Geophysical Research Letters

RESEARCH LETTER
10.1029/2019GL086480

Key Points:
- We estimate that annual floods increase by 3.3%, on average, for each percentage point increase in impervious basin cover
- This is the first study to apply a panel regression design to estimate the causal effect of impervious cover on floods
- Our approach demonstrates how to leverage temporal and spatial variation to isolate a causal effect, separate from other drivers of change

Supporting Information:
- Supporting Information S1

Correspondence to:
A. G. Blum, annaliseblum@gmail.com

Citation:
Blum, A. G., Ferraro, P. J., Archfield, S. A., & Ryberg, K. R. (2020). Causal effect of impervious cover on annual flood magnitude for the United States. Geophysical Research Letters, 47, e2019GL086480. https://doi.org/10.1029/2019GL086480

Abstract
Despite consensus that impervious surfaces increase flooding, the magnitude of the increase remains uncertain. This uncertainty largely stems from the challenge of isolating the effect of changes in impervious cover separate from other factors that also affect flooding. To control for these factors, prior study designs rely on either temporal or spatial variation in impervious cover. We leverage both temporal and spatial variation in a panel data regression design to isolate the effect of impervious cover on floods. With 39 years of data from 280 U.S. streamgages, we estimate that a one percentage point increase in impervious basin cover causes a 3.3% increase in annual flood magnitude (95% CI: 1.9%, 4.7%) on average. Using 2,109 streamgages, some of which have upstream regulation and/or overlapping basins, we estimate a larger effect: 4.6% (CI: 3.5%, 5.6%). The approach introduced here can be extended to estimate the causal effects of other drivers of hydrologic change.

1. Introduction
Humans have modified nearly all of the world’s rivers (Vörösmarty et al., 2010), leading to changes in water availability, floods, and droughts. To predict and manage risks associated with these changes, we need to quantify the causal links between human activities and hydrological patterns. The effect on floods of impervious surfaces, such as roads, roofs, and parking lots, is of particular interest, as demonstrated by a large body of research spanning the last half-century (Hollis, 1975; Leopold, 1968; Oudin et al., 2018; Sheng & Wilson, 2009; Shuster et al., 2005).

Despite a general consensus that growth in impervious cover reduces infiltration and increases runoff and flood magnitudes (Villarini & Slater, 2017), estimated effect sizes range from zero (Dudley et al., 2001; Wilben, 1976; Wright et al., 2013) to increases of up to a factor of 10 (Hollis, 1975; Leopold, 1968). Most studies estimate a range of effects falling between these extremes (Debbage & Shepherd, 2018; Hawley & Bledsoe, 2011; Hejazi & Markus, 2009; Jones & Grant, 1996; Ogden et al., 2011; Salavati et al., 2016; White & Greer, 2006). (See Supporting Information Table S1.)

Quantifying the effect of changes in impervious cover on floods is difficult because many rivers are responding to concurrent changes in climate, irrigation practices, land use, water withdrawals and transfers, and regulation of flows at dams (Allaire et al., 2015; Hejazi & Markus, 2009; Viglione et al., 2016). As a result, many confounding variables exist that may mask or mimic a causal relationship between impervious cover and flooding. Furthermore, flood series often have large natural variability, due to cycles of wet or dry years, and a low signal-to-noise ratio (Cohn & Lins, 2005; Hirsch, 2011; Merz et al., 2012), making statistical inference from observable data a challenge.

As noted in Villarini and Slater (2017), studies of impervious effects on floods generally follow one of two designs: (1) case studies of a single river basin using temporal data (Dudley et al., 2001; Sheng & Wilson, 2009; Smith et al., 2002; White & Greer, 2006) or (2) paired catchment studies comparing an urbanized basin to a similar basin with less impervious cover (Changnon & Demissie, 1996; Jones & Grant, 1996; Prosdocimi et al., 2015; Salavati et al., 2016). A third approach using simulation is also common (Du et al., 2012; Harrigan et al., 2014; Hejazi & Markus, 2009; Ogden et al., 2011).

The case study design relies on variation in impervious cover across time: Flooding in an early period, with minimal impervious cover, is compared to flooding in a later period, with increased impervious cover. This
design thus relies on the assumption that the early period’s flood pattern represents the counterfactual pattern of the later period, had impervious cover not increased. If, for example, precipitation increased at the same time that impervious cover increased, and the effects of precipitation are not accounted for, the estimator of the effect of impervious cover on flooding is confounded (in this case, biased upward).

In contrast, the paired catchment design relies on variation in impervious cover across space. In comparing flooding in an urbanized river basin to flooding in another basin with minimal impervious cover but otherwise similar characteristics, this design makes the assumption that the minimally impacted basin’s flood pattern represents the counterfactual pattern of the urbanized basin, had impervious cover not increased. Identifying a comparison catchment can be a challenge because large unimpaired catchments are uncommon, and impaired and unimpaired catchments often differ across other features that also affect flooding. Limited samples sizes mean that effects estimated from paired catchment designs can be sensitive to small changes in the data (Prosdocimi et al., 2015).

