Check for updates

Received: 27 July 2022 | Revised: 30 January 2024 | Accepted: 9 February 2024

DOI: 10.1111/ajps.12894

AJPS AMERICAN JOURNAL of POLITICAL SCIENCE

ARTICLE

Rain, rain, go away: 194 potential exclusion-restriction violations for studies using weather as an instrumental variable

Jonathan Mellon

Associate Professor, West Point Department of Systems Engineering, West Point, New York, USA

Correspondence
Jonathan Mellon, Associate Professor, West Point Department of Systems Engineering, 606 Thayer Rd, West Point, NY 10996, USA.
Email: Jonathan.mellon@westpoint.edu

Abstract

Instrumental variable (IV) analysis relies on the exclusion restriction—that the instrument only affects the dependent variable via its relationship with the independent variable and not via other causal routes. However, scholars generally justify the exclusion restriction based on its plausibility. I propose a method for searching for additional violations implied by existing social science studies. I show that the use of weather to instrument different independent variables represents strong prima facie evidence of exclusion-restriction violations for all weather-IV studies. A review of 289 studies reveals 194 variables previously linked to weather: all representing potential exclusion-restriction violations. Using sensitivity analysis, I show that the magnitude of many of these violations is sufficient to overturn numerous existing IV results. I conclude with practical steps to systematically review existing literature to identify and quantify possible exclusion-restriction violations when using IV designs.

Instrumental variable (IV) analysis aims to sidestep endogeneity. If an instrument (W) causally affects X_1 but is uncorrelated with the error term, the causal effect of X_1 on Y can be consistently estimated using techniques like two-stage least squares (2SLS). However, IV's power rests on strong assumptions. Most notably, the exclusion restriction requires that W affects Y only via the variable of interest X_1 (left panel of Figure 1) and not through other pathways (which can be represented by a mediating variable X_j as in the right panel of Figure 1).[1,2]

Despite the exclusion restriction's importance, there is no standard approach for finding other potential causal pathways from W to Y. Standard advice suggests trying to think of other causal pathways and seeing whether other people can think of any (e.g., Cunningham, 2018, p. 227). While theoretical consideration is important, it is unrealistic to expect scholars to independently think of every possible causal pathway. I propose using the existing academic literature to help find candidate exclusion-restriction violations. Specifically, we can search for social science variables (X_j) claimed to be caused by W. These variables are strong candidates for exclusion-restriction violations, as they usually only additionally require the existence of the $X_j \rightarrow Y$ relationship (exceptions are discussed later in Table 1). This approach leverages the huge amount of

[1] While I describe unconditional IV estimation, the same logic applies to conditional estimation except that the independence assumptions are conditional on exogenous covariates K.
[2] I follow Angrist et al. (1996) in separating exogeneity from the exclusion restriction.

Verification Materials: The materials required to verify the computational reproducibility of the results, procedures, and analyses in this article are available on the *American Journal of Political Science* Dataverse within the Harvard Dataverse Network, at: https://doi.org/10.7910/DVN/AQFTHK.

The Cornell Center for Social Sciences verified that the data and replication code submitted to the AJPS Dataverse replicates the numerical results reported in the main text of this article.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
Published 2024. This article is a U.S. Government work and is in the public domain in the USA. *American Journal of Political Science* published by Wiley Periodicals LLC on behalf of Midwest Political Science Association

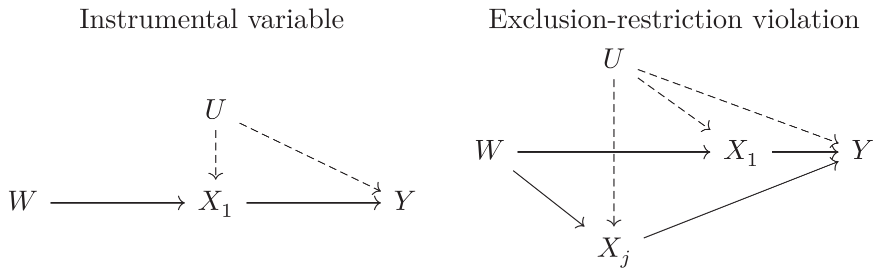


FIGURE 1 Directed Acylic Graphs (DAGs) illustrating instrumental variables (IVs) and endogeneity. *Note*: The figure illustrates the causal structure assumed by IV estimates and how exclusion-restriction violations undermine them.

work already done by other scholars, while still limiting the causal pathways to those that at least one scholar thought was publishable. While the exclusion restriction as a whole is still untestable,[3] this approach converts it into a series of testable substantive claims about particular social science variables and a weaker, untestable claim that there are no undocumented X_j variables violating the exclusion restriction.

I present a process to (1) search for exclusion-restriction variables in existing studies, (2) assess and document exclusion-restriction violations, and (3) conduct sensitivity analysis to assess the impact of exclusion-restriction violations on an IV result.

I demonstrate this approach on a widely used instrument: the weather. The weather is appealing as a source of exogenous variation because it seems obvious that the weather is essentially random and only related to human outcomes through narrow causal paths. Rather than independently discovering violating routes from weather to outcomes, I use the IV literature to identify alternative causal pathways. I review 159 weather-IV studies and an additional 130 studies of the direct effect of weather. In total, I identify 194 variables that social scientists have linked to weather.

Any exclusion-restriction violation will bias an IV estimate, but small violations may be acceptable. I assess the severity of these violations by using the first-stage of IV estimates of weather on X_j variables to conduct sensitivity analysis (Cinelli & Hazlett, 2022) on other weather IVs. This allows me to show how large a relationship would have to exist between an X_j variable and Y to make the $X_1 \rightarrow Y$ IV estimate insignificant. For some IV studies, X_j would have to explain as little as .01% of variance in Y to make the $X_1 \rightarrow Y$ IV estimate insignificant. However, in other cases, X_j variables could not make the IV estimate insignificant, even if X_j were perfectly correlated with Y. While both have exclusion-restriction violations, the sensitivity of the IV estimate in the former case should concern us more.

This paper's results challenge many weather-IV findings and raise questions about the validity of other widely used instruments. These results add to a growing body of work raising issues with the use of weather instruments. Sarsons (2015) identifies likely exclusion-restriction violations in *economic growth → conflict* rainfall-IV estimates, as rain predicts conflict even where incomes are unaffected; Horiuchi and Kang (2018) express concern that rainfall's effect on vote share is larger than can be explained by turnout; and Cooperman (2017) argues that rainfall suffers from spatial interdependence. Gallen and Raymond (2021) raise exclusion-restriction concerns for rainfall and other overused IVs in economics. This paper goes beyond their analysis by examining studies in other social science fields, and quantifying the strength of exclusion-restriction violations using sensitivity analysis.

The paper proceeds as follows. First, I discuss the implications of any two IV studies using the same instrument. Second, I outline my approach to reviewing literature on an instrument, conducting sensitivity analysis, and assessing potential exclusion-restriction violations. Third, I show the review's results, the extent of exclusion-restriction violations for weather IVs, and the sensitivity of IV results to these violations. This includes applying the full assessment method to a study examining the effect of economic activity on civil conflict. I conclude by discussing the implications of these results for the use of IVs by social scientists.

WHAT DOES ANOTHER STUDY USING THE SAME INSTRUMENT IMPLY?

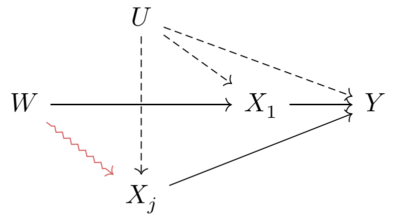
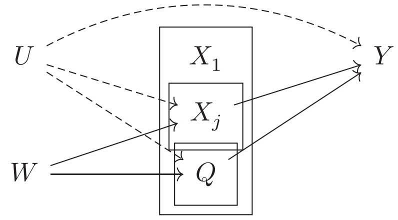
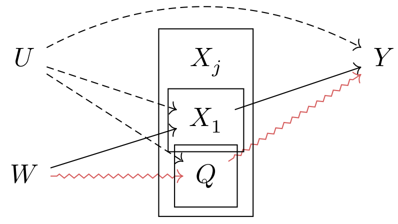
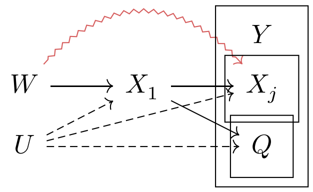
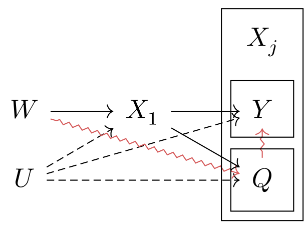
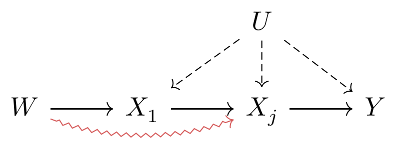
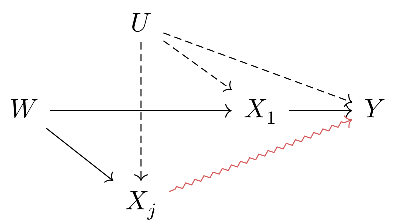
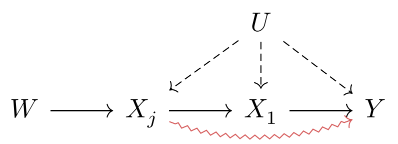
What does another study claiming a causal relationship between your instrument and another variable indicate? For instance, study 1 argues W is an instrument for X_1 to predict an outcome (Y) to avoid unobserved confounders (U). Study j proposes that W serves as an instrument for another variable X_j.

The simplest answer is that study j represents an exclusion-restriction violation where X_j represents an alternate pathway from W to Y that does not include X_1 (right panel of Figure 1). Other possibilities are presented in Table 1, with each row displaying a DAG compatible with study 1 and study j's IV estimate. The

[3] Statistical techniques can detect exclusion-restriction violations but not confirm their absence. The Sargan–Hansen test applies when there are more instruments than instrumented variables by testing whether the second-stage estimate's residuals (iteratively dropping instruments) correlate with the instruments (Kiviet, 2020). The zero-first-stage test verifies the treatment's reduced-form effect is zero in a subsample where it should not affect X_1 (Kippersluis & Rietveld, 2018).

15405907, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ajps.12894 by Cochrane Colombia, Wiley Online Library on [26/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 Situations where X_j is not an exclusion-restriction violation for an instrumental variable (IV) estimate of $X_1 \rightarrow Y$ despite another study finding $W \rightarrow X_j$.

