

# Direct and downstream health effects of pesticides: Evidence from the U.S. rollout of genetically modified crops and glyphosate

Preliminary: Do not cite or circulate without permission.

Edward Rubin    Emmett Saulnier\*

June 29, 2022

## Abstract

The advent of genetically modified (GM) crops quickly led to the widespread use of glyphosate within U.S. agriculture. Since glyphosate's introduction in 1974, policy-makers have debated its effects on human health. However, the scientific literature offers few generalizable studies on this significant technological innovation. Prior research has focused on lab studies on animals or epidemiologic studies of groups with high occupational exposure. Instead, we focus on the population at large. The general population could face exposure to glyphosate through water (runoff) or air (drift). To identify the causal effect of glyphosate exposure on health, we leverage (1) county-level variation in glyphosate use driven by (2) the timing of the release of GM technology for different crops and (3) differential suitability for these crops. Our results for runoff-driven exposure offer an additional dimension—comparing the effects of upstream pesticide applications to the placebo of downstream applications. Our early results suggest that glyphosate has a significant, negative effect on birth weight (in the same county and year), while upstream glyphosate does not affect birth weights. These estimates are likely lower bounds of the actual health consequences of glyphosate, as they are net of health benefits from reduced exposure to other pesticides (and decreased tillage). In addition to estimating the causal effects of glyphosate exposure on the general population, we contribute to a broader discussion on the efficiency and equity of environmental regulation and the challenges of evaluating these regulations.

**JEL Codes:** I18, Q15, Q53

**Keywords:** Pesticides, Health, Infant mortality, Water pollution, Agriculture

\*We thank Benjamin Hansen and Glen Waddell for excellent feedback. Errors are the authors'. Rubin: Assistant Professor of Economics, University of Oregon ([edwardr@uoregon.edu](mailto:edwardr@uoregon.edu)). Saulnier: Doctoral student in the Department of Economics at the University of Oregon. ([emmetts@uoregon.edu](mailto:emmetts@uoregon.edu)).

# 1 Introduction

The introduction of genetically modified (GM) corn, soy, and cotton has had a profound effect on agriculture in the United States. The key innovation provided by GM technology is resistance to the herbicide glyphosate. Farmers can spray glyphosate on their GM crops to kill weeds without damaging the crops themselves. While generating productivity benefits, GM technology has the potential to introduce health externalities. However, the sign of the health externality is unclear ex-ante. Glyphosate replaced other, more toxic herbicides used with non-GM crops, which could lead to positive health effects. Yet, GM crops allow for liberal use of glyphosate, leading to an overall increase in the volume of chemicals sprayed, which could cause negative health effects. We use quasi-experimental variation in GM adoption to quantify the sign and magnitude of the health externality resulting from GM crop adoption.

Our empirical strategy uses temporal variation in GM adoption from the commercial release of GM crops in the United States in 1996, paired with spatial variation in GM adoption driven by differences in the suitability of the land for growing the GM crops—corn, soy, and cotton. GM seeds became commercially available in the United States starting in 1996, and farmers rapidly adopted GM seeds and glyphosate in the following years. However, this technological change only directly affected locations where the GM crops are grown. Thus, we isolate variation in glyphosate use created by differences in the suitability of the environment for GM crops.

Regulatory agencies have offered differing guidance on the health effects of widespread glyphosate use. The World Health Organization released an article that labeled glyphosate as likely carcinogenic, leading many governments to consider a ban on the chemical (IARC, 2015). The EPA responded by conducting a review, concluding that it is not likely to be carcinogenic at relevant doses and not a threat to public health (EPA, 2020). These decisions are based on lab experiments performed on animals and observational studies that fail to identify the causal effect of glyphosate on human health. Additionally, the emerging literature exploring sub-clinical toxicity suggests that exposure to chemicals at low concentrations could still have harmful effects on long-term health (Landrigan, 2018). Thus, existing research may underestimate the

social costs associated with glyphosate use, leading to excessive use of the chemical relative to the social optimum. We seek to determine the causal effect of the introduction of glyphosate and GM crops on human health to inform the policy debate surrounding the technology.

Health effects are difficult to measure on a large scale, with many confounding factors making well-identified causal inference particularly important. While many potential health effects could arise from increased exposure to glyphosate, we focus on infant health metrics such as birth weight and infant mortality. These measures provide a close link between the timing of exposure and the health outcome since we can measure the effects of exposure during pregnancy on infant health within a year of birth. Compare this to the effect of glyphosate exposure on cancer, which may take decades to develop, is the result of many different factors, and remains a rare occurrence among the population as a whole — thus making the estimation of an effect on cancer very difficult.

We use a difference-in-differences approach to estimate the causal effect of glyphosate on health. Specifically, we compare infant health outcomes before and after 1996 in high and low GM suitability counties in the rural US. We find strong evidence of increases in glyphosate between our designated treatment and control groups, where high suitability counties increase their glyphosate use by relative to low suitability counties after the introduction of GM crops. We then document adverse effects on infant health outcomes, showing that birth weight has decreased by 10-15 grams in high suitability counties relative to low suitability counties. These estimates imply that increasing glyphosate use from the median to 90th percentile causes a 34.5 gram decrease in birth weight. These results are robust to various specifications.

This research extends a recent string of literature measuring the effects of pesticides on infant health. Frank (2021) and Taylor (2022) find adverse effects of insecticides on infant mortality using variation in insecticide use driven by exogenous changes in the load of pests. Dias, Rocha, and Soares (2019) use the release of GM soy to measure health effects of glyphosate in Brazil, finding that increases in glyphosate cause increases infant mortality. Camacho and Mejia (2017) use a campaign by the US and Colombian governments to spray coca fields with glyphosate by plane in order to reduce cocaine supply, also finding that glyphosate causes increases in infant

mortality. Other papers have found adverse effects of pesticides on infant health (Garry et al., 2002; Regidor, 2004; Larsen, Gaines, and Deschenes, 2017; Brainerd and Menon, 2014), but they are focused on populations that have high levels of exposure, most frequently as a result of being an agricultural worker.

The rest of the paper proceeds as follows. Section 2 gives additional institutional details on GM crops and glyphosate. Section 3 introduces the data that we use throughout the analysis. Section 4 details our empirical strategy, covering the estimation technique as well as identifying assumptions. Section 5 presents our main results, with Section 6 covering robustness checks and extensions. Finally, Section 7 concludes.

## 2 Institutions

### 2.1 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 (Duke and Powles, 2008). However, it is water-soluble, which means that the part that does not bind to soil enters the water supply (Baer and Marcel, 2014). 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 (Muller, 2020). It also breaks the nutrient cycle, forcing farmers to increase their dependence on fertilizer to feed their crops (Barrows, Sexton, and Zilberman, 2014). Over time, weeds grow resistant to glyphosate, forcing farmers to either increase the concentration of glyphosate or get "stacked" GM seeds, which provide resistance to multiple chemicals that the farmers then spray on their crops. Farmers in the US spend nearly \$8 billion on pesticides each year (Grube et al., 2011), applying glyphosate to 298 million acres of crops annually (EPA, 2019).

There are a few different ways in which glyphosate could impact human health. A growing

body of studies document the carcinogenic potential of glyphosate exposure over long periods, as reviewed in Portier, 2020. Additionally, research links glyphosate to endocrine disruption, which can result in developmental and reproductive health issues, among others (Munoz, Bleak, and Calaf, 2020). Finally, there is research into the effect of glyphosate on the microbiome. Many the organisms that make up our gut have the same pathway that glyphosate inhibits in plants, thus making these organisms sensitive to the chemical's presence in our bodies (Leino et al., 2021). These studies do not seem to be controversial on their internal validity. However, they are questioned on their external validity, as it is unclear how lab experiments extend to the real world.

Possible exposure mechanisms include direct contact, aerial drift, water pollution, and food residue. Direct exposure is known to be harmful, and the EPA releases guidance on best practices for safely applying glyphosate. We do not measure the effect of food residue, which is suspected to be a significant source of exposure (Landrigan, 2018). We attempt to measure the effect of exposure through water pollution and aerial drift.

## 2.2 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. 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 (Fernandez-Cornejo et al., 2014).

## 3 Data

We create a panel of counties between 1992 and 2017 with measures of glyphosate exposure, which we can then merge to micro-data on all births in the U.S. over the same period using the county and year of birth.

### 3.1 Pesticide Use Estimates

Our measure of glyphosate use comes from the United States Geological Survey's Pesticide National Synthesis Project (Thelin and Stone, [2013](#); Baker and Stone, [2015](#)). The USGS surveyed farmers to calculate pesticide use rates per acre of different crops planted at the crop reporting district level. They then multiply these usage rates by the total acreage of each crop within the crop reporting district to estimate the total amount of each pesticide used, measured in kilograms. Each pesticide has two estimates, high and low, where the high value assumes crop reporting districts with a missing usage rate for a pesticide applied the pesticide at the same rate as their neighbors on each crop. The low value assumes a missing usage rate for a pesticide means that farmers did not apply that pesticide. We use the high value throughout our analysis. Additionally, we normalize by the area of each county, thus our measure of glyphosate and other pesticides are in kilograms of active ingredient per square kilometer.

### 3.2 Infant health data

We have the universe of births and deaths in the United States between 1990 and 2019 from the National Center for Vital Health Statistics. The birth data contain information recorded on the birth certificates, including date of birth, birth weight, APGAR score, and demographics of the mother and father. We have access to the restricted versions of these files, which identify the county of birth for all births, compared to the publicly available data, which hides the county of birth for counties with less than 100,000 residents. We create various health outcome metrics from this data, focusing on birth weight and internal infant mortality rate, where internal infant mortality excludes deaths that occur from external causes such as accidents (Chay and Greenstone, [2003](#)). Since we cannot directly link infant deaths from the mortality data to births

in the natality file, we calculate annual infant mortality rates by county.

