Visualizing environmental econometrics using directed acyclic graphs

Pierce Donovan[†]

October 2023

Abstract

Environmental economists have gravitated toward writing empirical papers with an emphasis on causal inference. Despite this development, there has not been much progress in the way of adopting an explicit framework for communicating causal hypotheses—prior beliefs about the structure of a data generating process. The shortfall reduces the transparency and accessibility of the assumptions underlying effect identification and makes the testing of causal hypotheses impossible. This article explains why an explicit framework is worthwhile and demonstrates how Directed Acyclic Graphs can augment and standardize the communication of causal knowledge.

Keywords: causal models, identification, research design, specification testing, communication, directed acyclic graphs (JEL: A20, C12, C51, C52, Q50)

[†]Assistant Professor, Department of Economics, University of Nevada, Reno
email: `pierce.donovan@unr.edu`
web: `piercedonovan.github.io`

A framework for communicating causal knowledge

Most empirical papers in environmental economics today are structured around a single statistical inference of interest and written to emphasize how an identification strategy supports a measurement with a causal interpretation (Segerson, 2019). In many cases, the discussion appears to take more effort than the statistics, because designing clever identification strategies for causal inference is difficult. It is now the thoughtful communication of research design that defines "rigor" in empirical papers, rather than a barrage of specification tests (Angrist and Pischke, 2010).

This article considers the clarity, transparency, and testability of *causal hypotheses*—prior beliefs about the structure of a data generating process. This is thought to be addressed when testing a statistical inference's sensitivity to varying empirical assumptions about functional form or control variables, but this exercise in showing consistency across regression specifications is not undertaken with any explicit causal model in mind. While sensitivity analyses are likely to give researchers confidence in their results, they are unable to determine whether an approach estimates an effect of interest.

The spirit of specification testing is to learn whether a statement like "the effect of X on Y is likely positive" is permissible. A test of a result's sensitivity to an assumption about a data generating process is suitable to bolster this claim because it tests if a researcher is estimating the effect of interest. This change to a causal model will naturally motivate a specific change in an empirical strategy and highlight a critical prior belief; in contrast, the functional permutation of a regression specification fails to imply anything about the correctness of the underlying causal model. In essence, the typical regression specification test assumes that the result is always causally-interpretable—once the identification assumptions have been mentioned elsewhere in the article. When causal models are not explicit, the testing of causal hypotheses becomes impossible.

Directed Acyclic Graphs (DAGs) (Pearl, 1995) systematize the creation and testing of causal hypotheses. DAGs display a researcher's priors and empirical assumptions, motivate identification strategies, and indicate the conditions under which a regression yields a causally-interpretable measurement. They provide a way to choose conditioning variables, sources of data, and empirical methods in a manner that can be easily understood and validated by a wider audience of researchers, students, and stakeholders. The purpose of this article is to provide a complete introduction to this visualization technique, explore how it can facilitate causal hypothesis testing over a wide swath of settings, provide a fair assessment of the costs and benefits of DAGs, and share novel econometric insights discovered while investigating their use. This article lowers the costs of adoption for an emerging method of broad applicability in environmental economics.

Several articles compare the DAG framework to competing foundations for conveying causal information. The potential outcomes framework is already widely-used in applied economics to validate identification strategies, and has the advantage over DAGs in terms of the scope of empirical assumptions and implications that can be represented (e.g. monotonicity and the local average treatment effect) (Imbens, 2020). In contrast, the DAG framework is better at making the historically "ad hoc" facets of causal inference in economics more transparent and systematic (e.g. covariate selection and the description of sources of bias) (Schneider, 2020; Huntington-Klein, 2022b). Heckman and Pinto (2022) claims that the scope of these frameworks are limited [outside of empirical economics], and promotes structural equation modeling to address causal questions in non-empirical settings (e.g. general equilibrium).

The next two sections provide a primer on DAGs and several applications to illustrate their utility. The penultimate section demonstrates a review of a recent research article from the environmental economics literature with the aid of a DAG. The final section discusses the merits of integrating this innovation into future presentations, publications, peer reviews, and pedagogy.