Exception	Claim	Dag
1	W does not cause X_j in this context	
2	X_j is part of X_1	
3	X_1 is the only part of X_j caused by W	
4	X_j is part of Y	
5	Y is part of X_j	
6	W only causes X_j via X_1	
7	X_j does not cause Y	
8	X_j only causes Y via X_1	

Note: ⇝ indicates causal links that must be absent. Where there are multiple such links, at least one must be absent. A boxed variable inside a larger box indicates that the variable is a component of the variable represented by the larger box.

15405907, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ajps.12894 by Cochrane Colombia, Wiley Online Library on [26/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

15405907, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ajps.12894 by Cochrane Colombia, Wiley Online Library on [26/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

DAG uses wavy red arrows to show causal links that must be absent for study 1′s IV estimate to be causally identified.

First, we can reject the claim that $W \rightarrow X_j$ if the original claim is flawed or inapplicable (exception 1). For example, using prevailing winds (W) to estimate the effect of grain shipments (X_j) on 15th-century urbanization (Y) is irrelevant to a study using W to predict 21st-century pollution (X_1).

If X_j is part of X_1 (exception 2), it does not generally violate the exclusion restriction. Study 1 may use natural gas deposits (W) to instrument pollution exposure's (X_1) effect on mortality (Y), while study j uses natural gas deposits to instrument sulfur dioxide exposure (X_j). Sulfur dioxide exposure is part of pollution exposure, so this is not an exclusion-restriction violation for study 1.

If X_1 is part of X_j, X_j *may* not be an exclusion-restriction violation provided no other part of X_j (Q) is caused by W *and* causes Y (exception 3). For example, using elevation (W) as an instrument for heart attacks (X_1) would not necessarily be undermined by a claim that elevation causes health (X_j), if the effect of elevation on health is solely due to its effect on heart attacks. If elevation affects another component of health (Q) such as diabetes, X_j *would* represent an exclusion-restriction violation.

If X_j is part of Y (exception 4), it is not an exclusion-restriction violation unless W affects X_j through a pathway that does not include X_1. Study 1 may use proximity to college (W) as an instrument for education's (X_1) effect on income (Y), while study j uses proximity to college as an instrument for self-employment income (X_j). In this scenario, X_j is not an exclusion-restriction violation for study 1.

If Y is part of X_j, it may not be an exclusion-restriction violation (exception 5). The necessary assumption is either that W does not directly affect any part of X_j apart from Y, or Q (another subcomponent of X_j) does not affect Y. Reworking the previous example, study 1 uses proximity to college (W) to instrument education's (X_1) effect on self-employment income (Y), and study j uses W to instrument income (X_j). Study j does not violate study 1′s exclusion restriction provided that proximity to college does not influence self-employment income through income components unrelated to education. For instance, if self-employment income is crowded out by salary income earned from working at local colleges, income (X_j) may reflect an exclusion-restriction violation for study 1 via its subcomponent of salary income (Q).

Next, X_j may mediate $X_1 \rightarrow Y$ (exception 6). For example, drought (W) may cause agricultural incomes (X_1) to fall, leading to competition for resources (X_j), ultimately resulting in violence (Y). The existence of study j that uses droughts as an instrument for competition for resources does not necessarily invalidate study 1 that uses droughts as an instrument for agricultural income. However, study 1 must make a strong assumption (indicated by the wavy red arrow on the DAG): that there is no other causal pathway between droughts (W) and competition for resources (X_j) other than via agricultural income (X_1).

It is also possible that X_j does not cause Y (exception 7). Suppose that study 1 uses military draft lottery number (W) as an instrument for education's (X_1) effect on income (Y), while study j uses draft lottery number as an instrument for military service (X_j). If military service does not affect earnings, then study j does not violate study 1′s exclusion restriction.

Finally, X_j can mediate $W \rightarrow X_1$ (exception 8). For example, droughts (W) may cause lower crop yield (X_j), reducing agricultural income (X_1), leading to violence (Y). Study j's use of droughts as an instrument for crop yield does not automatically invalidate study 1′s use of droughts as an instrument for agricultural income's effect on violence. However, it requires study 1 to make an additional assumption: There are no other routes from crop yield (X_j) to violence (Y) other than through agricultural income (represented by the red wavy line). If an alternate pathway exists, such as via starvation or migration, this represents an exclusion-restriction violation and invalidates study 1′s use of W as an instrument for X_1.[4]

If none of these exceptions apply, then X_j is an exclusion-restriction violation with the DAG shown in the right panel of Figure 1.[5]

This discussion has focused on identifying whether the estimate is strictly causally identified. In practice, exclusion-restriction violations vary in magnitude, with corresponding levels of bias in the causal estimate. I address the magnitude of biases from exclusion-restriction violations using sensitivity analysis.

EXCLUSION-RESTRICTION VIOLATION ASSESSMENT METHODOLOGY

Having discussed how we should interpret other studies' claims about our instrument, I now outline a methodology for evaluating exclusion-restriction violations for an instrument using existing literature. I explain the general methodology at each step and also provide detail on how I apply this methodology to the weather-IV example in this paper. This approach can be applied retrospectively to an existing IV claim or prospectively in new IV studies.

[4] If a route from W to X_1 that does not go via X_j exists, W can be a valid instrument by conditioning on X_j and all confounders between X_j and Y. Though technically valid, the assumptions stretch credulity.

[5] Gallen and Raymond (2021) recommend conditioning on X_j to recover the causal effect of X_1 on Y but this entails strong data requirements and assumptions (see online Appendix E p. 7).

15405907, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ajps.12894 by Cochrane Colombia, Wiley Online Library on [26/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The methodology involves four steps. First, the literature is used to identify X_j variables linked to study 1′s instrument W. Second, potential exclusion-restriction variables are filtered for relevance using broad filters, such as country income. Third, each X_j variable is assessed using theory and literature to determine whether it falls under any of the exceptions in Table 1. Any remaining X_j variables are likely to indicate an exclusion-restriction violation for study 1.

However, the violation's importance depends on the strength of the $W \rightarrow X_j$ and $X_j \rightarrow Y$ causal links. Thus, the fourth step examines the bias in study 1′s estimate through sensitivity analysis, using estimates of relationship strength from applicable literature.

While I apply most of these steps to all weather-IV studies, I apply the more specific assessment of X_j variables to an *economic activity → civil conflict* weather-IV estimate.

Weather example

Weather IVs are widely used by social scientists, so provide a rich example for the exclusion-restriction methodology. I test the methodology in two ways in this paper. I test most of the stages of the process on all weather-IV estimates with the required information. This helps to understand how vulnerable a typical weather-IV estimate is to exclusion-restriction violations.

The exception-search stage is too intensive to perform on every IV estimate. I therefore also show the full process on a single IV estimate of the effect of economic activity on civil conflict (using rainfall as an instrument). This relationship has been repeatedly studied using weather instruments (Jensen & Gleditsch, 2009; Miguel & Satyanath, 2011; Miguel et al., 2004), with various criticisms (Ciccone, 2011; Sarsons, 2015) leveled at those studies. I examine a study using subnational data for African countries (Hodler & Raschky, 2014) with information available for sensitivity analysis. Hodler and Raschky (2014) find an effect of −.303 (SE = .111) on conflict probability for a 1-unit increase in the log of economic activity (measured using light intensity).

Enumerate potential exclusion-restriction violations

The method aims to identify a wide set of potential exclusion-restriction violations (documented causal links $W \rightarrow X_j$) for an IV study (study 1). The researcher enumerates possible linkages between an instrument W and variables that could affect humans (X_j) by searching the literature about W and examining what variables W has been used to predict. IV studies, which also use W as an instrument, are a particularly valuable source of X_j variables because they make explicitly causal claims and often provide estimates that can be used in sensitivity analysis. Researchers should also search the literature for X_j variables that are causally affected by closely related variables to W. For example, a researcher using elevation as an instrument should search for studies using variables such as air pressure or mountainous terrain as instruments or predictors. The researcher should also include any plausible X_j variables they can think of that are not covered in the existing literature.

Since a single exclusion-restriction violation can undermine the IV strategy, researchers may choose to examine an initial batch of X_j variables rather than attempting to fully enumerate all X_j variables. If this initial batch does not reveal exclusion-restriction violations that cause unacceptable levels of bias, researchers should extend the literature search until they either find an unacceptable exclusion-restriction violation or have exhausted the literature.

Application to weather instrumental variables

I apply the same method for enumerating potential exclusion-restriction variables to both the broad analysis of weather-IV studies and the narrow *economic activity → conflict* case. This literature review serves the dual purpose of identifying (1) the weather-IV studies under examination for the broad analysis and (2) the potential exclusion-restriction variables (X_j) for those studies.

Although these studies use various weather phenomena as instruments, different weather types are part of a complex causal network and are generally substantially correlated[6] both spatially[7] and longitudinally (Nicholls et al., 1997). This means the causal relationship *rain* → X_j implies some form of causal relationship between temperature and X_j, and wind speed and X_j, although the strength of that relationship may be attenuated.

To generate potential exclusion-restriction variables for weather-IV studies, I reviewed the first 500 results for the search term "weather 'instrumental variable'" on Google Scholar (while the search returned 12,700 results in total, the results after 500 were less relevant). I also included the first 100 results from "rain 'instrumental variable,'" some previously identified studies, and relevant studies cited in other papers. This yielded 159 weather-IV studies, and 130 studies using

[6] These correlations do not account for measurement error. Correlations between different weather phenomena are not much lower than the correlations between different measurements of the same phenomenon (Auffhammer et al., 2013).

[7] Annual precipitation (World Bank, 2014) correlates with 2014 yearly average temperature (World Bank, 2020) at the country level at $r = .28$.

weather as a key independent variable. These studies were primarily in economics, political science, and international development.

I record the instrument, instrumented variable, and outcome variable for each study. If a study contains more than one combination of these, I include a row for each combination. I also record the frequency of weather measurement, and IV, reduced-form, and first-stage estimates where available.

For analysis, I recode the variables into coarser theoretical categories. For example, I coded GDP, household income, average income, economic growth, permanent income, income, agricultural income, growth, per-capita income, crop revenue per hectare, poverty, transitory income, agricultural growth, and farm financial performance as "income." Less coarse categories would better distinguish concepts but would increase the count of potential exclusion-restriction violations.