A third common approach is to run multiple simulations with differing levels of impervious cover and compare predicted flood magnitudes (Du et al., 2012; Harrigan et al., 2014). This approach, however, cannot estimate a causal effect of impervious cover on floods as findings reflect assumptions made about the model structure and the effect of impervious cover. Such simulation approaches are also limited by small samples, uncertainties associated with model formulation, parameterization, selection of a baseline period, and data (Gyawali et al., 2015; B. Merz et al., 2012; Zhang et al., 2018).

Here we introduce a panel regression approach that improves upon previously used designs by leveraging variation in observed data across both time and space. Temporal variation comes from using a time series at a given stream gage. Spatial variation comes from using data across different basins in the same year. Borrowed from econometrics, the design focuses on isolating and quantifying the causal effect of a single driver—impervious cover—by controlling for other drivers of change (confounders). Recent reviews detail a variety of ways econometric approaches can improve our understanding of hydrologic change (Müller & Levy, 2019) and of coupled human-natural systems more broadly (Ferraro et al., 2019).

The panel regression design has several advantages. First, by exploiting repeated observations of annual floods and impervious cover across many basins, the approach reduces bias due to time-varying factors (e.g., regional weather events) and time invariant, but spatially varying factors (e.g., topography) that might confound identification of the effect of impervious cover on floods. In addition, the design eliminates the need to set subjective “pre” and “post” periods. Finally, it is unnecessary to identify a single reference catchment to serve as a counterfactual because the panel regression compares all basins and time periods, capturing the relevant variation in time and space.

While a few hydrology studies have used panel regression designs, none, to our knowledge, have applied the design to estimate a causal effect of impervious cover on streamflow or floods. Previously, panel approaches have been used to estimate the effect of forest cover on flood frequency (Ferreira & Ghimire, 2012), deforestation on streamflow (Levy et al., 2018), rainfall on low streamflow (Bassiouni et al., 2016), and urbanization on annual runoff coefficients (Steinschneider et al., 2013). The only study that has applied panel models in the context of urbanization and flooding is Over et al. (2016). However, rather than estimating a causal effect of urbanization on floods, the study applies fitted parameters to adjust the historical flood records for northeastern Illinois.

We focus on the effect of impervious cover on annual flood magnitude, defined as the largest streamflow in a given year. Annual flood magnitudes are one of the most common metrics of flooding and are used in flood frequency analysis (Hodgkins et al., 2019; R. Merz & Blöschl, 2008; Vogel et al., 2011). Impervious cover has been reported to have a larger effect on frequent floods, such as the annual flood, compared to infrequent extreme flood events, such as the 100-year flood. This difference has been attributed to soil behaving similarly to an impervious surface after prolonged rainfall and saturation (Hawley & Bledsoe, 2011; Hollis, 1975; Sauer et al., 1983). In a study of national flood trends, Hodgkins et al. (2019) found that changes in the magnitude of annual floods were consistent with change in the number of peaks-over-threshold, another commonly used metric of flooding.

In addition to our panel regression design, this study has two other novel attributes. First, we estimate an average causal effect of changes in impervious cover on annual flood magnitude for a large sample from
the whole United States, rather than a single basin or small set of basins. Second, we show how a causal graph can be useful for identifying causal relationships, assessing sources of bias, and developing an appropriate regression model structure.

2. Panel Regression Design for Causal Inference

2.1. Causal Graph

Causal graphs (or causal diagrams) are a tool to help identify the sources of variation in a causal variable and its outcome, thereby highlighting potential sources of bias (Morgan & Winship, 2015). Used across a range of fields, these graphs summarize our knowledge and communicate our assumptions about a hypothesized causal relationship (Pearl & Mackenzie, 2018). By laying out these assumptions, causal graphs help us avoid mistaking correlation for causation, as well as inform design of the most appropriate model structure for causal inferences. We interpret estimated effects as causal when we have controlled for all plausible confounding variables. Although uncommon to hydrology, a causal graph was previously used to illustrate the relationship between changes in vegetation and streamflow (Levy et al., 2018).

The causal graph in Figure 1a illustrates a hypothesized causal effect of changes in impervious cover on changes in annual flood magnitude. Confounders are variables that affect both the quantity of impervious cover and the magnitude of annual floods and thus can mask or mimic a causal relationship between the two. For example, upstream regulation of streamflow has been reported to reduce flood magnitude (FitzHugh & Vogel, 2010). By protecting downstream areas from floods, upstream regulation may also increase development and thus impervious cover. Regulation of flows is thus a potential confounding variable because it can affect both floods and impervious cover. The causal graph also illustrates variables that are not confounders. For example, changes in peak precipitation or precipitation intensity are likely to affect annual floods. However, unlike regulation, they are not confounders because they do not also affect changes in impervious cover.

