Supporting Information for

Glyphosate exposure and GM seed rollout unequally reduced perinatal health

Emmett Reynier and Edward Rubin[1]

[1] **To whom correspondence should be addressed. E-mail: edwardr@uoregon.edu**

This PDF file includes:

Supporting text
Figs. S1 to S38
Tables S1 to S9
SI References

Supporting Information Text

A. Background.

Genetically modified crops Monsanto developed the first genetically modified crops, releasing GM soy, corn, and cotton in 1996 in the United States. These plants are resistant to glyphosate, allowing farmers to spray their fields with glyphosate to kill weeds but not harm their crops. The pairing of GM seeds with glyphosate provides a simple and effective method for controlling weeds—previously, farmers had to use different pesticides, each effective on a smaller subset of weeds at different points in the cultivation process. This herbicide portfolio was supplemented by mechanical tilling. glyphosate previously had to be used sparingly since it would also kill the crops themselves. Farmers adopted GM seeds rapidly in the United States. The USDA provides data on national GM adoption rates starting in 1996 and for specific states beginning in 2000. Figure S2a shows the time series of adoption rates for the entire country (dark, bold line) and particular states (light, gray lines). In 2000, just four years after their release, GM seeds constituted 54 percent of soy acres, 61 percent of cotton acres, and 25 percent of corn acres. By 2010, adoption rates were around 90 percent for all three crops. Adoption of GM corn was generally slower and more heterogeneous across states than for either soy or cotton. Figure S2b shows spatial variation in adoption rates in 2000, 2005, and 2010. For corn and soy, states further west adopted slightly faster than states further east. Meanwhile, California and Texas adopted GM cotton slower than the Southeast.

Glyphosate and health Glyphosate is a broad-spectrum herbicide discovered and commercialized by Monsanto in the 1970s. Its popularity grew over the next twenty years because of its relatively favorable properties. glyphosate has a low toxicity relative to other chemicals used on farms. It breaks down fairly quickly and binds to the soil, decreasing runoff (1). However, it is water-soluble, which means that the part that does not bind to soil enters the water supply (2). It is an effective weed killer, working on a broad spectrum of plants. However, glyphosate does not just kill weeds, it also kills fungi and microorganisms in the soil, which can lead to the crops being susceptible to disease (3). It also breaks the nutrient cycle, forcing farmers to increase their dependence on fertilizer to feed their crops (4). Farmers in the US spend nearly $8 billion on pesticides each year (5), applying glyphosate to 298 million acres of crops annually (6).

Regulatory oversight The US EPA's current approval process for pesticides provides ample opportunities for applicants to steer the process toward approval.

A central critique in the pesticide-regulation literature is the EPA's tendency to rely upon regulated entities to design, run, and analyze non-peer-reviewed experiments to test chemicals' safety—ignoring clear conflicts of interest (7–9). This approach stands in stark contrast to the International Agency for Research on Cancer (IARC)'s reliance on published, peer-reviewed research (7, 8).

EPA's reliance on applicant-generated tests grants applicants the opportunity to influence evidence in the review process in several ways. First, this process places regulated entities in a position to effectively selectively report their way to approval by running multiple tests and only reporting studies that show no harm (8). Benbrook (2019) illustrates considerable differences between regulatory assays versus assays from peer-reviewed journals (10)—leading the EPA and IARC to conflicting conclusions. Several recent reports highlights similar concerns around regulation in the EU—especially in differing conclusions between the IARC and the European Food Safety Authority (EFSA) (11, 12).

Among reported tests, regulated entities may test exposure levels lower than (1) levels encountered in occupational settings and (2) levels where adverse health effects may be observed—especially when chemicals may have nonlinear dose-response relationships (13). Benbrook (2020) also highlights that the EPA's approval process focuses more on dietary consumption than on occupation exposure—a potentially large shortcoming for Roundup's high dermal penetration and possible wand application (8). Benbrook suggests that the regulatory process surrounding glyphosate has largely failed "to add common-sense worker protection provisions" (8).

Similarly, the review process may conclude with *no harm* due to flawed research designs (14), implementation (15), and/or analyses (9). Despite its potential impacts on public health, the data involved in these studies is often withheld (13)—further restricting review and oversight. Compounding the issue: scant post-approval monitoring of chemicals exposure in populations or the environment (7). Finally, there is the issue of regulatory capture, which we leave for future research.

While the conflict of interest is clear, these mistakes may also follow from the studies' lack of legitimate peer review. Ultimately, the EPA's reliance on industry-conducted studies, in conjunction with the effective null hypothesis of *no harm*, opens the door for avoidable public health risks.

B. Data.

Fertilizer data We get fertilizer data from the USGS (16), which estimates county-level nitrogen and phosphorous every five years between 1950 and 2017 using data from the USDA Census of Agriculture. They report separate estimates for farm commercial fertilizer applications, non-farm commercial fertilizer applications, and nutrient loads from manure. Since the rest of our data are annual, we interpolate for the years between the years of the Ag Census. We fit splines separately for each county and fertilizer type using all available years of data and then generate annual predictions for each county-year-fertilizer type. Figure S3 shows the results for a few example counties, where the points are the actual data and the lines are the fitted splines. The raw data only separates farm and non-farm commercial fertilizer use from 1987. Non-farm commercial fertilizer use is small, only making less than three percent of commercial fertilizer use. Therefore, we use four variables as controls:

phosphorous from commercial uses, nitrogen from commercial uses, phosphorous from manure, and nitrogen from manure. We normalize all of these variables by the total size of the county.

Crop Acreage and Yield Data We get crop acreage and yield data from the USDA NASS. Specifically, we get acres planted and yield per acre by county from 1985 to 2017 from their annual survey. These data include barley, several types of beans, canola, chickpeas, corn, upland cotton, flaxseed, lentils, mustard, oats, peanuts, peas, peppers, potatoes, rice, safflower, sorghum, soybeans, sugarbeets, sweet corn, sweet potatoes, tomatoes, and wheat. We aggregate all crops besides corn, soy, and cotton into an "other" category. The USDA masks data for counties where there are small amounts of farms producing a particular crop. They report acres and yield from those counties as an aggregate for the agricultural statistical district. To account for this, we allocate any production reported at the aggregate district level to counties in that district with missing data proportional to the size of the county.

Accuracy of FAO-GAEZ data New work highlights several issues with the GAEZ suitability data, calling into question the ability of the GAEZ suitability measures to predict historical yield for different crops (17). While we share many of the concerns raised—notably the lack of clarity around data sources and model validation—we think that the use of the GAEZ data is still appropriate in our setting. We are not using the GAEZ data to predict historical yields but rather as an instrument for glyphosate use. Thus, we only need the GAEZ suitability to be correlated with increases in glyphosate and for the exclusion restriction to hold. See *Methods* for an in-depth discussion of our identifying assumptions. Table S2 demonstrates that the GAEZ suitability measures correlate with observed acreage and yield of corn, soy, and cotton prior to the introduction of GM crops.

Other Data We collect data on county-level, annual estimates of total population by age and race from SEER (18). Age shares are done by decade from ages 0 to 70, with everyone older than 70 grouped into the final bin. Race shares are calculated for the white, Black, and hispanic populations.

C. OLS results.
Panels A and B of Table S7 contain results for our main specifications, but estimated with OLS rather than 2SLS. We find precise null effects across all outcomes, demonstrating the importance of isolating exogenous varaiation in glyphosate using our instruments.

D. Comparing rescaled reduced-form DID results with 2SLS results.
Our reduced-form DID results effectively provide estimates of 2SLS's first stage and reduced form—just in a DID framework without interacting the instrument with indicators for year. By dividing the reduced-form DID results for health outcomes by the results, we can get comparable estimates for the health effects due to glyphosate exposure. Table S4 provides these calculations—first repeating the results from the main text in Figure 3 and then rescaling the reduced-form DID health effects by the reduced-form DID glyphosate effect. The resulting estimates are very close to the 2SLS estimates.

E. Shift-share specification.
We can recast our identifying variation to be used similarly to that of a traditional "shift-share" specification, where the "shift" is national glyphosate use and the "share" is attainable yield in each county. Thus, the identifying variation is very similar to our main results—we get temporal variation driven by the nation-wide increase in glyphosate use after the release of GM crops, and we get spatial variation from the suitability of the land in each county for corn, soy, and cotton. In the shift-share specification, our instruments are national glyphosate and national glyphosate interacted with the attainable yield percentile for corn, soy, and cotton. Thus, the difference between this specification and our main specification is that the first stage uses the national glyphosate trend directly, rather than interacting attainable yield with year dummies. When calculating the national glyphosate for each county, we exclude glyphosate applied within 100km of the county and any glyphosate applied upstream of the county to ensure that the national glyphosate instrument satisfies the exclusion restriction—that national glyphosate only affects perinatal health through its effect on local glyphosate. Table **??** shows the results, which are generally similar, but smaller in magnitude than our main results.

F. Robustness of first-stage and reduced form results.
Figure S18 estimates our model for births to mothers with rural and non-rural residences separately. There is a small but largely insignificant decrease in birthweight after the release of GM seeds in 1996 in high GM attainable yield counties relative to low GM attainable yield counties. However, this effect is gone by 2010. We attribute this difference largely to measurement error in exposure—we do not think that exposure is very high for urban mothers, who are unlikely to be in contact with drift, dust, or water contaminated with glyphosate applied within that county. The lack of direct measurement of exposure to glyphosate is a weakness of our study, as all we know is the amount of glyphosate used in a county each year.

G. Robustness of 2SLS results for other outcomes.
We provide specification charts for non-birthweight outcomes in Figures S19 and S20. Additionally, we test robustness to various spatial subsets in Figure S21.

H. Prediction performance.
Figure S8 depicts how predicted birthweight percentiles maps to predicted (dark blue) and actual (light orange) birthweight. Given a predicted percentile, the mean predicted birthweight is, on average, quite close to the true mean birthweight—suggesting the ML approach indeed captures informative variation in birthweight.

Table S6 describes the performance of the random-forest model for predicting birthweight (the outcome for which the model trained) and birthweight percentile, decile, and quintile. Panel A evaluates predictions' performances on two metrics—mean absolute error (MAE: $|y - \hat{y}|$) and mean absolute percent error (MAPE: MAE/y). All results in the table focus on the set of infants born to rural-residence mothers, as this group matches the population of interest in the paper. The three rows of Panel A use this full sample of rural-residence mothers and then drop 1% and 5% of the top and bottom birthweights. Finally, Panel B provides two null models for comparison. The *Means* model always predicts the sample mean; the *At random* model predicts at random.

For predicting birthweight, the ML predictions offer improvements over both null models on both metrics (error level and percentage), and this dominance is even stronger when not evaluating the tails of the distribution. While the *Means* null model slightly edges out the full-sample model in terms of MAE for the percentile-based outcomes, the ML-based model is *much* stronger in terms of percent error. Percentage error is an important metric in our application, as we want match births to expected birthweight groups (as in Fig S8) rather than simply a low-error on-average prediction.

I. 2SLS results and the shape of the glyphosate damage function. Our estimates for the effect of glyphosate exposure on perinatal health make comparisons between communities with higher and lower levels of glyphosate exposure (instrumented communities' GM-crop suitability, compared to pre-GM-rollout comparisons, and conditional on fixed effects). Following the GM rollout, many low-suitability communities also experienced increased glyphosate exposure (for example, as shown in Figure 1c). As a result, our estimates for glyphosate's health damages compare infants with higher glyhposate exposure to infants with lower glyphosate exposure. Generally, infatns with *lower* levels of glyphosate exposure are generally still exposed to some non-zero amount glyphosate. How this non-zero exposure for the lower group affects our 2SLS estimates depends upon two items: the shape of glyphosate's damage function and the interpretation of the 2SLS estimates.

For the moment, ignore the temporal variation in our instrument and estimation. The reduced-form results from our 2SLS estimates effectively compare glyphosate's damages between in high and low GM-suitability areas, i.e., $\mathcal{D}(g_h) - \mathcal{D}(g_\ell)$. The first stage estimates the difference in glyphosate intensity between these areas, i.e., $g_h - g_\ell$. Finally, second-stage scales these differences by the differences in glyphosate intensity.

$$\begin{aligned}
&\textit{Reduced form: } \text{Damages from high vs. low glyphosate} && = \mathcal{D}(g_h) - \mathcal{D}(g_\ell) \\
&\textit{First stage: } \text{Differences in glyphosate intensity} && = g_h - g_\ell \\
&\textit{Second stage: } \text{Scaling damage difference by glyphosate difference} && = \frac{\mathcal{D}(g_h) - \mathcal{D}(g_\ell)}{g_h - g_\ell}
\end{aligned}$$

There are three main cases to consider for the shape of the damage function (with respect to its second derivative).

- **Linear damages** If glyphosate's damages for perinatal health are approximately linear/affine in glyphosate (i.e., $\alpha + \beta g$), then our second-stage estimates are unaffected by whether low-suitability areas have zero or non-zero levels of glyphosate—as long as low-suitability areas apply less glyphosate than high suitability areas (our first-stage estimates and event studies confirm this requirement is satisfied).

  $$\frac{\mathcal{D}(g_h) - \mathcal{D}(g_\ell)}{g_h - g_\ell} = \frac{\alpha + \beta g_h - \alpha - \beta g_\ell}{g_h - g_\ell} = \frac{\beta(g_h - g_\ell)}{g_h - g_\ell} = \beta$$

  Accordingly, if the damage function is linear, our second-stage will recover the marginal damages from a one-unit increase regardless of the level of glyphosate in low-suitability counties.

- **Concave damages** For concave damage functions, as g_ℓ increases from zero and approaches g_h from below, the reduced-form difference $\mathcal{D}(g_h) - \mathcal{D}(g_\ell)$ shrinks faster relative to the first-stage difference $g_h - g_\ell$. Consequently, higher levels of g_ℓ will generate lower second-stage (and reduced-form) estimates. If the parameter of interest is the average per-unit health damages due to moving from $g = 0$ to $g = g_h$, then the 2SLS will understate the actual damages. However, 2SLS estimates the (weighted) average marginal damage of increasing glyphosate from concentrations found in the US. More formally,

  $$\frac{d}{dg_\ell} \frac{\mathcal{D}(g_h) - \mathcal{D}(g_\ell)}{g_h - g_\ell} = \left[\frac{\mathcal{D}(g_h) - \mathcal{D}(g_\ell)}{g_h - g_\ell} - \mathcal{D}'(g\ell) \right] \frac{1}{(g_h - g_\ell)^2} \leq 0$$

  due to $\mathcal{D}$ being concave (strict inequality will follow from strict concavity).

- **Convex damages** Convex damage functions simply 'reverse' the results of concave functions: 2SLS will overstate the average damage of moving from $g = 0$ to $g = g_h$ when $g_\ell > 0$ but will provide a (weighted) average marginal damage for glyphosate concentrations common in US counties.

Because we are not in a position to take a strong stand on the shape of glyphosate's damage function, it likely makes more sense to consider our estimates as the marginal damages relative to glyphosate concentrations commonly encountered in the United States—a policy-relevant parameter requiring weaker assumptions.

J. Demographic trends. One concern for identification in our model is that the underlying composition of the population is changing in high vs. low GM attainable yield counties during the period of our study. Figure S22 shows event studies where we use demographics of the mother as outcomes with county and year-by-month fixed effects and no other controls to test whether demographics are changing over time. We find that births in high-yield counties are less likely to come from black mothers after the release of GM crops—this would otherwise be concerning for our main estimates, however, we (1) control for race and other demographics in our main estimation and they do not meaningfully impact the results, (2) *predicted* birthweight does not change over the time period of the study, and (3) we find significant effects of glyphosate on birthweight for babies with both white and non-white mothers.

K. Other forms of heterogeneity.

Mother's race Based on heterogeneity in predicted birthweight, we expect there to be differences in effect by mother's race. Fig S23 shows reduced from event studies for different outcomes by mother's race. Births to non-white mothers have a noisy, but generally larger effect than briths to white mothers.

Heterogeneity by month of birth These results do not exhibit consistent heterogeneity by month of birth, as seen in Figure S25. There are slightly higher effects during the first months of the year—which means that their gestational period began in the spring and early summer the time when the most glyphosate is applied.

Rural *vs* urban We compare results for rural and non-rural counties in *SI* Figure S18. As expected, the first stage is much weaker in non-rural than for rural counties since non-rural counties have more land uses competing with agriculture. We note that non-rural counties still grow GM crops and apply glyphosate—30% of corn, soy, and cotton acres and 29% of glyphosate applications are in non-rural counties. However, the mismeasurement of glyphosate exposure for infants in non-rural counties is likely to be considerably worse than in rural counties. Mothers residing in urban portions of a non-rural county will have lower glyphosate exposure than mothers residing near that same county's agricultural production. Yet, data constraints force us to assign all infants in a county the same level of glyphosate exposure. We estimate the effect of GM suitability in non-rural counties to be attenuated relative to rural counties, consistent with the non-rural counties having more measurement error in exposure.

L. Effect of GM on acreage and yield. Changes in agricultural activity unrelated to glyphosate that result from GM seed adoption could also affect infant health, threatening our identified effect of glyphosate on birth weight. For example, GM technology could lead farmers to bring marginal, not previously farmed land into agricultural production. This additional production could be associated with increased runoff into water or air pollution from dust or drift. Additionally, if yield increased with GM seeds, farmers could see an economic boost that could affect infant health. In order to rule out these as mechanisms for the observed effect of GM attainable yield on birth weight, we explore the effect of GM attainable yield on crop acreage and actual yield.

We use USDA NASS data on annual, county-level crop acreage and yield, regressing these variables on the max GM attainable yield percentile interacted with year. Figure S26a shows reduced-form event study results regressing our suitability measure on total crop acreage as a share of the county area. Unfortunately, 1995 seems to be a low outlier year, making the event study more challenging to interpret—however, total acreage appears to stay around the same level after 1995 as it was prior to 1995. We estimate a difference-in-difference model comparing before vs after 1995 that results in a small and not statistically significant difference in total acreage.

Figure S26b shows corn acreage as a percent of the county area. As with total acreage, the effect of GM suitability is noisy in the pre-period and seems unchanged until around 2007, at which point there does seem to be an increase in corn acreage for high-suitability counties relative to low-suitability counties. This timing coincides with when the renewable fuel standard increased incentives for farmers to plant corn (19). Figure S26c shows the results for soy acres as a percent of the county area. There was an initial bump in soy acreage after 1995, followed by a return to pre-1995 averages, consistent with the fact that GM seed varieties were available for soy before corn.

Meanwhile, Figure S26d shows the effect of GM suitability on cotton acreage. Cotton seems to have had a high outlier year in 1995 but remains consistent with the other pre-period years until 2006, after which it decreases in high relative to low suitability counties. In summary, we find that there does not seem to be an effect of GM suitability on total acreage, but this masks some substitution between crops.

We estimated models adding both fertilizers and acreage as controls. Figure S13 shows both the first stage and reduced form event studies under various iterations of controls—adding fertilizer and/or acreage controls does not result in meaningful or statistically significant differences of estimates in either the first stage or reduced form. Notably, adding acreage controls to all of the other controls does flatten out the upward pre-period trend in the first stage. The spec charts in Figures S6, S19, and S20 show our estimated coefficient on glyphosate from 2SLS on birthweight, gestation length, and the health index. Adding fertilizer controls slightly increases the magnitude of glyphosate's effect on birthweight, but by considerably less than when we add pesticide controls. This is true when adding just fertilizers as controls relative to no additional controls and when adding fertilizers to a specification with pesticides and unemployment as controls.

M. Other socioeconomic outcomes. Here, we explore the relationship between our attainable yield instrument and some socioeconomic outcomes in order to rule them out as mechanisms for the measured birth weight effect. We regress the socioeconomic variables on GM attainable yield interacted with year dummies with county and year fixed effects. The sample is a county-year panel of rural counties in the US between 1990 and 2013. Figure S28 shows the results. There is no change in farm or non-farm income, however there do appear to be changes in employment. The unemployment rate jumps after 2000—thus, we control for unemployment in our main regression, but note that this is four years after the release of GM seeds, thus the timing does not align to have been caused by GM. Meanwhile, farm employment is also declining, however there is a clear pre-trend. The release of GM seeds does not appear to affect this trend.

N. Effects of upstream glyphosate in water.

N.1. Predicting glyphosate in water with machine learning. To measure spillover effects from glyphosate applied upstream, we must have some measure of glyphosate exposure in water. Ideally, this would come from extensive monitoring, which consistently reports pesticide concentrations in water for a comprehensive set of water sources. Unfortunately, such a monitoring network does not exist, so we must create an alternative methodology to estimate glyphosate exposure from upstream spraying. We train a machine learning model to predict glyphosate concentrations using the limited glyphosate monitoring in water, along with water flow and other environmental characteristics.

Data preparation Our training data come from recent work by Medalie *et al.* (20), who took 3204 samples of glyphosate and its main degradate AMPA from 70 sites in the National Water Quality Network (NWQN), a nationally representative set of water bodies, between 2015 and 2017. Both chemicals are nearly omnipresent, with glyphosate detected in 75 percent of samples and AMPA detected in 90 percent. We link these measurements to data on glyphosate use, soil type, slope, and rainfall upstream from the sampling location.

We use a spatial water model to aggregate the amount of glyphosate applied upstream and downstream of each sampling location. Specifically, we use the level 8 HydroBASINS product from HydroSHEDS (21). These data are watershed polygons that delineate water basins across the globe in a standardized way. Importantly, they are assigned codes in a way that makes it possible to find all watersheds upstream and downstream from any given watershed.

We begin with the pesticide data. As in our local analysis, one may be concerned with the endogeneity of glyphosate use. Our estimates will be biased if spraying upstream of a sampling location correlates with other factors affecting health outcomes. We deal with this issue by using only exogenous variation in glyphosate use driven by the same instruments from our local analysis, namely that driven by the timing of the release of GM seeds and the suitability of a county for corn, soy, and cotton. We regress glyphosate on the GM attainable yield percentile interacted with year dummies, with year and county fixed effects to generate county-year level predictions of glyphosate. To disaggregate these county-level predictions into watersheds, we assume that spraying is uniform across the county and multiply the glyphosate prediction for each county by the portion of the county's total area covered by the watershed. Figure S29 shows the spatial distribution of predicted glyphosate by watershed across the United States in 2004.

Additionally, we collect several other variables that affect the runoff of glyphosate in a method loosely following the commonly used universal soil loss equation (USLE). This soil loss equation multiplies the erodibility of the soil, the slope of the land, rainfall, and two measures associated with land use. We aggregate soil erodibility and slope from the gridded soil survey to the watershed level by taking the average over all 30-meter cells in each watershed (22). Similarly, we use gridded, monthly precipitation from PRISM to help inform the potential for glyphosate to run into water (23). We aggregate the 4-kilometer cells to the watershed level by taking the simple average of cells within a watershed. Additionally, we aggregate to the annual level by taking the sum over the growing season, April through September, when most glyphosate is applied. Figure S30 shows national percentiles of soil erodibility, slope, and precipitation by watershed.

We then utilize the "Pfafstetter" watershed coding system used by the HydroBASINS data to find all watersheds upstream from each watershed. We have selected an example watershed in Washington County, Illinois, just east of St. Louis, for demonstration purposes. Figure S31 shows the example watershed in red and then highlights all of the watersheds upstream, which reach further north into Illinois, and all of the watersheds downstream, which follow the Mississippi River to the Gulf of Mexico.

When linking upstream and downstream watersheds, we calculate the distance between any two watersheds by summing the distance between centroids of each watershed that lies along the water flow between the two watersheds. We then aggregate the variables described above into 50-kilometer distance bins from -100 to 350, where negative values denote values for downstream watersheds. Figure S31 demonstrates the distance bins for our example watershed. The final dataset contains 2,142 water samples, where we removed 1064 samples from sites with no upstream watersheds entirely outside the site's county. We remove these to ensure that our measure of upstream spraying does not capture non-water mechanisms of glyphosate exposure, such as dust, drift, or direct contact.

Training the water concentration ML model We train LASSO and Random Forest (RF) models using the above mentioned dataset. We generate a fully saturated set of interaction terms between glyphosate, soil erodibility, slope, and rainfall as predictors in the LASSO model. The month of the sample is the only other predictor variable. Since the model's primary goal is to predict glyphosate concentrations back in time, we train the model on 1,385 observations from after October 2015 and validate performance with 757 observations from before October 2015. Within the training set, we tune parameters using 4-fold cross-validation, where each fold trains on 15 months of data and then tests performance on the preceding six months of data. Then, we select the parameter with the lowest average RMSE across folds to estimate the model on the entire training set. Figure S32 shows the cross-validation results.

We then assess performance of the tuned models using the 757 held out observations. Figure S33 shows the out-of-sample predictions versus their actual values. Both models predict AMPA concentrations much better than glyphosate concentrations, with an R-squared of 0.59 and 0.31 for the random forest and LASSO models respectively. Figure S34 shows the density of the out-of-sample predictions for each model, as well as actual values. Generally, the models slightly over-predict at low values, moreso for glyphosate than AMPA.