Filter potential exclusion-restriction violations for relevance

While the literature is a useful source of information about potential exclusion-restriction violations, not all existing $W \rightarrow X_j$ claims will be relevant to study 1′s estimate. Because I only focus on whether the link exists at this stage, I do not need to worry about factors that change the strength of the $W \rightarrow X_j$ link across contexts, but only factors that could make the link *exist* in the context of study j but not study 1. If the enumeration step produced a large number of potential exclusion-restriction variables, it may be useful to apply heuristics to limit variables to those most likely to affect study 1. The exact filters will depend on the study: Political regime might be relevant for a voting behavior study, whereas economic development could be relevant for a study of economic growth.

Application to weather instrumental variables

I consider two factors that could lead a causal relationship between weather and X_j to exist in one context but not another: (1) the temporal frequency of the weather variable and (2) the economic context of each study.

Weather can be measured at different levels of temporal aggregation, for example, daily versus yearly rainfall variation. Shorter periods are generally relevant to longer periods but not the inverse. An effect of daily sunlight levels on happiness might generalize to an effect of yearly sunlight effects. However, it is less likely that an effect of yearly rainfall on income would generalize to daily variation in rainfall. Long-period weather-IV studies (differences on the order of years) are therefore vulnerable to exclusion-restriction violations from all other weather-IV studies. However, if short-term studies control for long-term weather, it should be possible to remove exclusion-restriction violations working via longer term weather. For example, turnout weather-IV studies often control for the average weather on that date but this does not account for the correlation between election-day weather and weather patterns in the area over the past week or month, so fully controlling for other time periods is challenging. I split the studies into short term (weather measured monthly or more frequently) and long term (weather measured less than monthly).

The relevance of exclusion-restriction variables may also vary because the societies studied are different. Although societies vary along many dimensions, the most relevant factor for weather appears to be economic development. Poorer countries tend to have economies based on agriculture and raw-material extraction, while richer countries have more diversified economies. Consequently, the effects of weather may differ across these settings. To mitigate this problem, I categorize the countries in each study based on the World Bank's current development levels. High- and higher middle–income countries are grouped as high income, while lower middle- and lower income countries are grouped as low income. This compromises between only considering exclusion-restriction violations from identical contexts and allowing estimates from poor agricultural contexts to be relevant to high-income countries.

Search for exceptions where X_j is not a potential exclusion-restriction violation

Each X_j variable is assessed using the steps in Table 1 to assess eight possible exceptions for why it may not be an exclusion-restriction violation. The subsequent step is only taken if the previous step fails to identify an exception. Literature searches and theoretical consideration should inform the decision at each stage. If an exception is found for X_j, the evidence and/or theory should be documented so readers can assess its validity.

If no exceptions apply, X_j violates study 1′s exclusion restriction.

Application to weather instrumental variables

I do not attempt to search for exceptions for the broad set of weather IVs, as this would require investigating 10,744 possible causal relationships (158 Y variables and 68 X_j variables). However, an individual study would only need to assess 68 X_j variables for causal links to the Y variable for that study.

15405907, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ajps.12894 by Cochrane Colombia, Wiley Online Library on [26/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

To demonstrate this step, I apply the exception-search method to the *economic activity → civil conflict* IV estimate.

Sensitivity analysis

The methodology so far aims to gather examples of exclusion-restriction violations for an IV study (study 1). However, the presence of exclusion-restriction violations does not necessarily imply large biases in study 1′s IV estimates. Instead, we must examine whether X_j is a large enough violation of the exclusion-restriction assumption to alter study 1′s conclusions. I assess the IV estimate's sensitivity by looking at the reduced form (Cinelli & Hazlett, 2022), with an estimate being sensitive if the bias would make it nonsignificant or change the sign of the effect. While I focus on exclusion-restriction violations reducing an effect's magnitude, they could also arbitrarily increase its magnitude, depending on the sign of the $W \rightarrow X_j$ and $X_j \rightarrow Y$ relationships. I treat the X_j variables as confounders, although this means the sensitivity analysis could be affected by posttreatment bias if X_j and Y share unmeasured confounders. Therefore, the sensitivity estimates indicate the likely level of bias but should not be considered the final substantive answer. The sensitivity analysis currently only applies to the IV regression reduced form in the case of a single instrument. However, multiple-instrument studies are not systematically less susceptible to exclusion-restriction violations.

The impact of an exclusion-restriction violation on an IV study's conclusions depends on two quantities: the partial-R^2 of the relationship between the instrument W and an exclusion-restriction variable X_j ($R^2_{X_j \sim W|K}$), and the partial R^2 of the relationship between X_j and the dependent variable Y ($R^2_{Y \sim X_j|W,K}$), where K is the set of exogenous predictors in the original study.

Based on a pair of partial-R^2 values for a particular X_j, I compute the bias of the coefficient (Cinelli & Hazlett, 2022) using the *sensemakr* package (Cinelli et al., 2020) (where n is the sample size and p is the number of predictors):

$$\text{bias} = \text{SE}\sqrt{\frac{R^2_{Y \sim X_j|W,K} R^2_{X_j \sim W|K}}{1 - R^2_{X_j \sim W|K}}\,(n - p - 1)}$$

and the corrected standard error:

$$\text{SE}_{\text{corrected}} = \text{SE}\sqrt{\frac{1 - R^2_{Y \sim X_j|W,K}}{1 - R^2_{X_j \sim W|K}} \cdot \frac{n - p - 1}{n - p - 2}}$$

These are used to calculate a new central estimate and confidence interval, which can be used to assess whether the revised estimate is unacceptably different from the original. In this paper, I consider an unacceptable change to be one that makes the result no longer significant with the same sign (or a change of equivalent magnitude in the same direction as the original effect).

But how large are the relevant partial-R^2 values? These can be derived in various ways including with new data analysis or from studies, which provide the estimates in this format. The IV literature provides useful estimates of $R^2_{X_j \sim W|K}$ through its first-stage estimates of the relationship between W and various X_j variables measured via F-statistics ($F_{X_j \sim W|K}$).[8] For study 1, estimates of relationships between W and the X_j in other studies (j) indicate the strength of possible exclusion-restriction violations.

If both partial-R^2 values are available, the revised estimate can be directly calculated. However, often one or both will be unavailable. If one partial-R^2 value is missing, it is possible to ask how large the other partial-R^2 value would need to be in order to change the original result unacceptably. If the second partial-R^2 would have to be huge to overturn a result, then the exclusion-restriction violation is unlikely to greatly change study 1′s substantive conclusion. If the second partial-R^2 value could be minuscule yet still overturn the result, we should be wary of trusting study 1′s conclusions.

If partial-R^2 values are unavailable, researchers can calculate pairs of values that could invalidate study 1′s result and consider whether they can confidently rule out all of those combinations.

Filter fit statistics for transportability

For an $R^2_{X_j \sim W|K}$ estimate from study j to be relevant to study 1, three requirements must be met: (1) Studies 1 and j must use the same instrument,[9] (2) studies 1

[8] The $R^2_{Y \sim X_j|W,K}$ for a model predicting Y using a set of q instruments (X_j) and p exogenous predictors (K) is derived from the F-statistic $F_{Y,X_j|K}$: $R^2_{Y \sim X_j|W,K} = \frac{F_{Y,X_j|K} \cdot q}{(n-p-q)+q \cdot F_{Y,X_j|K}}$, where n is the number of observations. The degrees of freedom for the F-statistic are estimated from the available information about the number of observations, predictors, and instruments. Some studies likely accounted for clustering in their first-stage F-statistic but this information is rarely available, so I conservatively calculate degrees of freedom as $n - p - q$. Since the more robust approaches will return a smaller F-statistic, this assumption means I will tend to underestimate $R^2_{X_j \sim W|K}$ values and—consequently—the impact of exclusion-restriction violations. When researchers apply this technique, they may want to consider calculating $R^2_{Y \sim X_j|W,K}$ from replication data for crucial X_j variables to avoid these approximations.

[9] This is a stricter standard than at the enumeration stage where only the presence of a causal link is required.

15405907, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ajps.12894 by Cochrane Colombia, Wiley Online Library on [26/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

and j must model the same variance type (e.g., within-person),[10] and (3) study 1 must not control for X_j.[11]

Application to weather instrumental variables

For the broad analysis of weather IVs, I am only able to examine reduced-form estimates that use single weather instruments and provide sufficient information to derive the reduced-form equation.

While all the studies I consider use some form of weather as an instrument, the exact variable used varies across studies (see Table C1, online Appendix p. 5). Rainfall is the most common followed by temperature. Different weather types will not have the same strength of relationship to a particular X_j. A study that demonstrates a strong link between rainfall and X_j does not imply that X_j is as strongly related to sunlight or snowfall. I therefore only consider study j's $R^2_{X_j \sim W|K}$ estimate if it makes use of the same weather variable as study 1.

The variance being modeled differs across weather studies (see Table D1, online Appendix p. 6). For instance, some weather-IV studies use individual-level data at a single time (between person) while others use district-level data over time (within subnational). The most common variance modeled is within subnational. One example of within-subnational variance is a yearly panel of US states with year and state fixed effects. Within-person (or household) variation is the next most modeled form of variance: using person fixed effects, lagged-dependent variable, or first-differencing strategies. Studies that use within-unit variation often refer to the strategy as modeling the effect of weather shocks (because the effect of average weather is removed by unit fixed effects).

Few weather-IV estimates are excluded for controlling for other causal pathways. Of the 45 X_j variables with an $R^2_{X_j \sim W|K}$, only eight are used as a control variable in one of the 63 reduced-form estimates. Income is the most commonly controlled exclusion-restriction variable, controlled for in 14% of IV estimates.

[10] Different variance types make it difficult to compare goodness-of-fit due to different DGMs, uneven sizes of geographical units, and different amounts of variance.

[11] Controlling for X_j is inadvisable since it involves conditioning on a collider but would make the sensitivity analysis mischaracterize the bias. I therefore exclude first-stage estimates for j variables controlled for in study 1. Other exogenous controls (K) may affect partial-R^2 transferability but additional controls (beyond fixed effects) typically have a small effect on variance left to explain.

RESULTS

Enumerating potential exclusion-restriction violations for weather instrumental variables

In the assumed DAG in the left panel of Figure 1, weather only affects the causal network through X_1. In reality, the literature review of 289 studies finds that weather enters the social science causal network through 124 different variables, and indirectly reaches 194 variables of interest to social scientists. Each of these variables represents a potential exclusion-restriction violation for the IV DAG (including the *economic activity* → *civil conflict* DAG). This is illustrated in Figure 2, which shows the causal web for the most common weather type: rainfall. This is derived from 192 studies using rainfall and shows the wide variety of causal pathways the literature claims are linked to rain.