In addition, the causal graph includes two other factors: moderators, which are off the causal path and moderate the magnitude of the causal effect; and mechanisms, which lie on the causal path and represent the intermediate outcomes through which impervious cover affects flooding. Interacting the causal variable with moderators can shed light on the heterogeneity of the causal effect. Controlling for mechanisms would be inappropriate because it would mask part or all of the effect of impervious cover on flooding. For example, some studies have found that urbanization increases flooding not only by reducing infiltration rates, but also by increasing total storm rainfall (Zhang et al., 2018). This suggests that precipitation would be considered a mechanism; thus including precipitation in the model could mask part of the effect of impervious cover on flood change and bias our estimated effect.

2.2. Modeling Approach

Based on insights from the causal graph (Figure 1a), we developed a panel regression model to eliminate the effects of possible confounding variables that could bias our estimated effect of impervious cover on annual flood magnitude (Figure 1b). In contrast to prior studies that control for only specific confounding variables, our design controls for all possible confounders that are time invariant at the basin level or time varying at the national or regional level.

First, intercept terms ($\alpha_i$) specific to each streamgage $i$ control for any streamgage-specific characteristics that are constant over time and may be associated with both impervious cover and flood magnitude (e.g., elevation). Representing average annual flood magnitude across 1974–2012 for a given streamgage, these intercepts would be referred to as “fixed effects” in econometric literature. Here, the term “streamgage-specific” is equivalent to “basin-specific” because each streamgage measures streamflow from one basin.

Our panel design also controls for time-varying confounders at both national and regional scales, accounting for temporal trends in floods due to changes other than impervious cover. For example, if water diversions increased across the entire country during the study, this increase would be considered a national time-varying confounder. To control for time varying, nationwide confounders, we use annual dummy variables ($D_t$) that take the value of 1 to indicate a given year, and 0 otherwise. For time-varying regional confounders, we also include regional annual dummy variables ($D_{rt}$). Regions were defined as the physiographic regions of the United States (Figure 2), which have similar topography, rock structure, and geologic and geomorphic
For example, if peak precipitation increased in region \( r \) in year \( t \), the coefficient \( \gamma_{t,r} \) would reflect and control for that increase.

Although our study design controls for a wide range of confounding variables, an unobserved time-varying confounder at the subregional level could bias our inferences. This is because this “fixed effects” model relies on the assumption that there are no time-varying variables in \( \varepsilon_{i,t} \) that are also correlated with impervious cover (and we have already controlled for national and regional-level confounders). Fixed effects panel regression models are similar to random effects models (Verbeek, 2003), but differ in one key way: Fixed effects models permit unbiased estimation of the causal relationship (\( \beta \)) even when the streamgage-intercept terms (\( \alpha_i \)) are correlated with the error term \( \varepsilon_{i,t} \) (i.e., even when time-invariant attributes of the basin are correlated with both flooding and impervious cover). In other words, streamgage-specific intercepts are a parameter to be estimated in fixed effect models, whereas, these intercepts are part of the error term in random effects models. For more information, see Gelman (2005, p. 20), Verbeek (2003, pp 356–359), and the Supporting Information Text S1.

We fit models with the plm package (Croissant & Millo, 2008) in R (R Core Team, 2019). Because model errors at a given streamgage over time (\( \varepsilon_{i,t} \)) are likely to be serially correlated, we cluster the standard errors at the streamgage-level (Bertrand et al., 2004). Our clustered estimation of the variance allows for arbitrary serial correlation within each streamgage, as well as heteroskedasticity across streamgages (Wooldridge, 2002). See Supporting Information Text S1 for code. Finally, because we are focused on identifying a causal effect, rather than predicting future flood flows, the causal diagram is used to determine the correct model structure. Model validation methods typically used for predictive models, such as goodness of fit, are not appropriate for models aimed at causal inference because the goal is not prediction (Ferraro et al., 2019).