Generating predictions We use the model to predict county-month-level glyphosate and AMPA concentrations. We do this by making predictions for every watershed for each month between January of 1992 and December of 2017. We then take the weighted average of the predictions, where the weights are the proportion of the county's population that lives in the watershed. Our population estimates come from SEDAC's 2010 population grid (24). This grid estimates the population for one square kilometer pixels across the United States. We add the population counts for pixels within each watershed and then divide by the total population count for cells within the county to obtain the population weights. Figure S35 shows predicted AMPA in July of 2004 from the LASSO model from each watershed touching Washington County on the right and the population weight for those watersheds on the left. Figure S36 shows predicted AMPA in water for each county in July of 2004. We can then link the county-month-level predictions of glyphosate and AMPA to the birth certificate data.

N.2. Results: Effect from upstream glyphosate in water. Before using the machine learning predictions of glyphosate and AMPA in water, we first regress perinatal health outcomes on aggregate suitability over distance bins upstream or downstream from the mother's county of residence. These are of the form,

$$\text{Health}_{ijt} = \sum_{\tau\neq 1995} \gamma^{l}_{\tau}\text{GM}^{l}_{j} \times \mathbb{1}(t=\tau) + \sum_{\tau\neq 1995}\sum_{d} \gamma^{u}_{\tau d}\text{GM}^{u}_{jd} \times \mathbb{1}(t=\tau) + \Gamma X_{ijt} + \alpha_j + \lambda_t + \varepsilon_{ijt},$$

where GM^u_{jd} is the average GM suitability percentile in distance bin d upstream (or downstream) from county j. We now use GM^l_j to denote local GM suitability.

Figure S37 displays event study plots illustrating the effect of max GM attainable yield in upstream watersheds on birthweight, categorized into 50-kilometer distance bins. These results suggest that having land more suitable for GM crops *upstream* of a county does not lead to a change in birthweight after the release of GM seeds in 1996.

As Dias, Rocha, and Soares (25) emphasize, the potential effects of upstream glyphosate spraying would be strongest in places where there is more runoff from farms. We estimate the event study allowing for heterogeneity by high-soil-erodibility and high-precipitation, two factors that could increase runoff of glyphosate into surface water. Figure S38 shows the results for both high- and low-erodibility and precipitation. Neither demonstrate a consistent effect on birthweight.

Finally, we estimate the effect of predicted glyphosate and AMPA in water. We do this by running regressions of the form,

$$\text{Health}_{ijt} = \beta^{l}\widehat{GLY}^{l}_{jt} + \beta^{u}\widehat{\text{GLY}}^{u}_{jt} + \Gamma X_{ijt} + \alpha_j + \lambda_t + \varepsilon_{ijt}, \qquad [1]$$

where $\widehat{GLY}^{l}_{jt}$ represents local glyphosate exposure, predicted from the first stage Eq. (2). $\widehat{\text{GLY}}^{u}_{jt}$ denotes predicted exposure to glyphosate or AMPA from glyphosate applied upstream of county j in year t, where we generate predictions from the machine learning model described above. These predictions are plausibly exogenous, as the models are trained only on exogenous data. Table S9 shows the results of regressing these predictions of glyphosate or AMPA in water on birthweight. All four estimates, coming from either a LASSO or random forest model predicting either AMPA or glyphosate concentrations demonstrate a null effect of glyphosate or AMPA on birthweight.

We approach these findings cautiously; however, they suggest the absence of substantial downstream health spillovers resulting from glyphosate runoff. The lack of effect may be reasonably expected in the US relative to Brazil, as drinking water treatment in the US is more robust than that in Brazil (26). However, we cannot definitively exclude water exposure as a potential mechanism driving the local results. glyphosate runoff into the water could be causing issues within a county but not downstream of a county if the chemicals degrade quickly enough. Additionally, given the inherent measurement error in this process and the absence of a more refined chemical transport model, we refrain from making definitive claims about the existence of downstream spillovers from glyphosate use.

	Rural counties				Non-rural counties	
	High GM yield		*Low GM yield*			
Variable	Mean	Stnd. Dev.	Mean	Stnd. Dev.	Mean	Stnd. Dev.
Infant characteristics						
Female infant	0.488	0.500	0.488	0.500	0.488	0.500
Birthweight (g)	3,327.0	610.1	3,360.1	583.8	3,333.5	604.5
Gestation (wk)	39.050	2.788	39.180	2.565	39.015	2.660
Preterm	0.185	0.388	0.166	0.373	0.178	0.383
C-section	0.243	0.429	0.223	0.416	0.217	0.412
Health index	0.029	0.251	0.031	0.244	0.021	0.242
Low birthweight	0.075	0.263	0.063	0.243	0.072	0.258
Very low birthweight	0.013	0.115	0.010	0.099	0.013	0.114
Maternal characteristics						
Black	0.199	0.400	0.024	0.152	0.172	0.378
Non-White	0.217	0.412	0.074	0.262	0.217	0.412
Hispanic	0.030	0.171	0.103	0.304	0.179	0.383
Married	0.683	0.465	0.746	0.435	0.692	0.462
High-school grad.	0.749	0.433	0.763	0.425	0.769	0.421
College grad.	0.115	0.319	0.131	0.337	0.207	0.405
Local agriculture						
Glyphosate (kg/km^2)	0.003	0.004	0.001	0.002	0.002	0.004
Max. GM crop suit. pctl.	0.759	0.144	0.248	0.138	0.494	0.285
Corn acres/km^2	24.303	31.789	8.586	20.261	7.774	17.128
Soy acres/km^2	25.314	30.369	4.019	12.280	6.105	14.826
Cotton acres/km^2	3.872	11.910	1.258	8.937	1.199	6.268
Other acres/km^2	15.015	22.160	11.386	23.691	6.156	13.409
Local economy						
Unemployment rate	0.070	0.026	0.079	0.041	0.066	0.027
Pct. farm employment	0.078	0.057	0.077	0.071	0.014	0.027
Farm employment per capita	0.036	0.027	0.037	0.036	0.006	0.011
Total population	43,742	28,666	41,164	31,005	1,347,402	2,122,955
Farm income per capita	0.491	0.592	0.600	1.236	0.104	0.257
Non-farm income per capita	15.843	2.580	15.884	3.313	22.165	5.312
Employment rate	0.480	0.092	0.486	0.117	0.559	0.167
Totals						
Number of births	1,981,619		1,734,005		20,031,928	
Number of counties	898		1,125		1,085	

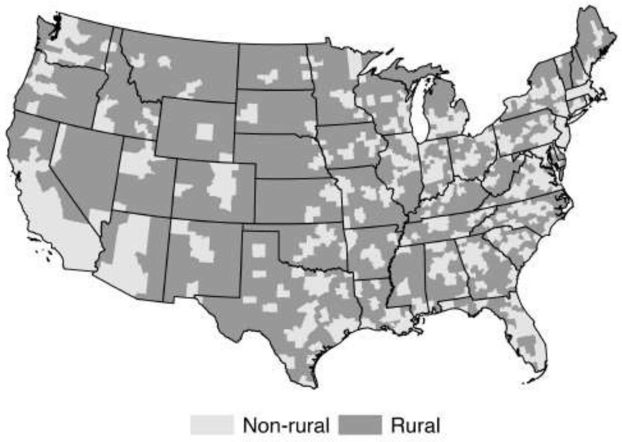
Table S1. Summary statistics for high- and low-yield rural counties and urban counties. Means and standard deviations are calculated at the birth level for the county group in the years 1992–1995. GM yield grouping is based upon being above or below the 50th percentile of maximum attainable yield for GM crops. Rural/non-rural split uses USDA rural-urban continuum codes from 2003.

Table S2. Correlation between GAEZ suitability measures and pre-period acerage for GM crops.

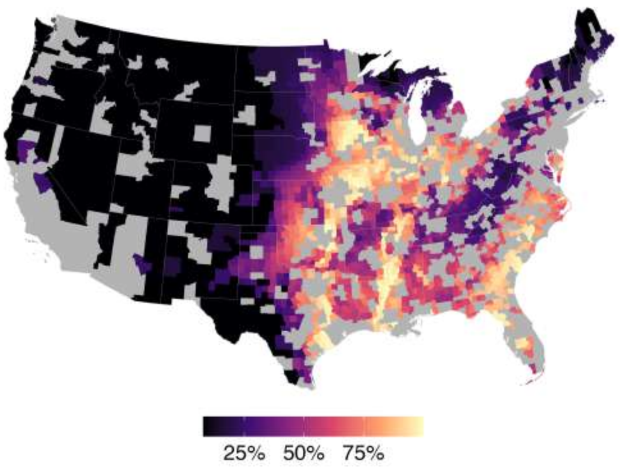
	Acreage Percentiles				Yield Percentiles			
Dep Var:	GM	Corn	Soy	Cotton	GM	Corn	Soy	Cotton
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	0.227	0.240	0.141	0.263	0.214	0.326	0.178	0.264
	(0.009)	(0.009)	(0.007)	(0.008)	(0.009)	(0.010)	(0.008)	(0.008)
GM GAEZ Yield Percentile	0.546				0.571			
	(0.015)				(0.015)			
Corn GAEZ Yield Percentile		0.521				0.348		
		(0.015)				(0.017)		
Soy GAEZ Yield Percentile			0.719				0.643	
			(0.012)				(0.014)	
Cotton GAEZ Yield Percentile				0.474				0.472
				(0.013)				(0.013)
Fit statistics								
Observations	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109
R^2	0.30	0.27	0.52	0.29	0.33	0.12	0.42	0.28

IID standard-errors in parentheses

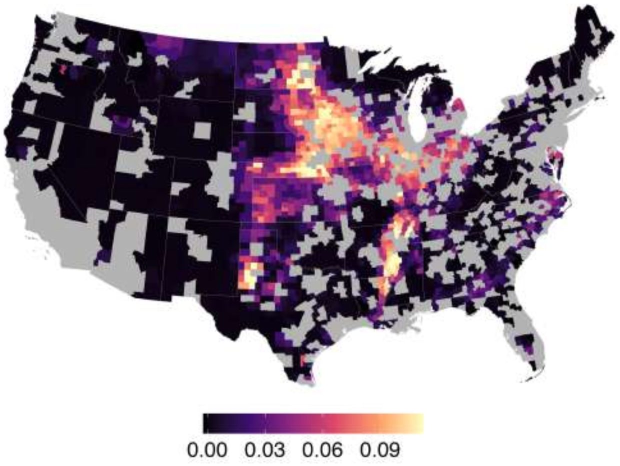
We first calculate the county-level 1990 to 1995 average planted acreage and yield for each of corn, soy, cotton, and the aggregate of all three for GM. We divide the acreage values by the total size of the county. We then convert the acreage share and yield values into a percentile relative to all counties in the continental US. The GAEZ yield percentiles are calculated as described in *Methods*.



(a) Rural vs non-rural counties

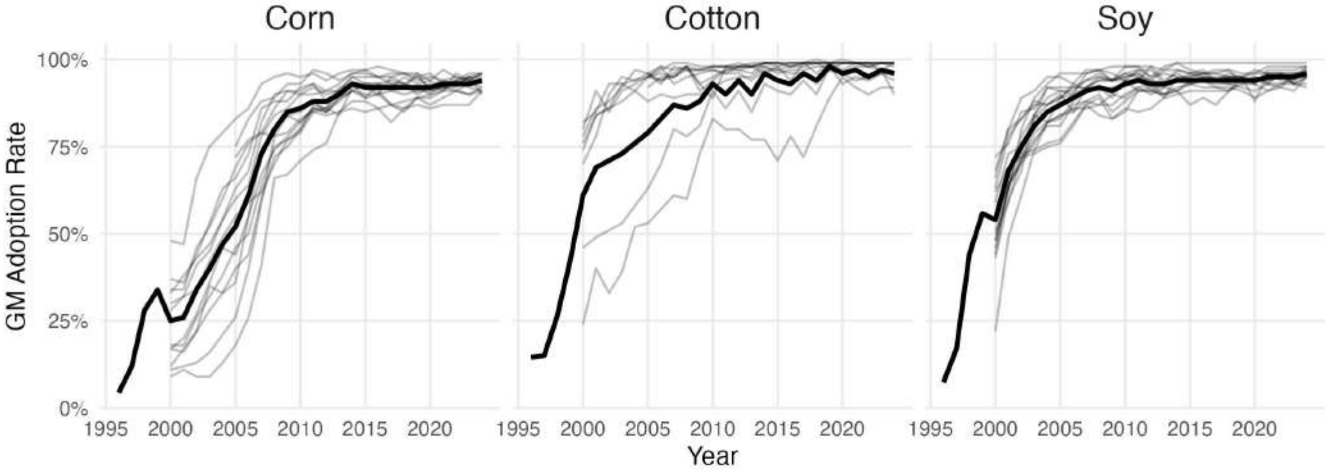


(b) GM crop suitability, Max GM attainable yield pctl in rural counties

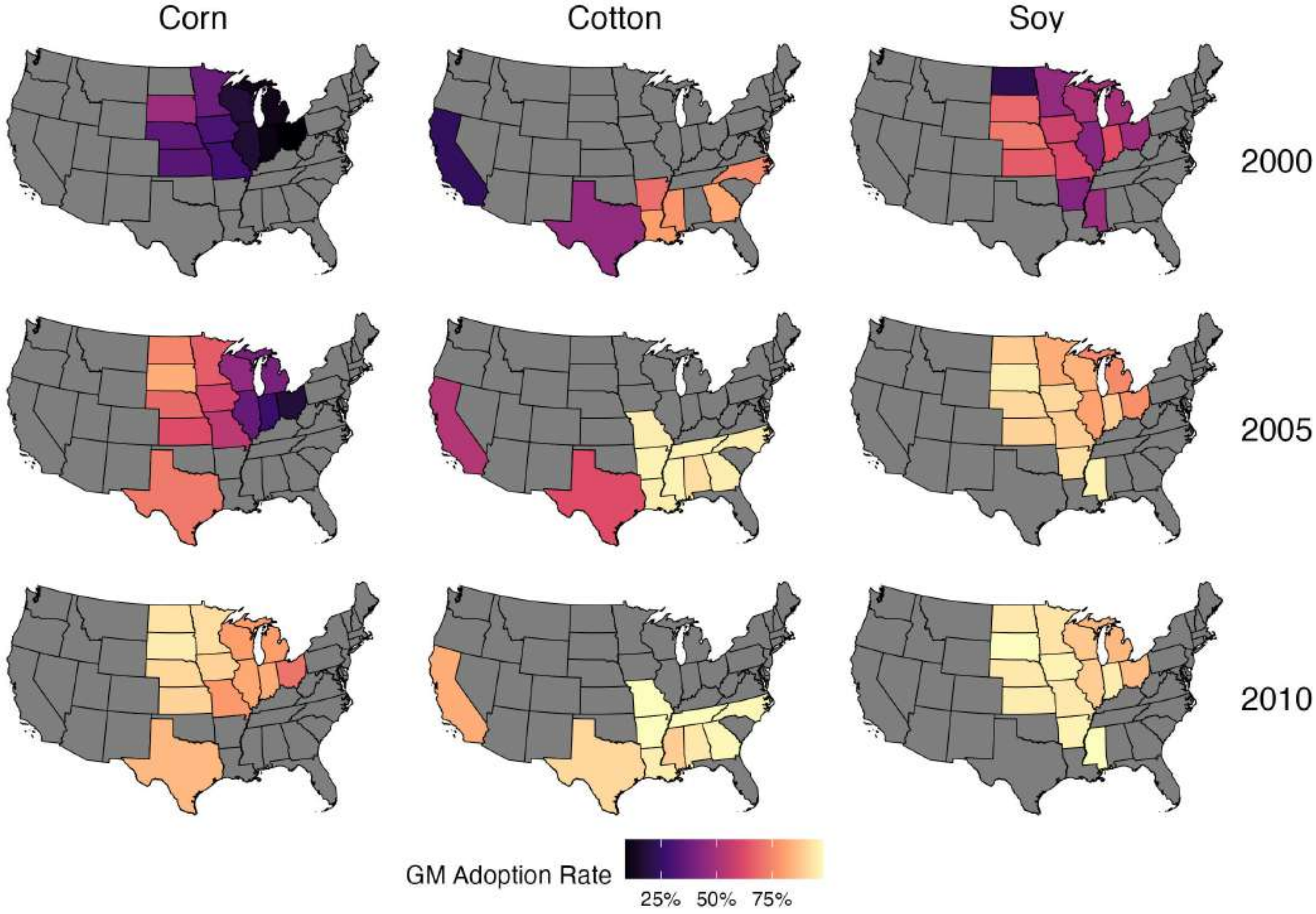


(c) Increase in glyphosate, 1995–2012 (kg/km^2) in rural counties

Fig. S1. GM crop suitability and increases in glyphosate for rural counties. **(a)** Rural counties using 2003 Rural-Urban Continuum codes from the US Department of Agriculture (USDA) to classify counties as rural. A rural county is any non-metro county, where the USDA defines a metro county as, "broad labor-market areas that include central counties with one or more urban areas with populations of 50,000 or more people. **(b)** Percentile of attainable yield for GM crops equals the difference in attainable yield between high- and low-input scenarios from FAO GAEZ (27) for corn, soy, and cotton. We rescale each crop to be a national percentile, take the maximum over the three crops, and finally scaling again to be a national percentile. Here we filter to only rural counties. **(c)** Change in glyphosate censored at the 1st and 99th percentiles and then filtered to only rural counties.



(a) Time series of GM crop adoption.



(b) Spatial variation in GM crop adoption.

Fig. S2. GM seeds were rapidly adopted after their 1996 release. (a) Shows the percent of crops with any GM technology by year. The bold line is the entire United States and the grey lines are specific states. **(b)** Shows spatial variation across states in GM adoption rates in 2000, 2005, and 2010. Data from the USDA (28).

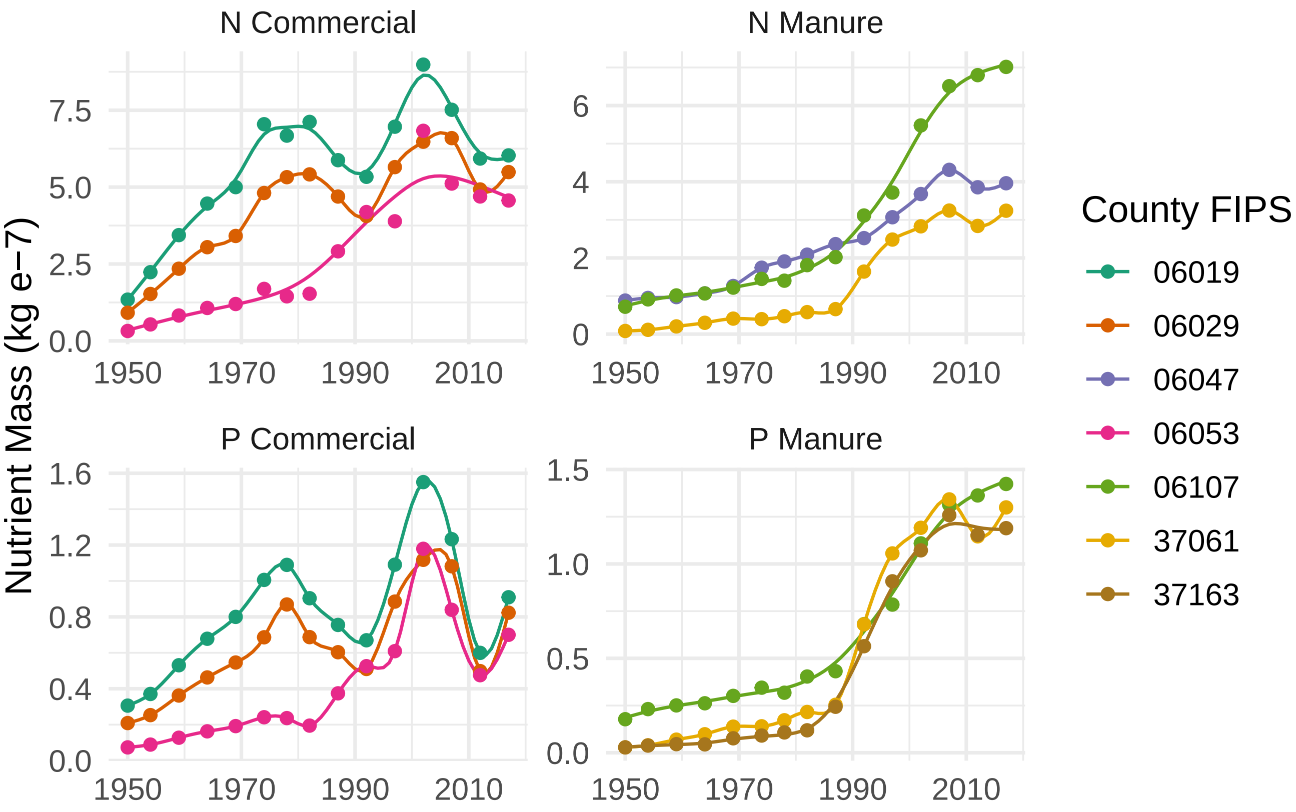


Fig. S3. Interpolation of fertilizer data. This figure shows the fitted spline values and raw data for seven example counties. The dots are raw data and lines are the fitted spline functions. We fit a separate spline for each county and fertilizer type, using all data available—a value every five years between 1950 and 2017 from USGS (16). We then generate annual-county-fertilizer type predictions to use with the rest of our data.

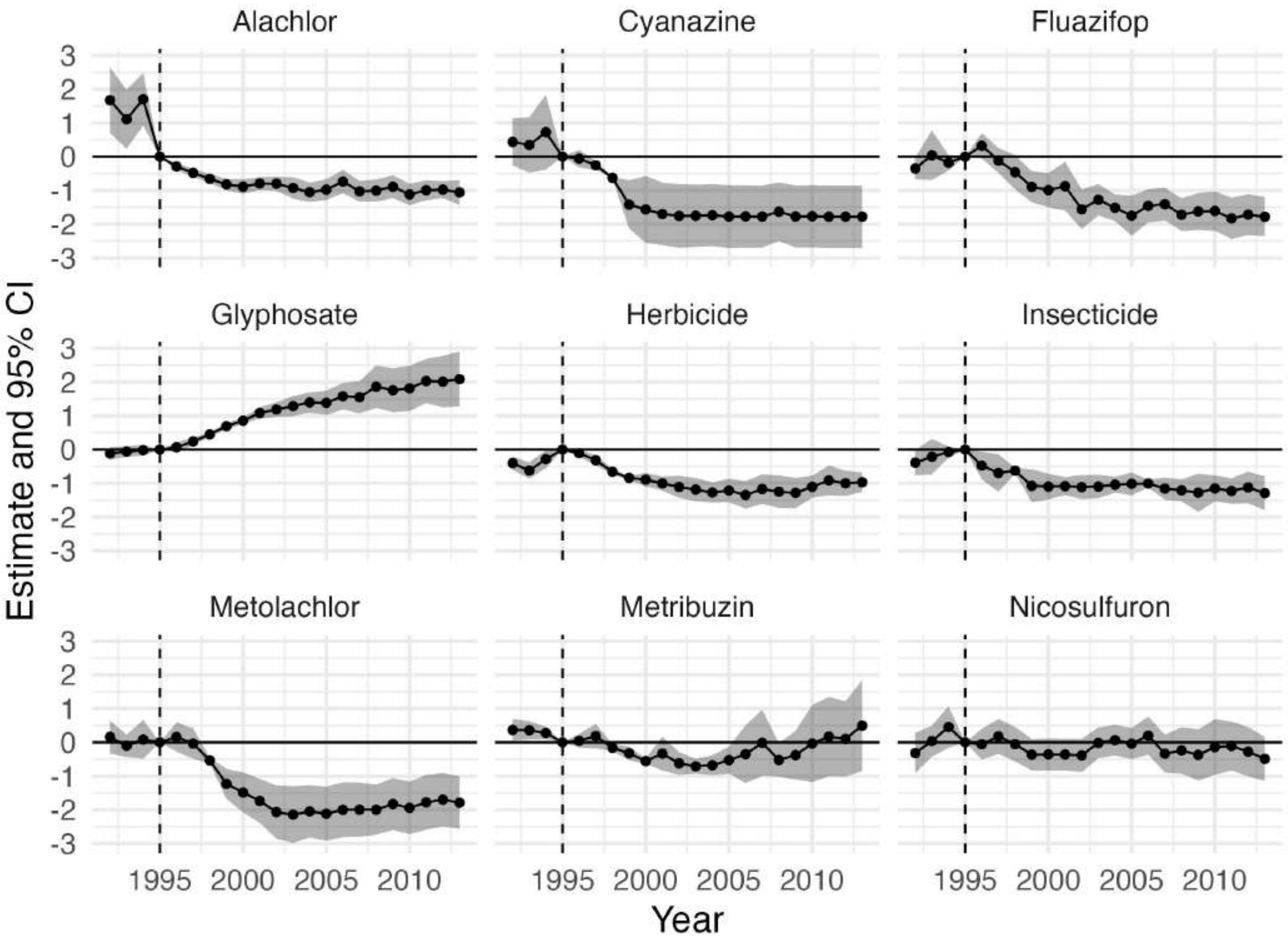


Fig. S4. Counties with high suitability for GM crops increased glyphosate intensity and reduced non-glyphosate pesticides with the introduction of glyphosate-resistant seeds. Each event study come from separate regressions where the given pesticide is regressed on local GM max attainable yield percentile interacted with year dummies with year and county fixed effects. All coefficients are scaled by the standard deviation of their respective variables. *Herbicide* and *Insecticide* each aggregate all other herbicides and insecticides not individually analyzed. Results from rural US counties. Standard errors are clustered by state and year. A unit of observation is county by year; regressions are weighted by total number of births.