A primer on Directed Acyclic Graphs

Figure 1 displays a representative DAG. When constructing a DAG, the initial task is to center thinking around a particular relationship of interest, e.g. $T \rightarrow Y$. Directed arrows like $T \rightarrow Y$ convey statements like "the outcome Y is in part determined by the status of some treatment T." Additional causal relationships between T and Y may be mediated by other variables (e.g. M). A causal model is expanded to explicitly consider any variables which could distort the observed relationship between T and Y away from a causal interpretation. The level of complexity of a model is ultimately up to the researcher, and every inclusion or omission of a variable or arrow marks an explicit assumption about the underlying data generating process.

Spurious relationships between T and Y are created by "confounding" variables—those which influence both T and Y—like X or Q. If the effects of these confounding variables are not mitigated, an estimate of the treatment effect will be biased. Conversely, variables like H will not influence the causal interpretation of a measured correlation between T and Y, since they do not contribute to a spurious relationship. Thus the DAG provides a way to differentiate between malignant and benign sources of variation.

Ideally, there would exist some experimental source of variation for T. This would sever any arrows leading towards T in Figure 1. But this ideal isn't necessary. If data on X and Q are available, a simple matching strategy can control for these confounding

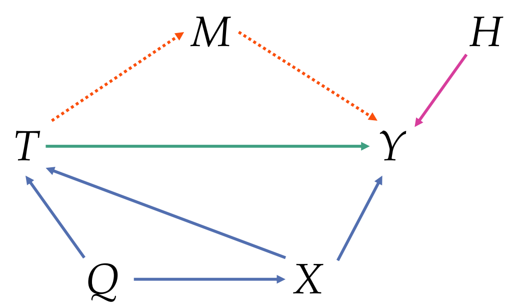

Figure 1: A representative DAG. Some variables—and the causal links between them—may not be observable, and these can be emphasized with dashed arrows. The use of color directs focus towards certain links of interest like $T \to Y$, although this is merely a stylistic preference.

variables. Regression will automatically isolate the causal relationship between T and Y from the variation induced by changes in X or Q. In the language of DAGs, this "closes" the non-causal paths between T and Y. The DAG does not keep track of the paths that are closed by a researcher's analysis—or the paths that remain "open"—it only signals whether a path is a problem that an identification strategy needs to address.

Sometimes a confounding variable like Q isn't observable, but the DAG above suggests that it does not need to be included in a regression if data on X is available. This is because of a restriction imposed on the model via the term "acyclic." Disallowing loops enables the endogenous variables in a model to be determined recursively, which in turn allows a DAG to yield statements about conditional independence, e.g. "Y is independent of Q conditional on X." Since variation in Q can only impact Y through an intermediate effect on X, controlling for the variation in X also halts the pass-through of information from Q to Y. Thus the requirement for closing a spurious path between T and Y can be relaxed to either controlling for a confounding variable—or one of its "descendants"—along that path. Here, once X is controlled for, the status of Q becomes immaterial with respect to the identification of the effect of T on Y.

When reporting regression output, the marginal "effect" on Y attributable to T reflects all remaining open paths from T to Y. The researcher determines whether this variation generates a measurement of a causal effect, and the causal claim relies on an implicit assumption that the DAG accurately models the data generating process. Thus the addition of an explicit model does not remove the potential for misspecification. However, if someone disagrees with a prior, they may make changes to the DAG and review whether the new DAG admits the same identification strategy or a completely different one.

In the present case, a viable control strategy for estimating $T \to Y$ would be to control for X, since X additionally closes the path involving Q. The decision to include or exclude H (or Q) in the regression will not affect the causal interpretation, although including these controls may improve the precision of the $T \to Y$ estimate. The direct effect cannot be disentangled from the indirect effect here, as the variable M is unobservable.

Applications of DAGs to environmental economics

This section demonstrates how to implement DAGs through the use of several examples from environmental economics. Despite their simplicity, the following examples would be fairly difficult to explain without the use of DAGs. This suggests that discussions of research design can be made more clearly and succinctly with the aid of the graphical approach. In this way, DAGs are hardly redundant, but augmentative.

Matching, mechanisms, and over-conditioning

Matching strategies are relatively easy to follow. Isolating an effect of interest involves partitioning observations based on observed characteristics, then comparing treated and control observations within these partitions. In the case that the desired control variables are observable, there is still more nuance to the matching strategy that must be considered. It is fairly straightforward to ascertain the variables that may be correlated to both treatment and outcome, but the *role* of the covariates must also be taken into consideration. This is something that the DAG makes explicit. To illustrate, Figure 2 warns of a potential pitfall involving over-conditioning while addressing the pathways through which an urban tree canopy may improve housing values.