Figure 1. (a) Causal graph illustrating the relationship between impervious cover and changes in annual flood magnitude. Arrows illustrate causal relationships, not physical pathways. (b) Panel data regression estimator used to estimate the average causal effect of impervious cover on annual flood magnitude. Regions are the 24 physiographic regions illustrated in Figure 2.
As in previous studies, we use a logarithmically linear model because annual flood values are strongly skewed (Vogel et al., 2011). By taking the natural logarithm of our response variable, we can interpret \( \beta \) as an estimate of the average causal effect of a one percentage point change in basin impervious cover on annual flood magnitude. Changes in impervious cover are reported in percentage points because they

Figure 2. (a) A map illustrating the location of each of the 2,109 streamgages and the physiographic regions defined by Fenneman and Johnson (1946). Colors indicate streamgages with significant trends in annual flood magnitude for 1974–2012 (Mann-Kendall test \( p < 0.1 \)). A larger circle indicates a basin increase in impervious cover by more than one percentage point 1974–2012. (b) Boxplots illustrating annual flood magnitude and drainage areas for the 2,109 streamgages for 1974–2012. (c) Same as (a) but for the 280 independent, unregulated streamgages. (d) Same as (b) but for the set of 280 streamgages.
reflect a difference in percentages. For example, an increase from 5% impervious cover to 10% impervious cover is a five percentage point increase, not a 5% increase. The units of annual flood magnitude are m$^3$/s (average for the day). Thus, our estimated effect ($\beta$) is reported as a percent change in flood magnitude from a one percentage point change in impervious cover.

We expect the effect of impervious cover on flooding to vary based on certain basin characteristics ("Moderators" in Figure 1a). High soil permeability and low basin slope have been reported to serve as a buffer against hydrologic changes like increased impervious cover (Hopkins et al., 2015). To test whether those attributes moderate the effect of impervious cover on flooding in our sample, we reestimate the model with interaction terms between impervious cover and soil permeability and between impervious cover and slope. We then test the null hypotheses that the estimated coefficient on each interaction term is equal to zero. Note that it is possible that these moderators could also be confounders. However, because they are both time-invariant characteristics of basins, their confounding effects are controlled for through the streamgage-specific intercept terms.

We also check for a nonlinear relationship (quadratic and cubic) between impervious cover and floods. In the region of Chicago, Illinois, Allen and Bejcek (1979) reported a decreasing effect of impervious cover on floods while Over et al. (2016) reported an increasing effect. Thus, we do not have a clear hypothesis about the sign of a nonlinear relationship. See Supporting Information for equations.

3. Development of Panel Data Set

We created a spatial and temporal data set that includes annual flood magnitude and annual percent impervious cover in the upstream basin. The time series of percent impervious cover in each basin came from the enhanced Geospatial Attributes of Gages for Evaluating Streamflow (GAGES) II data set based on the National Water-Quality Assessment Wall-to-Wall Anthropogenic Land use Trends data set (Falcone, 2017). These data were available for the years 1974, 1982, 1992, 2002, and 2012. During this period, impervious cover at a given basin either stayed constant or increased, but never decreased. This pattern allowed us to use linear interpolation to estimate impervious cover between the 5 years of observations, resulting in 39 years of data. As a robustness check, we also refit our model with no interpolated data using only the 5 years with observed impervious cover and under the more conservative assumption of a stepwise increase in impervious cover.

From Dudley et al. (2018), we obtained annual maximum streamflows at 2,683 US Geological Survey (USGS) streamgages in the conterminous United States. Given that streamgage siting and funding may be prioritized for locations with more human alteration, a design that includes incomplete streamgage records could result in a biased estimator. We thus selected the 2,109 streamgages with annual maximum streamflow available for every year from 1974 through 2012 (Figures 2a and 2b). To eliminate possible confounding, we also selected the 280 streamgages (Figure 2c) that were unaffected by upstream storage or overlapping basins for our main analysis. Upstream storage was calculated by dividing total upstream dam storage by annual average inflow (Falcone, 2011; Hodgkins et al., 2019) and streamgages with more than 1 day of upstream storage eliminated. To eliminate overlapping basins, we included only the most upstream basins. We use this subset of 280 streamgages for our main analysis and the larger data set as a robustness check. All basin characteristics, including storage and annual average inflow, were obtained from the GAGES II data set (Falcone, 2011).

For the main analysis of 280 study streamgages during 1974–2012, flood magnitude ranged from 0.01 to 6,513 m$^3$/s and basin sizes from 5–1,910 km$^2$ (Figure 2d). Percent impervious cover ranged from 0 to nearly 50%, but most basins had small amounts of impervious cover, with a median of less than 0.44%. A summary of study streamgage and basin characteristics, including data used to estimate the regression model, are provided in the Supporting Information Table S2.

Consistent with previous studies (Archfield et al., 2016; Hirsch & Ryberg, 2012; Slater & Villarini, 2016), we do not see a cohesive spatial pattern in trends in annual floods for the United States. Results from the non-parametric Mann-Kendall test ($p < 0.1$) indicate a range of positive and negative trends in annual floods scattered across the country. Among the streamgages in Figure 2c, increasing annual floods (positive trends) were detected at 12 basins overall, 3 of which experienced greater than 1 percentage point increase in
impervious cover during 1974–2012. However, at most streamgages, no trend was detected. Although we expect increases in impervious cover to increase flood magnitudes, it is not surprising that we find relatively few gages with statistically significant positive trends. Many factors cause changes in floods and these factors may serve to confound the relationship between impervious cover and flood magnitudes. Even at streamgages with negative trends in floods, floods may still be larger than they would have been in the absence of increases in impervious cover.