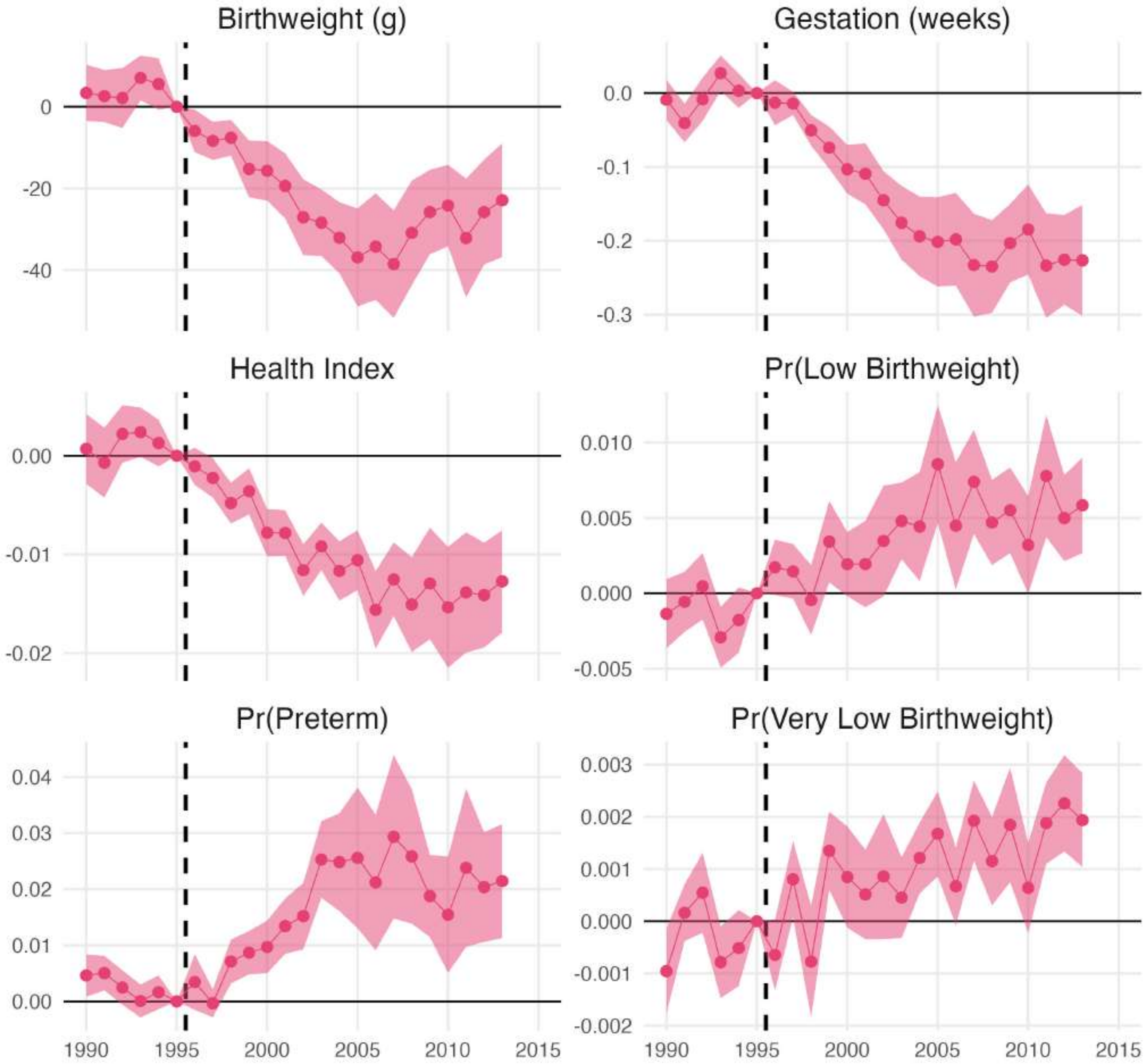


Fig. S5. Perinatal health declined in GM-crop suitable counties after the introduction of glyphosate-resistant seeds The subfigures extend Figure 2b to additional health outcomes—i.e., the estimated effect of local GM max attainable yield percentile on perinatal health outcomes relative to 1995. All regressions include county and year by month fixed effects and cluster errors by state and year. All regressions also control for family demographics, including mother's age, race, education, marital status, birth facility, resident status, previous births, sex of infant, and father's age and race. Sample restricted to births occurring in rural counties or to mothers residing in rural counties.

	BW	LBW	VLBW	Gestation	Preterm	C-section	Health Index
Panel A: Policy effect							
GLY/km^2	-843.5	0.167	0.047	-6.13	0.606	0.324	-0.398
	(329.8)	(0.079)	(0.017)	(1.73)	(0.223)	(0.208)	(0.119)
Controls							
(No additional controls)							
Panel B: GLY effect							
GLY/km^2	-1,280.7	0.279	0.087	-9.14	0.920	0.375	-0.549
	(496.5)	(0.103)	(0.029)	(2.47)	(0.309)	(0.374)	(0.195)
Controls							
Pesticides	Y	Y	Y	Y	Y	Y	Y
Fertilizers	Y	Y	Y	Y	Y	Y	Y
Employment	Y	Y	Y	Y	Y	Y	Y
Income	Y	Y	Y	Y	Y	Y	Y
Age Shares	Y	Y	Y	Y	Y	Y	Y
Race Shares	Y	Y	Y	Y	Y	Y	Y
Population	Y	Y	Y	Y	Y	Y	Y
Fixed-effects *(Both panels)*							
Family Demog.	Y	Y	Y	Y	Y	Y	Y
County	Y	Y	Y	Y	Y	Y	Y
Yr × Mo	Y	Y	Y	Y	Y	Y	Y
Summaries *(Both panels)*							
N (millions)	10.73	10.73	10.73	10.71	10.71	9.510	10.71
2012 mean	3,271.1	0.081	0.014	38.6	0.207	0.278	

Table S3. 2SLS estimates of the policy and direct GLY effects on perinatal health. Each coefficient estimate (column-panel combination) provides results from a separate 2SLS regression. The six outcomes are birthweight (BW), the probabilities of low birthweight (LBW; BW $<$ 2500g) and very low birthweight (VLBW; BW $<$ 2500g), gestation length, and the probability of a preterm birth (gestation $<$ 37 weeks). Both panels include family demographic, county, and year by month fixed effects. GLY effect (Panel B) additionally controls for other pesticides and unemployment. Sample restricted to births occurring in rural counties or to mothers residing in rural counties. Instruments are the attainable yield percentile for GM crops in each county interacted with year. Family demographic controls include mother's age, mother's race, mother's origin, mother's education, sex of child, total birth order, mother's residence status, and birth facility. Pesticide controls include alachlor, atrazine, cyanizine, fluazifop, metolachlor, metribuzin, and nicosulfuron. GLY/km^2 is kg/km^2. Standard errors in parentheses. We two-way cluster errors by year and state.

	Reduced-form DID results				*Two-stage results*	
	Raw reduced form		*(1–2) scaled by gly.*			
Outcome (*unit*)	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: *Health outcomes*						
Birthweight (*g*)	−29.42	−23.79	−29.2	−48.9	−19.6	−29.8
Gestation (*wk*)	−.148	−.119	−.146	−.244	−.143	−.213
LBW (*%pt*)	.507	.469	.502	.964	.399	.650
VLBW (*%pt*)	.111	.085	.110	.176	.110	.204
Preterm (*%pt*)	1.64	1.07	1.62	2.19	1.41	2.14
Health index	−.011	−.010	−.011	−.020	−.009	−.013
Panel B: *Glyphosate*						
Glyphosate (*kg/km^2*)	.024	.011				
Controls						
Ag. and econ.		Yes		Yes		Yes
Family demog.	Yes	Yes	Yes	Yes	Yes	Yes
County	Yes	Yes	Yes	Yes	Yes	Yes
Yr. × Mo.	Yes	Yes	Yes	Yes	Yes	Yes

Table S4. Comparing difference-in-differences and 2SLS results Sample restricts to births from mothers residing in a rural county. Instruments are the maximum attainable yield percentile for GM crops in each county (interacted with year in the 2SLS results in columns 5-6). All regressions include county and month-of-sample fixed effects and control for family and infant demographic controls (mother's age, mother's race, mother's origin, mother's education, sex of child, total birth order, mother's residence status, and birth facility). Glyphosate effect include additional *Ag. and econ.* controls (unemployment rate, employment rate, percent farm employment, and farm employment per capita, non-farm income per capita and farm income per capita, population, age shares, race/ethnicity shares, non-glyphosate pesticides, and fertilizer).

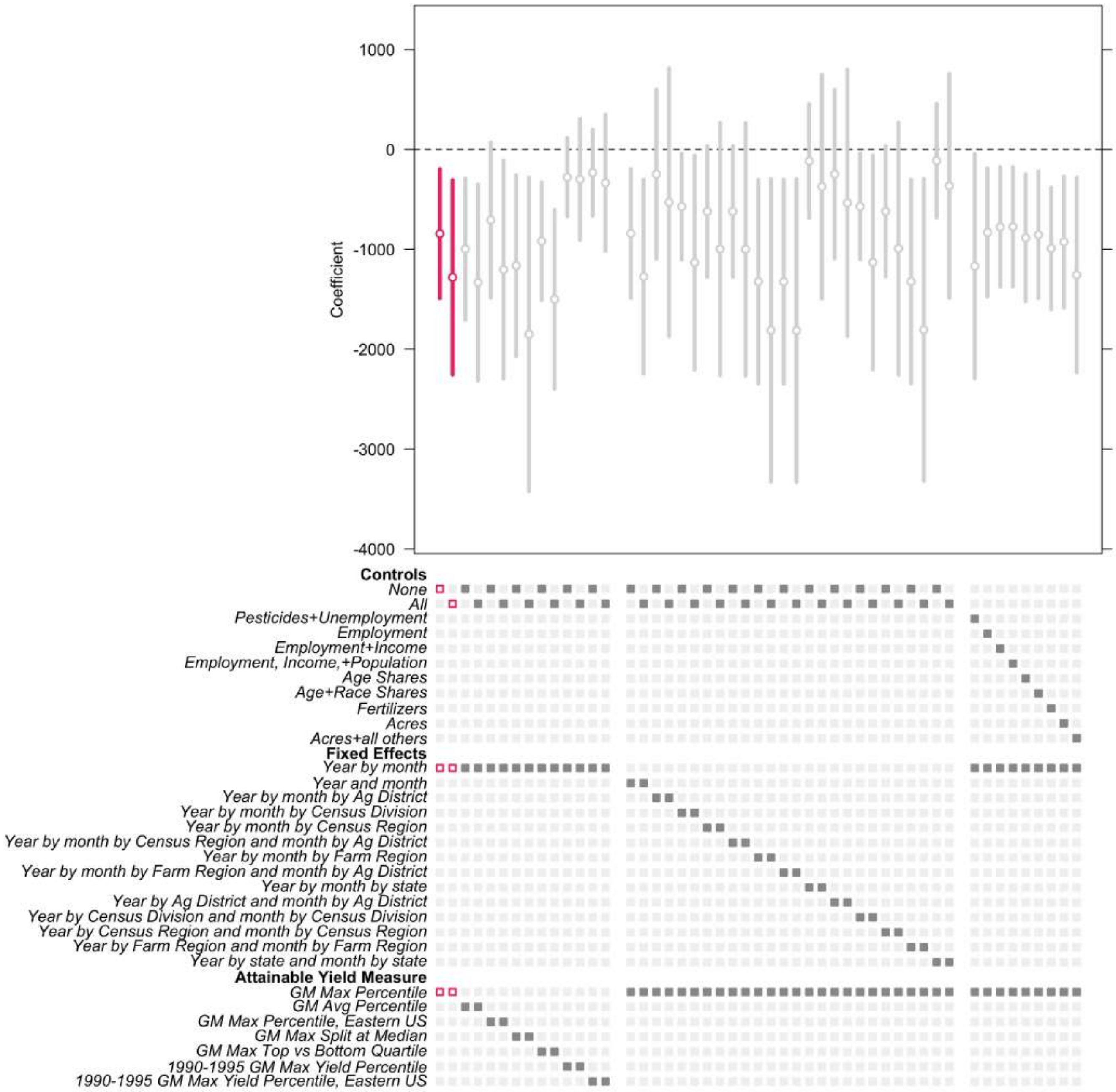


Fig. S6. The estimated effect of glyphosate on birthweight is robust to alternative specifications. Coefficients are the estimated marginal effect of glyphosate (kg/km^2) on birthweight. Our main specifications are highlighted. All regressions include county of residence, county of occurrence, and family demographic fixed effects, standard errors are clustered by state and year. Pesticide controls include alachlor, atrazine, cyanazine, fluazifop, metolachlor, metribuzin, and nicosulfuron. Employment controls include unemployment rate, employment rate, farm employment per capita, and farm employment share. Income controls include farm and nonfarm income per capita. Age shares controls are share of population in seven decade wide bins from ages 0 to 70, with over 70 as the omitted category. Race share controls are proportion of the population white, Black, and Hispanic. Fertilizer controls are commercial nitrogen, commercial phosphorous, manure nitrogen, and manure phosphorous. Acre controls are corn, soy, and cotton acres, as well as an aggregate of all other crop acreage. Family demographic FEs include mother's age, race, education, marital status, birth facility, resident status, previous births, sex of infant, and father's age and race. We vary the construction of GM attainable yield: "GM Max Percentile" is our main specification, "GM Avg Percentile" takes the average standardized attainable yield among corn, soy, and cotton (rather than the average) before re-scaling into a percentile, "GM Average, Split at Median" uses a binary high vs low attainable yield, where a county is high attainable yield if they are above the median attainable yield, "GM Max Top vs Bottom Quartile" is also binary, but only compares the top and bottom quartiles, omitting the middle group, and "1990-1995 GM Max Yield Percentile" is the percentile of observed yield in each county for corn, soy, and cotton between 1990 and 1995 using data from USDA NASS. "Eastern US" measures filter to counties east of the 100th meridian then calculate percentiles. Sample restricted to births occurring in rural counties or to mothers residing in rural counties.

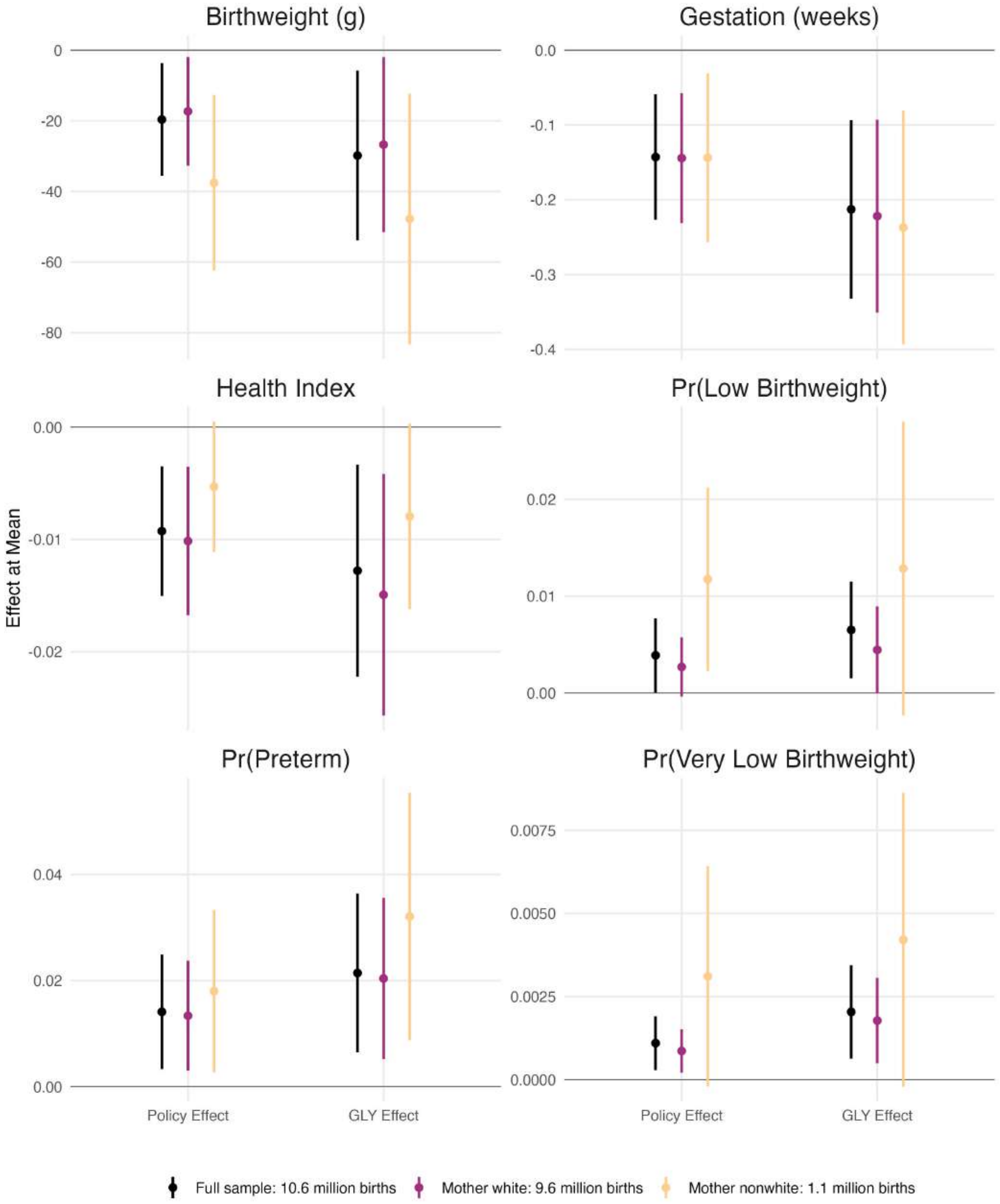


Fig. S7. glyphosate effects for infants born to non-white mothers are larger for birthweight and for the probabilities of preterm birth, LBW, and VLBW. Policy and Elyphosate effects for all outcomes at the mean level of glyphosate in 2012, estimated separately by mother's race. All regressions include county and year by month fixed effects, and control for family demographics. Standard errors are clustered by state and year. The Glyphosate Effect adds controls for other pesticides, employment, income, population, age and race shares, and fertilizers. The sample is restricted to births occurring in rural counties or to mothers residing in rural counties.

	(1)	(2)	(3)
Panel A: *Birthweight*			
Glyphosate/km^2	-843.5	-885.4	-855.0
	(329.8)	(324.8)	(323.2)
Panel B: *Gestation*			
Glyphosate/km^2	-6.13	-5.69	-5.65
	(1.73)	(1.67)	(1.66)
Panel C: *Health index*			
Glyphosate/km^2	-0.398	-0.361	-0.347
	(0.119)	(0.118)	(0.116)
Control sets			
Age Shares		Yes	Yes
Race Shares			Yes
Fixed-effects			
Family demog.	Yes	Yes	Yes
County	Yes	Yes	Yes
Yr. × Mo.	Yes	Yes	Yes

Table S5. Effect of glyphosate on birthweight, gestation, and health index: Robustness to age and race share controls Sample restricted to births from mothers residing in a rural county. Instruments are the maximum attainable yield percentile for GM crops in each county interacted with year. Family demographic controls include mother's age, mother's race, mother's origin, mother's education, sex of child, total birth order, mother's residence status, and birth facility. Age shares include the share of population in each county in seven 10 year age bins from age 0 to 70. We omit the over 70 category. Race shares include the share of the population in each county that is black, share white, and share hispanic. Glyphosate/km^2 is kg/km^2.

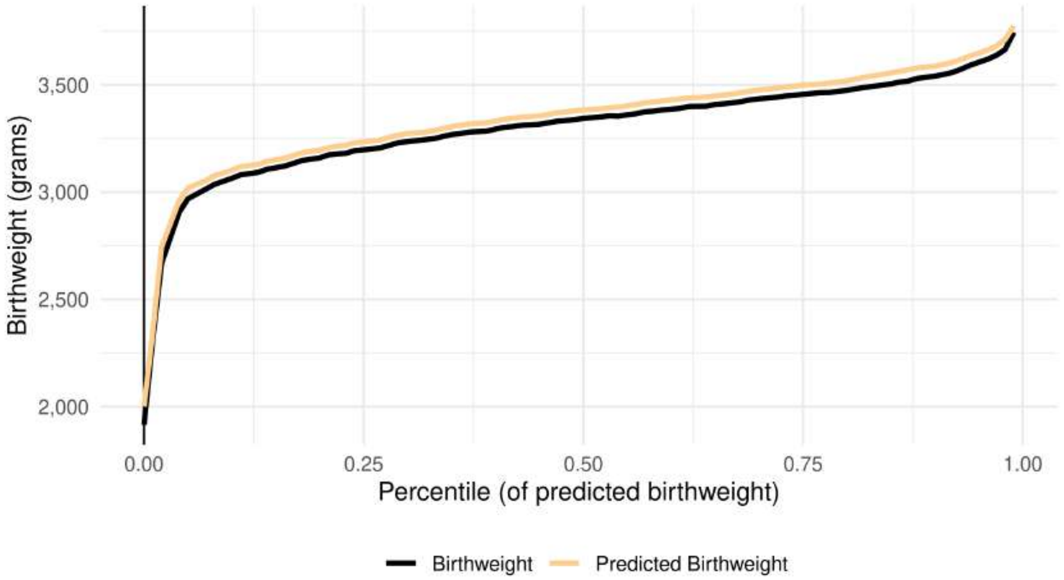
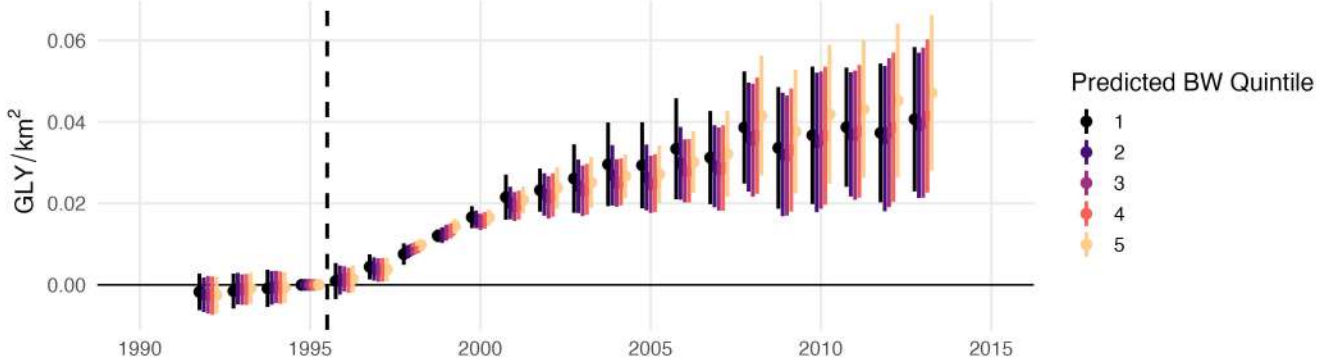


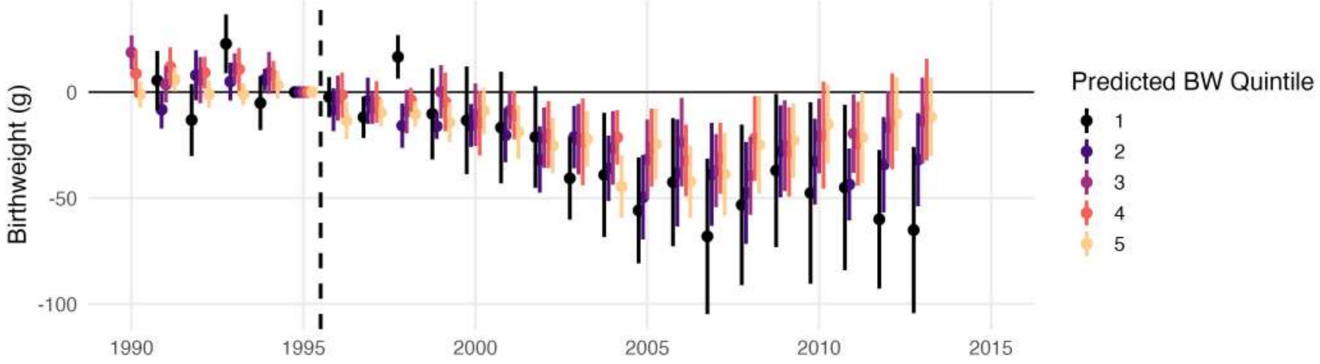
Fig. S8. Predicted birthweights closely match actual birthweights across the predicted birthweight distribution. At each predicted birthweight percentile (x-axis), we take the average actual birthwight and average predicted birthweight, which are both plotted in the y-axis. Sample includes births to mothers with rural residences from 1990 to 2013.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome (*units*):	**Birthweight** (*g*)		**Percentile** (*pp*)		**Decile** (*1–10*)		**Quntile** (*1–5*)	
Metric:	MAE	MAPE	MAE	MAPE	MAE	MAPE	MAE	MAPE
Panel A: *Actual predictions (across three samples)*								
Sample								
Full	406.7	15.7	26.1	290.7	2.58	79.0	1.25	59.1
1–99 pctl.	381.0	12.6	26.1	127.9	2.58	78.3	1.26	59.2
5–95 pctl.	329.0	10.3	26.1	91.7	2.59	73.0	1.27	58.8
Panel B: *Null models (on the full sample)*								
Approach								
Means	443.4	17.2	24.8	488.3	2.48	93.3	1.19	64.2
At random	643.6	23.4	33.3	324.5	3.30	106.1	1.60	77.0

Table S6. ML prediction performance *Panel A* describes the performance of the random-forest model across three samples—the full sample, births in percentiles 1–99, and births in percentiles 5–95. We evaluate the predictions on two metrics: mean absolute error (MAE: $|y - \hat{y}|$) and mean absolute percent error (MAPE: MAE$/y$). The model trained to predict birthweight (columns 1–2). We also evaluate the performance for predicting birthweight percentile (3–4), decile (5–6), and quintile (7–8). *Panel B* provides two *null* models that (1) predict the sample *means* or (2) predict *at random*. The table uses births to rural-residence mothers to match the main analyses throughout the paper.