Filter potential weather exclusion-restriction variables for relevance

The overall weather DAG lists all potential exclusion-restriction violations from the literature, but scholars may prefer a narrower selection of relevant studies. Figure 3 displays causal links asserted in weather studies (of all weather types) by country income and temporal frequency. Few studies examine short-term weather variation in low-income countries, so only 27 variables are linked to weather including pirate attacks (Cook & Garrett, 2013), COVID-19 (Shen et al., 2020), exercise (Aral & Nicolaides, 2017), and protest (Ritter & Conrad, 2016).

The other time/income combinations show a large network of potential exclusion-restriction violations. Short-term weather variation in high-income countries has 84 variables, long-term weather variation in low-income countries has 79 variables (this is the relevant set of X_j variables for the *economic activity* → *civil conflict* example), and long-term variation in high-income countries has 109 variables. Most weather-IV studies have many relevant potential exclusion-restriction violations.

If all of these $W \rightarrow X_j$ relationships exist, the typical weather-IV study requires *hundreds* of additional identifying assumptions, which quickly becomes implausible.

Which variables have been linked to weather?

The causal graphs reveal myriad potential exclusion-restriction violations, encompassing crucial social science variables.

15405907, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ajps.12894 by Cochrane Colombia, Wiley Online Library on [26/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

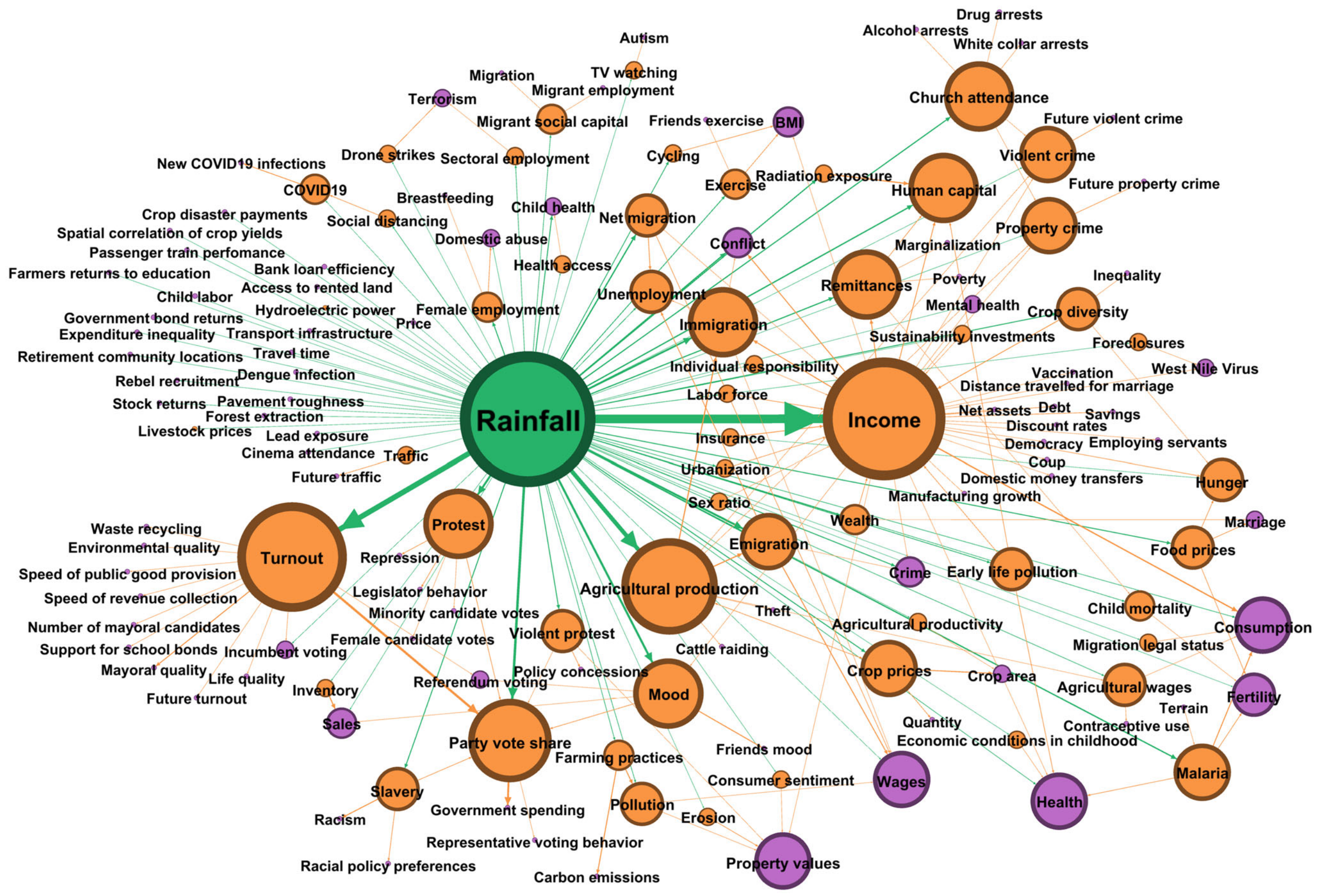


FIGURE 2 Causal web of rainfall from 192 papers. *Note*: This figure illustrates the plethora of variables causally linked to weather. These can represent exclusion-restriction violations. Node and tie sizes are proportional to appearances in literature. Colors: weather (green), instrumented variable (orange), and outcome (purple).

Income is the most widely instrumented variable in weather-IV studies of high-income and low-income countries. The link between weather and income is usually attributed to crop yield. But crop yield has also been proposed as a mechanism for weather effects on other X_j variables such as food prices and food production for personal consumption (e.g., Kubik & May, 2018). This means IV studies may attribute the effects of price levels, hunger, or malnutrition to income, thereby overestimating its effect.

Weather causally affects many variables unrelated to agriculture, but the relevance of these variables will vary. Most researchers can assume "helicopter flying conditions" and "pirate attacks" are not possible exclusion-restriction violations. However, many causal pathways are widely relevant to social science outcomes.

The first causal pathway is mood. Several studies claim weather influences mood (e.g., Bassi, 2019; Meier et al., 2019). In one weather-IV study, mood was linked to decision-making speed, investment decisions, inflation expectations, risk aversion, financial decisions, and perceived life expectancy (Guven & Hoxha, 2015). Almost any social science variable could be causally influenced by at least one of these variables. Mood threatens both long-term and short-term weather-IV studies.

Another major causal route is pollution. Pollution has been instrumented by weather variables including wind speed (e.g., Bondy et al., 2020; Peet, 2020), wind direction (e.g., Fan & Wang, 2020), atmospheric inversions (Bondy et al., 2020; Sager, 2019), and rainfall (e.g., Peet, 2020). Pollution—in turn—has been linked to variables including crime (Bondy et al., 2020), mental health (Gu et al., 2020), mortality (Fan & Wang, 2020), road accidents (Sager, 2019), house prices (Fontenla et al., 2019), cognition (Peet, 2020), and mood (Zheng et al., 2019). Pollution threatens both long-term and short-term weather-IV studies.

Skin tone is another important variable. Katz et al. (2020) find that some people tan when exposed to sunlight while others do not. Using a weather-IV approach, they argue the gap in labor–market outcomes between those who do and do not tan widens during sunny periods. This suggests that weather may impact social science outcomes through racial prejudice.

15405907, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ajps.12894 by Cochrane Colombia, Wiley Online Library on [26/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

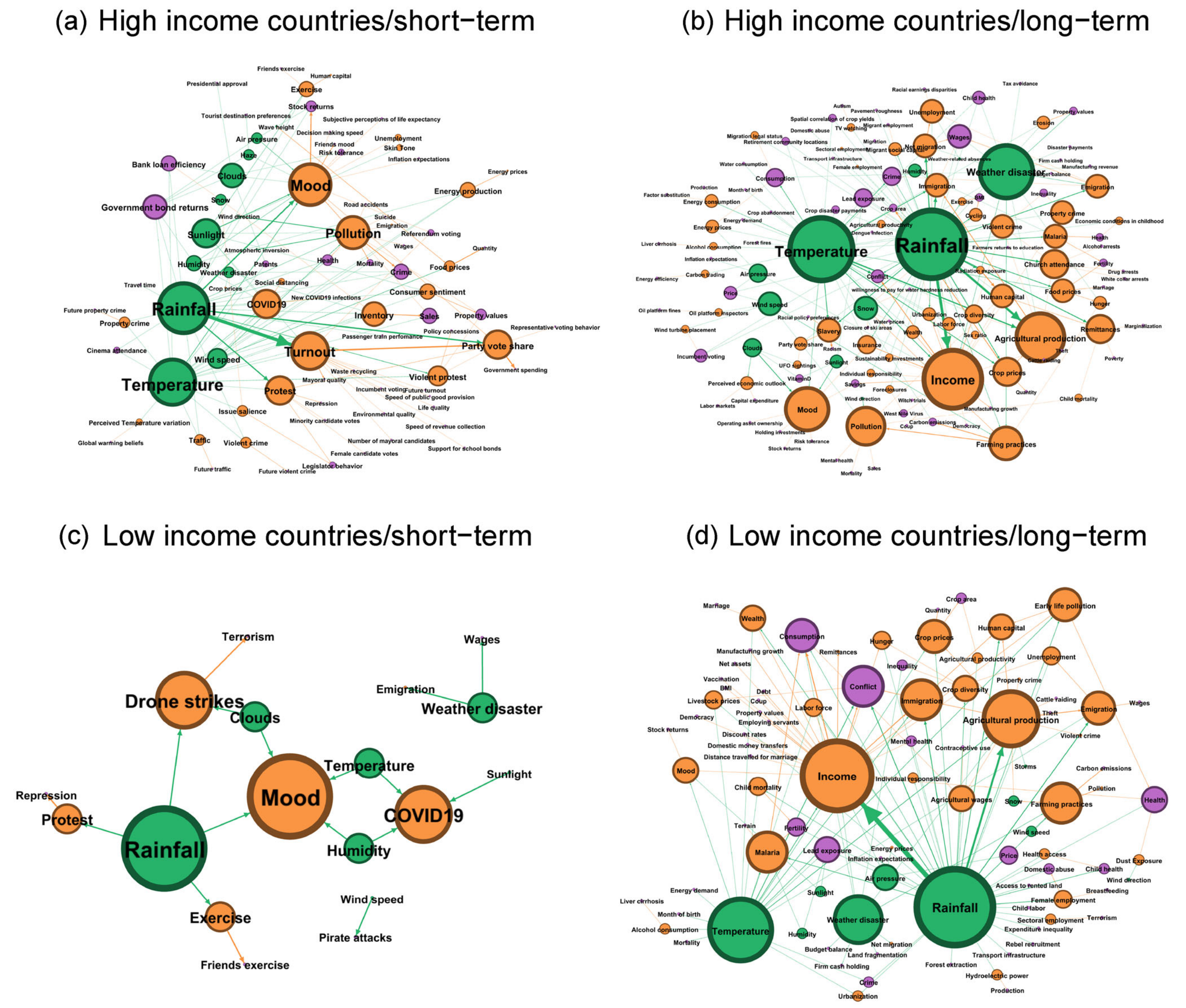


FIGURE 3 Causal webs of weather derived from 289 papers. *Note*: This figure shows that weather has been linked to many variables within time/development categories. Broken down by country-income classification and temporal frequency. Node and tie sizes are proportional to appearances in literature. Colors: weather (green), instrumented variable (orange), and outcome (purple).