4. Estimated Effect of Impervious Basin Cover on Flood Magnitude

4.1. National Average Estimated Effect

We estimate that a one percentage point increase in impervious basin cover causes a 3.3% increase in annual flood magnitude, on average. The 95% confidence interval (CI) of this estimate is (CI: 1.9–4.7%). This estimated effect size is based on our data set of 280 streamgages without upstream regulation or basin overlap. We also reestimate the model with the expanded data set of 2,109 streamgages to find a larger effect 4.6%, with a CI that overlaps with the CI of the main analysis (CI: 3.7%, 5.4%).

In Figure 3, we contrast our estimated effect from the 280 streamgages with estimates from two national-scale studies that both estimated the effect of impervious cover on floods as well as reported their data. In a review of studies from around the United States, Hollis (1975) reports large, but widely varying, estimated effects. Applying both paired catchment and model residual designs, Salavati et al. (2016) estimated smaller and less variable effects, with some being negative (albeit the confidence intervals cross zero). While these studies were the most appropriate comparison we could find, they do not provide an exact comparison; see Supporting Information for more details. To facilitate comparison, we added a fitted regression line to each prior study’s data in Figure 3. The plot extends up to a 25 percentage point change in impervious cover because this is the largest difference in impervious cover found at a single streamgage within our study.

4.2. No Evidence of Nonlinear Effects or Moderation by Slope or Soil Permeability

Although we expected the effect of impervious cover on flood magnitude to vary by soil permeability and/or basin slope, we do not find evidence to support these moderating effects (Supporting Information Table S3). Neither the coefficient on the interaction term between impervious cover and soil permeability (p value = 0.69) nor the coefficient on the interaction between impervious cover and slope (p value = 0.85) is statistically significant. In the nonlinear version of our main model, we also do not find evidence of nonlinear effects of impervious cover based on the statistical significance of estimated coefficients on the quadratic (p value = 0.93) and cubic (p value = 0.75) impervious cover variables (Supporting Information Text S3 and Table S3).

4.3. Robustness Checks

In our main analysis, we linearly interpolate impervious cover in between five decadal observations. As a robustness check, we reestimate our model using only the 5 years of observed impervious data with no interpolation and find a larger estimated effect: 6.1% (CI: 3.8%, 8.4%). We also reestimate the model assuming impervious cover changes as a step function (i.e., remains constant until the next decadal observation). With that assumption, we estimate a slightly smaller effect size of 2.6% (CI: 1.5%, 3.7%), which is not surprising as we expect this assumption to bias our estimate downward (Supporting Information Table S4).

5. Conclusions and Limitations

To predict and prepare for future floods, we need to understand the causes of observed changes in flooding. Scholars and policymakers have long been interested in the effect of impervious cover on flood magnitude, but estimated effects vary widely. Quantifying this link has been challenging due to small sample sizes and modeling approaches that fail to control for a range of confounding factors. With a large data set of streamgages across the United States, our panel regression design controls for confounding variables by leveraging both temporal and spatial variation in flooding and impervious cover to estimate a causal effect. Greater use of such temporal, spatial, and causal information to inform flood frequency has been called “essential” (Merz & Blöschl, 2008).
There are several limitations to this study, which can be explored in future work. First, although our panel design requires fewer assumptions for causal inference than previous approaches, it is not assumption-free. A causal interpretation of our estimated effect requires one to assume that there are no unobserved, time-varying, subregional variables that systematically affect changes in both impervious cover and flooding. That assumption might not be satisfied if, for example, previous floods affected both future flood magnitudes and future impervious cover. Second, we include a comparison of our main analysis (280 streamgages) to an expanded data set of 2,109 streamgages that are impacted by upstream regulation and/or basin nesting. Accounting for upstream regulation and nesting structure was beyond the scope of the study, but may have biased this estimate. Third, the effects of impervious cover on flooding may vary depending on the placement of impervious cover within the basin (Debbage & Shepherd, 2018; Oudin et al., 2018), but we did not account for the spatial distribution of impervious cover. This was because our goal was to estimate a national-level average effect that could inform policy and management decision-making.

Our finding that annual floods increase by 3.3%, on average, for each percentage point increase in impervious cover, falls within the wide range of previously estimated effects (Figure 3). Although we were not able to detect moderating effects of soil permeability or basin slope, further exploration of basin characteristics which moderate the relationship between impervious cover and floods is warranted. Estimation of the causal effect of impervious cover on the 100-year flood, or on changes in variability of floods (Hecht & Vogel, 2020), would provide a more nuanced understanding of the ways in which impervious cover affects flooding and extremes. Engineering and policy-making communities increasingly seek details on the magnitude of the effects of environmental change on water management and infrastructure. Our study design provides an approach to estimating the causal effect of imperviousness on floods that can be extended to other drivers of environmental change.