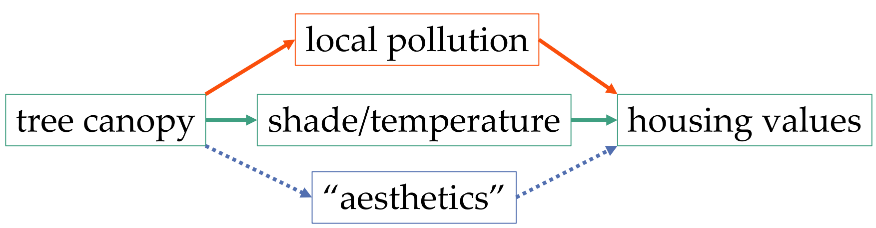

Figure 2: Controlling for a mediator will close a causal path, preventing the estimation of an indirect effect or the inclusion of the indirect effect in a total effect estimate.

The above DAG states that there are three mechanisms through which increasing tree cover conceivably impacts housing values, to wit, trees filter local air and groundwater pollutants, lower energy bills and provide passive cooling, and increase landscape aesthetic. In general, the extent to which a list like this is exhaustive is something that must be assumed by the researcher, as with any structural causal model. For expositional clarity, this model also assumes that there are no confounding factors (e.g. the size of a parcel) between tree canopy and housing values.

In considering the design of a hedonic analysis linking tree cover to housing values, a popular suggestion might be to control for local pollution levels—the logic being that tree cover and pollution are likely correlated, and both things impact housing values. But this is a blunder, as the logic has not determined how the pollution variable enters into the data generating process. Currently, it looks more like a "mediator" than a confounder because it facilitates a causal path from the tree canopy variable to housing values. Controlling for local pollution therefore closes a causal path that should have been left open. In the present setting, an unconditional regression is best for uncovering the tree canopy effect.

Bad controls, and the collider variable

Some research design strategies can inadvertently introduce bias to estimators. To illustrate, the following example concerns the long run impact of cumulative wildfire smoke exposure on respiratory health using hypothetical data from hospital admissions. Figure 3 provides a model of a data generating process that takes into account how the sample is being collected. Because smoke exposure also has negative long run impacts on immune system health, it increases the likelihood of a hospital visit through a second causal channel that is unrelated to respiratory health.

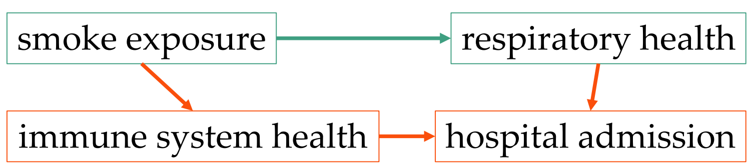

Figure 3: Collecting data using a selected sub-sample will not only result in a lack of external validity, but a lack of internal validity as well, due to collider bias.

Certain variables create spurious correlations between treatment and outcome only once they are included as controls—or in this case, baked into the data collection. In the present example, weakened immune system health and weakened respiratory health are both sufficient conditions for being more likely to showing up in the hospital data. But the sufficiency of either condition implies that among hospital admits, those with poorer immune system health are more likely to have better respiratory health (and vice versa). This is not a causal claim, but a spurious correlation created by the data generating process. Since people impacted by wildfires are more likely to have worsened immune system health than the unaffected, the spurious link will contribute to an under-estimation of the negative respiratory health effect.

DAGs alert researchers to potential identification pitfalls that would be difficult to explain without a graphical aid. Whenever a variable invites a collision of two arrows, the variable is called a "collider." Stratifying on a collider variable—through inclusion in the control set or the data collection process—will open an otherwise closed path. The abstraction of the identification problem to a graphical one makes the detection and discussion of a collider bias simple. Without the concept of a collider, these biases are easily missed.

A common mistake is to reduce this example to an external validity concern. However, the measured relationship between smoke exposure and respiratory health won't even be *internally*-valid here. The collider bias will reduce the measured smoke exposure effect for the hospitalized sub-sample. However, since the second mechanism through which smoke exposure increases the likelihood of hospital admission is known, an identification strategy that conditioned on immune system health would close the non-causal path from smoke exposure to respiratory health, even when the collider is in play. This allows the estimation of an unbiased respiratory health effect for the hospitalized population.