Migration follows a similar pattern. Migration studies generally justify the use of weather as an IV by weather's effect on income, as lower incomes provide a push factor for emigrants and higher incomes are a pull factor for immigrants (e.g., Strobl & Valfort, 2015). However, there is evidence for noneconomic links between migration patterns and weather, with Moretti (1998) finding good weather predicts the location of US retirement communities, and Rappaport (2007) finding the same pattern for US internal-migration more generally.

The next causal route is via the legacy of US slavery. Acharya et al. (2016) use cotton-growing climate (which includes weather variables) as an instrument for slavery. Slavery's legacy is therefore a potential exclusion-restriction violation for US IV studies using long-term weather.

There are hundreds of other variables including autism, government spending, drone strikes, unidentified flying object (UFO) sightings, global warming beliefs, alcohol consumption, electricity demand, recycling, malaria, sex ratios, education, insurance, lead exposure, female labor force participation, witch trials, cinema attendance, child labor, birth month, cattle theft, democracy, urbanization, and bicycle usage.

Searching for exceptions where X_j is not an exclusion-restriction violation

I demonstrate how to identify potential exceptions for X_j being an exclusion-restriction violation using an

15405907, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ajps.12894 by Cochrane Colombia, Wiley Online Library on [26/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

economic activity → *civil conflict* IV estimate (Hodler & Raschky, 2014).

Table 2 shows the 50 relevant X_j variables other studies claim are caused by weather. The assumption column shows the outcome for the variables in the X_j column. The justification and evidence for each categorization is shown in greater detail in Table A1 (online Appendix p. 2). My goal is not to advocate for these specific classifications, but to illustrate how the substantive assumptions behind the exclusion restriction can be shown explicitly to readers (who can decide whether they are convinced by the reasoning in Table A1 [online Appendix p. 2]).

I categorize six variables as not being violations because the proposed $W \rightarrow X_j$ relationship does not apply to the current context (exception 1). For example, the claim that rainfall causes female labor force participation is based on studies of India (Chin, 2012) and Vietnam (Thai & Myrskylä, 2012), where rice farming is considered women's work. Rice is a smaller part of African agriculture. I also reject X_j variables where the claim is based on flawed work. For instance, the human capital and urbanization links to rainfall are based on a single cross-sectional working paper (Lagerlöf & Basher, 2006) that used rainfall to instrument multiple different X_j variables. The published version of that paper omits this analysis.

Next, 13 X_j variables are not considered exclusion-restriction violations because they are part of X_1 (exception 2). "Economic activity" is broad, so many other variables such as "crop production" and "income" can be considered components of it. Weather-IV studies with a narrower X_1 variable (such as crop income) would not be able to exclude as many X_j variables at this stage.

Two variables are eliminated because X_j is part of Y (exception 4). Rebel recruitment can be reasonably treated as part of civil conflict rather than a separate process. Similarly, some scholars argue there is no strict boundary between crime and civil conflict (Collier, 2000).

I eliminate nine cases where X_1 mediates the relationship between W and X_j and, crucially, where there is no reason to believe that W has causal pathways to X_j that do not run via X_1 (exception 6). I used the original studies linking W and X_j to establish whether multiple mechanisms were posited or only an economic activity mechanism. For instance, Ajefu and Abiona (2019) link weather shocks to child labor, but explicitly treat weather as an income-based shock meaning child labor falls under exception 6. Conversely, although weather's effect on emigration is often described as a response to income shocks (Kleemans & Magruder, 2018; Pugatch & Yang, 2011), others describe it as

TABLE 2 Instrumental variable (IV) identification assumptions for the causal effect of economic activity (X_1) on civil conflict (Y) for potential exclusion-restriction variables (X_j) related to weather instruments (W) in relevant literature.

Exception	Assumption	X_j
1	W does not cause X_j in this context	Female employment Early life pollution Individual responsibility Human capital Urbanization Month of birth
2	X_j is part of X_1	Agricultural productivity Income Agricultural production Labor force Wealth Crop prices Sectoral employment Agricultural wages Price Consumption Hydroelectric power Unemployment Energy demand
3	X_1 is the only part of X_j caused by W	
4	X_j is part of Y	Rebel recruitment Crime
5	Y is part of X_j	
6	W only causes X_j via X_1	Forest extraction Access to rented land Expenditure inequality Child labor Budget balance Firm cash holding Inflation expectations Inequality Livestock prices
7	X_j does not cause Y	Farming practices Child mortality Crop diversity Child health Fertility Health Dust exposure Mortality
8	X_j only causes Y via X_1	Energy prices
None	X_j is an exclusion-restriction violation	Immigration Emigration Health access Mental health Malaria Lead exposure Transport infrastructure Hunger Alcohol consumption Mood Land fragmentation

Note: Remaining exclusion-restriction violations shown in bottom row.

15405907, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ajps.12894 by Cochrane Colombia, Wiley Online Library on [26/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

working via direct displacement such as Hurricane Maria's effect on Puerto Rican emigration (DeWaard et al., 2020), meaning emigration does not fall into this exception.

There are also situations where I assert that the $X_j \rightarrow Y$ relationship does not exist (exception 7). In each case, I search the literature for evidence of the relationship and consider plausible links not discussed in the literature. For example, no studies claim child mortality causes civil conflict in the near term, and it is difficult to imagine why this relationship would exist. Of 20 X_j variables examined for causal links to civil conflict, 12 appear to cause civil conflict and eight do not. This highlights that most endogenous social science variables are causally connected.

Finally, there was one case where $X_j \rightarrow$ *civil conflict* is fully mediated by economic activity (exception 8). Natalini et al. (2020) only posit economic mechanisms for the link between energy prices and civil conflict.

No cases were found where X_1 is the only relevant part of X_j (exception 3); or Y is part of X_j (exception 5).

This leaves 11 X_j variables, which are likely exclusion-restriction violations for *economic activity* $\rightarrow$ *civil conflict*. These include malaria (which prolongs civil wars according to Bagozzi, 2016), emigration (which provides an exit option for participants and victims of conflict according to Peters & Miller, 2022), and hunger (which drives civil conflict independently of economic effects according to Koren & Bagozzi, 2017).

These exclusion-restriction violations imply the *economic activity* $\rightarrow$ *civil conflict* IV model is not causally identified. However, the extent of bias in the estimate depends on the strength of the $W \rightarrow X_j$ and $X_j \rightarrow Y$ causal links.

Multiple pathways to the same dependent variable

While I do not apply the full exception-search methodology to the broad set of weather-IV estimates, the literature review uncovered some sets of studies implying multiple pathways between weather and the same dependent variable.

Ten weather-IV studies have crime as a dependent variable. However, the X_j variables that crime is claimed to be causally influenced by include income (e.g., Mehlum et al., 2006), crop production (Papaioannou, 2017), air quality (Bondy et al., 2020), church attendance (Moreno-Medina, 2021), and immigration (Chalfin, 2014), as well as directly through heat-induced aggression (Blakeslee & Fishman, 2018; Jacob et al., 2007; Ranson, 2014).

Party vote share is the dependent variable in 17 studies, with X_j most commonly being turnout (Arnold & Freier, 2016; Artés, 2014; Gomez et al., 2007; Hansford & Gomez, 2010; Horiuchi & Kang, 2018; Kang, 2019; Knack, 1994). However, other studies argue that election-day weather affects voters' mood (Bassi, 2019; Duhaime & Moulton, 2016) or risk tolerance (Bassi, 2019), or that previous weather discouraged protest attendance (Madestam et al., 2013; Wasow, 2020).

Finally, 15 studies have conflict as a dependent variable, with income being the most common X_j variable (Hodler & Raschky, 2014; Hsiang et al., 2011; Jensen & Gleditsch, 2009; Miguel & Satyanath, 2011; Miguel et al., 2004). However, other conflict studies use weather to instrument agricultural production (Caruso et al., 2016), water scarcity (Landis et al., 2017), livestock prices (Maystadt & Ecker, 2014), and immigration (Bhavnani & Lacina, 2015).

These examples demonstrate that some weather-IV studies have directly contradictory identifying assumptions. But even without directly observing $X_j \rightarrow Y$, the existence of $W \rightarrow X_j$ changes the IV strategy from only requiring claims about the narrow domain of weather, to making causal claims about the complex social world.

Sensitivity analysis of weather instrumental variables

The literature review contained 57 weather-IV estimates with reduced-form estimates or sufficient information to reverse-engineer the reduced form. The literature also produced 121 first-stage estimates that I converted into $R^2_{X_j \sim W|K}$ values.

On average, I identified 32 potential exclusion-restriction violations per reduced-form equation (93 if dependent variables are included). However, on average, I only found a suitable $R^2_{X_j \sim W|K}$ estimate for three X_j variables for each reduced-form equation and in 15 cases I found none.

This means the sensitivity analysis covers only a small fraction of the potential exclusion-restriction violations documented in the literature. Importantly, the broad analysis of weather IVs has not searched for exclusion-restriction exceptions, so some of these apparent violations will fall into one of those exceptions. When applying this method to a particular IV claim, it would be important to include this step.

Figure 4 shows each reduced-form estimate's sensitivity (y-axis) to the X_j variables with relevant $R^2_{X_j \sim W|K}$ estimates. The shading and number indicate the strength of the $R^2_{Y \sim X_j|W,K}$ relationship required to make the $X_1 \rightarrow Y$ IV estimate insignificant. Smaller values indicate that a weak $X_j \rightarrow Y$ relationship could make study 1′s result insignificant.