(a) *First-stage* effect of GM suitability on glyphosate intensity by predicted birthweight quintile



(b) *Reduced-form* effect of GM suitability on birthweight by predicted birthweight quintile

Fig. S9. First-stage event study coefficients are similar across predicted BW quintiles, reduced form shows larger effects in lower quintiles. **(a)** Estimated event-study coefficients for the effect of local GM max attainable yield percentile on glyphosate by year relative to 1995 by predicted birthweight quintile. Pesticide data only go back to 1992—there are no coefficients in 1990–1991. **(b)** Similar event study but with birthweight as outcome. Estimates from each predicted birthweight quintile come from separate regressions. All regressions include county, year by month, and family demographic fixed effects. Standard errors are clustered by state and year. Family demographics include mother's age, race, education, marital status, birth facility, resident status, previous births, sex of infant, and father's age and race. Sample restricted to births occurring in rural counties or to mothers residing in rural counties.

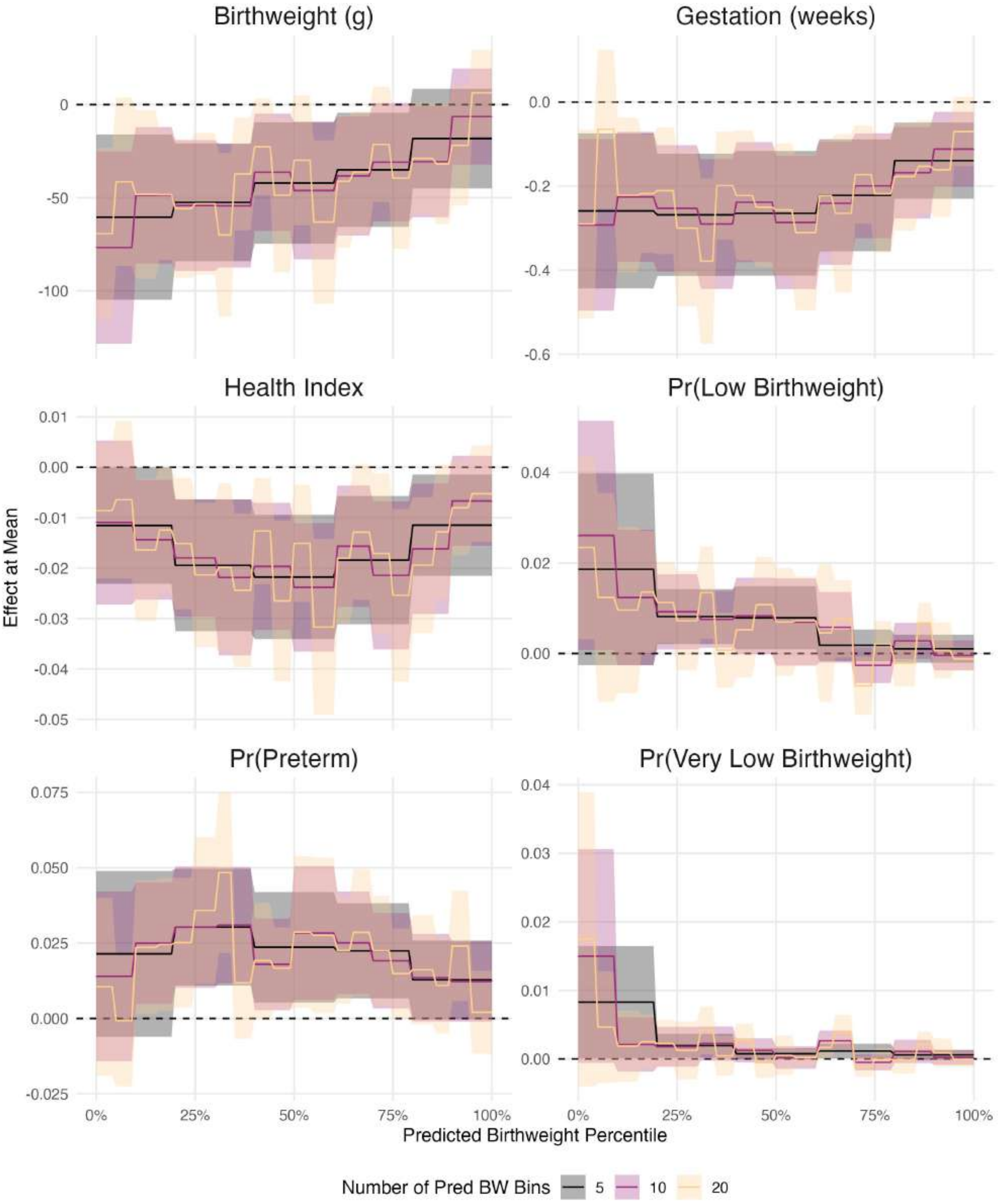


Fig. S10. Heterogeneity in Glyphosate Effect is consistent across various predicted birthweight bin sizes, greater disparities among birthweight outcomes. Estimated Glyphosate Effect at mean of glyphosate/km^2 on various perinatal health outcomes instrumented with GM attainable yield interacted with year. All regressions include county, year by month, and family demographic fixed effects, and control for other pesticides, employment, income, population, age and race shares, and fertilizers. Standard errors are clustered by state and year. Family demographics include mother's age, race, education, marital status, birth facility, resident status, previous births, sex of infant, and father's age and race. Sample restricted to births occurring in rural counties or to mothers residing in rural counties.

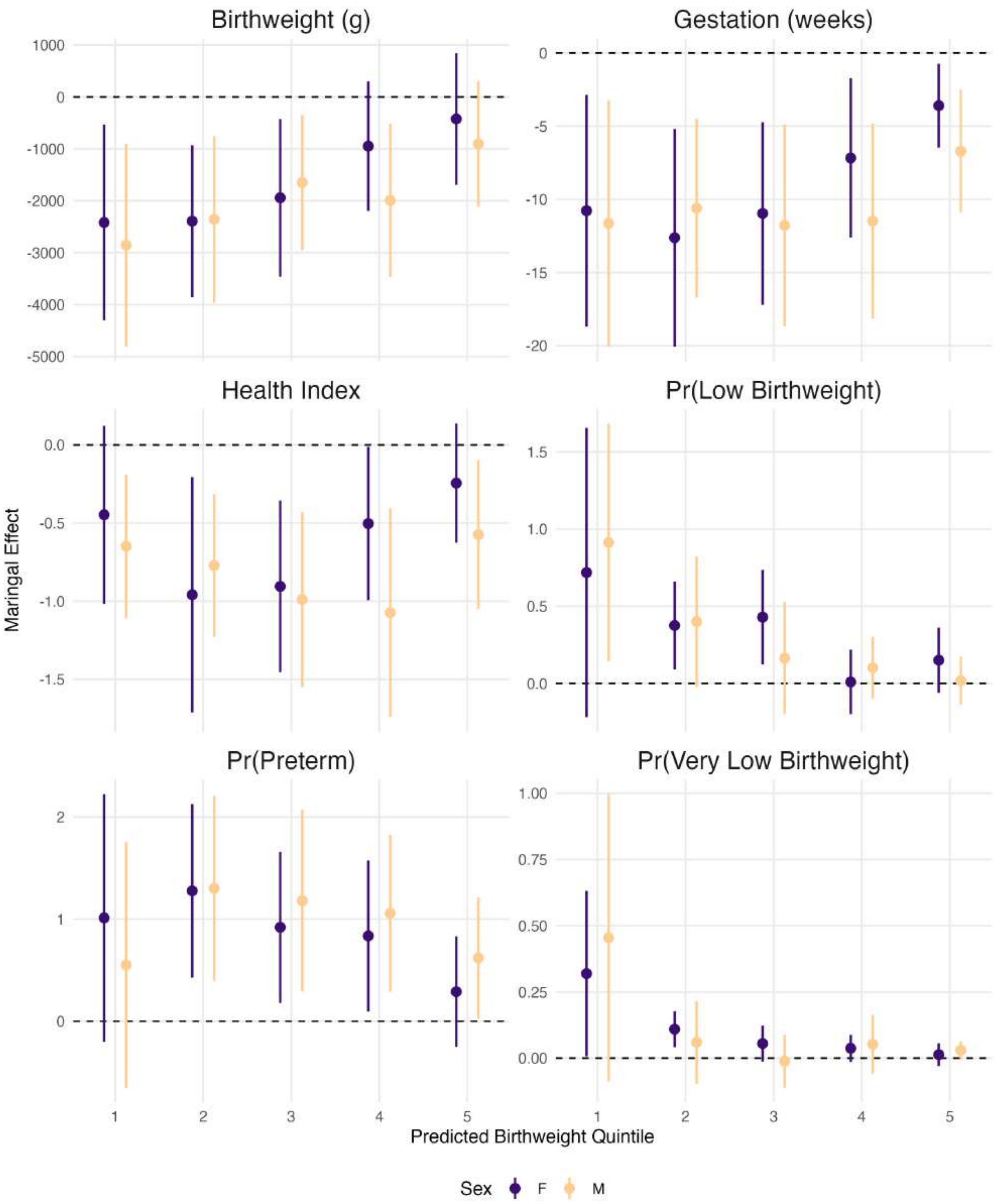


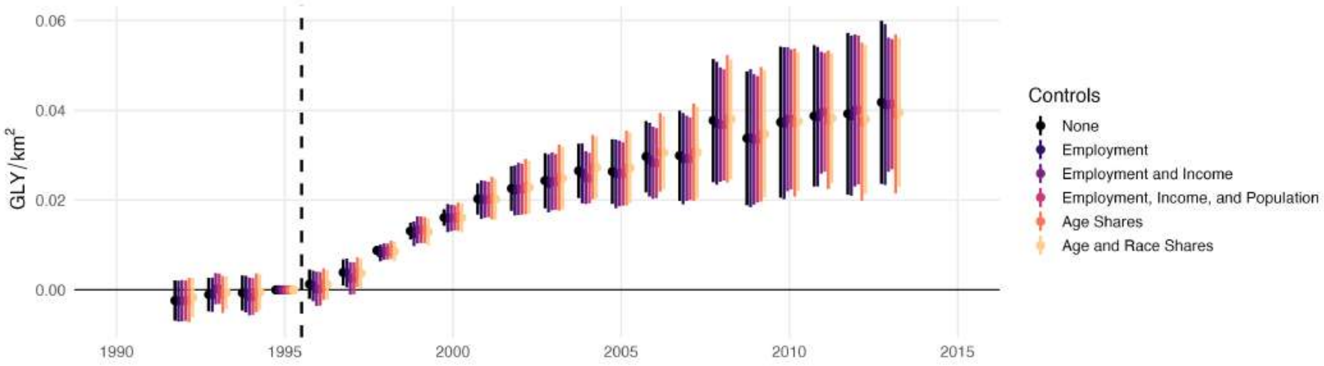
Fig. S11. Limited evidence of heterogeneous marginal effects by sex within predicted BW quintile. Estimated marginal effect of glyphosate/km^2 on various perinatal health outcomes instrumented with GM max attainable yield interacted with year. All regressions include county, year by month, and family demographic fixed effects and control for other pesticides, employment, income, population, age and race shares, and fertilizers. Standard errors are clustered by state and year. Family demographics include mother's age, race, education, marital status, birth facility, resident status, previous births, sex of infant, and father's age and race. Sample restricted to births occurring in rural counties or to mothers residing in rural counties.

	BW	LBW	VLBW	Gestation	Preterm	C-section	Health Index
Panel A: Policy effect							
GLY/km^2	24.6	-0.023	-0.003	-0.511	0.030	0.021	-0.047
	(75.9)	(0.022)	(0.006)	(0.373)	(0.050)	(0.068)	(0.029)
Controls							
(No additional controls)							
Panel B: GLY effect							
GLY/km^2	19.9	0.005	-0.002	-0.368	0.046	-0.011	-0.006
	(60.2)	(0.015)	(0.006)	(0.328)	(0.040)	(0.071)	(0.028)
Controls							
Pesticides	Y	Y	Y	Y	Y	Y	Y
Fertilizers	Y	Y	Y	Y	Y	Y	Y
Employment	Y	Y	Y	Y	Y	Y	Y
Income	Y	Y	Y	Y	Y	Y	Y
Age Shares	Y	Y	Y	Y	Y	Y	Y
Race Shares	Y	Y	Y	Y	Y	Y	Y
Population	Y	Y	Y	Y	Y	Y	Y
Fixed-effects *(Both panels)*							
Family Demog	Y	Y	Y	Y	Y	Y	Y
County	Y	Y	Y	Y	Y	Y	Y
Yr × Mo	Y	Y	Y	Y	Y	Y	Y
Summaries *(Both panels)*							
N obs. (millions)	10.73	10.73	10.73	10.71	10.71	9.51	10.73

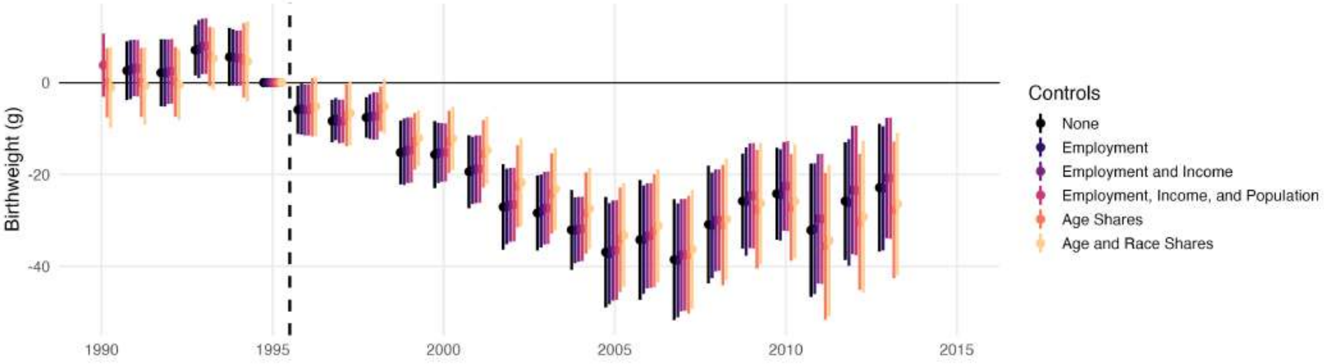
Table S7. OLS estimates of the policy and direct GLY effects on perinatal health. Each coefficient estimate (column-panel combination) provides results from a separate OLS regression. Both panels include family demographic, county, and year by month fixed effects. GLY effect (Panel B) additionally controls for other pesticides, employment, income, population, age and race shares, and fertilizers. Sample restricted to births occurring in rural countries or from mothers residing in rural counties. Family demographic controls include mother's age, mother's race, mother's origin, mother's education, sex of child, total birth order, mother's residence status, and birth facility. Pesticide controls include alachlor, atrazine, cyanizine, fluazifop, metolachlor, metribuzin, and nicosulfuron. GLY/km^2 is kg/km^2. Standard errors in parentheses. We two-way cluster errors by year and state.

	BW	LBW	VLBW	Gestation	Preterm	C-section	Health Index
Panel A: Policy effect							
GLY/km^2	-545.3	0.112	0.032	-4.06	0.418	0.174	-0.247
	(206.0)	(0.052)	(0.012)	(1.00)	(0.142)	(0.149)	(0.069)
Controls							
(No additional controls)							
Panel B: GLY effect							
GLY/km^2	-933.3	0.212	0.062	-6.21	0.664	0.100	-0.353
	(329.1)	(0.082)	(0.024)	(1.61)	(0.213)	(0.268)	(0.119)
Controls							
Pesticides	Y	Y	Y	Y	Y	Y	Y
Fertilizers	Y	Y	Y	Y	Y	Y	Y
Employment	Y	Y	Y	Y	Y	Y	Y
Income	Y	Y	Y	Y	Y	Y	Y
Age Shares	Y	Y	Y	Y	Y	Y	Y
Race Shares	Y	Y	Y	Y	Y	Y	Y
Population	Y	Y	Y	Y	Y	Y	Y
Fixed-effects *(Both panels)*							
Family Demog	Y	Y	Y	Y	Y	Y	Y
County	Y	Y	Y	Y	Y	Y	Y
Yr × Mo	Y	Y	Y	Y	Y	Y	Y
Summaries *(Both panels)*							
N obs. (millions)	10.73	10.73	10.73	10.71	10.71	9.51	10.73

Table S8. Effect of GLY on perinatal health estimated with 2SLS shift-share instrument. Each coefficient estimate (column-panel combination) provides results from a separate 2SLS regression. Instruments are the measure of suitability in each county interacted with national glyphosate usage, excluding glyphosate from counties within 100km or upstream. The six outcomes are birthweight (BW), the probabilities of low birthweight (LBW; BW $<$ 2500g) and very low birthweight (VLBW; BW $<$ 2500g), gestation length, and the probability of a preterm birth (gestation $<$ 37 weeks). Both panels include family demographic, county, and year by month fixed effects. GLY effect (Panel B) additionally controls for other pesticides and unemployment. Sample restricted to births occurring in rural counties or to mothers residing in rural counties. Instruments are the attainable yield percentile for GM crops in each county interacted with year. Family demographic controls include mother's age, mother's race, mother's origin, mother's education, sex of child, total birth order, mother's residence status, and birth facility. Pesticide controls include alachlor, atrazine, cyanizine, fluazifop, metolachlor, metribuzin, and nicosulfuron. GLY/km^2 is kg/km^2. Standard errors in parentheses. We two-way cluster errors by year and state.

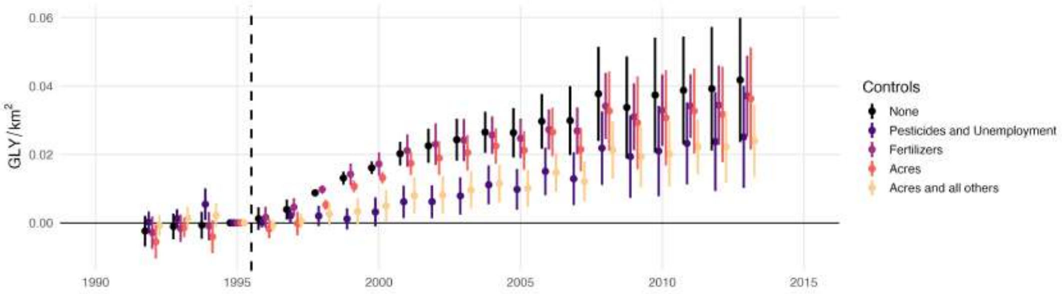


(a) First-stage effect of local GM attainable yield on glyphosate

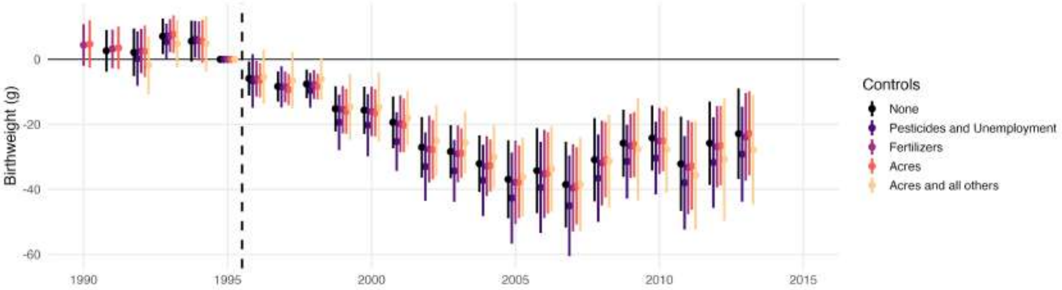


(b) Reduced-form effect of local GM attainable yield on birthweight

Fig. S12. Robustness of birthweight effect to alternative economic controls. Estimated effect of local GM max attainable yield percentile on birthweight relative to 1995. Employment controls include the unemployment rate, employment rate, farm employment per capita, and farm employment share. Income controls include farm and nonfarm income per capita. Age share controls are the share of population in seven decade wide age bins, with the over 70-population as the reference group. Race shares are the proportion of the population white, Hispanic, and Black. All regressions include fixed effects for family demographics, county, and year-month and standard errors are clustered by state and year. Family demographics include mother's age, race, education, marital status, birth facility, resident status, previous births, sex of infant, and father's age and race. Sample restricted to births occurring in rural counties or to mothers residing in rural counties.

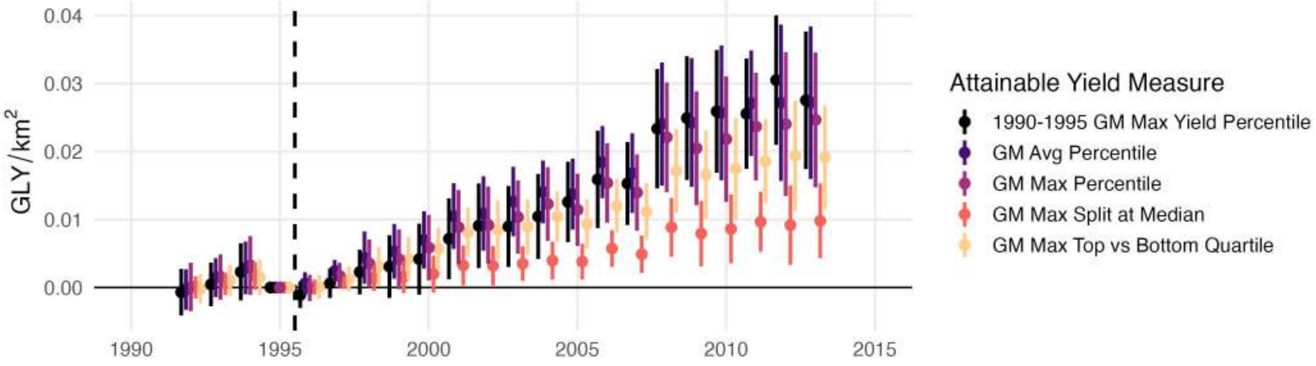


(a) First-stage effect of local GM attainable yield on glyphosate

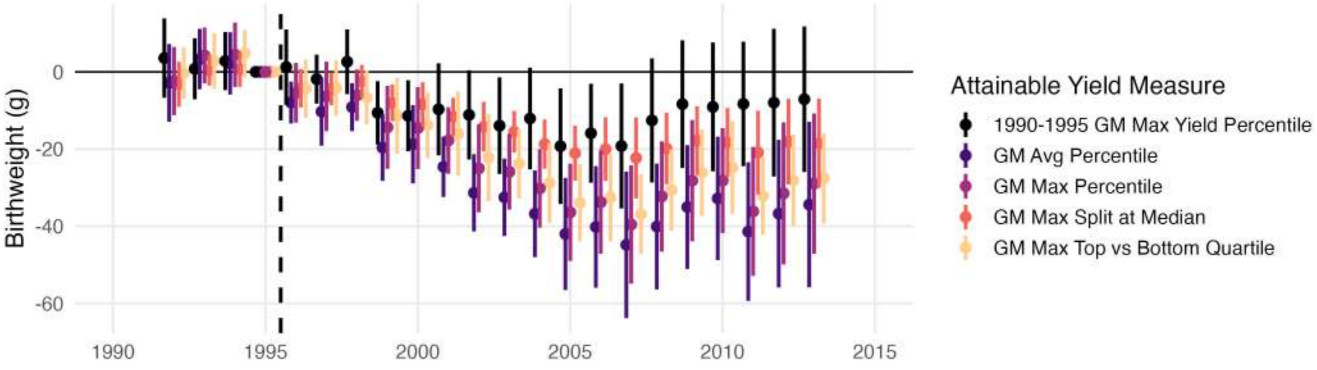


(b) Reduced-form effect of local GM attainable yield on birthweight

Fig. S13. Robustness of birthweight effect to alternative farm controls. Estimated effect of local GM max attainable yield percentile on birthweights relative to 1995. Fertilizer controls include commerical nitrogen, commercial phosphorous, manure nitrogen, and manure phosphorous. Acre controls include soy, corn, and cotton acres, as well as an aggregate of total acreage from other crops. "Acres and all others" specification uses all of the economic controls—pesticides, fertilizers, acres, employment, income, age and race shares, and population. All regressions include fixed effects for family demographics, county, and year-month and standard errors are clustered by state and year. Family demographics include mother's age, race, education, marital status, birth facility, resident status, previous births, sex of infant, and father's age and race. Sample restricted to births occurring in rural counties or to mothers residing in rural counties.