### 3.3 Attainable Yield

We use the FAO-GAEZ attainable yield for soybeans, corn, and cotton to measure the suitability of a county for genetically modified crops (FAO and IIASA, 2022). These data assign potential yield values to one square kilometer pixels based on environmental factors such as soil type, slope, and climate. We aggregate the pixels to the county by taking the average of all pixels within a county. We take the difference between the high-input attainable yield and the low-input attainable yield to focus on the counties with the largest incentive to adopt GM crops. The underlying model calculates the high-input scenario assuming that farmers have access to modern technology for crop management, including GM seeds. Whereas the model calculates the low-input scenario assuming more traditional farming methods.

We aggregate the three GM crops by standardizing the attainable yield difference so that each crop has a mean of zero and a standard deviation of one. We then take the simple average across the standardized yield differences. Figure 1 shows the spatial distribution of attainable yield difference for GM crops, showing that northern Missouri, Illinois, Indiana, and the Mississippi River are the places most suitable for growing GM crops. Additionally, parts of the south have higher yields due to their suitability for cotton.

We then classify counties east of the 100th meridian<sup>1</sup> as either high or low attainable yield, with the cutoff between the two groups being the median value of the average standardized yield difference for GM crops. High attainable yield counties are "treated," and low attainable yield counties are "untreated." Figure 4c shows assignment of counties to treatment using the median cutoff. We exclude urban counties since exposure to glyphosate is likely to be significantly smaller in cities.

We explore alternative definitions of this treatment assignment in section 6. First, we show similar results using the continuous attainable yield difference on the right-hand side of our

---

<sup>1</sup>The 100th meridian is considered the boundary between the dry west and humid east. 81% of the total glyphosate sprayed between 1992 and 2004 was east of the 100th meridian.

estimating equations rather than a binary treatment variable. We then check cutoff values for high and low attainable yield for every 5th percentile between the 25th and 75th percentile. We also add in varying "buffers", where we exclude counties close to the cutoff between high and low attainable yield. Finally, we use individual crops rather than the aggregate of corn, soy, and cotton, including some placebo crops that should not have seen changes in glyphosate use.

### 3.4 Upstream aggregation

We use a spatial water model to estimate glyphosate exposure based on the amount of glyphosate sprayed upstream of each county. We summarize the process here but leave the details in the appendix section [A.1](#). First, we disaggregate the county-level glyphosate data into watersheds. Then, we find all watersheds upstream of a particular watershed using the water model. We then take a weighted sum of glyphosate sprayed upstream of each watershed, where the weights are a function of distance. Glyphosate sprayed further upstream leads to less exposure than glyphosate sprayed closer, as the chemical degrades and becomes more diluted over time. Since we do not have a chemical transport model for glyphosate,<sup>2</sup> we assume a weighting function for our main results. We show that our results are robust to various specifications of the weighting function.

Once we estimate glyphosate sprayed upstream each watershed, we must aggregate it back up to the county level to analyze it with our health metrics. We take the weighted average of upstream glyphosate for all watersheds in a county, where the weights are the portion of the county's population that lives within that watershed. We use population estimates for one square kilometer pixels from SEDAC to calculate the population weights for each watershed.

We perform this same aggregation algorithm with additional variables—other pesticides, the attainable yield differences, and treatment indicators. The aggregated treatment indicators give us a weighted average of upstream or downstream treatment status. Additionally, we can

---

<sup>2</sup> Similar to the InMAP or HYSPLIT models for air particles

reverse the flow direction, aggregating glyphosate and other variables downstream of each county. Since glyphosate sprayed downstream of each county will not affect health outcomes in that county, we use it as a placebo test.

### 3.5 Other data

We collect annual acreage and yield of crops harvested within a county from the USDA's annual survey, obtained from the USDA NASS database. We get estimates of the total population for each county from the Census Bureau, using their intercensal population estimates. We also collect the number of farm and non-farm employees from the BEA.

### 3.6 The Spatial and Temporal Distribution of Glyphosate

Figure 2 shows the spatial distribution of glyphosate per square kilometer in 1995 and 2004. Over this time period, the total amount of glyphosate sprayed in the US more than quadrupled from 15 thousand kilograms to 65 thousand kilograms. In 2017, total glyphosate use was 128 thousand kilograms. We can see that spraying is spatially correlated, with most coming in the Midwest and along the Mississippi. There are also pockets of higher intensity in the Carolinas, Georgia, Florida, and California.

Figure 3 shows the spatial distribution of upstream glyphosate in 2004. Exposure to upstream glyphosate is highest for counties along major rivers, particularly in the Mississippi River drainage, which collects runoff from the entire Midwest.

Figure 4 shows the time series of glyphosate intensity and mean birth weight for different groups, which we will use to identify the effect of glyphosate on infant health. The growth in glyphosate increases starting in 1996 when the first genetically modified crops became available commercially. Meanwhile, birth weights decline across the country. However, high GM yield counties saw larger decreases in birth weight as compared to low GM yield counties, urban counties, or counties west of the 100th meridian.

### 3.7 Summary Statistics

Table 1 shows summary statistics comparing high and low GM attainable yield counties, urban counties, and counties west of the 100th meridian during the period before the release of GM crops. The high and low potential yield counties are similar on these observable characteristics.

## 4 Empirical strategy

We seek to estimate the causal effect of GM crop adoption and glyphosate on infant health. To isolate exogenous variation in glyphosate use, we leverage temporal variation due to the commercial release of GM crops in 1996 and spatial variation in GM adoption due to differences in the suitability of the environment for growing those crops.

### 4.1 Estimation

We first use difference-in-differences to estimate the effect of GM crop adoption on glyphosate use (first stage) and infant health (reduced form).

$$y_{ijt} = \gamma_l D_{jt}^l + \gamma_u D_{jt}^u + \alpha_j + \lambda_t + \varepsilon_{ijt}, \quad (1)$$

where  $y_{ijt}$  is glyphosate exposure for the first stage or infant health for the reduced form.  $D_{jt}^l$  is an indicator variable that denotes local "treatment," where  $D_{jt}^l$  equals one if county  $j$  is highly suitable for GM crops and  $t > 1995$ . We detail our approach to assigning counties to treatment in Section 3.3.  $D_{jt}^u$  is a measure of the proportion of treated counties upstream of county  $j$ . Section A.1 has our methodology for aggregating upstream variables.  $\alpha_j$  and  $\lambda_t$  are county and year fixed effects. Thus,  $\gamma_l$  measures the local effect of GM crop release on outcome  $y$  in high GM yield counties relative to the change in outcomes in low GM yield counties.  $\gamma_u$  tells us the effect of going from entirely low GM yield upstream to all high GM yield upstream.

A linear model for the effect ( $\beta_l$ ) of local glyphosate intensity and the effect ( $\beta_u$ ) of upstream

glyphosate intensity on infant health outcome  $y$  is

$$y_{ijt} = \beta_l G_{jt}^l + \beta_u G_{jt}^u + X'_{ijt} \eta + \alpha_j + \lambda_t + \varepsilon_{ijt}. \quad (2)$$

for individual  $i$ , residing in county  $j$ , in year  $t$ . Some measures of infant health are calculated at the county level (i.e. infant mortality), in which case we drop the  $i$  subscripts from the above model.  $G_{jt}^l$  is local glyphosate exposure, measured as the total amount of glyphosate sprayed in county  $j$  in year  $t$ .  $G_{jt}^u$  is an estimate of exposure to glyphosate from glyphosate spraying upstream of county  $j$  in year  $t$ .  $X'_{ijt}$  is a vector of controls.

Estimating equation (2) is unlikely to identify the true effect of glyphosate on health due to endogeneity and measurement error. To rectify those issues, we use instruments to isolate exogenous variation in both local and upstream glyphosate. Our instruments are the treatment variables from above,  $D^l$  and  $D^u$ , interacted with year dummies for each year before and after 1995.

## 4.2 Identification

Our model includes county and year fixed effects, and thus the identifying assumption is that changes in glyphosate use drive deviations from the trend in outcomes for high GM yield counties. We could also state this as the well-known parallel trends assumption—that if genetically modified crops had not been released, then the difference between high and low GM yield counties for glyphosate use and infant health would have remained constant. Below we discuss a few threats to our identification strategy.

There could be differential trends in high and low GM attainable yield counties prior to the release of GM crops. Additionally, there could be a single outlier year that drives the average result, despite the smooth increase in glyphosate over time. In order to assess these concerns, we estimate an event study model,

$$y_{ijt} = \alpha_j + \lambda_t + \sum_{\tau \neq 1995} \gamma_\tau^l D_{j\tau}^l + \gamma_\tau^u D_{j\tau}^u + \varepsilon_{ijt}, \quad (3)$$

where instead of estimating the treatment effect as an average over all years post treatment, we estimate different coefficients for each year prior to 1995 and each year after 1995.

The measured effects could be due to changes in other factors affecting infant health in treated counties. On the intensive margin, adopting GM crops requires farmers to substitute away from traditional crops. Thus, increases in glyphosate correlate with decreases in other herbicides used on traditional crops. We therefore interpret the reduced form model as the net effect of GM crops relative to non-GM crops.

Additionally, we carry out two placebo tests to support our identifying assumption. First, we use attainable yield differences for crops that do not have GM technology to define treatment. The placebo crops should not lead to any changes in glyphosate, nor any changes in health outcomes. However, they could pick up on other changes in other variables that would affect agricultural areas generally.

As a second placebo, we replace the upstream treatment variable with a measure of whether counties downstream are high or low GM yield. Downstream glyphosate use will not affect health outcomes directly but could reflect changes in unobserved variables correlated over space. For example, suppose we have an increase in income in the treated county that has spillover effects increasing income in upstream counties. These spillovers are also likely to affect downstream counties, which the downstream placebo test would capture.

For the 2SLS model given by Equation (2), we must assume the exclusion restriction holds to interpret the coefficient causally—that our instruments, high GM attainable yield interacted with year indicators, only affect infant health through changes in glyphosate use conditional on our controls. Thus, our IV approach may underestimate the effect of glyphosate on health because it measures the effect of glyphosate relative to the profile of herbicides farmers used before adopting genetically modified crops and glyphosate. We present estimates with and without controlling for other herbicides that farmers would use instead of glyphosate, but prefer the specification without these controls since it is policy-relevant.