15405907, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ajps.12894 by Cochrane Colombia, Wiley Online Library on [26/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

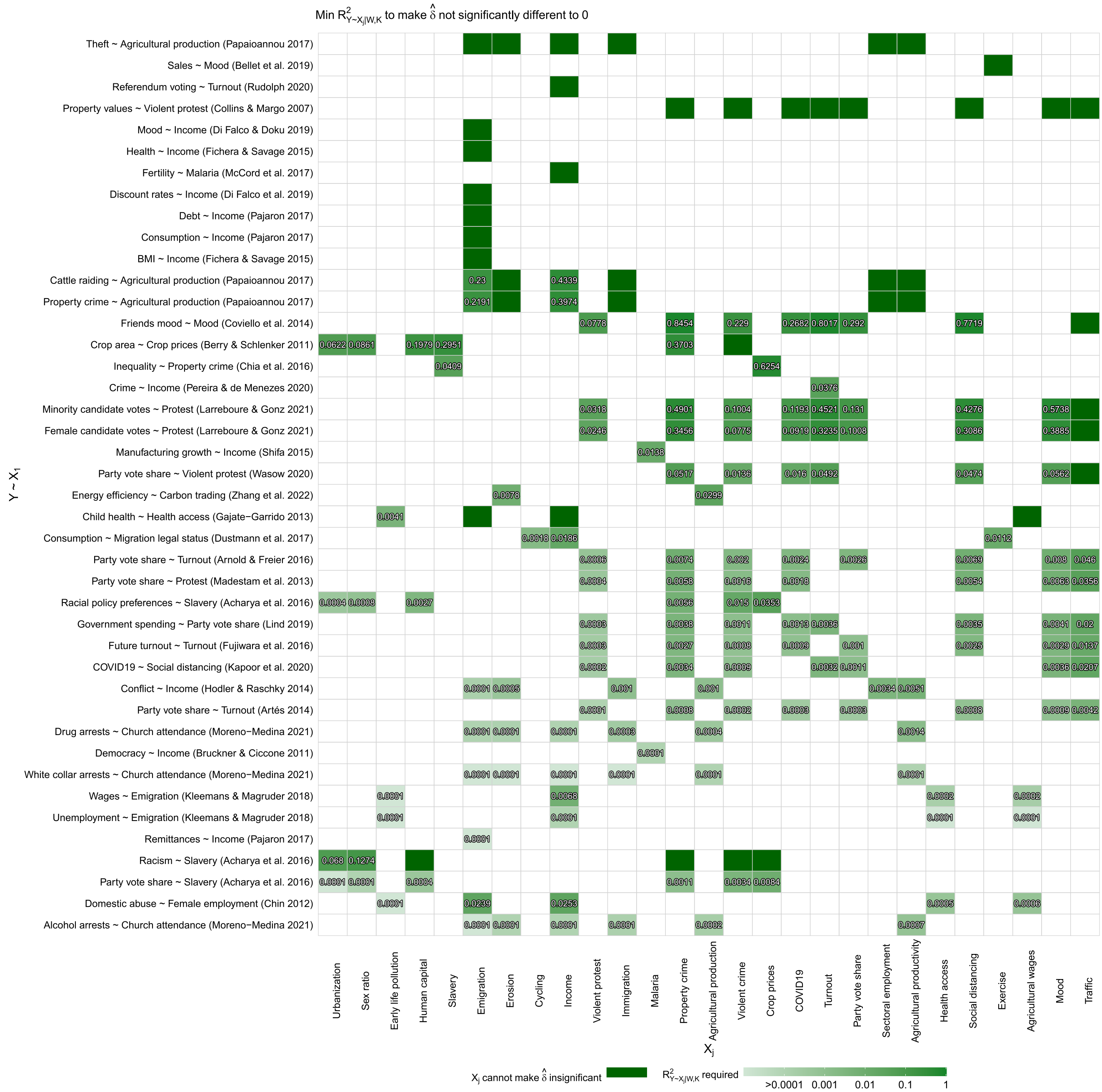


FIGURE 4 The minimum relationship between an exclusion-restriction violation variable, X_j, and dependent variable Y ($R^2_{Y\sim X_j|W,K}$) required to make the instrumental variable (IV) relationship between X_1 and Y ($\hat{\delta}$) insignificant. *Note*: This figure shows that many weather-IV estimates require only small relationships between X_j and Y to make the result insignificant. Strength of required relationship indicated using shading on log-10 scale. Dark green shading indicates X_j cannot make the relationship insignificant. Lowest $R^2_{Y\sim X_j|W,K}$ tested was .0001. Fifteen reduced-form equations without relevant $R^2_{X_j\sim W|K}$ estimates are omitted. Fit statistics are filtered using the broad criteria and fit-statistic criteria.

Despite exclusion-restriction violations, some studies are likely robust to large $X_j \rightarrow Y$ relationships. For instance, Collins and Margo (2007) estimate the effect of violent protests on property prices using a rainfall instrument. None of the eight variables with relevant first-stage estimates (including property crime, violent crime, mood, and traffic) could make this estimate insignificant even if they were perfectly correlated with property prices. While many additional potential exclusion-restriction violations do not have first-stage estimates, the result appears relatively robust to exclusion-restriction violations.

15405907, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ajps.12894 by Cochrane Colombia, Wiley Online Library on [26/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Other studies require only small $X_j \rightarrow Y$ relationships. Moreno-Medina (2021) examines the effect of church attendance on various types of crime by using the number of weeks where it rained between 9 a.m. and 1 p.m. on Sunday in a year as an instrument for church attendance. There are five relevant potential exclusion-restriction violations with first-stage fit statistics available: emigration, erosion, income, crop production, and agricultural productivity. For the weakest result (alcohol arrests), any of these variables would only need to explain .07% of variance in alcohol arrests to make the result insignificant. None of the paper's findings would require one of the exclusion-restriction variables to explain more than .14% of variance in a type of crime. While these variables are not major drivers of crime, it is plausible that they could explain tiny fractions of a percent of variation in crime. Even though the instrumented variable in the paper is rain on Sundays, the tiny proportion of variance needed to make the results insignificant implies almost anything could overturn them.

A high proportion of studies have at least one potential exclusion-restriction variable that could easily make the result insignificant. Of the 42 reduced-form estimates with at least one exclusion-restriction estimate to test, 20 had at least one X_j variable that would need to explain less than 1% of variance in Y to make the result insignificant, and 17 have at least one X_j variable that would need to explain less than .1% of variance in Y to make the result insignificant. Only 11 were sufficiently robust that no relationship between X_j and Y could make the estimate insignificant.

Since most X_j variables were not tested due to data limitations, these results represent a lower bound of weather-IV estimate sensitivity and researchers should make sure to consider plausible effect sizes for the other X_j variables when applying this method.

Figure 5 shows the strength of relationships required to make the *economic activity* → *civil conflict* IV estimate insignificant. In order to believe the results are acceptably biased (defined as the coefficient and standard error not changing enough to make the result insignificant or increasing in magnitude by an equivalent amount), we must assume that each pair of partial-R^2 values is below the curve in Figure 5. If a reader finds this plausible for the 11 exclusion-restriction variables, then it is reasonable to believe *economic activity* → *civil conflict*. If both $R^2_{X_j \sim W|K}$ and $R^2_{Y \sim X_j|W,K}$ were around .0045, the results would no longer be significant. A higher variance explained for one would reduce the size needed for the other.

I would personally be skeptical of basing an important empirical claim about the world on the assumption that none of the 11 exclusion-restriction variables lie above this curve. However, the important point is that readers can assess what they are required to believe to accept the result.[12]

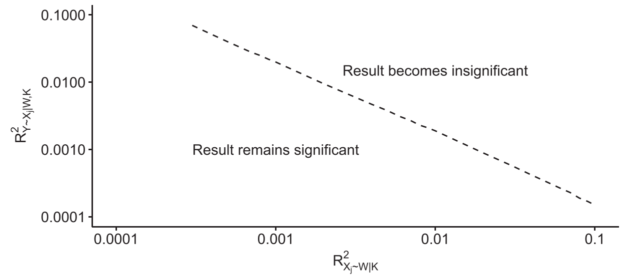


FIGURE 5 Combinations of $R^2_{Y \sim X_j|W,K}$ and $R^2_{X_j \sim W|K}$ that would make *economic activity* → *conflict* no longer significant. *Note*: This figure shows that various modest combinations of relationship strengths between the exclusion-restriction variable and instrument/dependent variable could make the *economic activity* → *conflict* insignificant. Pairs of values below the curve are insufficient to make the result insignificant. Pairs above the curve would make the result insignificant. Both axes are on a log 10 scale.

As discussed in the earlier sensitivity analysis, transferring partial-R^2 values across contexts requires meeting several criteria. Of the 11 exclusion-restriction variables, only emigration has an $R^2_{X_j \sim W|K}$ value that meets these requirements: Two relevant studies report a mean $R^2_{X_j \sim W|K}$ of .011 between rainfall and emigration (Badaoui et al., 2017; Corbi & Ferraz, 2018). If this value applies to the *economic activity* → *conflict* estimate, *emigration* → *conflict* would only need an $R^2_{Y \sim X_j|W,K}$ of .001 to render the *economic activity* → *civil conflict* estimate insignificant. Even this figure should be treated carefully. While some studies link rainfall to emigration through mechanisms other than income, many studies argue income is *a* relevant pathway. Consequently, emigration may need to explain somewhat more variance to overturn the result.

No applicable $R^2_{X_j \sim W|K}$ values exist for rainfall's effect on malaria. However, the literature provides three first-stage estimates based on other variance types: between-person (Chang et al., 2014), within-country (McCord et al., 2017), and between-subnational (Cutler et al., 2010), ranging from .0084 to .031. Even the smallest of these implies an $R^2_{Y \sim X_j|W,K}$ of .0021 would render the IV estimate insignificant.

Explicit sensitivity conditions allow readers to see what substantive claims are required for an IV result to hold. Here, the *economic activity* → *civil conflict* estimate is sensitive to small effects of variables that

[12] Since there are multiple exclusion-restriction violations that create unacceptable bias, I do not continue the exclusion-restriction variable search. However, if a researcher was willing to accept this level of bias, they would need to continue to enumerate j variables from the literature to ensure that they had not missed important exclusion-restriction variables.

15405907, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ajps.12894 by Cochrane Colombia, Wiley Online Library on [26/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

the literature demonstrates are exclusion-restriction violations. However, studies with stronger estimates or less problematic instruments may require weaker assumptions to accept the IV results. By explicating and quantifying the conditions for strict identification and sensitivity, readers can better assess the plausibility of IV claims.