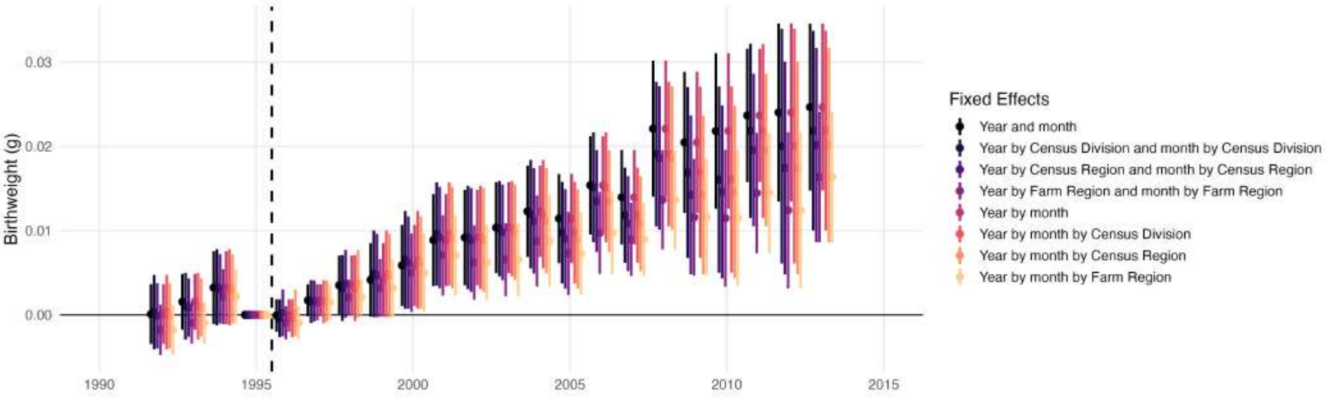


(a) First-stage effect of various instruments on glyphosate/km^2

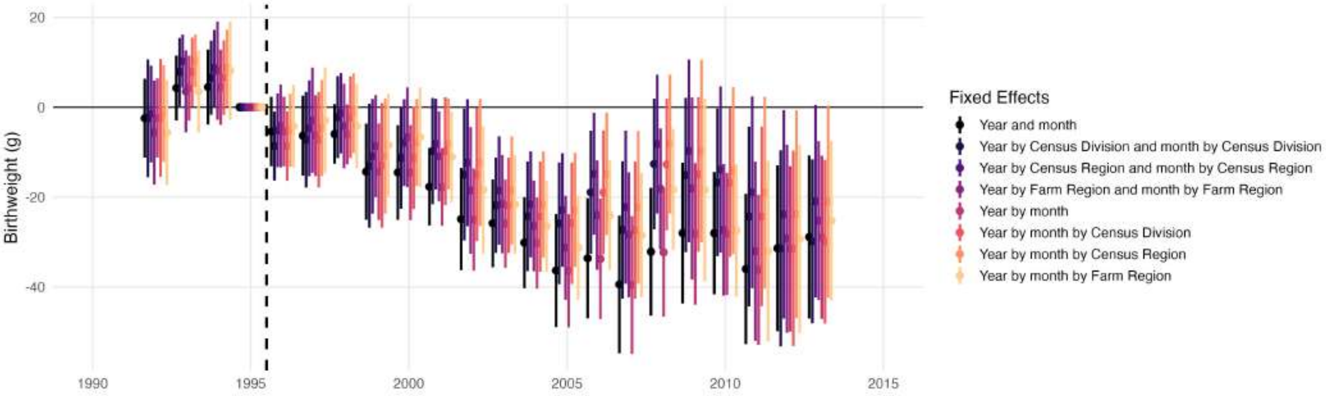


(b) Reduced-form effect of various instruments on birthweight

Fig. S14. Robustness of birthweight effect to alternative instruments. We vary the construction of our GM suitability measure: "GM Max Percentile" is our main specification, "GM Avg Percentile" takes the average standardized attainable yield among corn, soy, and cotton (rather than the max) before re-scaling into a percentile, "1990-1995 GM Max Yield Percentile" is constructed using pre-period realized yields for corn, soy, and cotton, "GM Average, Split at Median" uses a binary high vs low attainable yield, where a county is high attainable yield if they are above the median attainable yield, and "GM Max Top vs Bottom Quartile" is another binary treatment definition that compares just the top and bottom quartiles in GM Max Percentile, omitting the middle group. All regressions include family demographics, county, and year by month fixed effects and standard errors are clustered by state and year. The regressions control for fertilizers, other pesticides, employment, income, age and race shares, and population. Family demographic fixed effects include mother's age, race, education, marital status, birth facility, resident status, previous births, sex of infant, and father's age and race. Sample restricted to births occurring in rural counties or to mothers residing in rural counties.

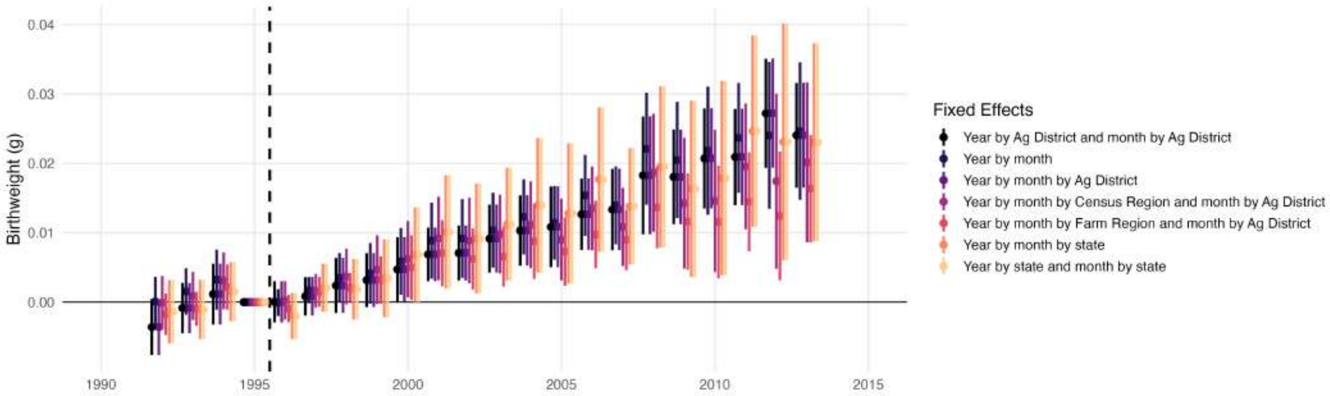


(a) First-stage effect of various instruments on glyphosate kg/km^2

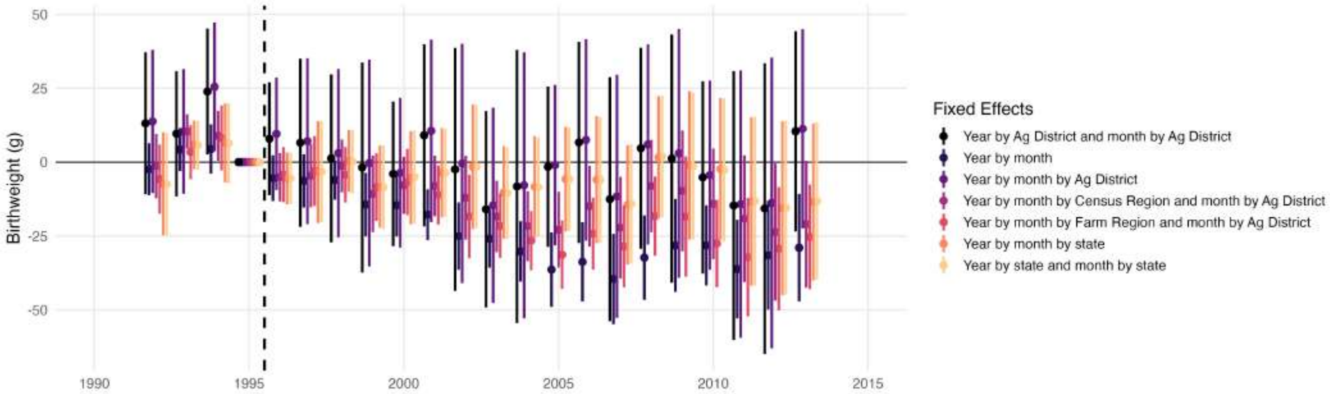


(b) Reduced-form effect of various instruments on birthweight

Fig. S15. Robustness of birthweight effect to alternative fixed effects. We vary the fixed effects included. All regressions include family demographics, county, and year by month fixed effects and standard errors are clustered by state and year. The regressions control for fertilizers, other pesticides, employment, income, age and race shares, and population. Family demographic fixed effects include mother's age, race, education, marital status, birth facility, resident status, previous births, sex of infant, and father's age and race. Sample restricted to births occurring in rural counties or to mothers residing in rural counties.

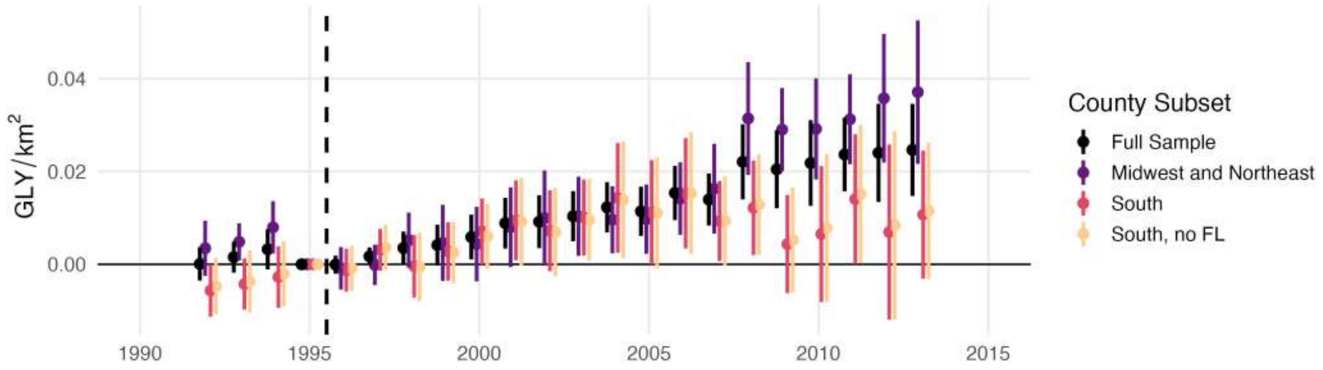


(a) First-stage effect of various instruments on glyphosate kg/km^2

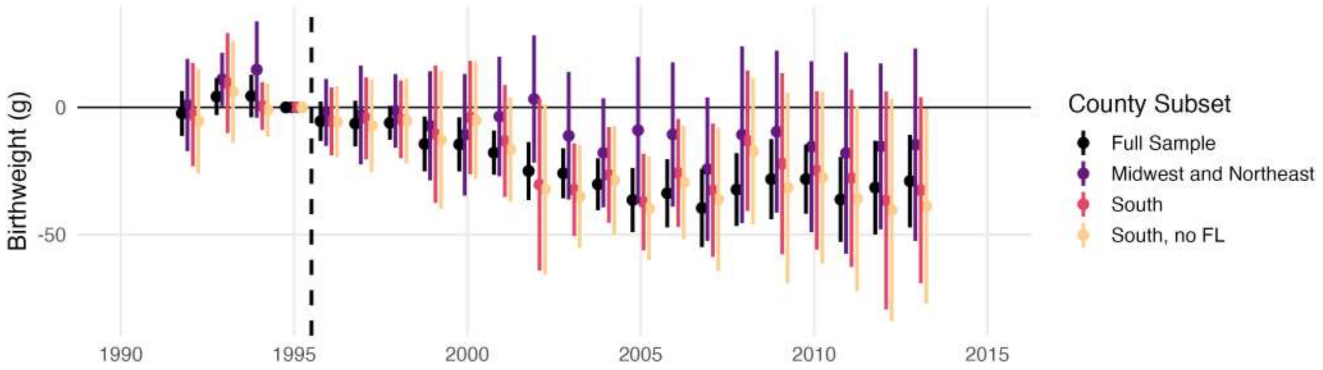


(b) Reduced-form effect of various instruments on birthweight

Fig. S16. Robustness of birthweight effect to alternative fixed effects—within state. We vary the fixed effects included. All regressions include family demographics, county, and year by month fixed effects and standard errors are clustered by state and year. The regressions control for fertilizers, other pesticides, employment, income, age and race shares, and population. Family demographic fixed effects include mother's age, race, education, marital status, birth facility, resident status, previous births, sex of infant, and father's age and race. Sample restricted to births occurring in rural counties or to mothers residing in rural counties.

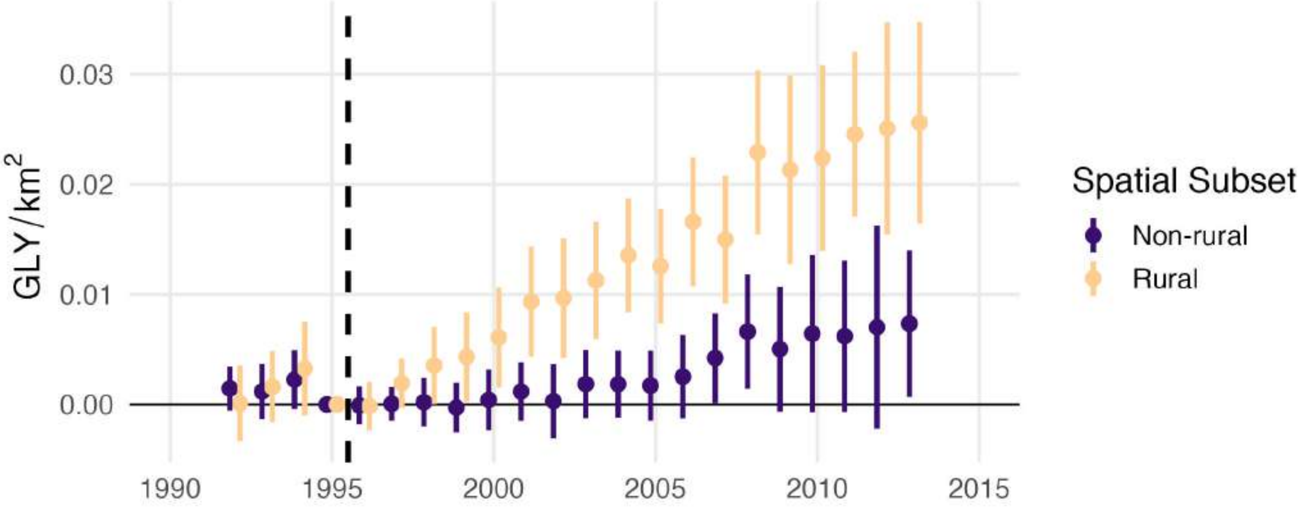


(a) First-stage effect on glyphosate kg/km^2

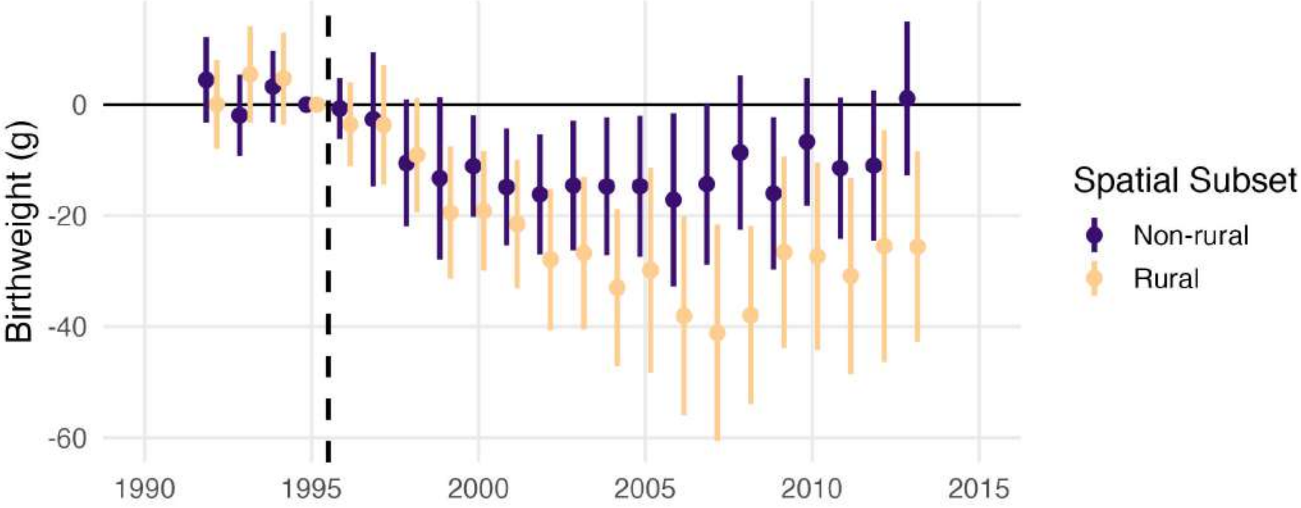


(b) Reduced-form effect on birthweight

Fig. S17. Heterogeneity in Birthweight Effect by Geographic Subsets. Estimated effect of local max GM attainable yield percentile on birthweight relative to 1995. The geographic subsets are primarily defined using census regions (Midwest, Northeast, or South). Fig S14 shows results with just the eastern US. All regressions include county and year by month fixed effects and standard errors are clustered by state and year. All regressions also control for other pesticides, employment, income, population, age and race shares, fertilizers and family demographics, including mother's age, race, education, marital status, birth facility, resident status, previous births, sex of infant, and father's age and race. Sample restricted to births occurring in rural counties or to mothers residing in rural counties.



(a) First-stage effect on glyphosate kg/km^2, rural and non-rural counties.



(b) Reduced-form effect on birthweight, rural and non-rural counties.

Fig. S18. Birthweight event studies by rural and non-rural counties. Estimated effect of local max GM attainable yield percentile on birthweight relative to 1995 for births to mothers residing and occurring in rural and non-rural counties. All regressions include county and year by month fixed effects and standard errors are clustered by state and year. All regressions also control for other pesticides, employment, income, population, age and race shares, and fertilizers, and family demographics, including mother's age, race, education, marital status, birth facility, resident status, previous births, sex of infant, and father's age and race.

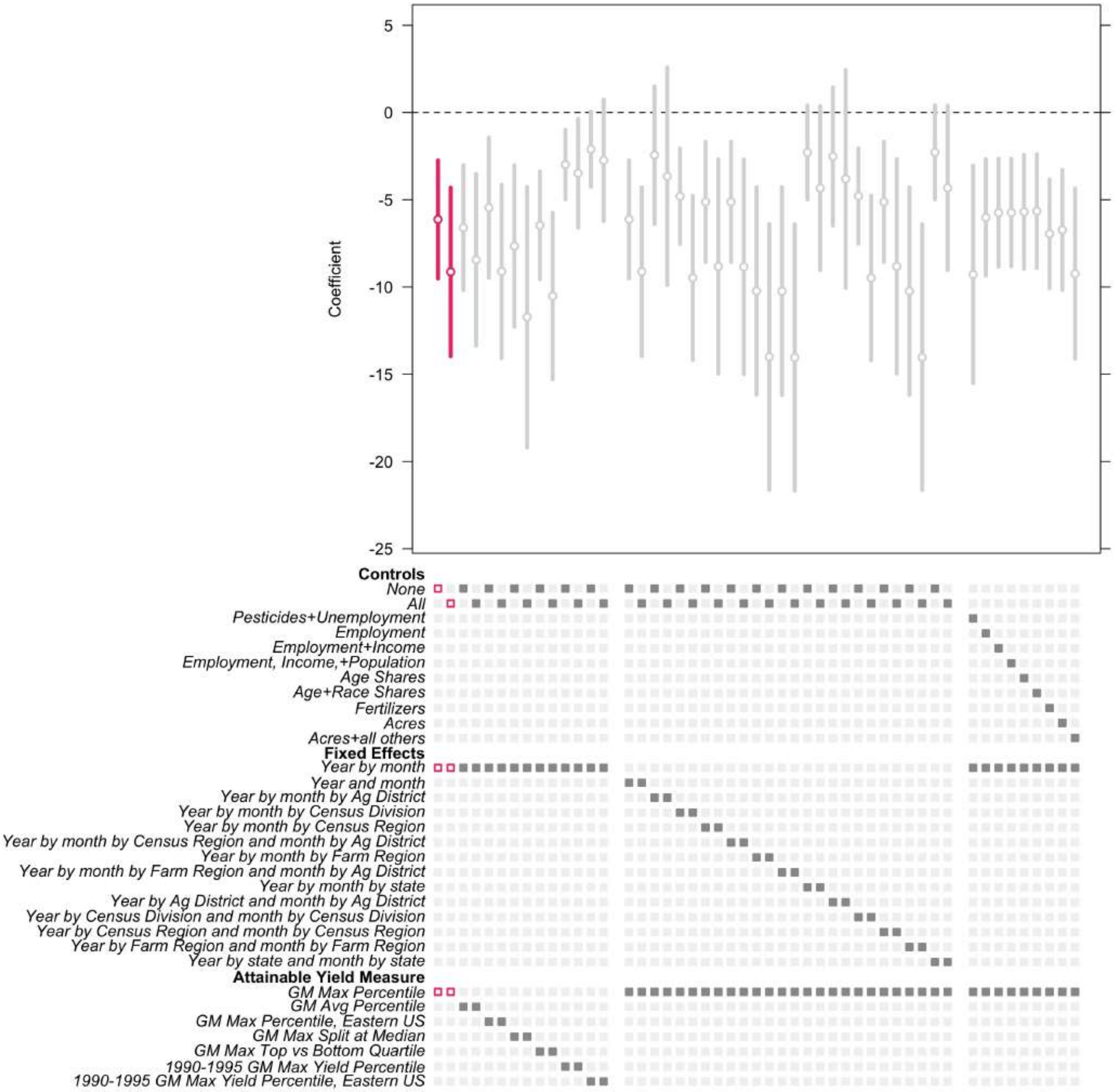


Fig. S19. The estimated effect of glyphosate on gestation is robust to alternative specifications. Coefficients are the estimated marginal effect of glyphosate (kg/km^2) on gestation. Our main specifications are highlighted. All regressions include county of residence, county of occurrence, and family demographic fixed effects, standard errors are clustered by state and year. Pesticide controls include alachlor, atrazine, cyanazine, fluazifop, metolachlor, metribuzin, and nicosulfuron. Employment controls include unemployment rate, employment rate, farm employment per capita, and farm employment share. Income controls include farm and nonfarm income per capita. Age shares controls are share of population in seven decade wide bins from ages 0 to 70, with over 70 as the omitted category. Race share controls are proportion of the population white, Black, and Hispanic. Fertilizer controls are commercial nitrogen, commercial phosphorous, manure nitrogen, and manure phosphorous. Acre controls are corn, soy, and cotton acres, as well as an aggregate of all other crop acreage. Family demographic FEs include mother's age, race, education, marital status, birth facility, resident status, previous births, sex of infant, and father's age and race. We vary the construction of GM attainable yield: "GM Max Percentile" is our main specification, "GM Avg Percentile" takes the average standardized attainable yield among corn, soy, and cotton (rather than the average) before re-scaling into a percentile, "GM Average, Split at Median" uses a binary high vs low attainable yield, where a county is high attainable yield if they are above the median attainable yield, "GM Max Top vs Bottom Quartile" is also binary, but only compares the top and bottom quartiles, omitting the middle group, and "1990-1995 GM Max Yield Percentile" is the percentile of observed yield in each county for corn, soy, and cotton between 1990 and 1995 using data from USDA NASS. "Eastern US" measures filter to counties east of the 100th meridian then calculate percentiles. Sample restricted to births occurring in rural counties or to mothers residing in rural counties.