Figure 3. Estimated causal effects of impervious cover on annual flood magnitude from three studies: (i) this study; (ii) Hollis (1975); and (iii) Salavati et al. (2016) which applied two approaches, Paired Catchments and Model Residuals. The right axis (Ratio relative to pre-urbanized flood) is the metric used in Hollis (1975) and is defined as

\[
\frac{\text{Peak Discharge after Urbanization}}{\text{Peak Discharge before Urbanization}}
\]

so a doubling of peak discharge, or annual flood, is equivalent to a 100% increase in flood magnitude (Supporting Information Text S2).

References

Allaire, M. C., Vogel, R. M., & Kroll, C. N. (2015). The hydromorphology of an urbanizing watershed using multivariate elasticity. Advances in Water Resources, 86, 147–154. https://doi.org/10.1016/j.advwatres.2015.09.022
Allen, H. E., & Bejeck, R. M. (1979). Effects of urbanization on the magnitude and frequency of floods in north-eastern Illinois. U.S. Geological Survey Water-Resources Investigations Report 79–36. https://doi.org/10.3133/wri7936
Archfield, S. A., Hirsch, R. M., Viglione, A., & Bloschl, G. (2016). Fragmented patterns of flood change across the United States. Water Resources Research, 43, 10,232–10,239. https://doi.org/10.1002/2016WR017936

Acknowledgments

We thank James Falcone, Dave Blodgett, Rich Vogel, Robert Hirsch, Glenn Moglen, Jory Hecht, Ben Hobbs, Marc Müller, Will Farmer, and three anonymous reviewers. A. G. B. is grateful for funding from the Howard & Glenadore Pim Postdoctoral Fellowship. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Data on floods, impervious cover, and basin characteristics can be accessed through the following websites, respectively (https://doi.org/10.5066/P9AEGXY0; https://doi.org/10.5066/F7HQ3XS4; and https://water.usgs.gov/GIS/metadata/usgswrd/XML/gagesII_Sept2011.xml).
Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? Quarterly Journal of Economics, 119, 249–275. Retrieved from https://academic.oup.com/qje/article-abstract/119/1/249/1876608

Changnon, S. A., & Demissie, M. (1996). Detection of changes in streamflow and floods resulting from climate fluctuations and land use–drainage changes. Climatic Change, 32(4), 411–421. https://doi.org/10.1007/BF00403543

Cohn, T. A., & Lina, H. F. (2005). Nature’s style: Naturally trendy. Geophysical Research Letters, 32, 1–5. https://doi.org/10.1029/2005GL024476

Croissant, Y., & Millo, G. (2008). Panel Data Econometrics in R: The plm Package. Journal of Statistical Software, 27(2), 1–43. https://doi.org/10.18637/jss.v027.i02

Debbage, N., & Shepherd, J. M. (2018). The influence of urban development patterns on streamflow characteristics in the Charlanta megaregion. Water Resources Research, 54, 3728–3747. https://doi.org/10.1002/2017WR021594

Du, J., Qian, L., Rui, H., Zuo, T., Zheng, D., Xu, Y., & Xu, C. Y. (2012). Assessing the effects of urbanization on annual runoff and flood events using an integrated hydrological modeling system for Qinhuihai River basin, China. Journal of Hydrology, 464-465, 127–139. https://doi.org/10.1016/j.jhydrol.2012.06.057

Dudley, R., Hodgkins, G., Mann, A., & Chisholm, J. (2001). Evaluation of the effects of development on peak flood discharges in Collyer Brook, Maine. U.S. Geological Survey Water-Resources Investigations Report 2001–4156. https://doi.org/10.3133/wri20014156

Dudley, R. W., Archfield, S. A., Hodgkins, G. A., Renard, B., & Ryberg, K. R. (2018). Basin characteristics and peak-streamflow trends and change-points for 2,683 USGS streamgages in the conterminous U.S. U.S. Geological Survey data release. https://doi.org/10.5066/P9AEGX0Y

Falcone, J. A. (2011). GAGES-II: Geospatial attributes of gages for evaluating streamflow (digital data set). USGS. Retrieved from http://water.usgs.gov/GIS/metadata/usgswrd/XML/gagesII_Sep2011.xml

Falcone, J. A. (2017). U.S. Geological Survey GAGES-II time series data from consistent sources of land use, water use, agriculture, timber activities, dam removals, and other historical anthropogenic influences. https://doi.org/10.1002/PHIQX54

Fenneman, N. M., & Johnson, D. W. (1946). Physiographic divisions of the conterminous U.S. Retrieved from https://water.usgs.gov/GIS/metadata/usgswrd/XML/physiolko.xml