## 5 Results

Here we present the results from estimating the main specifications for the models detailed in Section 4.

### 5.1 First stage effect of GM Adoption

Figure 5 shows the results from estimating (3) with glyphosate as the outcome variable. The reference year is 1995, the year before the release of GM crops. We can see a large increase in the difference in glyphosate use between high and low soy suitability counties starting in the year that GM soy was released. As adoption of GM soy increased over the late 1990s and early 2000s, the gap in glyphosate sprayed between high and low soy suitability counties grows. We see similar results when looking at upstream glyphosate—an increase in the share of high GM yield counties leads to an increase in upstream glyphosate. Table 2 reports the difference-in-differences estimates for the first stage. High GM yield counties increase glyphosate use by  $0.0048 \text{ kg/km}^2$  more than low GM yield counties after 1995. Similarly, going from no high GM yield counties upstream to all high GM yield counties upstream increases the amount of upstream glyphosate.

These first-stage results align with the intuition behind using attainable yield as a first stage—higher attainable yield leads to increased adoption of GM crops, which increases glyphosate since glyphosate and GM crops are complementary products. We do not include any additional controls, and the standard errors are clustered by county and year.

### 5.2 Reduced-form effect of GM Adoption

Figure 6 shows the results from estimating (3) with birth weight as the outcome variable. These estimates are the reduced form effect of GM adoption on birth weight. Before the release of GM crops, there was a consistent null effect of both the local and upstream high GM attainable yield treatment variables. These consistent pre-trends support our parallel trends assumption. When GM crops became available in 1996, birth weight in local high GM yield counties decreased

relative to the low soy suitability counties. In 2004, birth weight in rural, high GM yield counties decreased by 15 grams relative to the change in birth weights for rural, low GM yield counties over the same time period. When including controls for mother demographics, the effects are qualitatively unchanged, but slightly larger in magnitude.

However, there is a precisely estimated null effect of high GM attainable yield counties upstream. These results suggest that there are not downstream spillovers on infant health (at least as measured by birth weight) from spraying glyphosate.

Table 3 shows the difference-in-differences estimate of the reduced form effect of GM crop adoption on birth weight. Birth weight in rural, high GM yield counties decreased by 6 grams after the release of GM crops relative to the change in birth weights for rural, low GM yield counties. This specification underestimates the effect since it includes earliest years of GM when there were low adoption rates. Including mother demographic controls increases the magnitude of this estimate slightly. As with the event study, we have a precise null effect of upstream treatment on birth weight.

In interpreting this as the causal effect of GM crop adoption on birth weights, we must assume that if GM crops were not released, the difference in birth weights between high and low soy suitability counties would have remained constant. Section 6 explores and systematically rules out some alternative explanations for the decrease in birth weights in high GM yield counties relative to low GM yield counties, providing additional evidence supporting our identifying assumptions.

### 5.3 Effect of glyphosate with 2SLS

Since we have isolated exogenous variation in glyphosate, we can then use that exogenous variation to identify the health effect of glyphosate in equation (2) using two-stage least squares. We use the first stage event study from equation (3) to predict glyphosate, which we then use as an instrument for actual glyphosate use. We find negative and statistically significant effects of glyphosate on birth weights. To get a sense of the magnitude of the estimates, going from the median to the 90th percentile of glyphosate use in 2004 would cause a 34.5 gram decrease in

birth weight, using the coefficient from the model without any controls. Controlling for mother demographics increases the magnitude of the point estimate, though we cannot reject that the two estimates are equal.

An important caveat to these results is that interpreting the effect of glyphosate on health causally requires that we assume that our instruments only affect birth weight through their effect on glyphosate. Since farmers were substituting away from other pesticides, the decrease in other chemicals will bias our estimates positively, meaning that we can interpret the effects here as a lower bound on the true effects of glyphosate. We include a few common herbicides that farmers would have substituted away from in the regression, and the magnitude of our estimated effect of glyphosate on birth weight increases as expected.

## 6 Extensions/R robustness

Here we present various robustness checks to ensure that other factors are not driving our measured health effects of GM crops and glyphosate.

### 6.1 Alternate treatment criteria

Our institutional setup is not truly a binary treated or untreated situation. Every county has access to GM technology, but our identification relies on varying exposure to the release of these GM seeds. Our choice to use the median to split high and low GM attainable yield is arbitrary, thus we first check sensitivity to this decision by comparing results when we split at the 40th or 60th percentile. Additionally, we have results when we exclude the middle 10 percent and consider above 55th percentile high yield and below the 45th percentile low yield. Figure 7 shows the first stage and reduced form event studies using these alternative treatment criteria. Our results are qualitatively unchanged.

Since we use the average attainable yield across soy, corn, and cotton in our main specification, we can also use the crops individually to define treatment. Figure ?? shows the results when we use soy attainable yield to define treatment. The first stage is more precise and larger in

magnitude than when we use the GM average. However, the reduced form effect on birth weight is quite similar.

In our main analysis, we exclude counties west of the 100th meridian. This is because there are very few GM crops grown and little glyphosate used west of the 100th meridian. Figure 9 shows the results if we define treatment using the median GM or soy attainable yield across the entire country. Again, the results are similar to our main specification.

## 6.2 Effect on Acreage

It is possible that increases in agricultural activity on an extensive margin are driving our results. GM technology could lead to marginal, not previously farmed land into agricultural production. Here we explore whether being highly suitable for GM crops led to changes in planted crop acreage.

Figure 10 shows the event study estimation for the effect of high GM yield on total acreage planted in a county, revealing that GM technology has a null effect on total planted acres. However, there is a statistically significant decrease in soy acres starting in 2000, but no significant change in corn.

## 6.3 Placebo tests

**Coming soon!** There are a few placebo tests that we can run

- Define treatment based on non-gm crops
- Use downstream rather than upstream glyphosate

## 7 Conclusions

GM crops and glyphosate are technological innovations that changed agriculture in the US and globally. Here, we quantify an externality from those technological innovations on infant health. We use the timing of the release of GM seeds in 1996 along with spatial variation in the suitability of the land for growing GM crops to isolate exogenous variation in GM adoption and

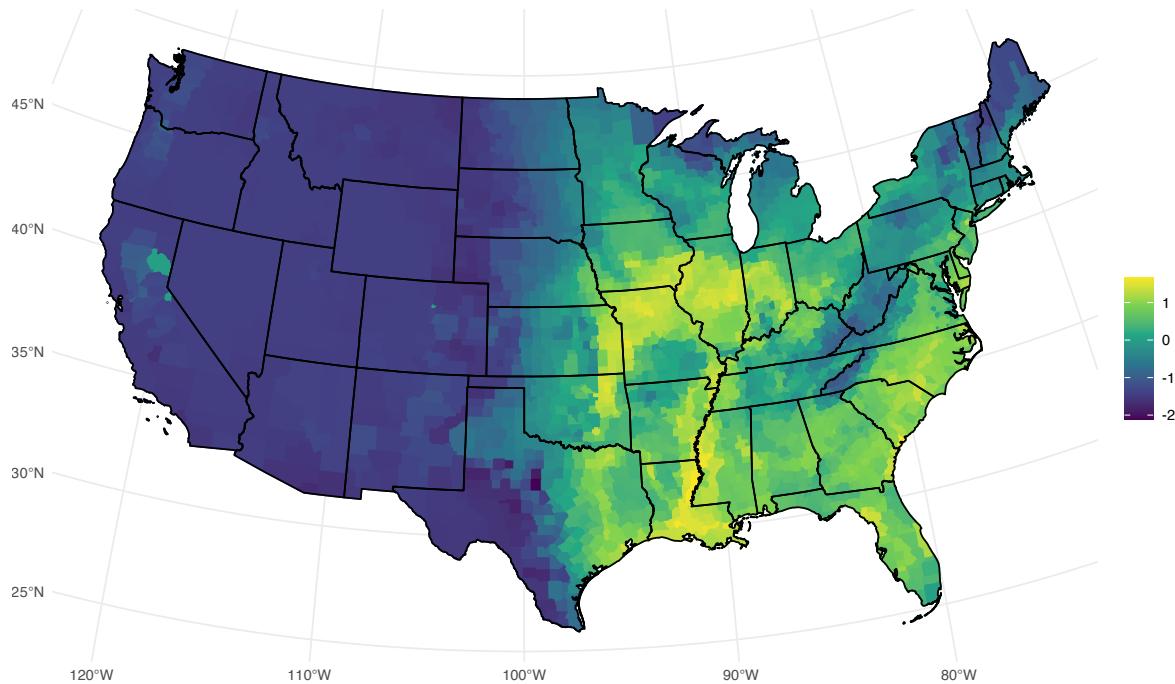
glyphosate use. Our main results show that high GM attainable yield decreased birth weight by 10-15 grams after 1995 relative to the change in birth weight in low GM attainable yield counties. If we assume that the effect of GM crops on birth weight is through glyphosate, this suggests that going from the 50th to 90th percentile in glyphosate use in 2004 leads to a 34.5-gram decrease in birth weight.

These results suggest that GM crops have a net negative health effect, in contrast to the widely held belief that GM crops are beneficial to health since glyphosate is less toxic than the herbicides it replaced. Given these negative health externalities, policy-makers must have additional discussions about the regulation of glyphosate and GM crops to determine the set of socially optimal policies around these technologies.