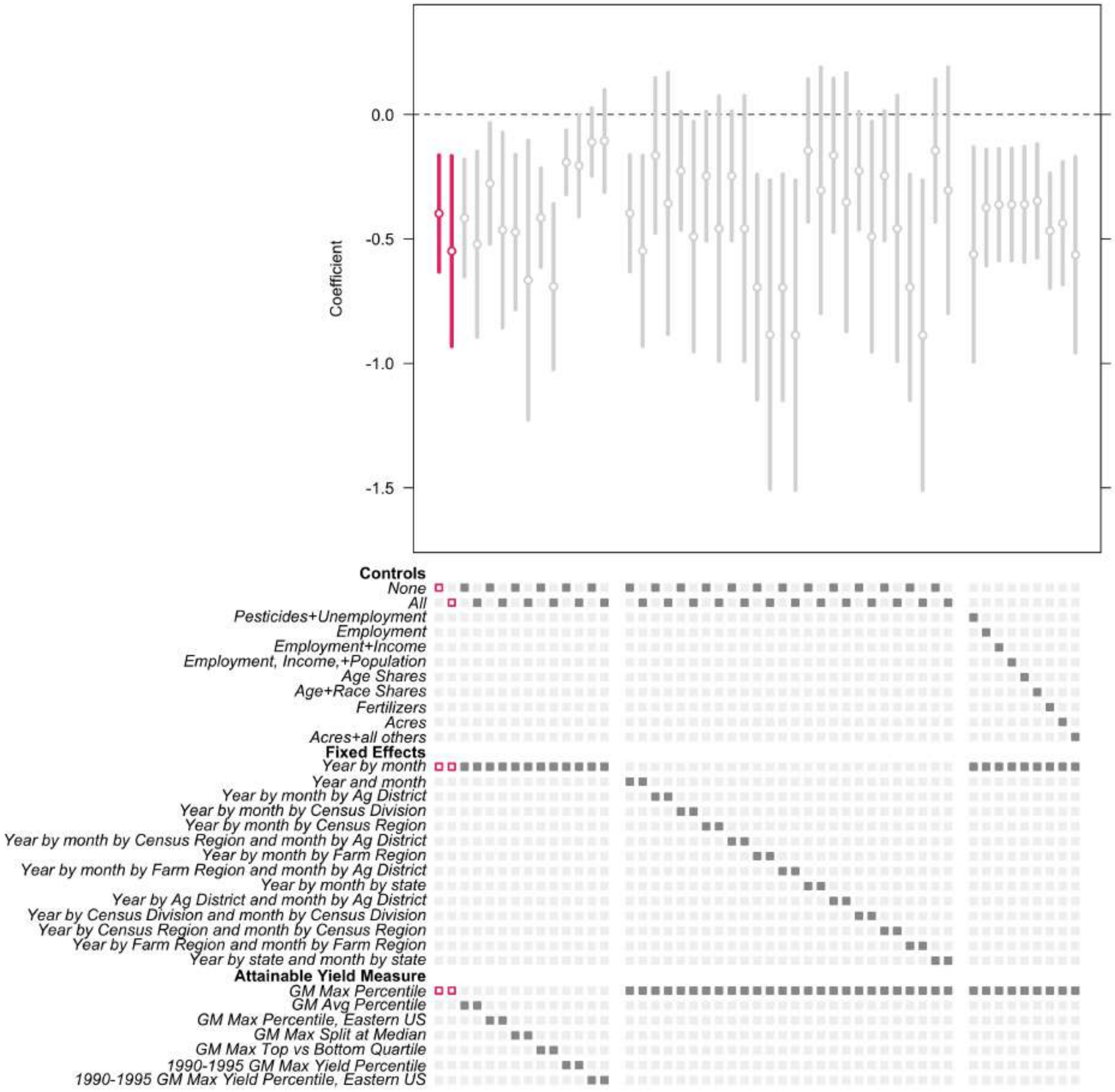


Fig. S20. The estimated effect of glyphosate on the health index is robust to alternative specifications. Coefficients are the estimated marginal effect of glyphosate (kg/km^2) on the health index. Our main specifications are highlighted. All regressions include county of residence, county of occurrence, and family demographic fixed effects, standard errors are clustered by state and year. Pesticide controls include alachlor, atrazine, cyanazine, fluazifop, metolachlor, metribuzin, and nicosulfuron. Employment controls include unemployment rate, employment rate, farm employment per capita, and farm employment share. Income controls include farm and nonfarm income per capita. Age shares controls are share of population in seven decade wide bins from ages 0 to 70, with over 70 as the omitted category. Race share controls are proportion of the population white, Black, and Hispanic. Fertilizer controls are commercial nitrogen, commercial phosphorous, manure nitrogen, and manure phosphorous. Acre controls are corn, soy, and cotton acres, as well as an aggregate of all other crop acreage. Family demographic FEs include mother's age, race, education, marital status, birth facility, resident status, previous births, sex of infant, and father's age and race. We vary the construction of GM attainable yield: "GM Max Percentile" is our main specification, "GM Avg Percentile" takes the average standardized attainable yield among corn, soy, and cotton (rather than the average) before re-scaling into a percentile, "GM Average, Split at Median" uses a binary high vs low attainable yield, where a county is high attainable yield if they are above the median attainable yield, "GM Max Top vs Bottom Quartile" is also binary, but only compares the top and bottom quartiles, omitting the middle group, and "1990-1995 GM Max Yield Percentile" is the percentile of observed yield in each county for corn, soy, and cotton between 1990 and 1995 using data from USDA NASS. "Eastern US" measures filter to counties east of the 100th meridian then calculate percentiles. Sample restricted to births occurring in rural counties or to mothers residing in rural counties.

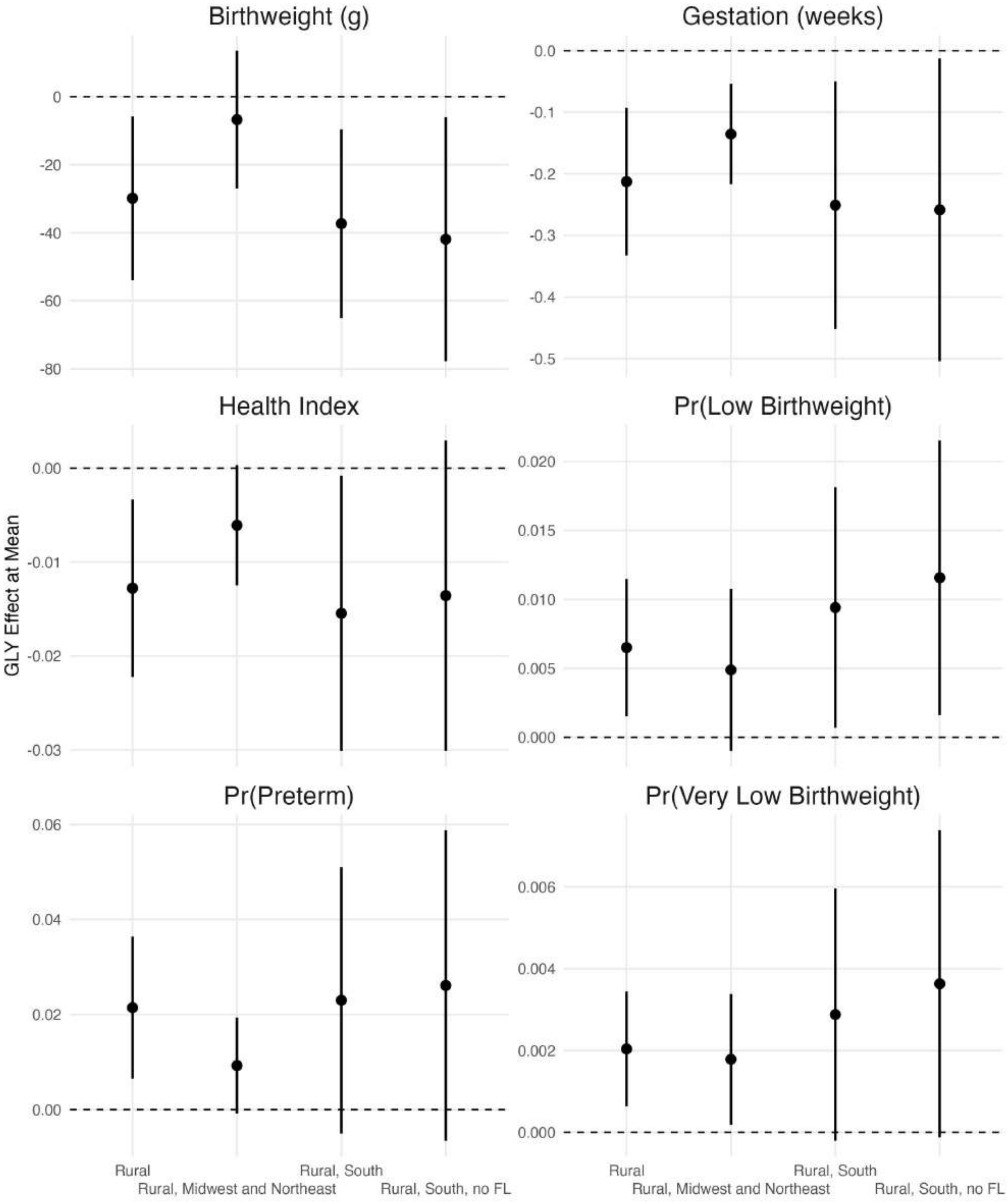


Fig. S21. Robustness to spatial subsets, all outcomes. Estimated effect of glyphosate/km^2 on various perinatal health outcomes instrumented with GM max attainable yield interacted with year. All regressions include county, year by month, and family demographic fixed effects and control for other pesticides, employment, income, population, age and race shares, and fertilizers. Standard errors are clustered by state and year. Family demographics include mother's age, race, education, marital status, birth facility, resident status, previous births, sex of infant, and father's age and race.

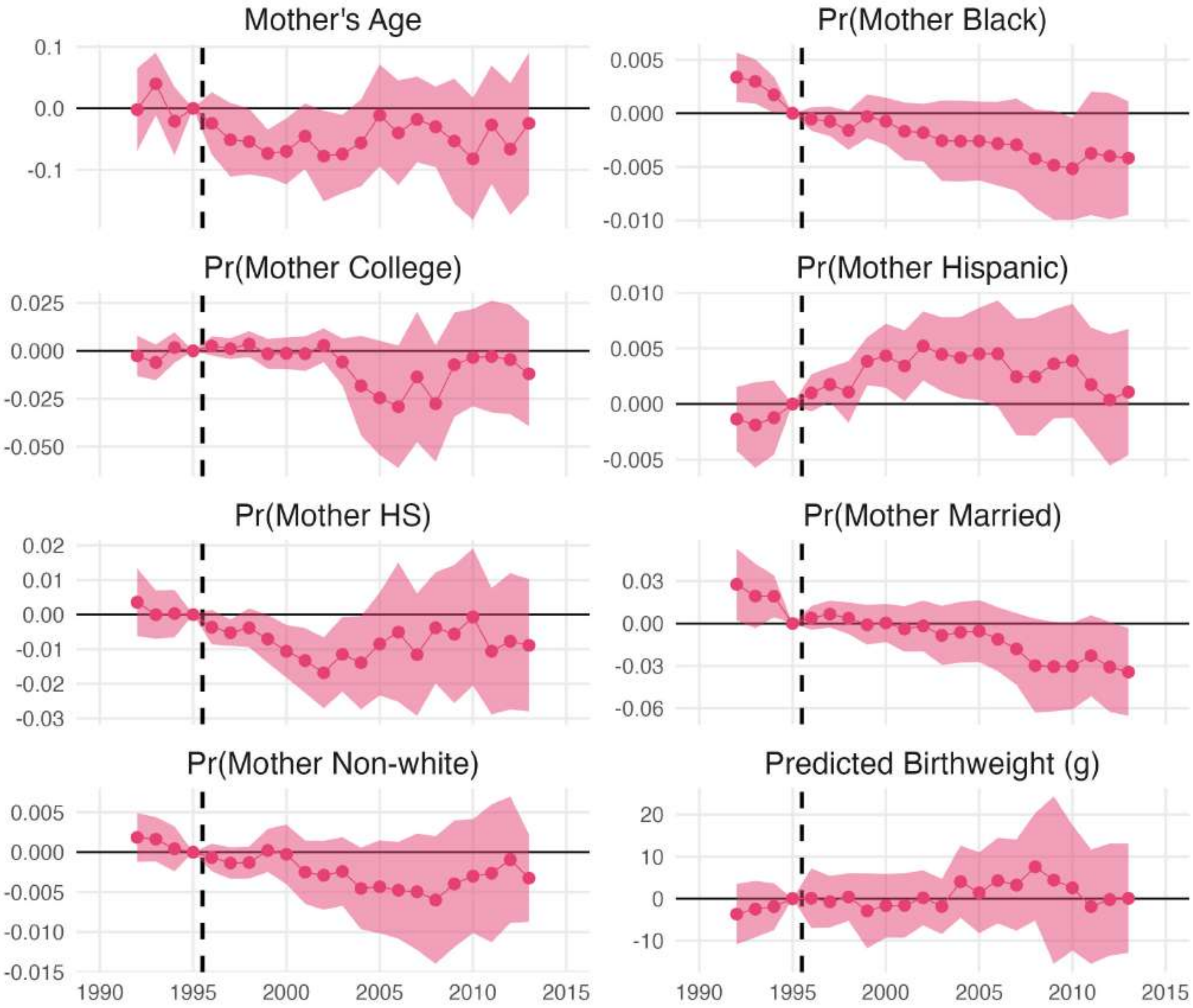


Fig. S22. Demographic event studies. Estimated effect of local GM max attainable yield percentile on various demographics outcomes relative to 1995. All regressions include county and year by month fixed effects and control for family demographics, including mother's age, race, education, marital status, birth facility, resident status, previous births, sex of infant, and father's age and race (controls exclude the outcome). Standard errors are clustered by state and year. Sample restricted to births occurring in rural counties or to mothers residing in rural counties.

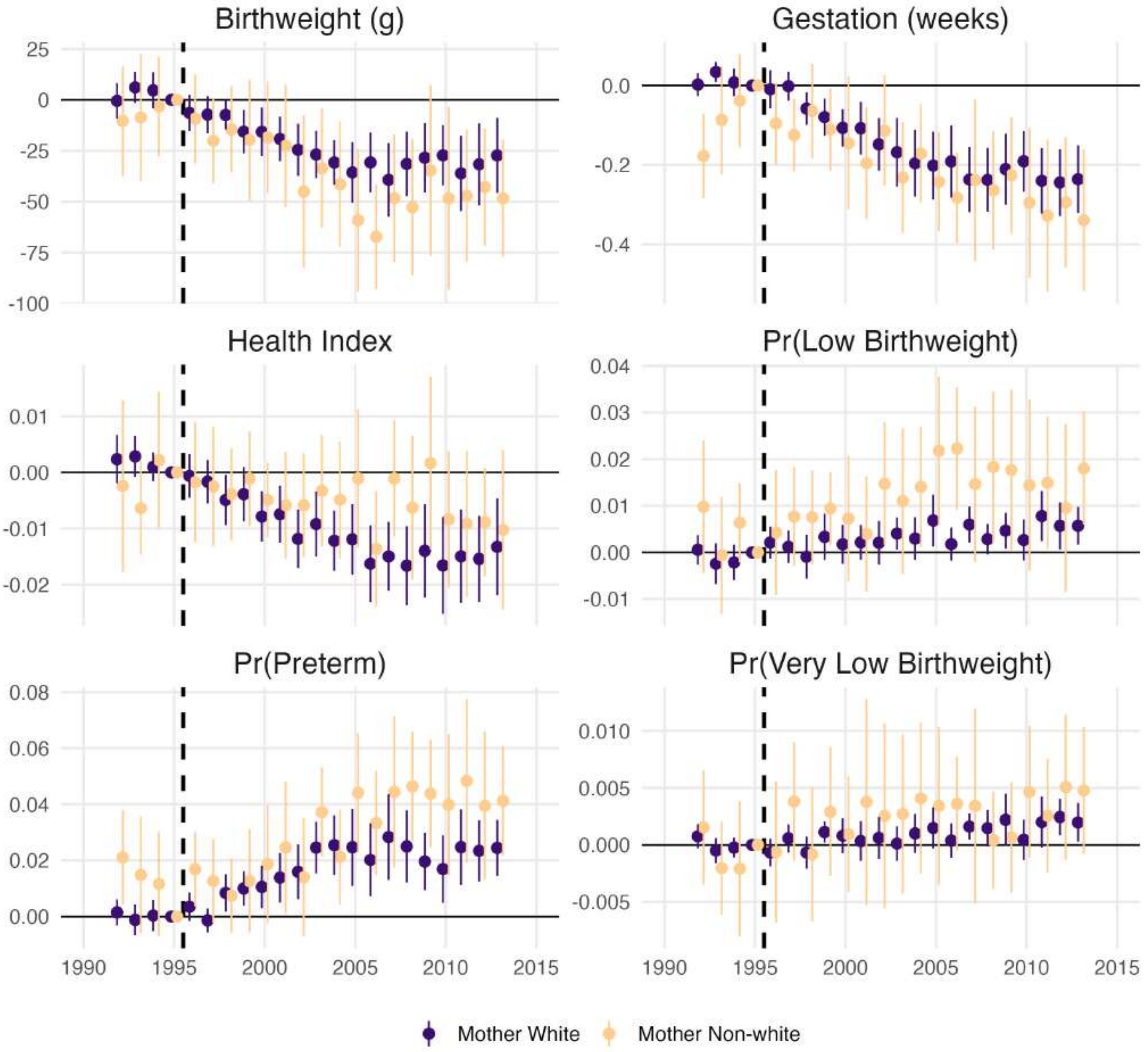


Fig. S23. Reduced form heterogeneity by mother's race. Estimated effect of local GM max attainable yield percentile on various perinatal health outcomes relative to 1995. All regressions include county and year by month fixed effects and standard errors are clustered by state and year. All regressions also control for other pesticides, employment, income, population, age and race shares, fertilizers, and family demographics, including mother's age, race, education, marital status, birth facility, resident status, previous births, sex of infant, and father's age and race. Sample restricted to births occurring in rural counties or to mothers residing in rural counties.

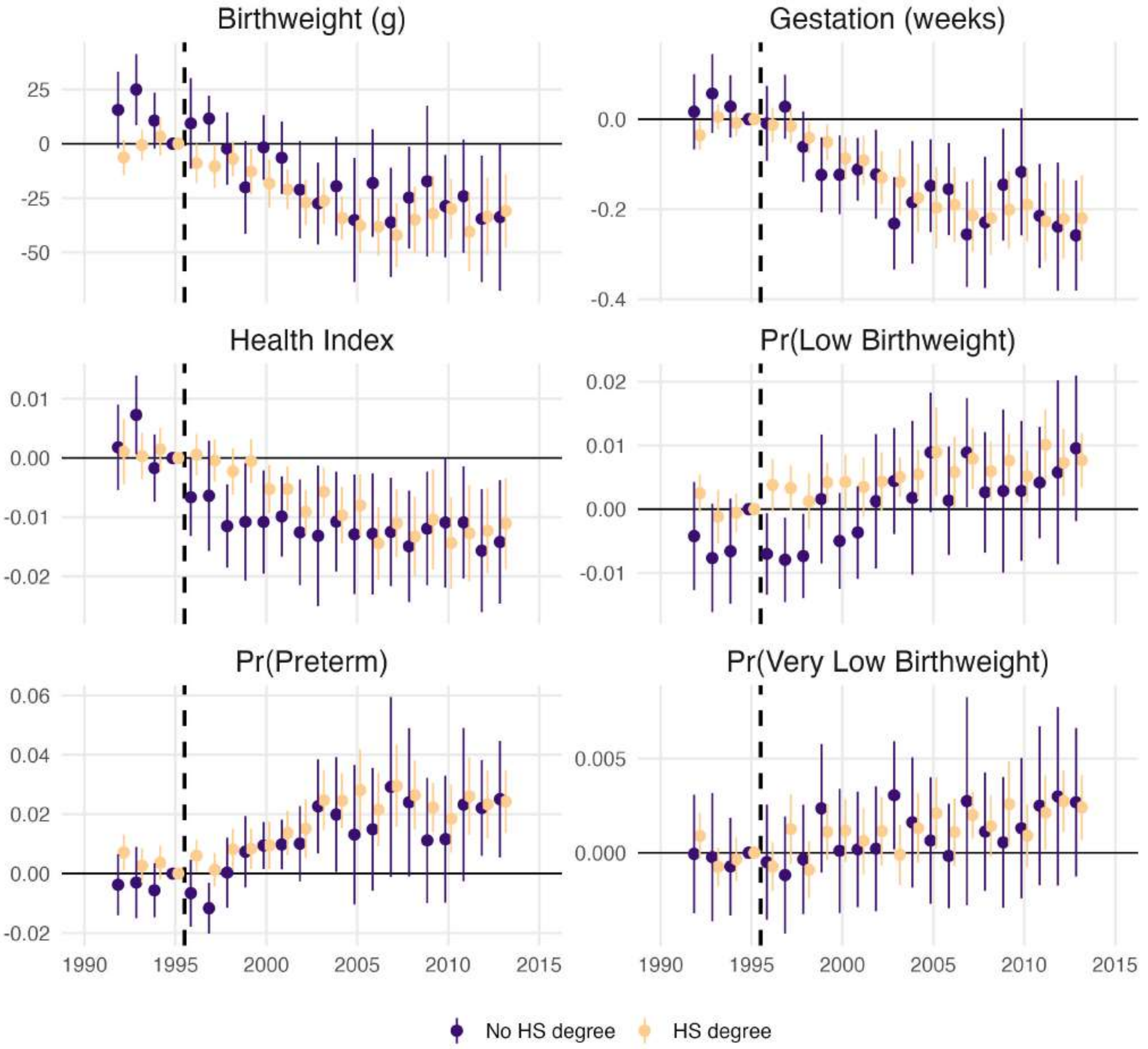


Fig. S24. Reduced form heterogeneity by mother's education. Estimated effect of local GM max attainable yield percentile on various perinatal health outcomes relative to 1995. All regressions include county and year by month fixed effects and standard errors are clustered by state and year. All regressions also control for other pesticides, employment, income, population, age and race shares, fertilizers, and family demographics, including mother's age, race, education, marital status, birth facility, resident status, previous births, sex of infant, and father's age and race. Sample restricted to births occurring in rural counties or to mothers residing in rural counties.