The robustness of an identification strategy

Does adopting an electric vehicle decrease household emissions? The direct effect is likely positive due to increased mineral extraction and fossil fuel-derived electricity demand. However, the bulk of the adoption effect is likely indirect and negative, through the replacement of a gas-powered car. As a straw-man, consider the economist who controls for the number of gas-powered cars in the household. This would clearly be a mistake, as they have closed a causal path. By controlling for the number of gas-powered cars, they removed the replacement channel.

But this economist wouldn't even measure the direct effect with their strategy, given the understanding of car buying preferences implied by Figure 4. Other relevant household characteristics will manifest a collider bias when controlling for the number of gas-powered cars. For example, households with car enthusiasts are drawn to gas-powered cars and driving more often, both for the sake of leisure. The control variable is a collider on the enthusiast path—conditioning on the number of gas-powered cars introduces a spurious correlation between electric vehicle adoption and enthusiasm.

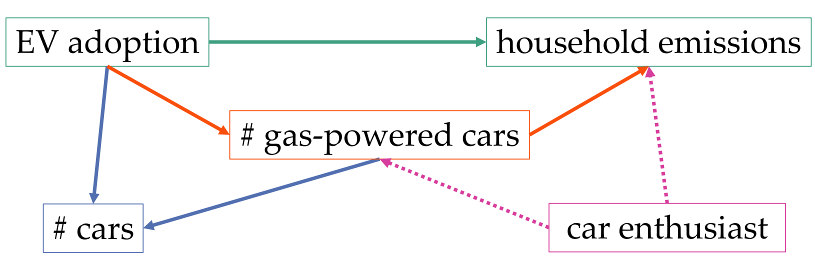

Figure 4: The true impact of electric vehicle adoption on household emissions is only revealed if controls for the number of cars in the household—gas-powered or otherwise—are omitted.

To many environmental economists, the replacement story is probably obvious, although the collider story is likely not. But the recommended solution might be to control for the *total* number of cars instead. According to Figure 4, this is also a blunder. In that causal model, the total number of cars acts as another collider. By stratifying on the number of cars in the household, it is assumed that there is a one-to-one replacement of a gas-powered car for every electric vehicle adopted—which is unlikely. Some households may move from two to three total cars with their adoption, but they will be compared to three [gas-powered] car households instead of two car households. The proposed solution moves the result from an underestimate to an overestimate of the decrease in emissions.

This model says that controls for the number of gas-powered cars and the number of total cars in the household should both be avoided. A different data source or setting may imply a different causal structure than the one in Figure 4. In that case, the no-control strategy can be re-validated by checking the new DAG for any threats to identification.

Efficient communication of identification assumptions

Fishing quota systems have revolutionized the organization of fisheries, ensuring the biological sustainability of the targeted species and maximizing the value of an authorized total allowable catch. In these systems, fishers explicitly own shares of this total allowable catch and are free to trade these fishing rights among themselves. The market for quota shares autonomously guides ownership into the hands of the most efficient fishers, therefore increasing the economic rents that accrue to each share.

How much of the variation in the price of quota is due to changes in fleet efficiency (i.e. catch per unit effort), as opposed to changes in the demand for fish? A potential problem with separating these two impacts is the fact that any market development that increases the value of the quota may additionally incentivize investment that leads to greater efficiency. Figure 5 shows how uncontrolled demand side factors create spurious relationships when interpreting the effect of changes in fleet efficiency on the price of quota.

Consider a potential regulation enacted due to concerns about the bycatch of an endangered species often found co-mingling with the targeted species. This could manifest as a restriction on a certain type of gear or fishing location. The restriction reduces the efficiency of the fleet, impacts the value of fishing quota through this efficiency mechanism, and is plausibly exogenous with respect to shifts in demand.