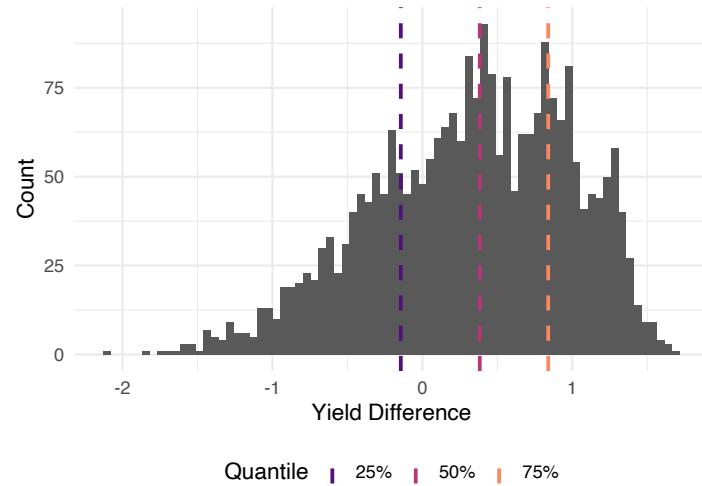
## 8 Figures

**Figure 1:** The distribution of attainable yield for GM crops

(a) Attainable yield for GM crops by county

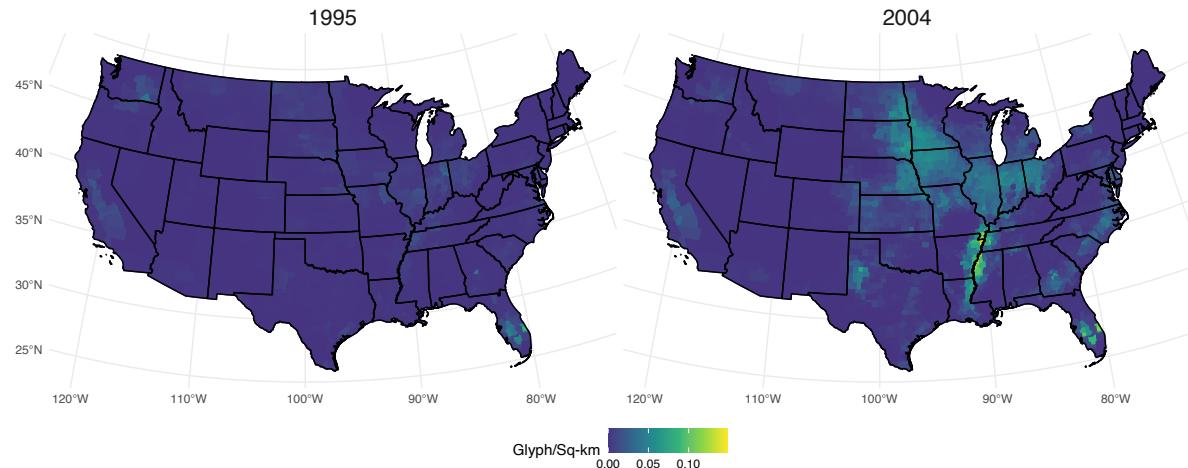


(b) Distribution of attainable yield for GM crops east of the 100th meridian

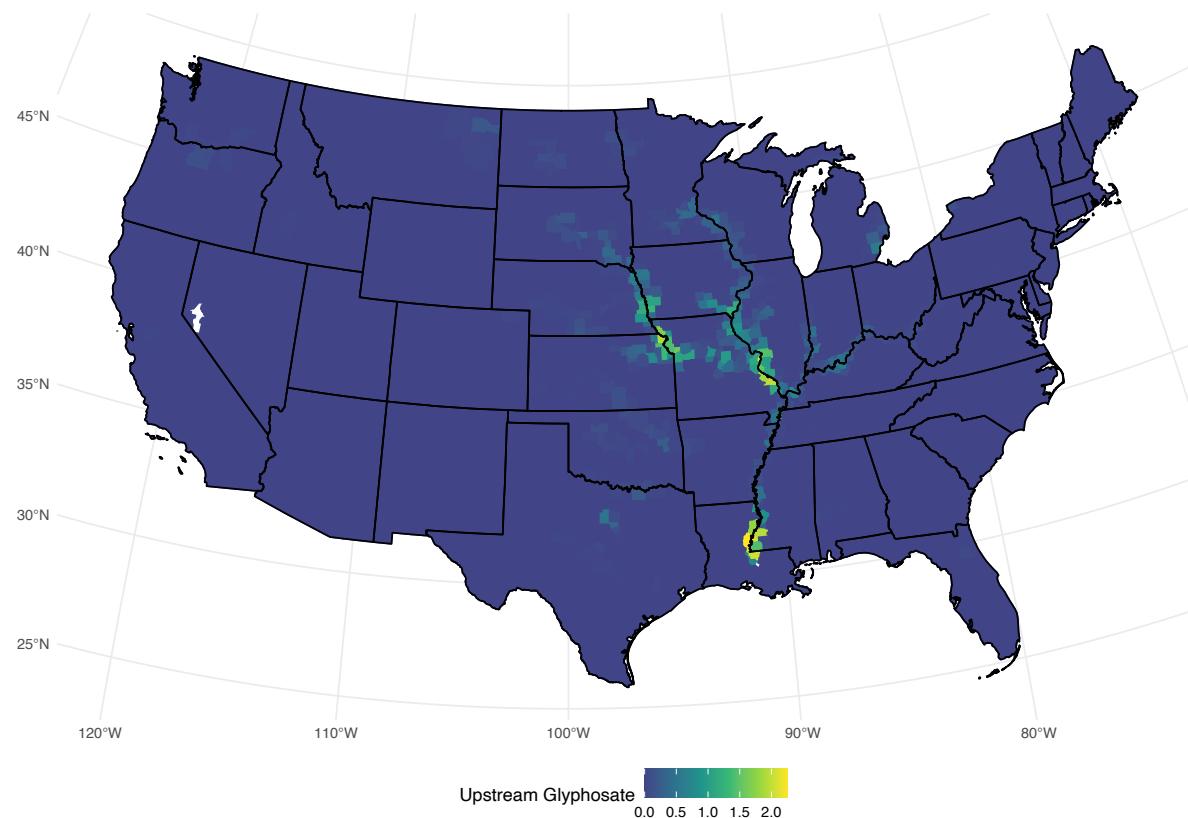


These figures show the difference between attainable yield for the high and low input scenarios averaged across soy, corn, and cotton, after the difference for each crop has been standardized. The histogram only includes counties east of the 100th meridian. Dashed lines represent options to split high and low suitability for each crop.

**Figure 2:** Glyphosate use by county in 1995 and 2004.

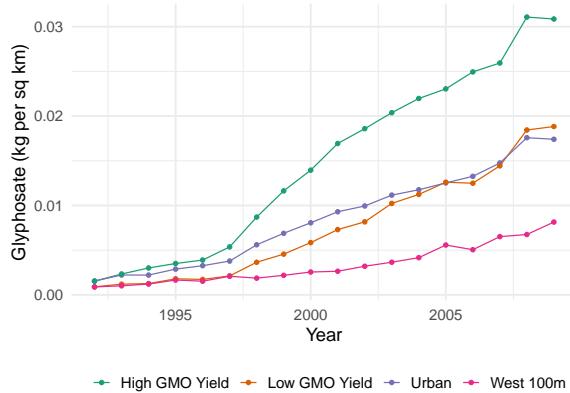


**Figure 3:** Upstream Glyphosate in 2004 by County

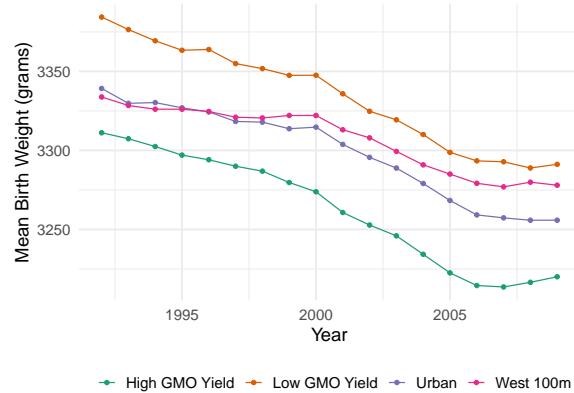


**Figure 4:** Identifying variation for reduced form estimates.

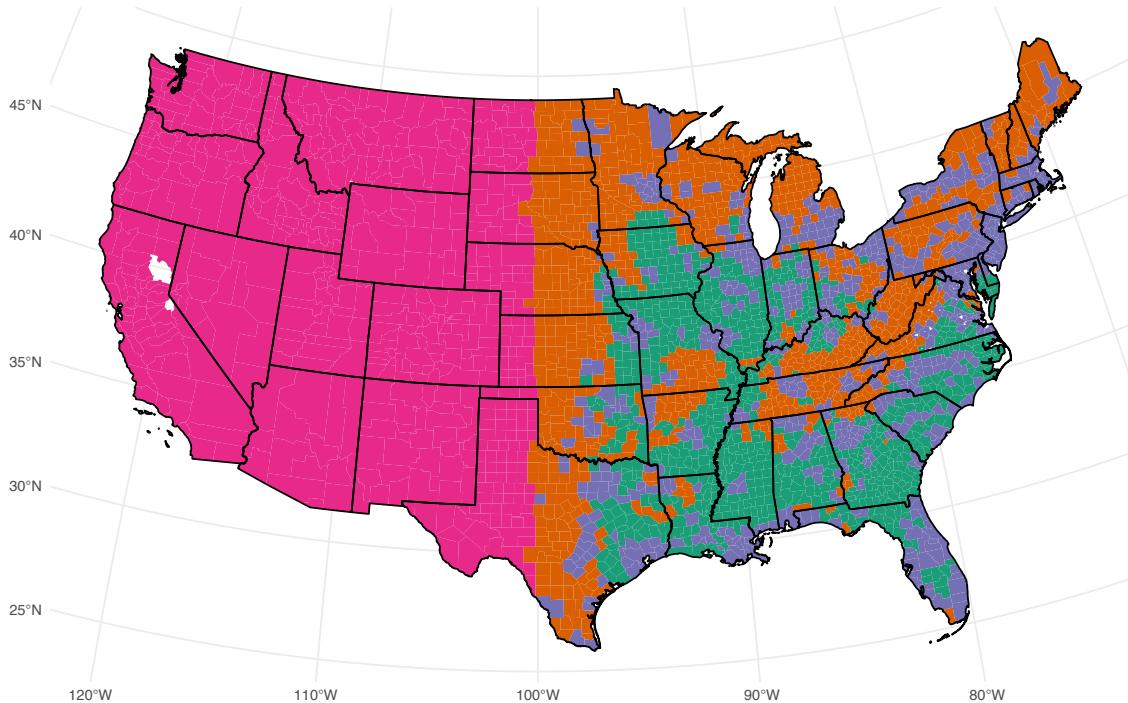
(a) Glyphosate intensity by year.