Ferraro, P. J., Sanchirico, J. N., & Smith, M. D. (2019). Causal inference in coupled human and natural systems. Water Resources Research, 55, 1215–1236. https://doi.org/10.1029/2018WR023476

Ferreira, S., & Ghimire, R. (2012). Forest cover, socioeconomic and reported activities, dam removals, and other historical anthropogenic influences. Journal of Hydrology, 415-416, 139–153. https://doi.org/10.1016/j.jhydrol.2012.06.057

FitzHugh, T. W., & Vogel, R. M. (2010). The impact of dams on streamflow in the United States. River Research and Applications, 27, 1192–1215. https://doi.org/10.1002/rra.1417

Gelman, A. (2005). Analysis of variance—Why it is more important than ever. Annals of Statistics, 33, 1–53. https://doi.org/10.1214/009053600000001048

Gaywali, R., Griffiths, V. W., Watkins, D. W., & Fennessey, N. M. (2015). Regional regression models for hydro-climate change impact assessment. Hydrological Processes, 29, 1972–1985. https://doi.org/10.1002/hyp.10312

Harrigan, S., Murphy, C., Hall, J., Wilby, R. L., & Sweeney, J. (2014). Attribution of detected changes in streamflow using multiple working hypotheses. Hydrology and Earth System Sciences, 18, 1935–1952. https://doi.org/10.5194/hess-18-1935-2014

Hayley, R. J., & Bledsoe, B. P. (2011). How do flow peaks and durations change in suburbanizing semi-arid watersheds? A southern California case study. Journal of Hydrology, 405(1–2), 69–82. https://doi.org/10.1016/j.jhydrol.2011.05.011

Hecht, J. S., & Vogel, R. M. (2020). Updating urban design flood discharges for changes in central tendency and variability using regression. Advances in Water Resources, 136, 103484. https://doi.org/10.1016/j.adwatre.2019.103484

Hejazi, M. I., & Markus, M. (2009). Impacts of urbanization and climate variability on floods in northeastern Illinois. Journal of Hydrologic Research, 10, 606–616. https://doi.org/10.1016/SASHE.HE1943-5584.000020

Hirsch, R. M., & Ryberg, K. R. (2012). Has the magnitude of floods across the USA changed with global CO2 levels? Hydrological Sciences Journal, 57(1), 1–9. https://doi.org/10.1080/02626667.2011.621895

Hirsch, R. M. (2011). A perspective on nonstationarity and water management. Journal of the American Water Resources Association, 47, 436–446. https://doi.org/10.1111/j.1752-1688.2011.00539.x

Hodgkins, G. A., Dudley, R. W., Archfield, S. A., & Renard, B. (2019). Effects of climate, regulation, and urbanization on historical flood trends in the United States. Journal of Hydrology, 578(August 2018), 697–709. https://doi.org/10.1016/j.jhydrol.2019.03.102

Hollis, G. E. (1975). The effect of urbanization on floods in different recurrence interval. Water Resources Research, 11(3), 431–435. https://doi.org/10.1029/WR011i03p00431

Hopkins, K. G., Morse, N. B., Bain, D. J., Better, N. D., Grimm, N. B., Morse, J. L., et al. (2015). Assessment of regional variation in streamflow responses to urbanization and the persistence of physiography. Environmental Science and Technology, 49, 2724–2732. https://doi.org/10.1021/es505389y

Jones, J. A., & Grant, G. (1996). Peak flow responses to clear-cutting and roads in small and large basins, western Cascades, Oregon. Water Resources Research, 32(4), 959–974. https://doi.org/10.1029/95WR03493

Leopold, L. B. (1968). Hydrology for urban land planning: A guidebook on the hydrologic effects of urban land use. U.S. Geological Survey Circular, Circular 5, 1–21. https://doi.org/10.1017/CBO9781107415324.004

Levy, M. C., Lopes, A. V., Cohn, A., Larsen, L. G., & Thompson, S. E. (2018). Land use change increases streamflow across the arc of deforestation in Brazil. Geophysical Research Letters, 45, 3520–3530. https://doi.org/10.1002/2017GL076526

Merz, B., Vorogushyn, S., Uhlemann, S., Delgado, J., & Hundecha, Y. (2012). HEWS opinions: “More efforts and scientific rigour are needed to attribute trends in flood time series”. Hydrology and Earth System Sciences, 16, 1379–1387. https://doi.org/10.5194/hess-16-1379-2012

Merz, R., & Blöschl, G. (2008). Flood frequency hydrology: I. Temporal, spatial, and causal expansion of information. Water Resources Research, 44(8), 1–17. https://doi.org/10.1029/2007WR006744

Morgan, S. L., & Winship, C. (2015). Counterfactuals and causal inference. New York, NY: Cambridge University Press.