CONCLUSIONS

Cunningham (2018) argues that a good instrument should have a “certain ridiculousness.” Until the causal pathway is explained, the link between the instrument and outcome seems absurd. In a world where forest fires prevent Australians and Californians from leaving their homes for months, and 1–3 billion people are projected to be left outside of historically habitable temperature ranges (Xu et al., 2020), the link between weather and the social world is not absurd enough.

This paper shows that a widely used instrument is flawed because it systematically violates the exclusion-restriction assumption for most plausible applications. These findings suggest our standards for accepting instruments are overly lenient and atheoretical. The causal revolution in social science has revealed the wishful thinking underlying selection-on-observables. The same rigor must be applied to IV assumptions.

Nothing in this paper disproves particular empirical claims. Many weather-IV papers provide independent sources of evidence, and robustness checks (e.g., placebo tests) and some results will be true by chance. Nevertheless, my results suggest the underlying assumptions of weather-IVs are not strictly true, and many results are likely wrong.

Evidence for widespread exclusion-restriction violations adds to growing evidence that IV regression studies are often poorly conducted. Brodeur et al. (2020) found IV studies showed substantial evidence of p-hacking. Similarly, Lee et al. (2021) find the standard first-stage F-test used in IV regression is underpowered, and that over half of IV papers in the *American Economic Review* were no longer significant after accounting for this. Lal et al. (2024) find that most top political science IV articles make statistical errors, including inappropriate standard error and F-statistic calculations.

It is tempting to interpret instruments as being “used up” by multiple uses. However, it is important to distinguish the *fact* of whether an exclusion-restriction violation exists from our *knowledge* of that fact (which new research can change). This may seem unfair to authors who used the best available information, but the goal of this method is to assess empirical claims not to cast judgment on researchers.

While a scholar cannot know whether a new instrument will turn out to have exclusion-restriction violations, there are some classes of instruments where W does not seem as likely to have multiple X_j pathways. Random assignment to a treatment with partial compliance is unlikely to suffer from exclusion-restriction violations (e.g., winning a charter school lottery as an instrument for attending charter school; Angrist et al., 2017). Narrow policy changes also seem less susceptible. For instance, Porzio et al. (2022) instrument a cohort’s average years of education with the school-leaving age when they were young. These narrow instruments are also likely to be much easier to assess with this paper’s exclusion-restriction method precisely because existing studies will not have linked them to huge numbers of other variables.

By contrast, other widely used instruments (e.g., population density, historical splits in countries, lead exposure, colonialism, distance from anything) likely suffer from the problems described in this paper. Indeed, many of these instruments have faced criticism for similar reasons (Gallen & Raymond, 2021; Morck & Yeung, 2011).

This paper offers a way forward for IV inference through a method for systematically finding and assessing exclusion-restriction violations from existing literature. This process unpacks the exclusion restriction into a series of substantive (and potentially testable) claims about the world. While we cannot prove the exclusion restriction, we can use the literature to identify and assess variables that have been linked to our instrument. This paper also demonstrates how sensitivity analysis (Cinelli & Hazlett, 2022, 2020) can be used to quantify the strength of an exclusion-restriction violation required to overturn a result. This approach allows researchers to acknowledge violations of their identifying assumptions without having to abandon analysis that is not perfectly identified.

Avoiding contradictions in the literature is important beyond weather IVs. Typical nonexperimental social science models implicitly assume a sparse causal graph, which is unrealistic given the volume of papers making causal claims (see online Appendix B p. 4). Unfortunately, this volume means constructing a realistic causal graph is likely beyond the means of individual social scientists. Infrastructure to systematically record, dispute, and qualify the literature’s causal claims could allow scientists to appropriately specify models with nonexperimental data.

Causal knowledge in social science has few easy shortcuts. Apparently simple solutions such as IVs still rely on strong substantive claims about the world. To build cumulative knowledge, we must engage

15405907, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ajps.12894 by Cochrane Colombia, Wiley Online Library on [26/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

directly with the messy causal web of the social world.

ACKNOWLEDGMENTS

Thanks to Tom Pepinsky, Jack Bailey, Chris Prosser, Carlos Cinelli, the University of Manchester Democracy and Elections Group, the Trade, Environment, Development and Energy Research Group at the University of Birmingham, and the McGill Department of Economics for feedback on drafts of this paper. The views expressed herein are those of the author and do not reflect the position of the United States Military Academy, the Department of the Army, or the Department of Defense.

ORCID

Jonathan Mellon https://orcid.org/0000-0001-6754-203X

REFERENCES

Acharya, Avidit, Matthew Blackwell, and Maya Sen. 2016. "The Political Legacy of American Slavery." *Journal of Politics* 78(3): 621–41.

Ajefu, Joseph, and Olukorede Abiona. 2019. "Impact of Shocks on Labour and Schooling Outcomes and the Role of Public Work Programmes in Rural India." *Journal of Development Studies* 55(6): 1140–57.

Angrist, Joshua, Peter Hull, Parag Pathak, and Christopher Walters. 2017. "Leveraging Lotteries for School Value-Added: Testing and Estimation." *The Quarterly Journal of Economics* 132(2): 871–919.

Angrist, Joshua, Guido Imbens, and Donald Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91(434): 444–55.

Aral, Sinan, and Christos Nicolaides. 2017. "Exercise Contagion in a Global Social Network." *Nature Communications* 8(1): 1–8.

Arnold, Felix, and Ronny Freier. 2016. "Only Conservatives Are Voting in the Rain: Evidence from German Local and State Elections." *Electoral Studies* 41(March): 216–21.

Artés, Joaquín. 2014. "The Rain in Spain: Turnout and Partisan Voting in Spanish Elections." *European Journal of Political Economy* 34(June): 126–41.

Auffhammer, Maximilian, Solomon Hsiangy, Wolfram Schlenker, and Adam Sobelz. 2013. "Using Weather Data and Climate Model Output in Economic Analyses of Climate Change." *Review of Environmental Economics and Policy* 7(2): 181–98.

Badaoui, Eliane, Eric Strobl, and Frank Walsh. 2017. "The Impact of Internal Migration on Local Labour Markets in Thailand." *Economic Development and Cultural Change* 66(1): 147–77.

Bagozzi, Benjamin. 2016. "On Malaria and the Duration of Civil War." *Journal of Conflict Resolution* 60(5): 813–39.

Bassi, Anna. 2019. "Weather, Risk, and Voting: An Experimental Analysis of the Effect of Weather on Vote Choice." *Journal of Experimental Political Science* 6(1): 17–32.

Bhavnani, Rikhil, and Bethany Lacina. 2015. "The Effects of Weather-Induced Migration on Sons of the Soil Riots in India." *World Politics* 67(4): 760–94.

Blakeslee, David, and Ram Fishman. 2018. "Weather Shocks, Agriculture, and Crime: Evidence from India." *Journal of Human Resources* 53(3): 750–82.

Bondy, Malvina, Sefi Roth, and Lutz Sager. 2020. "Crime Is in the Air: The Contemporaneous Relationship between Air Pollution and Crime." *Journal of the Association of Environmental and Resource Economists* 7(3): 555–85.

Brodeur, Abel, Nikolai Cook, and Anthony Heyes. 2020. "Methods Matter: P-Hacking and Publication Bias in Causal Analysis in Economics." *American Economic Review* 110(11): 3634–60.

Caruso, Raul, Ilaria Petrarca, and Roberto Ricciuti. 2016. "Climate Change, Rice Crops, and Violence: Evidence from Indonesia." *Journal of Peace Research* 53(1): 66–83.

Chalfin, Aaron. 2014. "What Is the Contribution of Mexican Immigration to U.S. Crime Rates? Evidence from Rainfall Shocks in Mexico." *American Law and Economics Review* 16(1): 220–68.

Chang, Simon, Belton Fleisher, Seonghoon Kim, and Shi Liu. 2014. "Long-Term Health Effects of Malaria Exposure around Birth: Evidence from Colonial Taiwan." *Economic Development and Cultural Change* 62(3): 519–36.

Chin, Yoo. 2012. "Male Backlash, Bargaining, or Exposure Reduction? Women's Working Status and Physical Spousal Violence in India." *Journal of Population Economics* 25(1): 175–200.

Ciccone, Antonio. 2011. "Economic Shocks and Civil Conflict: A Comment." *American Economic Journal: Applied Economics* 3(4): 215–27.

Cinelli, Carlos, and Chad Hazlett. 2022. "An Omitted Variable Bias Framework for Sensitivity Analysis of Instrumental Variables." SSRN.

Cinelli, Carlos, and Chad Hazlett. 2020. "Making Sense of Sensitivity: Extending Omitted Variable Bias." *Journal of the Royal Statistical Society. Series B: Statistical Methodology* 82(1): 39–67.

Cinelli, Carlos, Jeremy Ferwerda, and Chad Hazlett. 2020. "Sensemakr: Sensitivity Analysis Tools for OLS in R and Stata." SSRN.

Collier, Paul. 2000. "Rebellion as a Quasi-Criminal Activity." *Journal of Conflict Resolution* 44(6): 839–53.

Collins, William, and Robert Margo. 2007. "The Economic Aftermath of the 1960s Riots in American Cities: Evidence from Property Values." *Journal of Economic History* 67(4): 849–83.

Cook, Duncan, and Sally Garrett. 2013. "Somali Piracy and the Monsoon." *Weather, Climate, and Society* 5(4): 309–16.

Cooperman, Alicia. 2017. "Randomization Inference with Rainfall Data: Using Historical Weather Patterns for Variance Estimation." *Political Analysis* 25(3): 277–88.

Corbi, Raphael, and Tiago Ferraz. 2018. *Rainfall, Internal Migration and Local Labor Markets in Brazil*. Department of Economics-FEA/USP Working Paper Series No 2018-24

Cunningham, Scott. 2018. "Causal Inference: The Mixtape." https://tufte-latex.googlecode.com

Cutler, David, Winnie Fung, Michael Kremer, Monica Singhal, and Tom Vogl. 2010. "Early-Life Malaria Exposure and Adult Outcomes: Evidence from Malaria Eradication in India." *American Economic Journal: Applied Economics* 2(2): 72–94.

DeWaard, Jack, Janna Johnson, and Stephan Whitaker. 2020. "Out-Migration from and Return Migration to Puerto Rico after Hurricane Maria: Evidence from the Consumer Credit Panel." *Population and Environment* 42(1): 28–42.

Duhaime, Erik, and Taylor Moulton. 2016. "'Swingin' in the Rain: The Impact of Inclement Weather on Voting Behavior in U.S. Presidential Elections." SSRN.