(b) Mean birth weight by year.



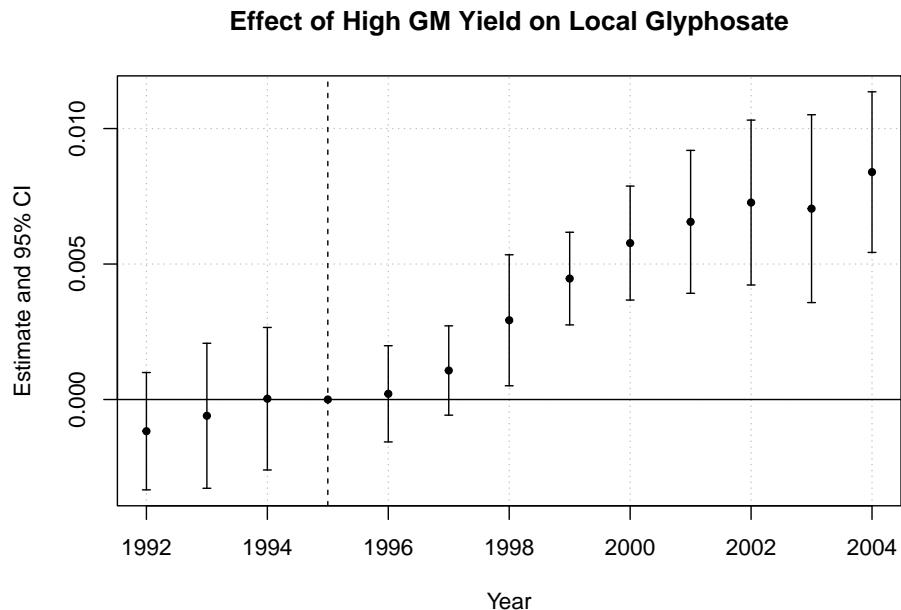
(c) High and low GM yield counties.



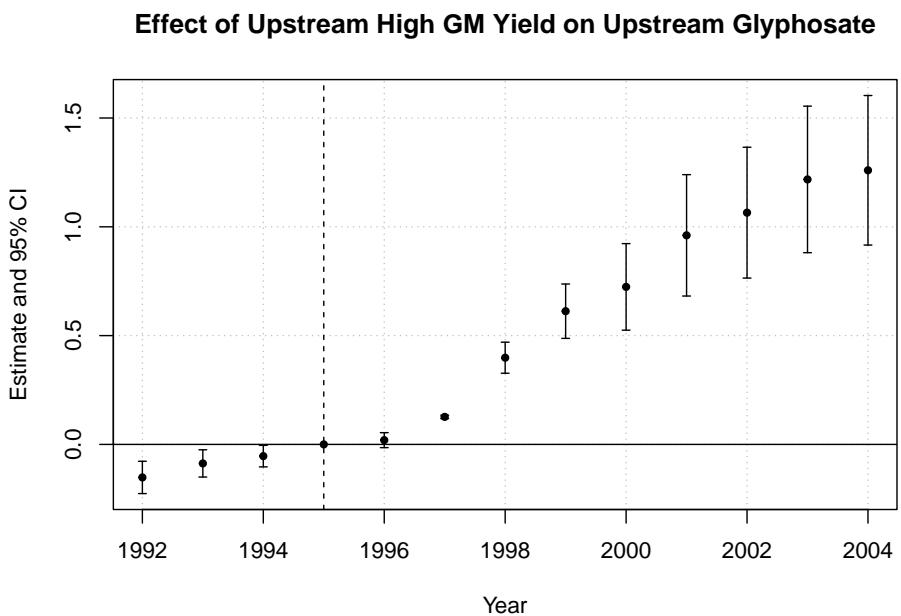
High (low) GM yield denote rural counties above (below) the median attainable yield difference, which is calculated as an average across soy, corn, and cotton, for counties east of the 100th meridian. Urban denotes urban counties east of the 100th meridian.

**Figure 5:** First stage effect of high GM yield on glyphosate

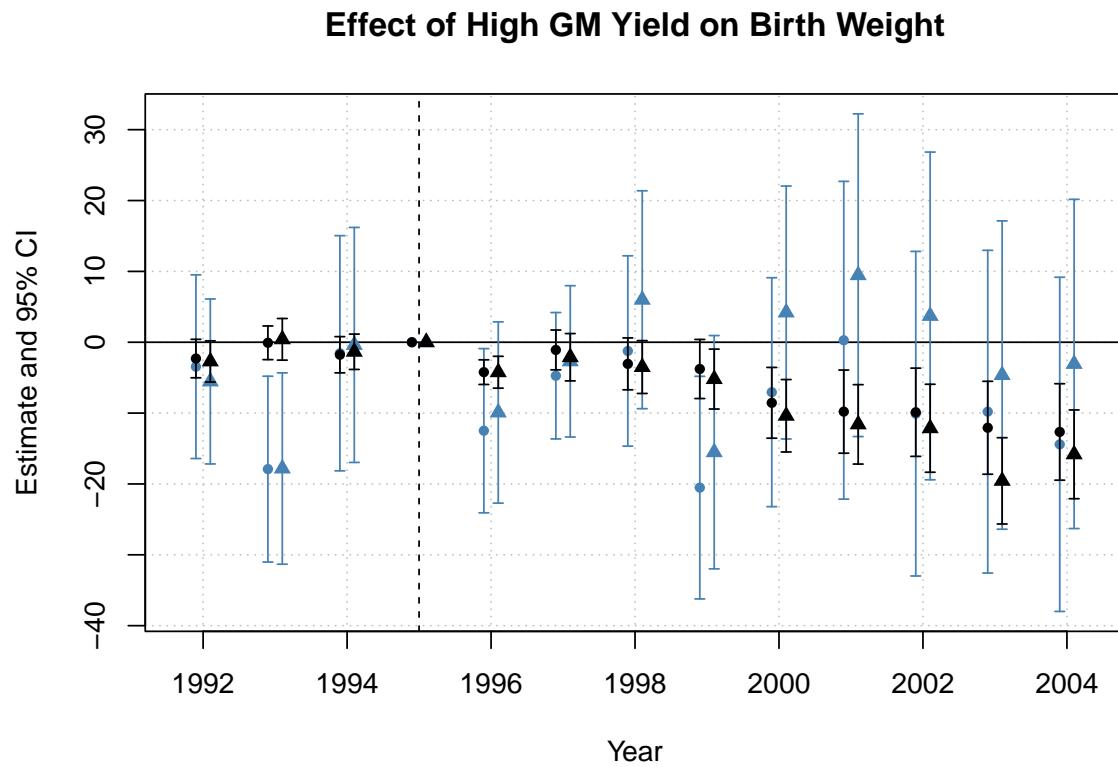
(a) First stage event study for local glyphosate



(b) First stage event study for upstream glyphosate



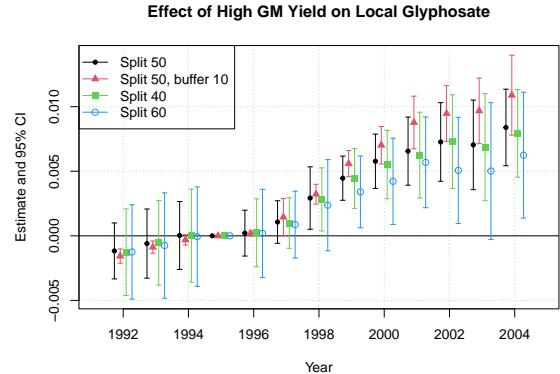
**Figure 6:** Reduced form event study



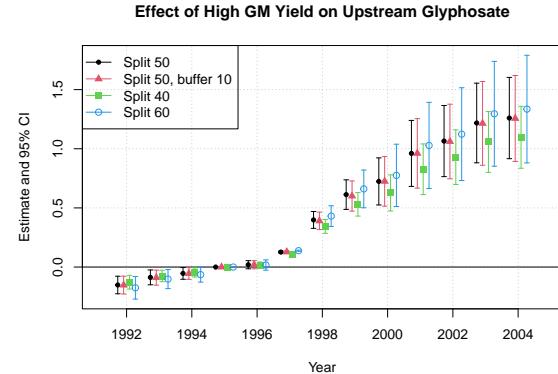
Blue is effect from upstream glyphosate and black is effect from local glyphosate. Dots are estimates from the model without controls, triangles are estimates from the model that includes controls.