Figure 5 is consistent with this fishery story, and it suggests using the gear restriction variable as a solution to the identification problem. The regulation turns fleet efficiency into a collider variable along the spurious path from gear restriction to the quota price, and it acts as a mediator on the causal path. Thus an unconditional regression of the

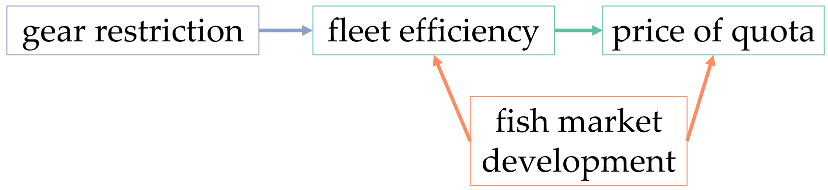

Figure 5: Given the mental model above, the corresponding DAG reflects the assumptions required for the gear restriction variable to be a useful instrument.

price of quota on the gear restriction variable provides a causally-interpretable estimate. A regression of fleet efficiency on the gear restriction is similarly unbiased. Dividing the two relevant regression coefficients yields the instrumental variables estimator for the effect of fleet efficiency on the price of quota—another causally-interpretable estimate.

The DAG above reveals the familiar identification assumptions for the instrumental variables approach—"relevancy, independence, and validity"—by checking for the presence or absence of three key arrows involving the gear restriction variable. The fishery story admits an arrow from gear restriction to fleet efficiency (relevance—though the strength of the instrument must be checked separately), but not to or from the price of quota or fish market development variables (validity and independence, respectively). Figure 5 transparently communicates what is believed to be true about the instrument, and readers can more easily follow (or dispute) these assumptions.

Discovering new methods

The previous section used a DAG to identify multiple links between variables that could be estimated without bias. The resulting strategy emulated the familiar instrumental variables approach. But each new DAG structure provides an opportunity to discover a novel identification strategy that would have otherwise remained hidden. This last example illustrates the "random filter" identification strategy (Donovan, 2023) for deriving an unbiased estimate of a treatment effect amid selection into treatment.

How beneficial are climate resilience-motivated crop insurance programs in developing nations? The more savvy farmers will likely have the most interest in an insurance program if made available, as well as the most sophisticated farming operations. Any naïve regression strategy would clearly pick up this selection bias. Figure 6 shows how to remedy this selection problem. Consider the mechanism through which insurance would lead to a benefit. If there is no adverse weather event (and thus no crop damage/insurance claim), then crop insurance will have no positive impact on income. If this crop damage variable isn't correlated with business savvy, the two links in the chain from crop insurance to income can be estimated separately, without bias.

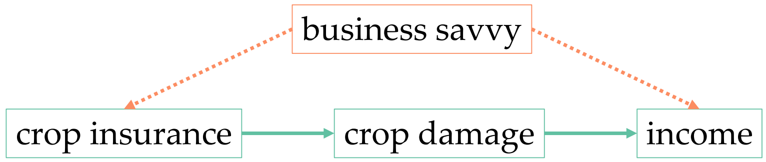

Figure 6: When an exogenous mediator exists between treatment and outcome, a researcher can separately identify the effects of treatment on the mediator and mediator on the outcome, then chain these effects together.

The first link in the causal chain can be identified with a simple regression of crop damage on crop insurance. The spurious path from crop insurance to crop damage is not an issue, since income assumes the role of a collider and omitting income closes this path. The second link is identified by regressing income on crop damage while controlling for crop insurance to close the other non-causal path. Multiplying these impacts together recovers a causally-interpretable estimate of the effect of crop insurance on income.

Figure 6 states three assumptions about this mediator. First, the mediator must intercept all causal paths from crop insurance to income. If another causal path existed outside of the mediator's reach, the identification strategy will close that path. The other two assumptions require the crop damage variable to have unconfounded relationships with crop insurance and income. This means the business savvy variable—or any other unobservable—may not create a non-causal path between treatment and outcome that involves the mediator. If these assumptions hold, i.e. if one can genuinely draw the data generating process like Figure 6, then the identification strategy above is valid.

The savviest of farmers could potentially mitigate their exposure to weather-related damages through siting or some other mechanism; this creates a confounding path involving the mediator. In this case, a modification to the DAG presents a more robust identification strategy. Conditional on the known risk of a disastrous weather event, the event itself is now plausibly-exogenous with respect to the business savvy of individual farmers. Adding a measurable control variable like "risk of damage" in between the business savvy and crop damage variables would signal that a control strategy is available to close the spurious path. A researcher should run the two aforementioned regressions while controlling for exposure risk.

An enhanced review of a recent JAERE article

In early July, 2015, the water utility serving the city of Burbank, California sent warnings to households who would be violating upcoming summer irrigation restrictions based on their behavior in late-June. The notices contained details about a new monitoring system that used real-time data to automatically inform the utility of irrigation outside of permitted times. This treatment resulted in a tenfold increase in the number of households ever notified of an irrigation violation—and signaled a change in the likelihood of enforcement and threat of financial penalties should overuse continue.