Fan, Maoyong, and Yi Wang. 2020. "The Impact of PM2.5 on Mortality in Older Adults: Evidence from Retirement of Coal-Fired Power Plants in the United States." *Environmental Health: A Global Access Science Source* 19(1): 1–13.

Fontenla, Matías, M. Goodwin, and Fidel Gonzalez. 2019. "Pollution and the Choice of Where to Work and Live within Mexico City." *Latin American Economic Review* 28(11): 1–17.

Gallen, Trevor, and Ben Raymond. 2021. "*Broken Instruments*". SSRN.

Gomez, Brad, Thomas Hansford, and George Krause. 2007. "The Republicans Should Pray for Rain: Weather, Turnout, and Voting in U.S. Presidential Elections." *Journal of Politics* 69(3): 649–63.

15405907, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ajps.12894 by Cochrane Colombia, Wiley Online Library on [26/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Gu, Hejun, Weiran Yan, Ehsan Elahi, and Yuxia Cao. 2020. "Air Pollution Risks Human Mental Health: An Implication of Two-Stages Least Squares Estimation of Interaction Effects." *Environmental Science and Pollution Research* 27(2): 2036–43.

Guven, Cahit, and Indrit Hoxha. 2015. "Rain or Shine: Happiness and Risk-Taking." *Quarterly Review of Economics and Finance* 57(August): 1–10.

Hansford, Thomas, and Brad Gomez. 2010. "Estimating the Electoral Effects of Voter Turnout." *American Political Science Review* 104(2): 268–88.

Hodler, Roland, and Paul Raschky. 2014. "Economic Shocks and Civil Conflict at the Regional Level." *Economics Letters* 124(3): 530–33.

Horiuchi, Yusaku, and Woo Kang. 2018. "Why Should the Republicans Pray for Rain? Electoral Consequences of Rainfall Revisited." *American Politics Research* 46(5): 869–89.

Hsiang, Solomon, Kyle Meng, and Mark Cane. 2011. "Civil Conflicts Are Associated with the Global Climate." *Nature* 476(7361): 438–41.

Jacob, Brian, Lars Lefgren, and Enrico Moretti. 2007. "The Dynamics of Criminal Behavior: Evidence from Weather Shocks." *Journal of Human Resources* 42(3): 489–527.

Jensen, Peter, and Kristian Gleditsch. 2009. "Rain, Growth, and Civil War: The Importance of Location." *Defence and Peace Economics* 20(5): 359–72.

Kang, Woo. 2019. "Liberals Should Pray for Rain: Weather, Opportunity Costs of Voting and Electoral Outcomes in South Korea." *Political Science* 71(1): 61–78.

Katz, Tamar, Tali Regev, Shay Lavie, Haggai Porat, and Ronen Avraham. 2020. "Those Who Tan and Those Who Don't: A Natural Experiment on Colorism." *PLoS ONE* 15(7): 1–14.

Kippersluis, Hans van, and Cornelius Rietveld. 2018. "Beyond Plausibly Exogenous." *Econometrics Journal* 21(3): 316–31.

Kiviet, Jan. 2020. "Testing the Impossible: Identifying Exclusion Restrictions." *Journal of Econometrics* 218(2): 294–316.

Kleemans, Marieke, and Jeremy Magruder. 2018. "Labour Market Responses to Immigration: Evidence from Internal Migration Driven by Weather Shocks." *Economic Journal* 128(613): 2032–65.

Knack, Steve. 1994. "Does Rain Help the Republicans? Theory and Evidence on Turnout and the Vote." *Public Choice* 79(April): 187–209.

Koren, Ore, and Benjamin Bagozzi. 2017. "Living off the Land: The Connection between Cropland, Food Security, and Violence Against Civilians." *Journal of Peace Research* 54(3): 351–64.

Kubik, Zaneta, and Julian May. 2018. "Weather Shocks, Food Prices and Food Security: Evidence from South Africa." In *The 56th Annual Conference of the Agricultural Economics Association of South Africa* 42.

Lagerlöf, Nils-Petter, and Syed Basher. 2006. "Geography, Population Density, and Per-Capita Income Gaps across US States and Canadian Provinces." Munich Personal RePEc Archive, no. 369.

Lal, Apoorva, Mackenzie Lockhart, Yiqing Xu, and Ziwen Zu. 2024. "How Much Should We Trust Instrumental Variable Estimates in Political Science? Practical Advice Based on 67 Replicated Studies." *Political Analysis*: 1–20.

Landis, Steven, Babak Rezaeedaryakenari, Yifan Zhang, Cameron Thies, and Ross Maciejewski. 2017. "Fording Differences? Conditions Mitigating Water Insecurity in the Niger River Basin." *Political Geography* 56(January): 77–90.

Lee, David, Justin McCrary, Marcelo Moreira, and Jack Porter. 2021. "Valid t-Ratio Inference for IV." NBER Preprint.

Madestam, Andreas, Daniel Shoag, Stan Veuger, and David Yanagizawa-Drott. 2013. "Do Political Protests Matter? Evidence from the Tea Party Movement." *Quarterly Journal of Economics* 128(4): 1633–85.

Maystadt, Jean, and Olivier Ecker. 2014. "Extreme Weather and Civil War: Does Drought Fuel Conflict in Somalia through Livestock Price Shocks?" *American Journal of Agricultural Economics* 96(4): 1157–82.

McCord, Gordon, Dalton Conley, and Jeffrey Sachs. 2017. "Malaria Ecology, Child Mortality & Fertility." *Economics and Human Biology* 24(February): 1–17.

Mehlum, Halvor, Edward Miguel, and Ragnar Torvik. 2006. "Poverty and Crime in 19th Century Germany." *Journal of Urban Economics* 59(3): 370–88.

Meier, Armando, Lukas Schmid, and Alois Stutzer. 2019. "Rain, Emotions and Voting for the Status Quo." *European Economic Review* 119(October): 434–51.

Miguel, Edward, and Shanker Satyanath. 2011. "Re-Examining Economic Shocks and Civil Conflict." *American Economic Journal: Applied Economics* 3(4): 228–32.

Miguel, Edward, Shanker Satyanath, and Ernest Sergenti. 2004. "Economic Shocks and Civil Conflict: An Instrumental Variables Approach." *Journal of Political Economy* 112(4): 725–53.

Morck, Randall, and Bernard Yeung. 2011. "Economies, History, and Causation." *Business History Review* 85(1): 39–63.

Moreno-Medina, Jonathan. 2021. "Sinning in the Rain: Weather Shocks, Church Attendance and Crime." *The Review of Economics and Statistics* 105(1): 1–46.

Moretti, Enrico. 1998. "Social Returns to Education and Human Capital Externalities: Evidence from Cities." Centre for Labour Economics Preprint.

Natalini, Davide, Giangiacomo Bravo, and Edward Newman. 2020. "Fuel Riots: Definition, Evidence and Policy Implications for a New Type of Energy-Related Conflict." *Energy Policy* 147 (August 2019): 111885.

Nicholls, Neville, Wasyl Drosdowsky, and Beth Lavery. 1997. "Australian Rainfall Variability and Change." *Weather* 52(3): 66–72.

Papaioannou, Kostadis. 2017. "'Hunger Makes a Thief of Any Man': Poverty and Crime in British Colonial Asia." *European Review of Economic History* 21(1): 1–28.

Peet, Evan. 2020. "Early-life environment and human capital: evidence from the Philippines." *Environment and Development Economics* 26(1): 1–25.

Peters, Margaret, and Michael Miller. 2022. "Emigration and Political Contestation." *International Studies Quarterly* 66(1): 1–39.

Porzio, Tommaso, Federico Rossi, and Gabriella Santangelo. 2022. "The Human Side of Structural Transformation." *American Economic Review* 112(8): 2774–814.

Pugatch, Todd, and Dean Yang. 2011. "The Impact of Mexican Immigration on U.S. Labor Markets: Evidence from Migrant Flows Driven by Rainfall Shocks." SSRN.

Ranson, Matthew. 2014. "Crime, Weather, and Climate Change." *Journal of Environmental Economics and Management* 67(3): 274–302.

Rappaport, Jordan. 2007. "Moving to Nice Weather." *Regional Science and Urban Economics* 37(3): 375–98.

Ritter, Emily, and Courtenay Conrad. 2016. "Preventing and Responding to Dissent: The Observational Challenges of Explaining Strategic Repression." *American Political Science Review* 110(1): 85–99.

Sager, Lutz. 2019. "Estimating the Effect of Air Pollution on Road Safety Using Atmospheric Temperature Inversions." *Journal of Environmental Economics and Management* 98(251): 1–20.

Sarsons, Heather. 2015. "Rainfall and Conflict: A Cautionary Tale." *Journal of Development Economics* 115(July): 62–72.

Shen, Xinyi, Chenkai Cai, and Hui Li. 2020. "Quantifying Socioeconomic Activities and Weather Effects on the Global Spread of COVID-19 Epidemic." Arxiv.

Strobl, Eric, and Marie Valfort. 2015. "The Effect of Weather-Induced Internal Migration on Local Labor Markets. Evidence from Uganda." *World Bank Economic Review* 29(2): 385–412.

15405907, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ajps.12894 by Cochrane Colombia, Wiley Online Library on [26/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Thai, Thuan, and Mikko Myrskylä. 2012. "Rainfall Shocks, Parental Behavior and Breast-Feeding: Evidence from Rural Vietnam." MPIDR Working Paper 2012-009.

Wasow, Omar. 2020. "Agenda Seeding: How 1960s Black Protests Moved Elites, Public Opinion and Voting." *American Political Science Review* 114(3): 638–59.

World Bank. 2014. "Table 3.2 World Development Indicators: Agricultural Inputs." World Bank.

World Bank. 2020. "World Bank Climate Knowledge Portal." World Bank.

Xu, Chi, Timothy Kohler, Timothy Lenton, Jens Svenning, and Marten Scheffer. 2020. "Future of the Human Climate Niche." *PNAS* 117(21).

Zheng, Siqi, Jianghao Wang, Cong Sun, Xiaonan Zhang, and Matthew Kahn. 2019. "Air Pollution Lowers Chinese Urbanites' Expressed Happiness on Social Media." *Nature Human Behaviour* 3(3): 237–43.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Mellon, Jonathan. 2025. "Rain, rain, go away: 194 potential exclusion-restriction violations for studies using weather as an instrumental variable." *American Journal of Political Science* 69: 881–98. https://doi.org/10.1111/ajps.12894

15405907, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ajps.12894 by Cochrane Colombia, Wiley Online Library on [26/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License