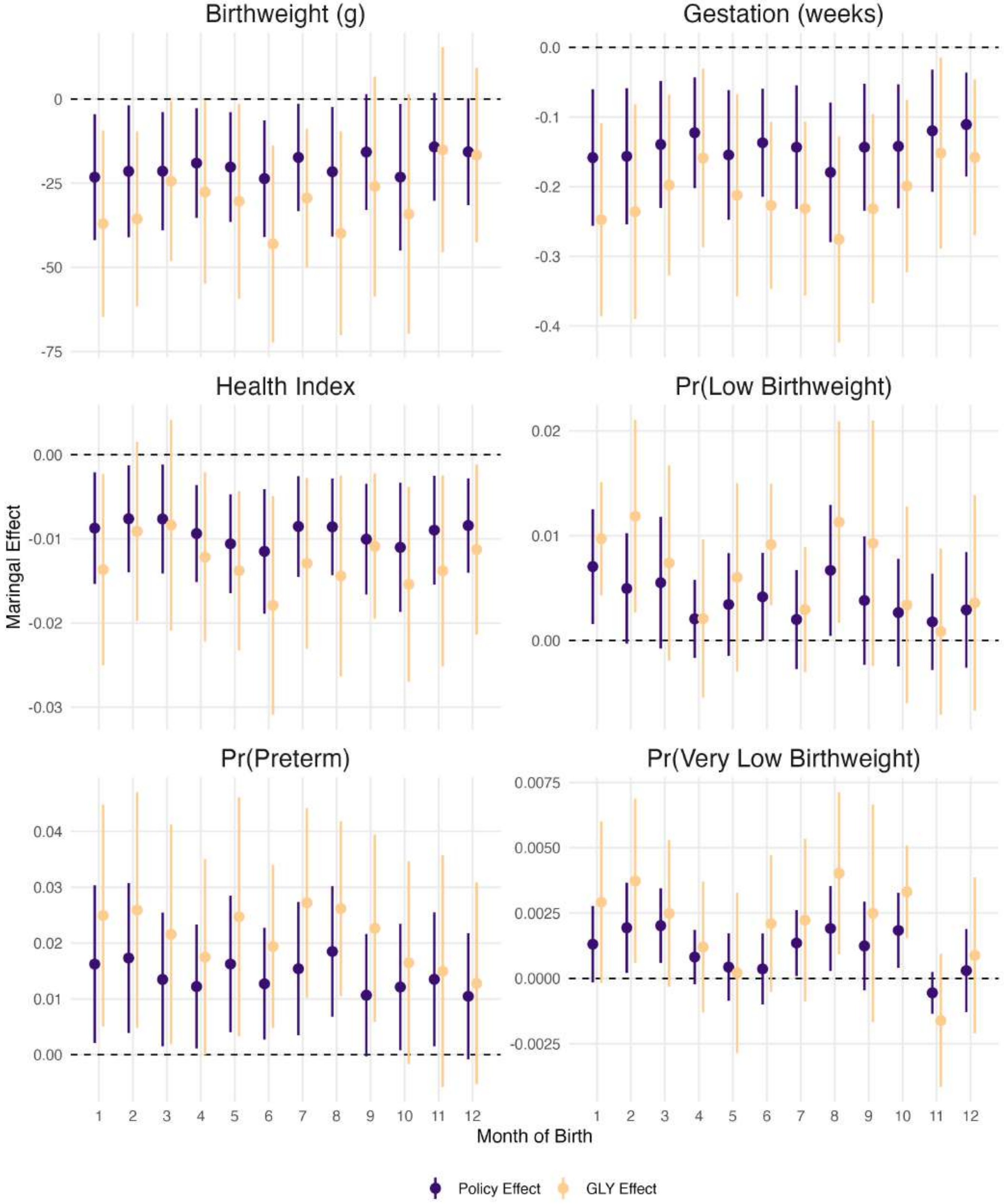
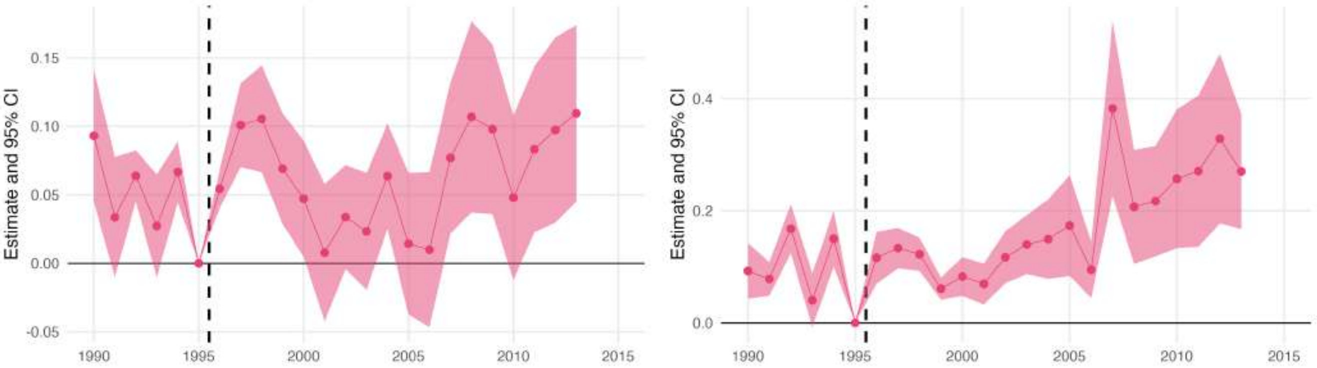
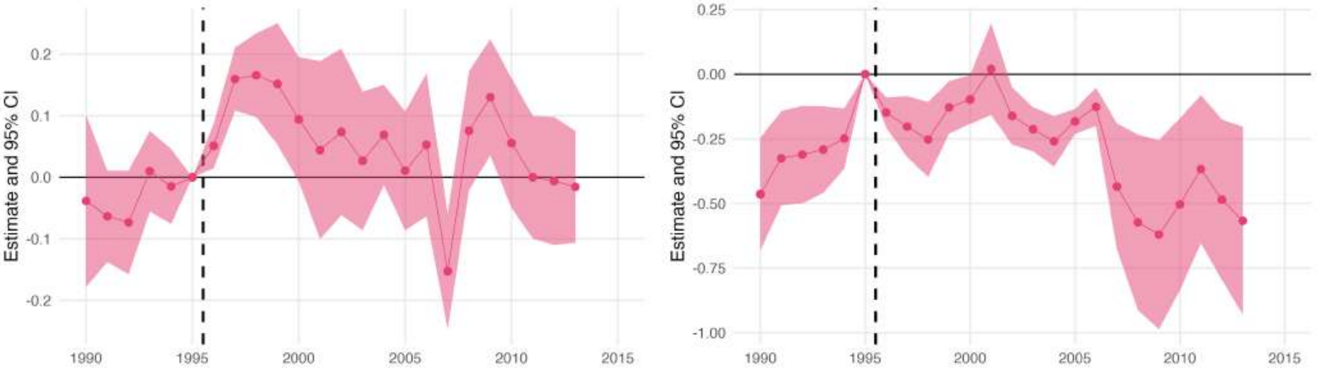


Fig. S25. Heterogeneity in effect by different month of birth, all outcomes. Estimated effect of glyphosate/km^2 on various perinatal health outcomes instrumented with GM max attainable yield interacted with year. All regressions include county, year, and month fixed effects and standard errors are clustered by state and year. Family demographics include mother's age, race, education, marital status, birth facility, resident status, previous births, sex of infant, and father's age and race.



(a) Total acreage

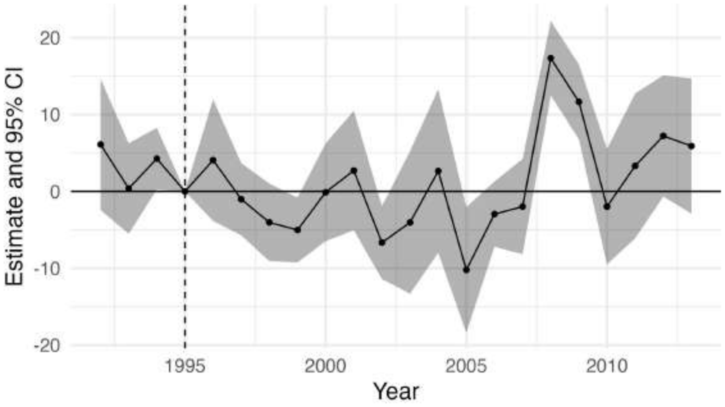
(b) Corn acreage



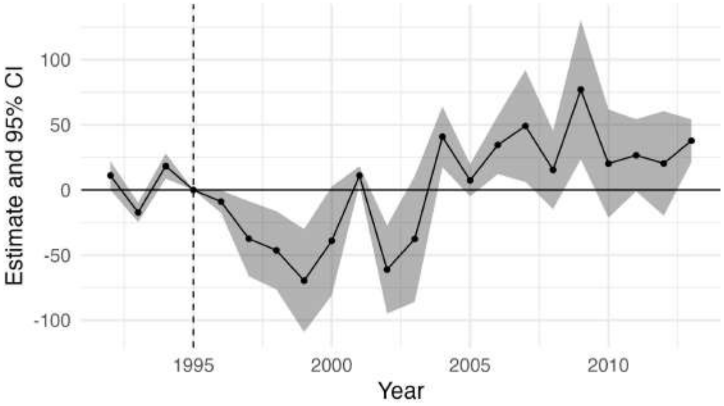
(c) Soy acreage

(d) Cotton acreage

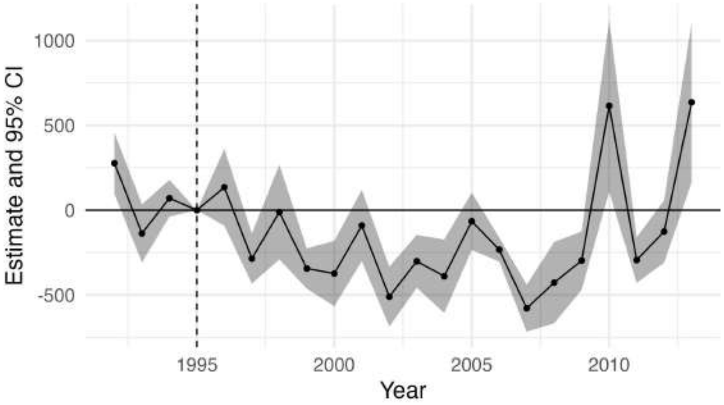
Fig. S26. Effect of local Max GM attainable yield on crop acreage. Standard errors are clustered by state and year. All regressions include county and year fixed effects with no other controls and are weighted by the number of births. Sample restricted to rural counties.



(a) Soy yield

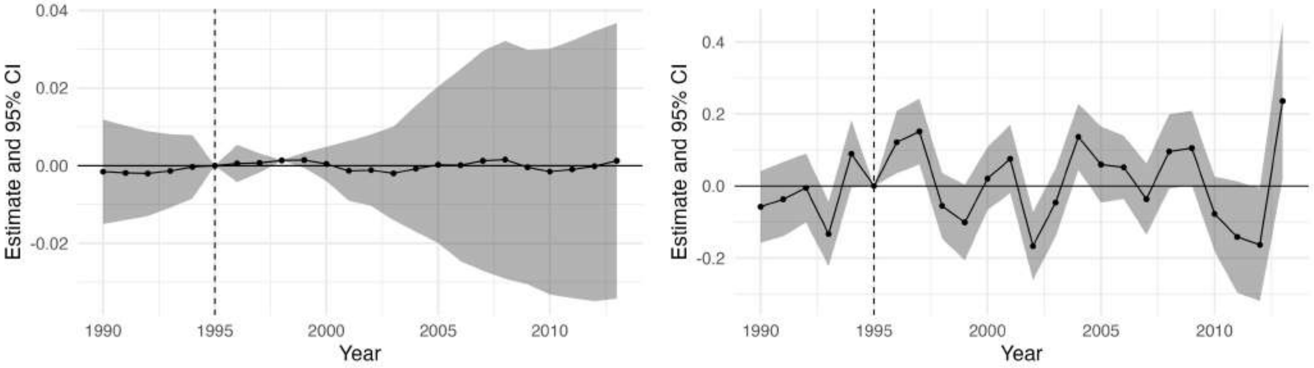


(b) Corn yield



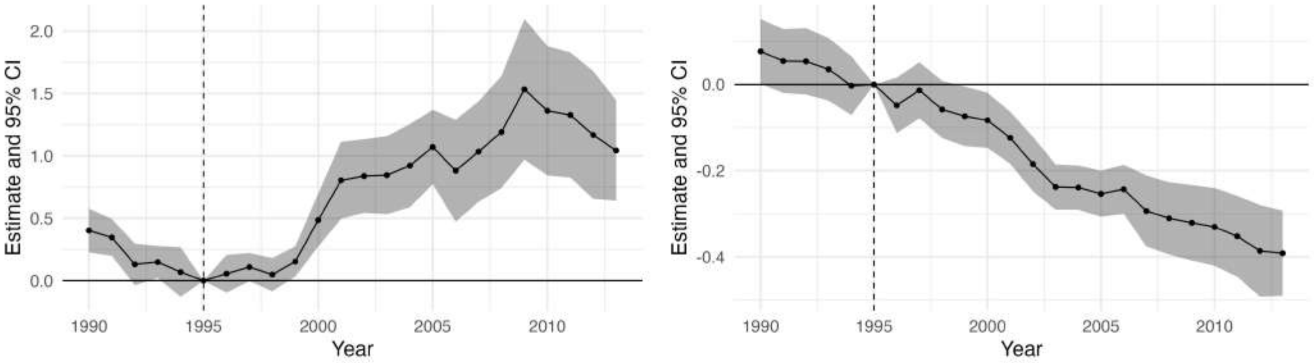
(c) Cotton yield

Fig. S27. Effect of local GM max attainable yield on crop yield. Soy and corn yield is measured in bushels/acre, while cotton is measured in lbs/acre. Standard errors are clustered by state and year



(a) Effect of GM suitability on non-farm income.

(b) Effect of GM suitability on farm income



(c) Effect of GM suitability on unemployment rate

(d) Effect of GM suitability on farm employment

Fig. S28. Coefficients from an event study regression of various socioeconomic variables on GM suitability for rural counties. Standard errors are clustered by state and year. Regressions are weighted by total births. Sample restricted to rural counties.

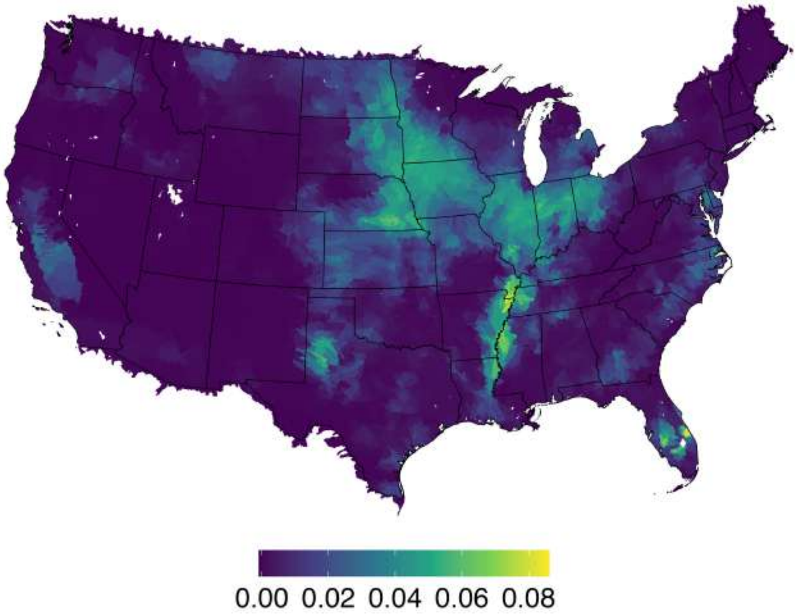
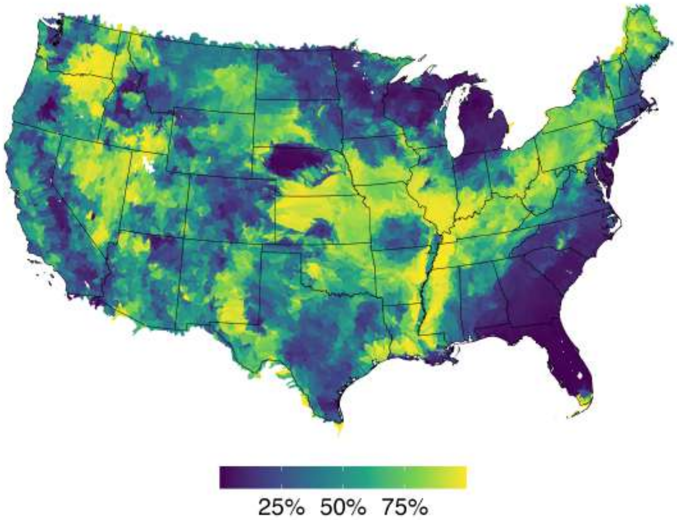
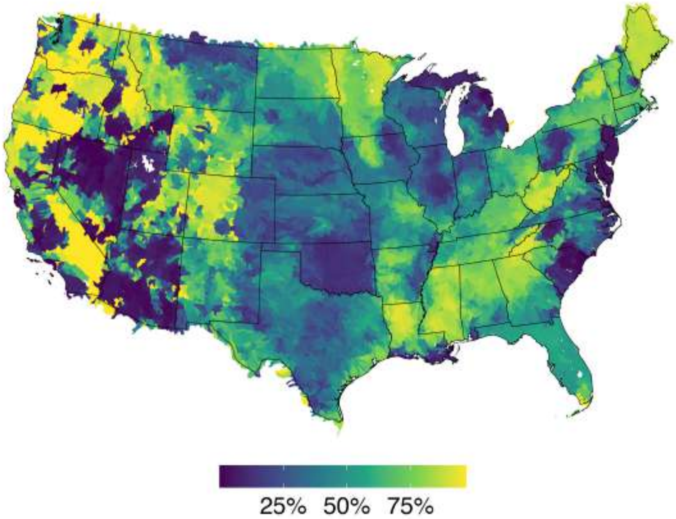


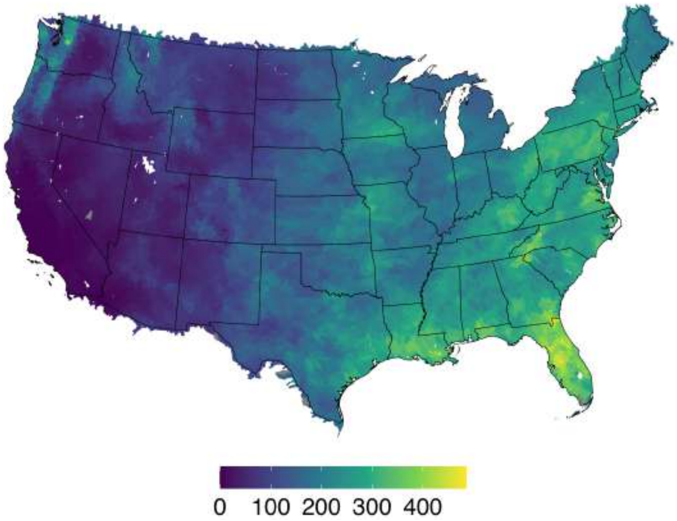
Fig. S29. Predicted glyphosate disaggregated into watersheds in 2004. These predictions come from our first stage model regressing glyphosate on local GM attainable yield percentile with county and year fixed effects. We disaggregate from county into watersheds using the portion of the county's area that is covered by each watershed. We generate predictions for each year, but only show 2004 to accompany the exposition.



(a) Soil erodibility (K factor) percentile.

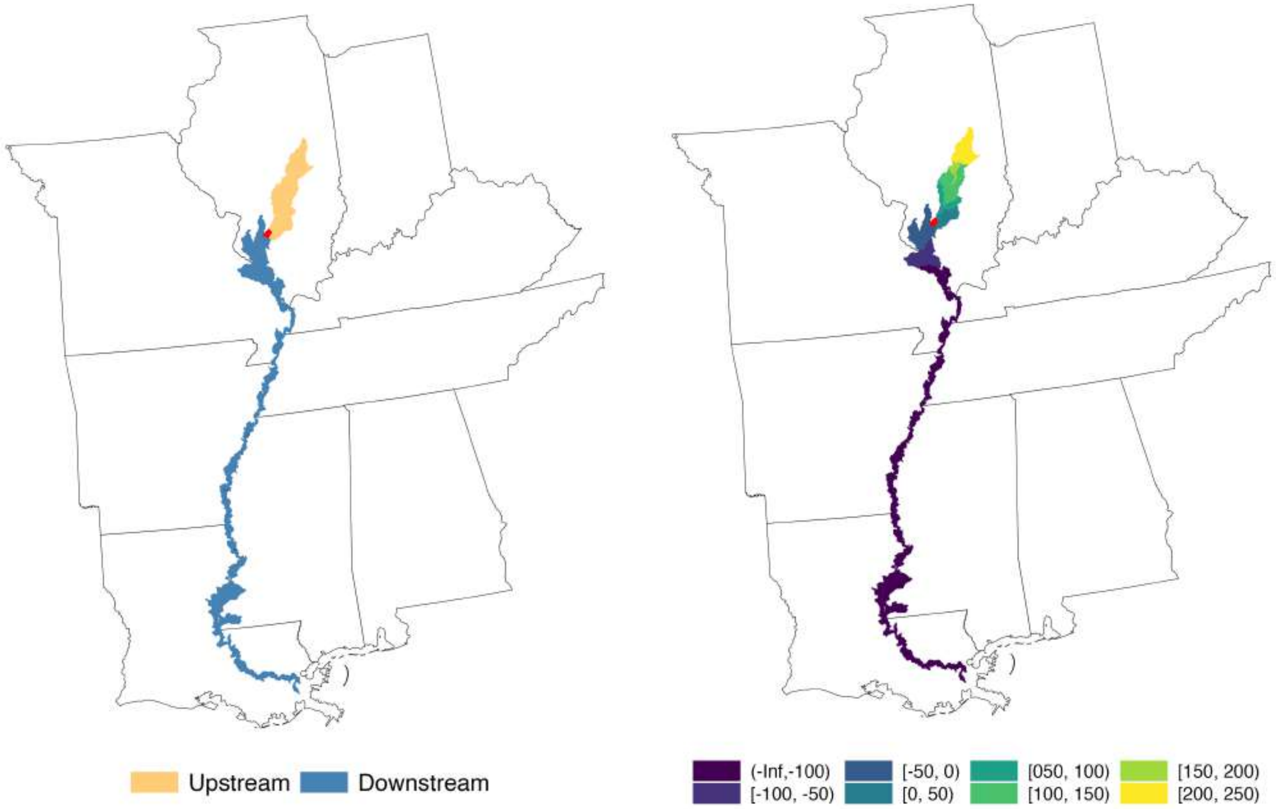


(b) Slope percentile.



(c) Growing season (Apr to Sep) precipitation in 2004.

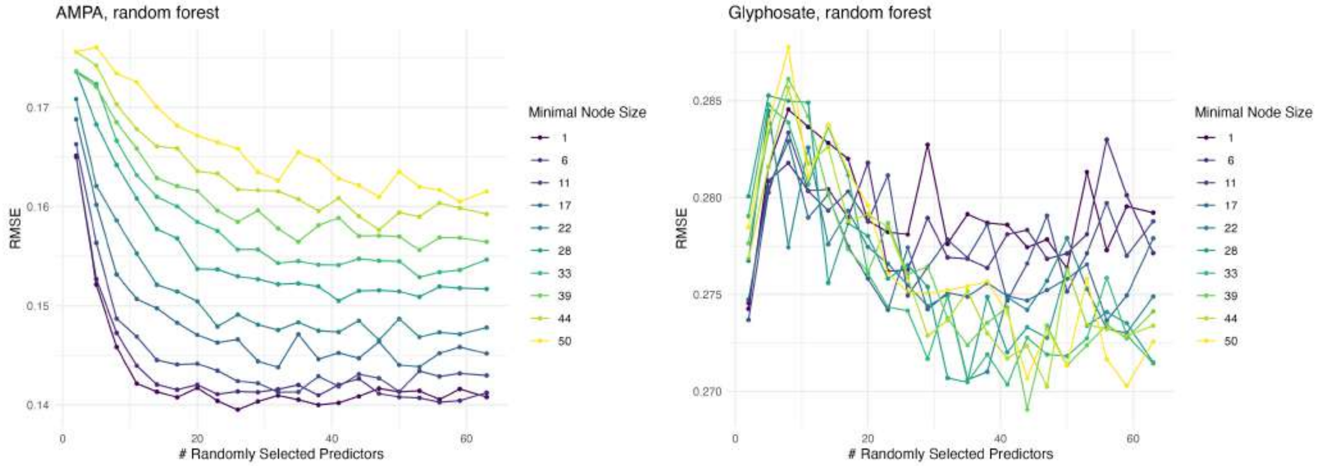
Fig. S30. Spatial variation in water ML predictors. Each map depicts a watershed-level average of the given variable. See text for details.



(a) Upstream and downstream watersheds.

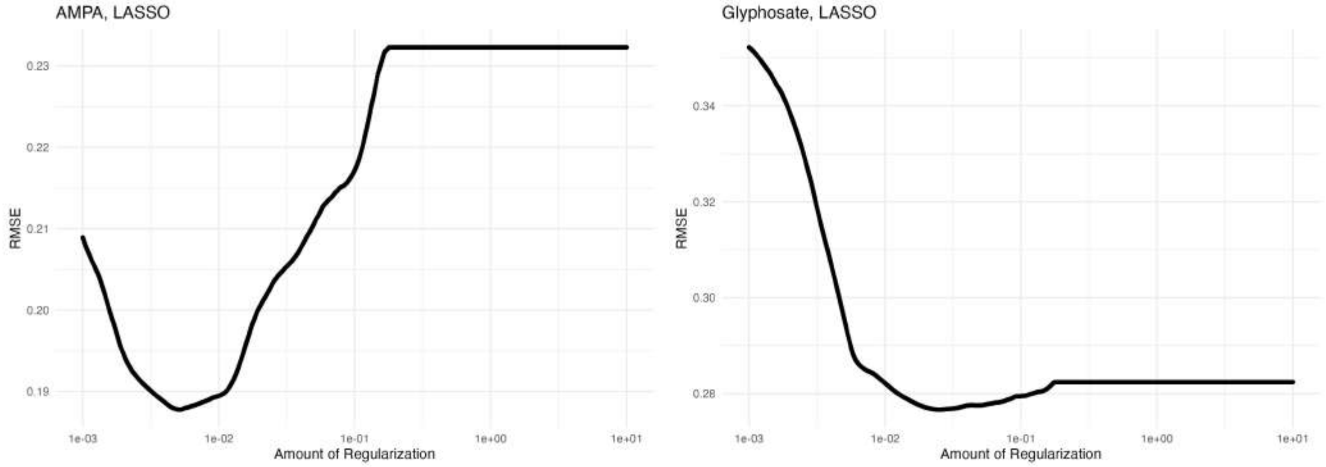
(b) Distance bins.