West, Fairlie, Pratt, and Rose (2021) reported a substantial decrease in water consumption in response to receiving a notice—roughly 600 gallons per week, or 31% of mean household use—using a fuzzy regression discontinuity design. Treatment was predominantly determined by whether a household was found to be non-compliant with upcoming summer restrictions during a week in late-June; the noncompliance algorithm counted the number of days with peak hourly water usage exceeding an arbitrarily-chosen threshold (by which point irrigation was evident). Since the summer restrictions only allowed for irrigation two days per week, when the third-highest daily peak consumption hour exceeded 125 gallons, a household would be deemed non-compliant.

The running variable of interest is household water use during the third-highest daily peak consumption hour. The utility allowed for some leniency for historically efficient customers, so crossing the above threshold only implied a jump in the *probability* of receiving a warning. The lower-volume customers were instead evaluated based on their fifth-highest daily peak consumption hour, but this "consumption tier" information was not available to the research team. The imperfect compliance in treatment assignment using the stricter rule supported a fuzzy regression discontinuity approach.

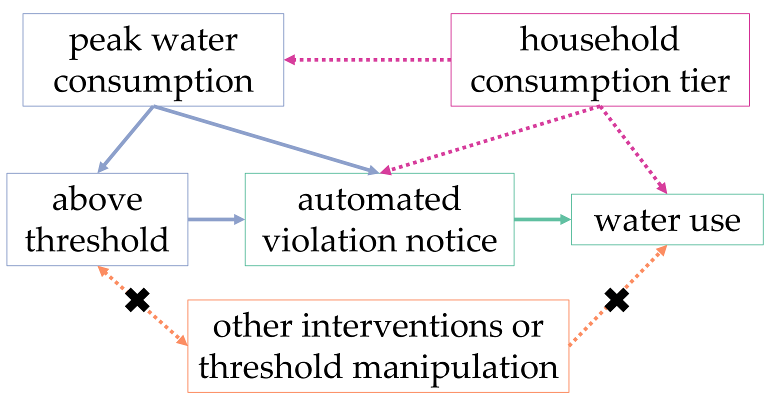

Figure 7: Sketching the DAG consistent with an article's narrative highlights key assumptions, barriers to identification, sources of potential bias, and the validity of the identification method.

The nature of the omitted variable bias becomes clearer with the DAG in Figure 7—drawn to represent the authors' understanding of the data generating process. Those who use less water historically would likely continue to do so and not receive a notice, while the higher baseline households would be treated. The unbalanced treatment and control groups biases the magnitude of the treatment effect downward.

Figure 7 aids a reviewer in running through the requirements for a satisfactory fuzzy regression discontinuity design. The previously mentioned instrumental variables assumptions are conditionally satisfied. Clearly, the instrument—indicating if the stricter peak consumption threshold is exceeded—partly determines whether a household receives an irrigation restriction violation notice (it was a necessary precondition). And because the crossing of the peak consumption threshold is defined entirely by a household's peak water consumption, conditional on the running variable, the instrument is independent of the unobserved confounder—the household consumption tier. Lastly, the instrument only impacts household water use through the delivery of a notice.

There are still the assumptions specific to the regression discontinuity approach to discuss. First, the goal of this regression discontinuity design is to separate the drop in water use due to receiving a notice from the [increasing] water use trend in the peak consumption running variable. Upon controlling for the running variable, any paths that utilize continuous variation in water use across the threshold are closed, leaving only the potential discontinuity to be measured.

Second, there shouldn't be an additional cause for a discrete change in water use when comparing households just over the threshold to those just below it. For expositional clarity, the "local randomization" assumption is symbolized here by a non-path from the threshold to the outcome using the bold 'X's in Figure 7. Because water use will be regressed on the peak consumption threshold—and not the notice status itself—the authors hope to attribute the threshold-derived jump in outcomes to the notice, rather than some other factor. They make this assertion easy to believe, as no concurrent intervention utilizing the threshold existed (eliminating the arrow from the threshold to a potential intervention mediator), and no households were privy to the threshold or the automated detection algorithm (eliminating the arrow from a potential sorting confounder to the threshold).