**Figure 7:** Event study with different high and low GM yield splits

(a) First stage event study for local glyphosate with different splits

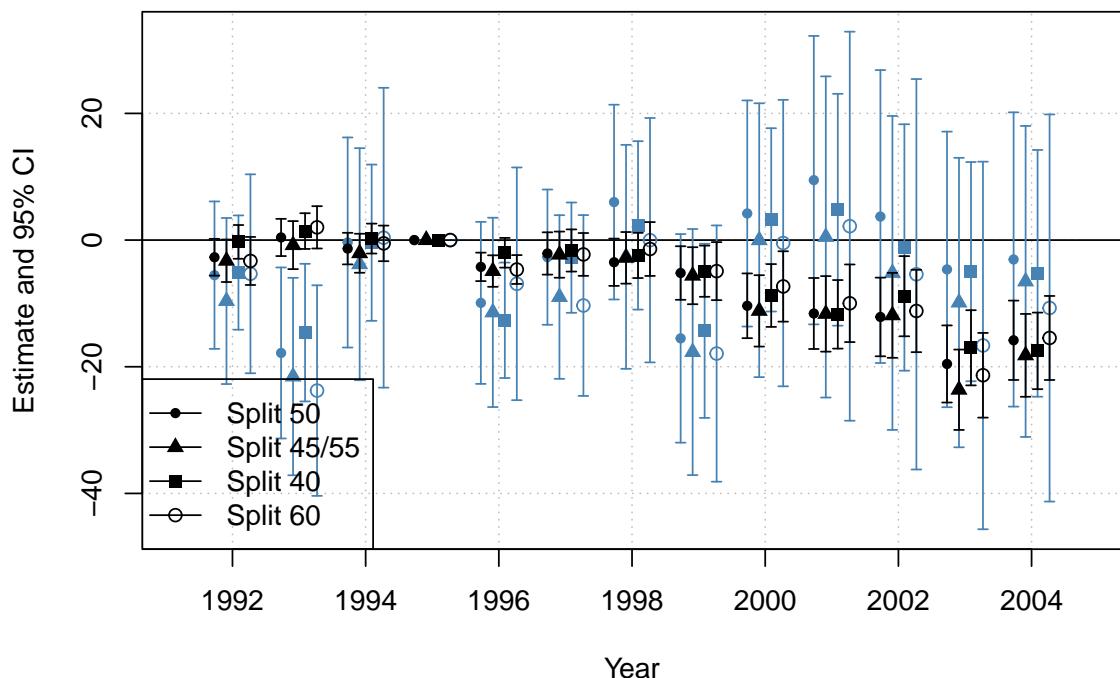


(b) First stage event study for upstream glyphosate with different splits



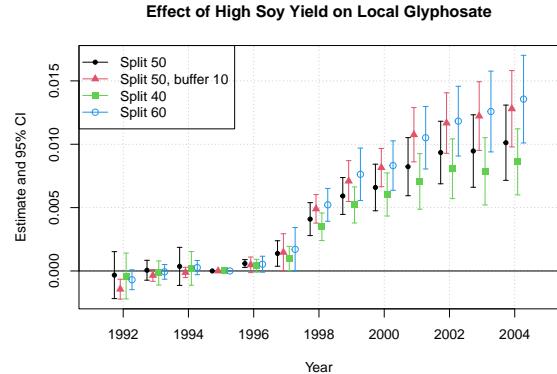
(c) Reduced from event study with different splits

### Effect of High GM Yield on Birth Weight

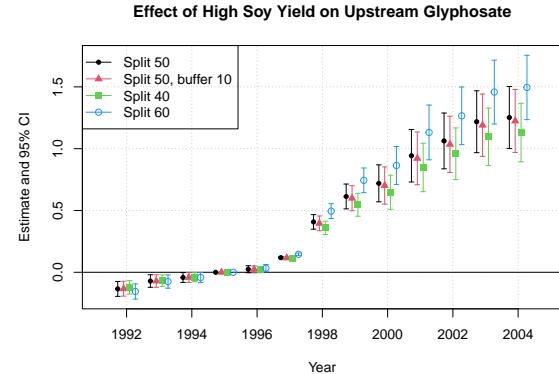


**Figure 8:** Event study using attainable yield for soy

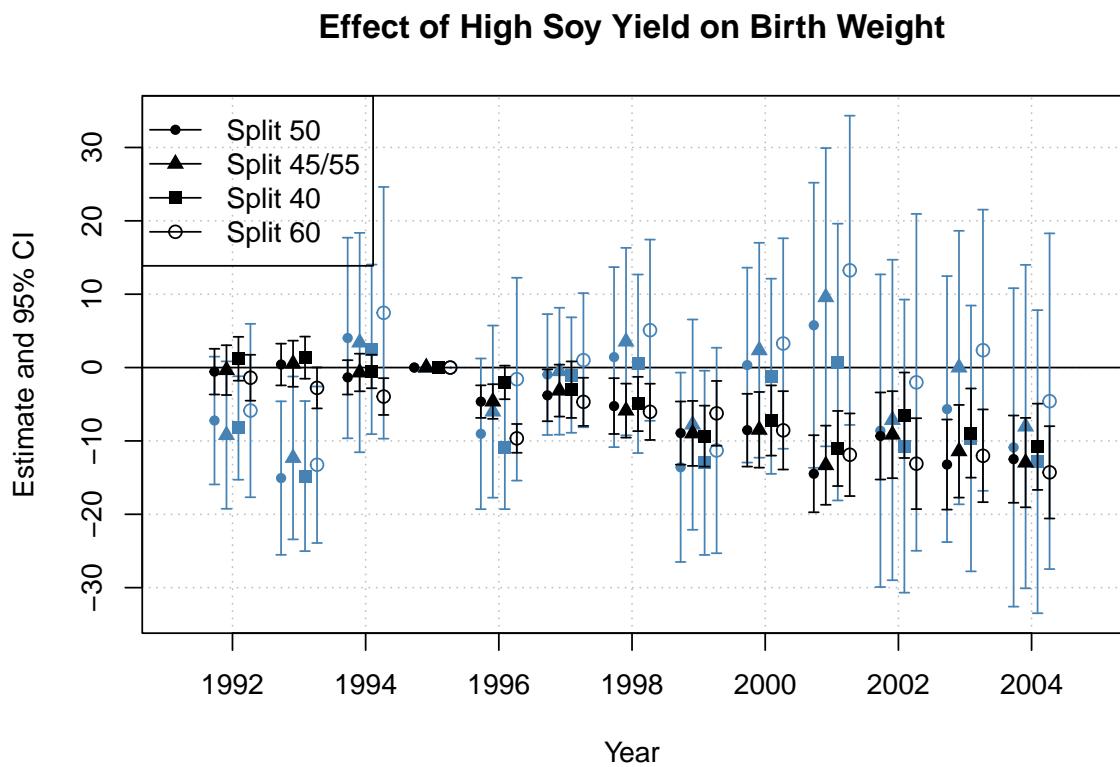
(a) First stage event study for local glyphosate using soy yield



(b) First stage event study for upstream glyphosate using soy yield

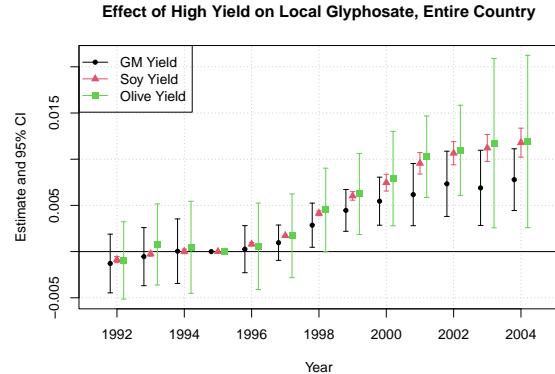


(c) Reduced from event study using soy yield

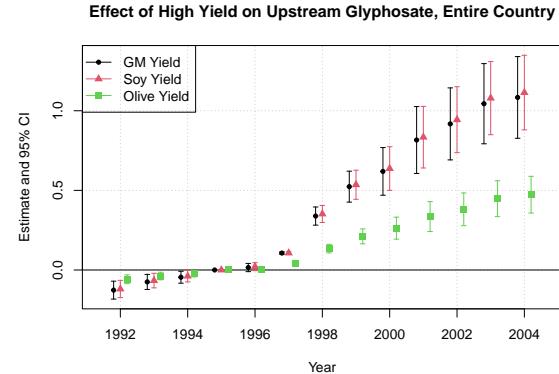


**Figure 9:** Event study using the entire country

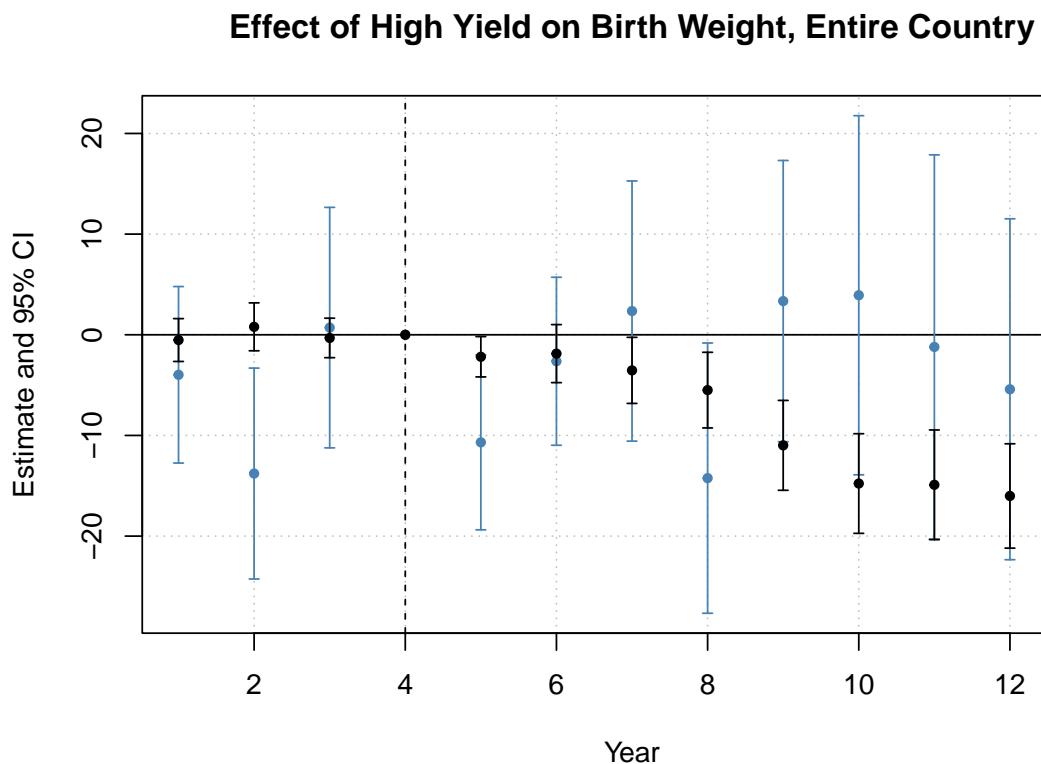
(a) First stage event study for local glyphosate using the entire country