Müller, M. F., & Levy, M. C. (2019). Complementary vantage points: Integrating hydrology and economics for sociohydrological knowledge generation. Water Resources Research, 55, 2549–2571. https://doi.org/10.1029/2019WR024786

Ogden, F. L., Raj Pradhan, N., Downer, C. W., & Zahn, J. A. (2011). Relative importance of impervious area, drainage density, width function, and subsurface storm drainage on flood runoff from an urbanized catchment. Water Resources Research, 47, 1–12. https://doi.org/10.1029/2010WR010550

Oudin, L., Salavatli, B., Furusho-Perco, C., Ribstein, P., & Saadi, M. (2018). Hydrological impacts of urbanization at the catchment scale. Journal of Hydrology, 559, 774–786. https://doi.org/10.1016/j.jhydrol.2017.08.064
Over, T. M., Saito, R. J., & Soong, D. T. (2016). Adjusting annual maximum peak discharges at selected stations in northeastern Illinois for changes in land-use conditions. 

Pearl, J., & Mackenzie, D. (2018). The book of why: The new science of cause and effect. New York, NY: Basic Books. 

Prosdocimi, I., Kjeldsen, T. R., & Miller, J. D. (2015). Detection and attribution of urbanization effect on flood extremes using nonstationary flood-frequency models. Water Resources Research, 51, 4244–4262. https://doi.org/10.1002/2015WR017065 

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria: R Core Team. Retrieved from https://www.r-project.org/ 

Salavati, B., Oudin, L., Furusho-Pecon, C., & Ribstein, P. (2016). Modeling approaches to detect land-use changes: Urbanization analyzed on a set of 43 US catchments. Journal of Hydrology, 538, 138–151. https://doi.org/10.1016/j.jhydrol.2016.04.010 

Sauer, V. B., Thomas, W. O., Stricker, V. A., & Wilson, K. V. (1983). Flood characteristics of urban watersheds in the United States. USGS Water-Supply Paper, 2207, 1–16. http://doi.wiley.com/10.1002/ra.1417 

Sheng, J., & Wilson, J. P. (2009). Impacts of impervious surface on watershed hydrology: A review. Urban Water Journal, 6(4), 263–275. https://doi.org/10.1080/15730620903086529 

Slater, L. J., & Villarini, G. (2016). Recent trends in U.S. flood risk. Geophysical Research Letters, 43, 12,428–12,436. https://doi.org/10.1002/2016GL071199 

Smith, J. A., Baeck, M. L., Morrison, J. E., Sturdevant-Rees, P., Turner-Gillespie, D. F., & Bates, P. D. (2002). The regional hydrology of extreme floods in an urbanizing drainage basin. Journal of Hydrometeorology, 3(3), 267–282. https://doi.org/10.1175/1525-7541(2002)003<0267:trhoef>2.0.co;2 

Steinschneider, S., Yang, Y.-C. E., & Brown, C. (2013). Panel regression techniques for identifying impacts of anthropogenic landscape change on hydrologic response. Water Resources Research, 49, 7874–7886. https://doi.org/10.1002/2013WR013818 

Vogel, R. M., Yaindl, C., & Walter, M. (2011). Nonstationarity: Flood magnification and recurrence reduction factors in the United States, 1–11. https://doi.org/10.1111/j.1752-1688.2011.00541.x, http://engineering.tufts.edu/cee/people/vogel/documents/floodMagnification.pdf 

Vörösmarty, C. J., McIntyre, P. B., Gessemer, M. O., Dudgeon, D., Prusevich, A., Green, P., et al. (2010). Global threats to human water security and river biodiversity. Nature, 467, 555–561. https://doi.org/10.1038/nature90440 

White, M. D., & Greer, K. A. (2006). The effects of watershed urbanization on the stream hydrology and riparian vegetation of Los Peñasquitos Creek, California. Landscape and Urban Planning, 74, 125–138. https://doi.org/10.1016/j.landurbplan.2004.11.015 

Wibben, H. C. (1976). Effects of urbanization on flood characteristics in Nashville-Davidson county, Tennessee. U.S. Geological Survey Water-Resources Investigations Report 76-121. https://doi.org/10.3133/wri76121 

Woodruff, J. (2002). Econometric analysis of cross section and panel data. Cambridge, MA: MIT Press. 

Wright, D. B., Smith, J. A., Villarini, G., & Baeck, M. L. (2013). Estimating the frequency of extreme rainfall using weather radar and stochastic storm transposition. Journal of Hydrology, 488(C), 159–165. https://doi.org/10.1016/j.jhydrol.2013.03.003 

Zhang, W., Villarini, G., Vecchi, G. A., & Smith, J. A. (2018). Urbanization exacerbated the rainfall and flooding caused by Hurricane Harvey in Houston. Nature, 563(7731), 384–388. https://doi.org/10.1038/s41586-018-0676-z