From this discussion, the authors' fuzzy regression discontinuity approach appears to generate an unbiased estimate of the effect of the irrigation restriction violation notices on summertime water use. But the DAG cannot fully cover the scope of an empirical review. Much of this is because a DAG is a non-parametric representation of a data generating process. Figure 7 cannot suggest a particular implementation of fuzzy regression discontinuity (two stage least squares, local linear regression, bandwidth, etc.), and these decisions involve a separate set of assumptions from the causal ones shown by the DAG. Importantly, the sensitivity of results to these additional assumptions should be analyzed across specifications derived from the same causal model.

The other shortcoming is that a DAG may miss some nuance relating to the estimate of interest. In the [common] case of heterogeneous treatment effects, an instrumental variables strategy typically makes use of a monotonicity assumption (trivially met here) to show that an estimate is a local average treatment effect (presently equivalent to an average treatment effect on the treated). The additional reliance on the extrapolation of counterfactuals in regression discontinuity means that this treatment effect is also limited in scope to settings where the peak consumption threshold for notice eligibility is around 125 gallons. The traditional counterfactual logic in these statements does not map to any feature in the graph.

Opportunities for implementation

Economics is one of the rare disciplines that uses presentations to improve work in progress—rather than report on finished products. But it is crucial to remember that identification strategies are nuanced and borne out from a place of deep familiarity with specific data and settings. The inclusion of a DAG increases the benefits of presenting research because it makes these things clear to others. This leads to higher-quality feedback. It is easy to evaluate whether a speaker's DAG represents the data generating process in their setting, turn any discrepancy into a question for the presenter, and test if a variation still admits the use of the proposed identification strategy.

DAGs provide an opportunity to make the research process more transparent and responsive to criticism. When reviewers ask for some sort of sensitivity analysis, they can use the graphical language for raising concerns about causal hypotheses. A new causal hypothesis generates a modification to a DAG, which then organically motivates a new regression. This first-principles style approach to testing model robustness is more academically rigorous than suggesting modifications to a regression with language concerning functional form alone. It matches the spirit of specification testing, which aims to build confidence that an effect of interest is being measured without bias.

The integration of DAGs in econometrics education creates tremendous value for pedagogy (e.g. Cunningham (2021); Huntington-Klein (2022a)). Graduate studies in economics today already highlight causal effect identification, yet many of the insights shared in this article are absent from applied econometrics instruction. Integration creates even more value at the undergraduate level; for example, the introductory econometrics course at the University of Nevada, Reno attracts a diverse group of students to economics through its primary emphasis on the causal storytelling skills for which econometricians are known. The use of DAGs only increases access to this class and further facilitates the building of a desirable competency.

This accessibility can be extended to communication with policymakers and stakeholders. By simplifying causal effect identification while retaining the rigor of an explanation, DAGs may increase the value of economics research for environmental policy.

References

Angrist, J. D. and Pischke, J.-S. (2010). The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics. *Journal of Economic Perspectives*, 24(2):3–30.

Cunningham, S. (2021). *Causal Inference: The Mixtape*. Yale University Press.

Donovan, P. (2023). A random filter identification strategy for addressing selection into treatment. *Working Paper*.

Heckman, J. J. and Pinto, R. (2022). The Econometric Model for Causal Policy Analysis. *Annual Review of Economics*, 14(1):893–923.

Huntington-Klein, N. (2022a). *The effect: an introduction to research design and causality*. CRC Press, Taylor & Francis Group.

Huntington-Klein, N. (2022b). Pearl before economists: the book of why and empirical economics. *Journal of Economic Methodology*, 29(4):326–334.

Imbens, G. W. (2020). Potential Outcome and Directed Acyclic Graph Approaches to Causality: Relevance for Empirical Practice in Economics. *Journal of Economic Literature*, 58(4):1129–1179.

Pearl, J. (1995). Causal diagrams for empirical research. *Biometrika*, 82(4):669–688.

Schneider, E. B. (2020). Collider bias in economic history research. *Explorations in Economic History*, 78:101356.

Segerson, K. (2019). Reflections—On the Role of Theory in Contemporary Environmental and Natural Resource Economics. *Review of Environmental Economics and Policy*, 13(1):124–129.

West, J., Fairlie, R. W., Pratt, B., and Rose, L. (2021). Automated Enforcement of Irrigation Regulations and Social Pressure for Water Conservation. *Journal of the Association of Environmental and Resource Economists*, 8(6):1179–1207.