of *M*). Solid arrows indicate presence while dashed arrows indicate absence of relationship between variables.


I
E
Y
X
C
M