Fig. S31. Capturing upstream and downstream watersheds. For an example watershed in Illinois (highlighted in red), we show all of the watersheds upstream and all of the watersheds downstream. We calculate distance upstream and downstream using the distance between the centroids of watersheds along the path, then categorize these into 50-kilometer distance bins.



(a) CV Results for AMPA Random Forest Model.

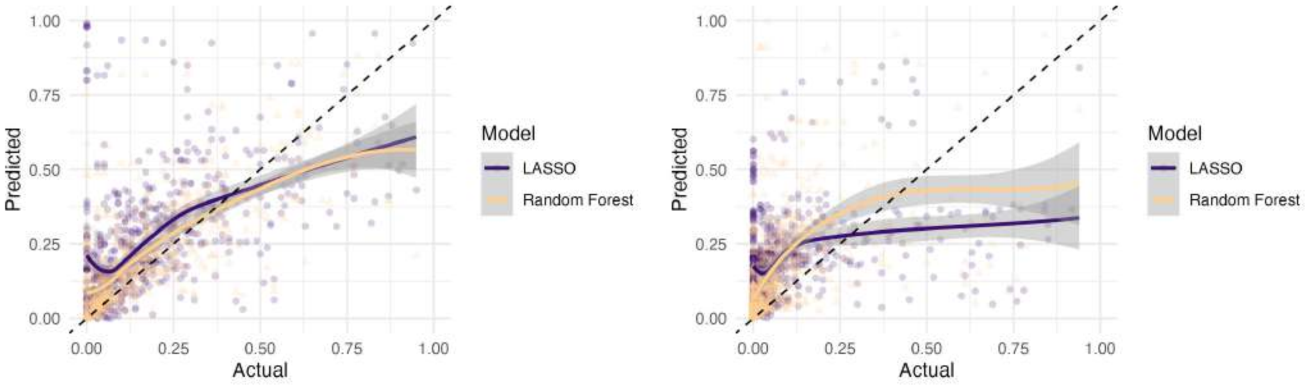
(b) CV Results for glyphosate Random Forest Model.



(c) CV Results for AMPA LASSO Model.

(d) CV Results for glyphosate LASSO Model.

Fig. S32. Cross Validation Results.



(a) Predicted vs Actual AMPA Concentrations.

(b) Predicted vs Actual glyphosate Concentrations.

Fig. S33. Out-of-sample prediction performance for LASSO and Random Forest models. Predictions are made on the 757 held-out observations in order to assess the model fit for LASSO and random forest models. Smooth lines are that of a generalized additive model.

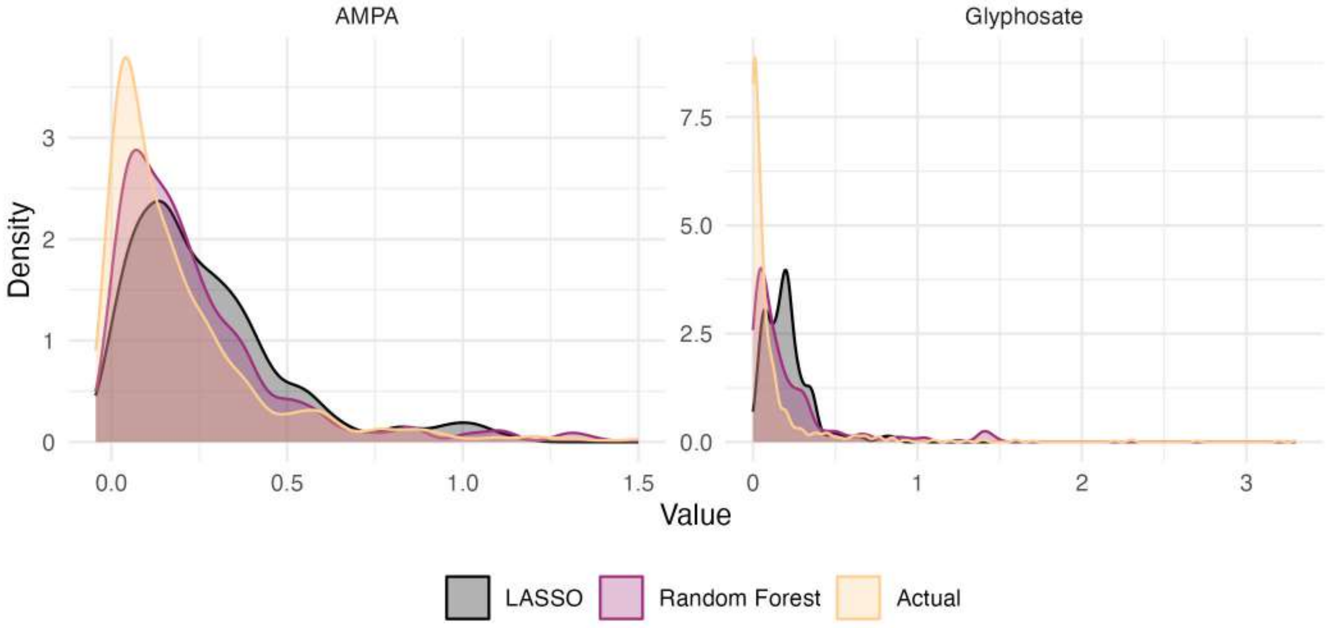


Fig. S34. Density of out-of-sample predictions relative to the actual values.

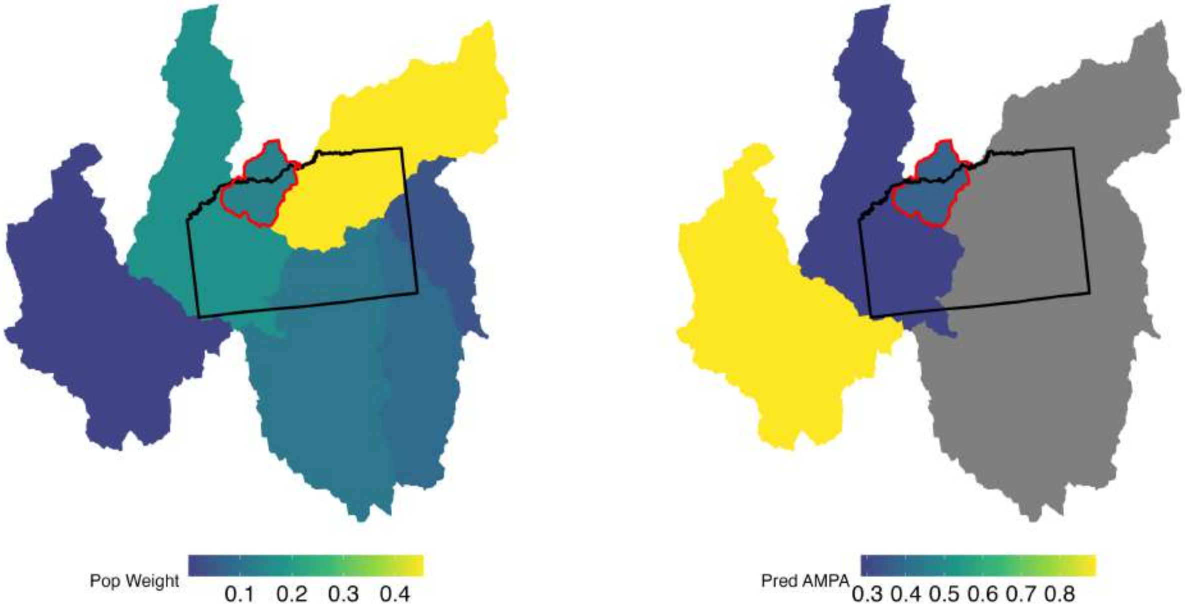


Fig. S35. Aggregating watersheds to counties. For the same watershed as in Fig S31 with the red border, on the left, we have population weights for Washington County (black outline). On the right, we have our predicted AMPA in July, 2004 using the LASSO model in each watershed touching Washington County. Thus, to generate county-month level predicted AMPA, we take the weighted average of predictions (left), where the weights come from the population in each watershed (right).

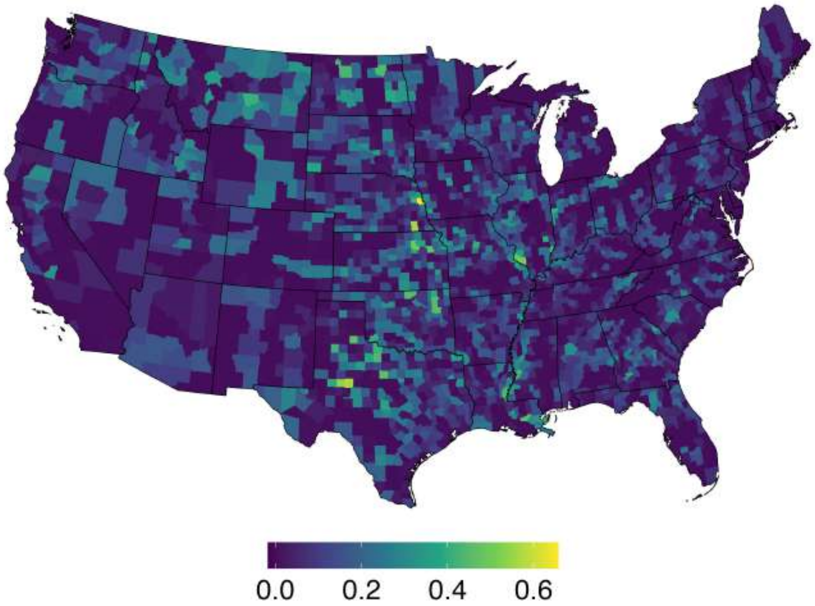


Fig. S36. Predicted county-level AMPA in July of 2004. This is a map for one month (July), in one year (2004), using one of four predictive models (LASSO predicting AMPA). We generate county level predictions like this for all months and years between 1992 and 2017 with LASSO and random forest models predicting glyphosate and AMPA.

Dep Var	BW			
Model:	(1)	(2)	(3)	(4)
Predicted AMPA (LASSO)	16.1			
	(11.2)			
Predicted AMPA (RF)		-3.14		
		(8.83)		
Predicted GLY (LASSO)			-6.01	
			(16.7)	
Predicted GLY (RF)				-5.94
				(6.69)
Local attainable yield	Yes	Yes	Yes	Yes
Local pesticides	Yes	Yes	Yes	Yes
Unemployment	Yes	Yes	Yes	Yes
Fixed-effects				
Family Demog	Yes	Yes	Yes	Yes
Yr x Mo + Cnty	Yes	Yes	Yes	Yes
Fit statistics				
N (millions)	7.910	7.910	7.910	7.910

Clustered (Year & State) standard-errors in parentheses

Table S9. Effect of predicted GLY or AMPA in water on birthweight.

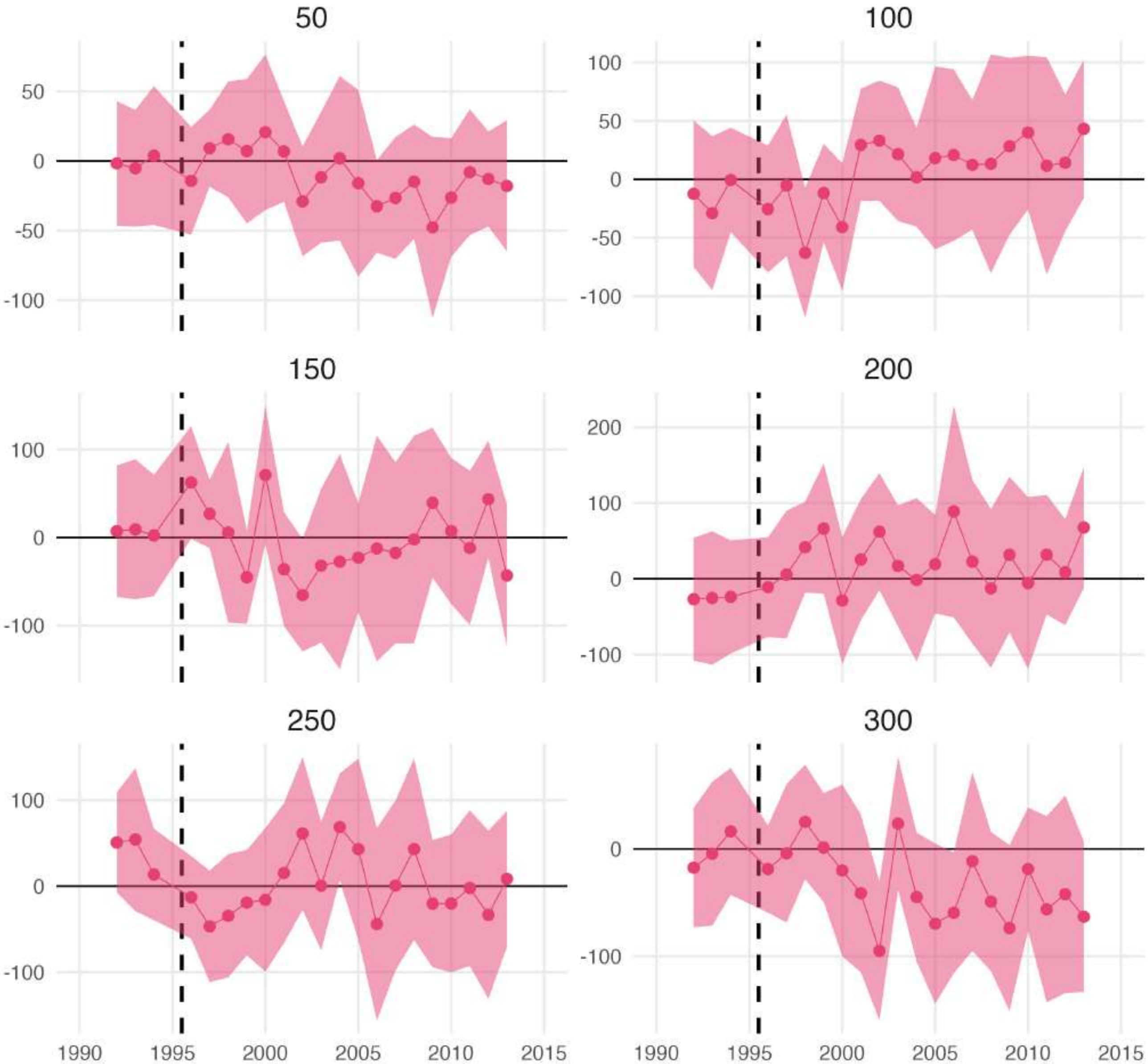


Fig. S37. Effect of upstream glyphosate by distance bin. Estimated effect of upstream GM max attainable yield percentile on various perinatal health outcomes relative to 1995. Bin labels represent the lower bound distance between the county and the upstream watershed, thus "50" is an aggregate of watersheds 50 to 100km upstream of a county. All regressions include county and year by month fixed effects and standard errors are clustered by state and year. All regressions also control for local attainable yield interacted with year, other pesticides, employment, income, population, age and race shares, fertilizers, and family demographics, including mother's age, race, education, marital status, birth facility, resident status, previous births, sex of infant, and father's age and race. Sample restricted to births from mothers with rural residence.

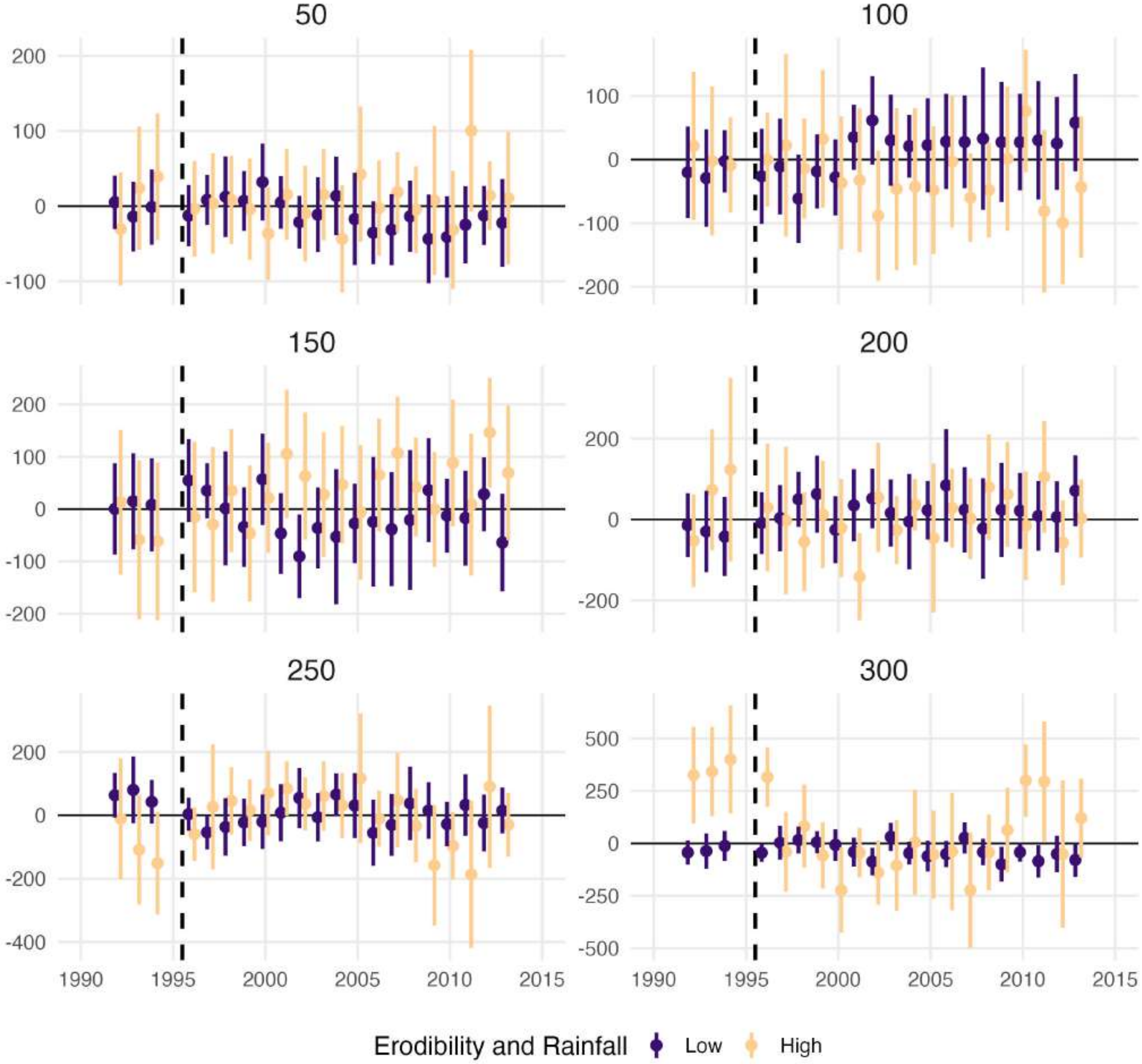


Fig. S38. Effect of upstream glyphosate by distance bin by high and low soil erodibility and precipitation. Estimated effect of upstream GM max attainable yield percentile on various perinatal health outcomes relative to 1995. Bin labels represent the lower bound distance between the county and the upstream watershed, thus "50" is an aggregate of watersheds 50 to 100km upstream of a county. All regressions include county and year by month fixed effects and standard errors are clustered by state and year. All regressions also control for local attainable yield interacted with year, other pesticides, employment, income, population, age and race shares, fertilizers, and family demographics, including mother's age, race, education, marital status, birth facility, resident status, previous births, sex of infant, and father's age and race. Sample restricted to births from mothers with rural residence.

References

1. SO Duke, SB Powles, Glyphosate: A once-in-a-century herbicide. *Pest Manag. Sci.* **64**, 319–325 (2008).
2. KM Baer, BJ Marcel, Glyphosate (2014).
3. B Muller, Glyphosate-A love story. Ordinary thoughtlessness and responseability in industrial farming. *J. Agrar. Chang.* **21**, 160–179 (2020).
4. G Barrows, S Sexton, D Zilberman, Agricultural Biotechnology: The Promise and Prospects of Genetically Modified Crops. *J. Econ. Perspectives* **28**, 99–120 (2014).
5. A Grube, D Donaldson, T Kiely, L Wu, Pesticides Industry Sales and Usage: 2006 and 2007 Market Estimates, (U.S. EPA), Technical report (2011).
6. U.S. EPA, Glyphosate: Response to Comments, Usage, and Benefits, Technical report (2019).
7. CM Benbrook, Why regulators lost track and control of pesticide risks: Lessons from the case of glyphosate-based herbicides and genetically engineered-crop technology. *Curr. Environ. Heal. Reports* **5**, 387–395 (2018).
8. C Benbrook, Shining a light on glyphosate-based herbicide hazard, exposures and risk: Role of non-hodgkin lymphoma litigation in the usa. *Eur. J. Risk Regul.* **11**, 498–519 (2020).
9. L Sheppard, S McGrew, RA Fenske, Flawed analysis of an intentional human dosing study and its impact on chlorpyrifos risk assessments. *Environ. Int.* **143**, 105905 (2020).
10. CM Benbrook, How did the US EPA and IARC reach diametrically opposed conclusions on the genotoxicity of glyphosate-based herbicides? *Environ. Sci. Eur.* **31** (2019).
11. CJ Portier, et al., Differences in the carcinogenic evaluation of glyphosate between the international agency for research on cancer (iarc) and the european food safety authority (efsa). *J. Epidemiol. Community Heal.* **70**, 741–745 (2016).
12. P Clausing, C Robinson, H Burtscher-Schaden, Pesticides and public health: an analysis of the regulatory approach to assessing the carcinogenicity of glyphosate in the european union. *J. Epidemiol. Community Heal.* **72**, 668–672 (2018).
13. MJ Davoren, RH Schiestl, Glyphosate-based herbicides and cancer risk: a post-iarc decision review of potential mechanisms, policy and avenues of research. *Carcinogenesis* **39**, 1207–1215 (2018).
14. FS vom Saal, et al., Flawed experimental design reveals the need for guidelines requiring appropriate positive controls in endocrine disruption research. *Toxicol. Sci.* **115**, 612–613 (2010).
15. A Mie, C Rudén, P Grandjean, Safety of safety evaluation of pesticides: developmental neurotoxicity of chlorpyrifos and chlorpyrifos-methyl. *Environ. Heal.* **17** (2018).
16. JA Falcone, Estimates of county-level nitrogen and phosphorus from fertilizer and manure from 1950 through 2017 in the conterminous United States, (U.S. Geological Survey), Technical Report 2020-1153 (2021).
17. P Rhode, How Suitable are FAO-GAEZ Crop Suitability Indices for Historical Analysis? in *SI 2024 Development of the American Economy.* (National Bureau of Economic Research, Boston, MA), (2024).
18. Surveillance, Epidemiology, and End Results (SEER) Program, U.S. County Population Data - 1969-2022 (2024).
19. TJ Lark, et al., Environmental outcomes of the US Renewable Fuel Standard. *Proc. Natl. Acad. Sci.* **119**, e2101084119 (2022).
20. L Medalie, et al., Influence of land use and region on glyphosate and aminomethylphosphonic acid in streams in the USA. *Sci. The Total. Environ.* **707**, 136008 (2020).
21. B Lehner, G Grill, Global river hydrography and network routing: Baseline data and new approaches to study the world's large river systems. *Hydrol. Process.* **27**, 2171–2186 (2013).
22. Soil Survey Staff, Gridded Soil Survey Geographic (gSSURGO) Database for the Conterminous United States (2021).
23. PRISM Climate Group, "Prism Gridded Climate Data" (2021).
24. CIESIN, U.S. Census Grids (Summary File 1), 2010 (2017).
25. M Dias, R Rocha, RR Soares, Down the River: Glyphosate Use in Agriculture and Birth Outcomes of Surrounding Populations. *The Rev. Econ. Stud.* **90**, 2943–2981 (2023).
26. United Nations, UN-Water SDG 6 Data Portal (https://sdg6data.org/) (2024) Online; accessed 11 June 2024.
27. FAO, IIASA, Global Agro Ecological Zones version 4 (GAEZ v4) (http://www.fao.org/gaez/en) (2021).
28. USDA, Adoption of Genetically Engineered Crops in the U.S. (2024).