(b) First stage event study for upstream glyphosate using the entire country

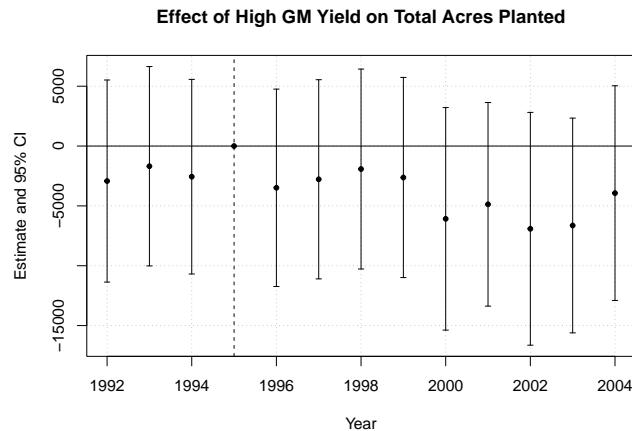


(c) Reduced from event study using the entire country

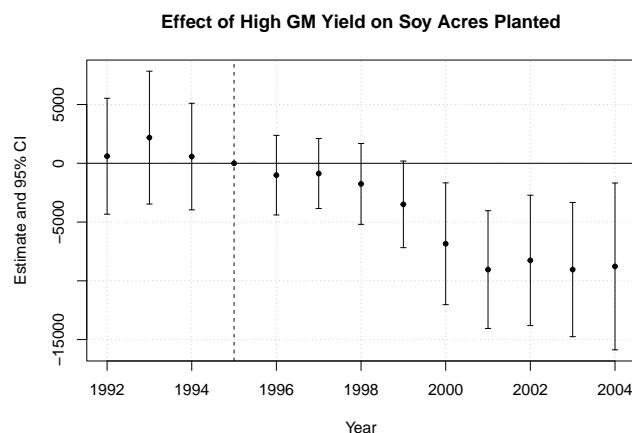


**Figure 10:** Effect of high GM yield on crop acreage

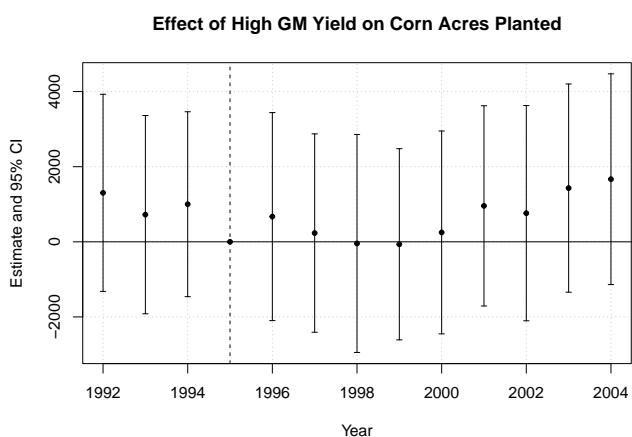
(a) Total crop acreage



(b) Soy acreage



(c) Corn acreage



## 9 Tables

**Table 1:** Summary Statistics by county between 1992 and 1995

Variable	High GM Yield		Low GM Yield		Urban		West 100m	
	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd
Number of Counties	838	0	935	0	799	0	488	0
Birth Weight (g)	3345.69	81.89	3407.69	72.02	3386.99	59.69	3354.3	101.83
Pct Low Birth Weight	7.91	2.17	6.2	1.7	6.91	1.44	6.38	2.36
Percent Male	51.11	1.85	51.19	2.41	51.26	0.9	51.53	3.38
Infant Mortality	3.71	3.26	3.39	5.37	5.88	6.83	3.31	3.65
Total Births	346.85	286.17	301.63	300.42	3765.97	9079.14	328.33	1516.54
Glyphosate ( $g/km^2$ )	2.59	3.09	1.29	1.46	2.25	3.72	1.05	1.49
Total Crop Area ( $km^2$ )	354.49	414.88	351.65	490.9	243.35	386.54	337.28	478.71
Total Pop (1000's)	25.26	19.37	24.01	23.11	241.71	485.73	17.65	22.76
Percent Hispanic	1.39	2.81	3.32	11.31	5.25	10.07	12.42	17.54
Unemployment Rate	7.04	2.57	6.78	3.49	5.95	2.43	6.7	4.06
Pct Some HS Degree	35.95	8.91	32.86	10.49	24.99	8.22	25.48	8.83
Pct HS Degree	35.43	5.9	35.86	6.16	32.64	6.15	32.86	4.87
Pct Some College	18.32	4.29	19.8	5.26	24.46	5.12	26.98	5.07
Pct College Degree	10.3	3.58	11.48	4.64	17.92	7.83	14.68	5.72
Income per Capita	16.38	2.18	16.82	2.73	20.79	4.18	17.76	3.89

Means and standard deviations are calculated on county level averages between 1992 and 1995, which is the period prior to the release of GM crops.

**Table 2:** Difference-in-Differences estimates for the first stage effect of GM crops on glyphosate

Dependent Variables:	Local Glyphosate (1)	Upstream Glyphosate (2)
<i>Variables</i>		
Local Treatment	0.0048** (0.0016)	
Upstream Treatment		0.7154*** (0.1753)
<i>Fixed-effects</i>		
Year	Yes	Yes
GEOID	Yes	Yes
<i>Fit statistics</i>		
Observations	21,994	28,083
R <sup>2</sup>	0.649	0.717
Within R <sup>2</sup>	0.040	0.215
<i>Clustered (Year &amp; GEOID) standard-errors in parentheses</i>		
<i>Signif. Codes: ***: 0.01, **: 0.05, *: 0.1</i>		

**Table 3:** Difference-in-Differences estimates for the reduced form effect of GM crops on birth weight

Model:	Dependent Variable: Birth Weight		
	(1)	(2)	(3)
<i>Variables</i>			
Local Treatment	-5.87* (2.95)	-5.68* (3.05)	-7.01** (3.06)
Upstream Treatment		-3.24 (9.30)	4.26 (9.54)
<i>Fixed-effects</i>			
Year	Yes	Yes	Yes
County	Yes	Yes	Yes
<i>Fit statistics</i>			
Observations	5,383,402	5,383,402	5,037,287
R <sup>2</sup>	0.028	0.028	0.070
Within R <sup>2</sup>	0.000	0.000	0.043
Controls	No	No	Yes

*Clustered (Year & County) standard-errors in parentheses*

*Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1*

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.

**Table 4:** Two Stage Least Squares Estimates for the effect of glyphosate on birth weight

Dependent Variable:	Birth Weight		
Model:	(1)	(2)	(3)
<i>Variables</i>			
Glyphosate	-1,000.08** (387.80)	-1,221.08*** (394.34)	-1,302.95* (601.22)
<i>Fixed-effects</i>			
Year	Yes	Yes	Yes
County	Yes	Yes	Yes
<i>Fit statistics</i>			
Observations	5,383,402	5,037,287	5,383,402
F-test (1st stage), Glyphosate	37,150.2	34,871.3	23,641.5
Controls	None	Demographic	Pesticides

*Clustered (Year & County) standard-errors in parentheses*

*Signif. Codes:* \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

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, cyanazine, fluazifop, metolachlor, metribuzin, and nicosulfuron (but no demographic controls).

## References

- Baer, KM and Marcel, BJ** (2014). *Glyphosate*. Edition: 3 Pages: 767-769 Publication Title: Encyclopedia of Toxicology Volume: 1.
- Baker, Nancy T. and Stone, Wesley W.** (2015). *Estimated annual agricultural pesticide use for counties of the conterminous United States, 2008-12*. Tech. rep. 907. ISSN: 2327-638X Publication Title: Data Series. U.S. Geological Survey. (Visited on 05/16/2022).
- Barrows, Geoffrey, Sexton, Steven, and Zilberman, David** (2014). “Agricultural Biotechnology: The Promise and Prospects of Genetically Modified Crops”. *Journal of Economic Perspectives* 28.1, 99–120.
- Brainerd, Elizabeth and Menon, Nidhiya** (2014). “Seasonal effects of water quality: The hidden costs of the Green Revolution to infant and child health in India”. *Journal of Development Economics* 107, 49–64.
- Camacho, Adriana and Mejia, Daniel** (2017). “The health consequences of aerial spraying illicit crops: The case of Colombia”. en. *Journal of Health Economics* 54, 147–160. (Visited on 10/04/2021).
- Chay, Kenneth Y. and Greenstone, Michael** (2003). “The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession\*”. *The Quarterly Journal of Economics* 118.3, 1121–1167. (Visited on 09/08/2021).
- Dias, Mateus, Rocha, Rudi, and Soares, Rodrigo R** (2019). “Glyphosate Use in Agriculture and Birth Outcomes of Surrounding Populations”.
- Duke, Stephen O and Powles, Stephen B** (2008). “Glyphosate: a once-in-a-century herbicide”. en. *Pest Management Science* 64.4, 319–325. (Visited on 06/13/2022).
- EPA** (2019). “Glyphosate: Response to Comments, Usage, and Benefits”. – (2020). “EPA Finalizes Glyphosate Mitigation”. Publisher: Environmental Protection Agency.
- FAO and IIASA** (2022). *Global Agro Ecological Zones version 4 (GAEZ v4)*. (Visited on 06/14/2022).
- Fernandez-Cornejo, Jorge, Wechsler, Seth, Livingston, Mike, and Mitchell, Lorraine** (2014). *Genetically Engineered Crops in the United States*. en. SSRN Scholarly Paper ID 2503388. Rochester, NY: Social Science Research Network. (Visited on 10/04/2021).

- Frank, Eyal** (2021). “Substitution Costs of Natural Inputs: Increased Pesticide Use Following Decline in Biological Pest Control”.
- Garry, Vincent F, Harkins, Mary E, Erickson, Leanna L, Long-Simpson, Leslie K, Holland, Seth E, and Burroughs, Barbara L** (2002). “Birth defects, season of conception, and sex of children born to pesticide applicators living in the Red River Valley of Minnesota, USA.” *Environmental Health Perspectives* 110.suppl 3, 441–449.
- Grube, Arthur, Donaldson, David, Kiely, Timothy, and Wu, La** (2011). “Pesticides Industry Sales and Usage: 2006 and 2007 Market Estimates”. Publisher: EPA.
- IARC** (2015). “Some Organophosphate Insecticides and Herbicides”. *Monographs* 112, 321–398.
- Landrigan, Philip J.** (2018). “Pesticides and Human Reproduction”. eng. *JAMA internal medicine* 178.1, 26–27.
- Larsen, Ashley E., Gaines, Steven D., and Deschenes, Olivier** (2017). “Agricultural pesticide use and adverse birth outcomes in the San Joaquin Valley of California”. *Nature Communications* 8.1.
- Lechner, Bernhard and Grill, GÄCEnther** (2013). “Global river hydrography and network routing: baseline data and new approaches to study the world’s large river systems”. en. *Hydrological Processes* 27.15. \_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/hyp.9740>, 2171–2186. (Visited on 06/06/2022).
- Leino, Lyydia, Tall, Tuomas, Helander, Marjo, Saloniemi, Irma, Saikkonen, Kari, Ruuskanen, Suvi, and Puigbo, Pere** (2021). “Classification of the glyphosate target enzyme (5-enolpyruvylshikimate-3-phosphate synthase) for assessing sensitivity of organisms to the herbicide”. *Journal of Hazardous Materials* 408.
- Muller, Birgit** (2020). “Glyphosate-A love story. Ordinary thoughtlessness and responseability in industrial farming”. *Journal of Agrarian Change* 21.1, 160–179.
- Munoz, Juan P., Bleak, Tammy C., and Calaf, Gloria M.** (2020). “Glyphosate and the key characteristics of an endocrine disruptor: A review”. *Chemosphere*.

- Portier, Christopher J.** (2020). “A comprehensive analysis of the animal carcinogenicity data for glyphosate from chronic exposure rodent carcinogenicity studies”. *Environmental Health* 19.1.
- Regidor, E** (2004). “Paternal exposure to agricultural pesticides and cause specific fetal death”. *Occupational and Environmental Medicine* 61.4, 334–339.
- Taylor, Charles A** (2022). “Cicadian Rhythm: Insecticides, Infant Health and Long-term Outcomes”.
- Thelin, Gale and Stone, Wesley** (2013). *Estimation of Annual Agricultural Pesticide Use for Counties of the Conterminous United States, 1992-2009*. en. Scientific Investigations Report 2013-5009. Series: Scientific Investigations Report. U.S. Geological Survey.
- University, Center For International Earth Science Information Network-CIESIN-Columbia** (2017). *U.S. Census Grids (Summary File 1), 2010*. Type: dataset. (Visited on 06/09/2022).

## Appendix A

### A.1 Appendix: Upstream and Downstream Glyphosate Estimation

To measure spillover effects from glyphosate sprayed upstream, we must have some sort of measure for glyphosate exposure in water. Ideally, this would come from extensive monitoring that 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 use a limited amount of glyphosate monitoring in water to help validate our results.

We use a spatial water model to aggregate the amount of glyphosate sprayed upstream and downstream of each county. Specifically, we use the level 8 HydroBASINS product from HydroSHEDS (Lehner and Grill, 2013). These data are watershed polygons that delineate water basins across the globe in a standardized way. Importantly, they are assigned codes in such a way that makes it possible to find all watersheds upstream and downstream from any given watershed.

Our pesticide data are at the county level, thus we begin by disaggregating the county-level pesticide use into watersheds within each county. We assume that spraying is uniform across the county and multiply the total pesticide used in a county by the portion of the county's total area covered by the watershed. Figure A3 shows the spatial distribution of glyphosate use per square kilometer by watershed across the United States in 2006.

We then utilize the "Pfafstetter" watershed coding system used by the HydroBASINS data to find all of the watersheds upstream and all of the watersheds downstream from every watershed. Spraying upstream has the potential to affect health outcomes in a given location, but downstream spraying should not. Thus, we can use the downstream spraying as a placebo test that should not return significant results. To demonstrate the process, we have selected an example watershed in Washington County, Illinois, which is just east of St. Louis. Figure A4 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. We then must aggregate all of the pesticides sprayed upstream of each watershed. Since there is no chemical transport model for glyphosate,<sup>3</sup> we choose a variety of weighting functions to aggregate upstream pesticides, where the weights are generally decreasing as the distance between the target watershed and the upstream watershed increases.

Once we have aggregated upstream and downstream pesticides to the watershed level, we must reaggregate to the county level to compare it to our health outcomes, which are all geo-

---

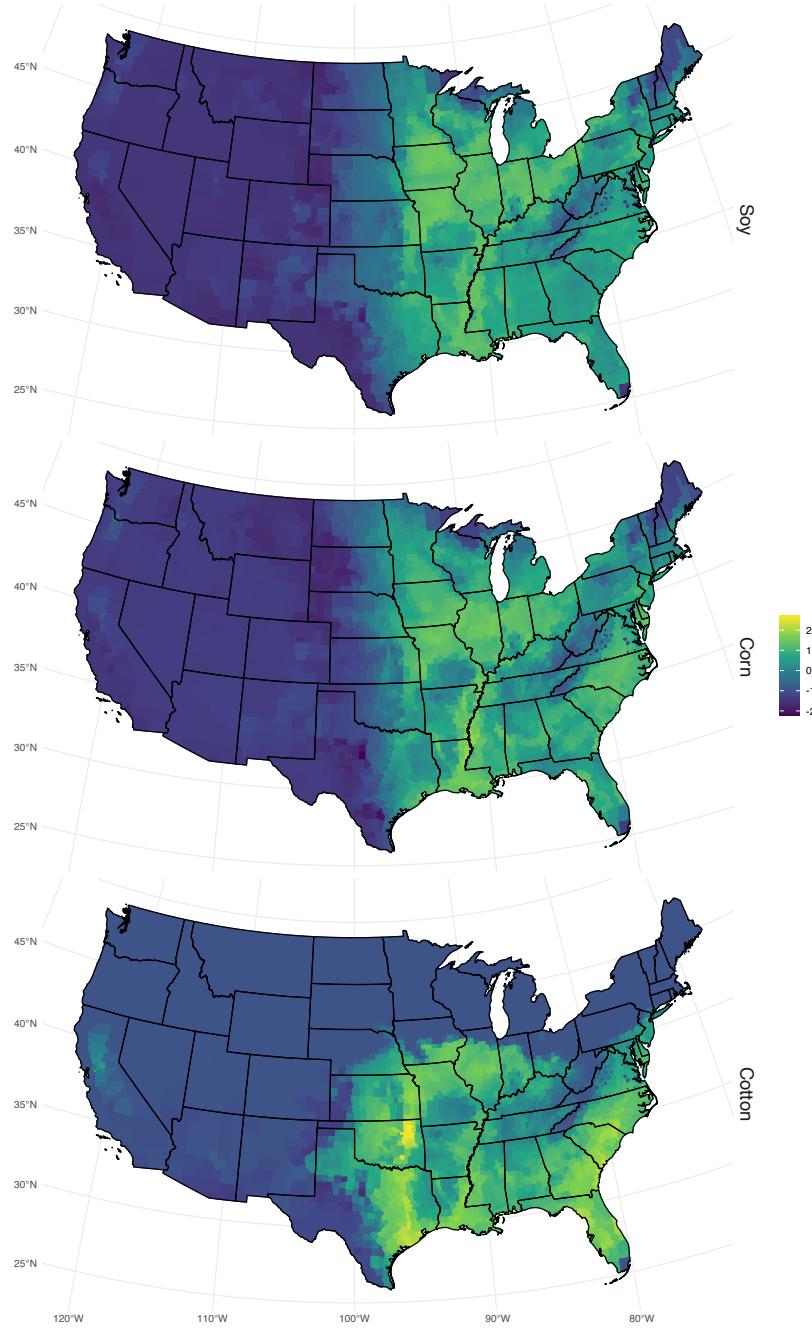
<sup>3</sup> Something like the InMAP or HySPLIT models for air particles

graphically identified at the county level. For this, we weight the pesticides sprayed upstream of each watershed by the proportion of the county's population that lives in the watershed. Our population estimates come from SEDAC's 1990, 2000, and 2010 population grids (University, 2017). These grids give estimates of the population for one square kilometer pixels across the United States. We add up the population counts for cells within each watershed and then divide by the total population count for cells within the county to obtain the population weights. Figure A5 shows the upstream glyphosate for each watershed touching Washington county on the right, and then on the left has the population weight for each of those watersheds.

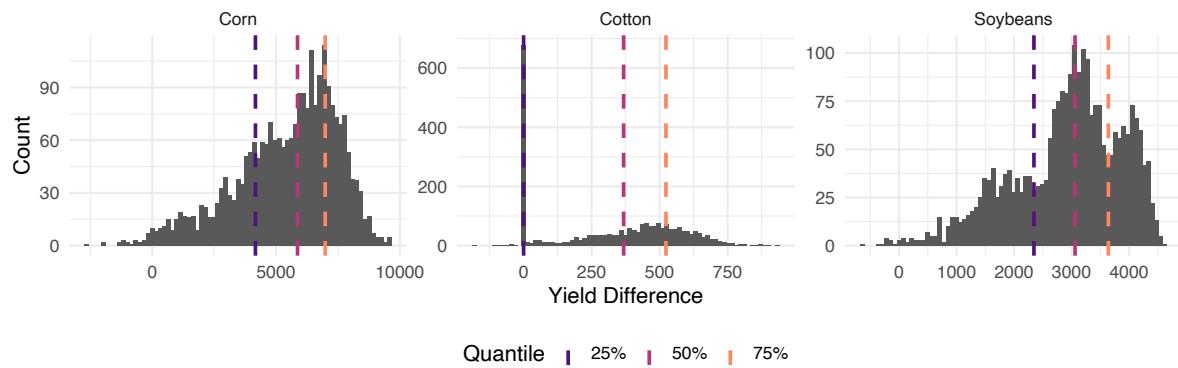
Finally, we take the weighted sum of upstream pesticides using the population weights described above. This gives us an estimate of the amount of pesticides in surface water for each county, which is the result shown in Figure 3.

## A.2 Appendix: Figures

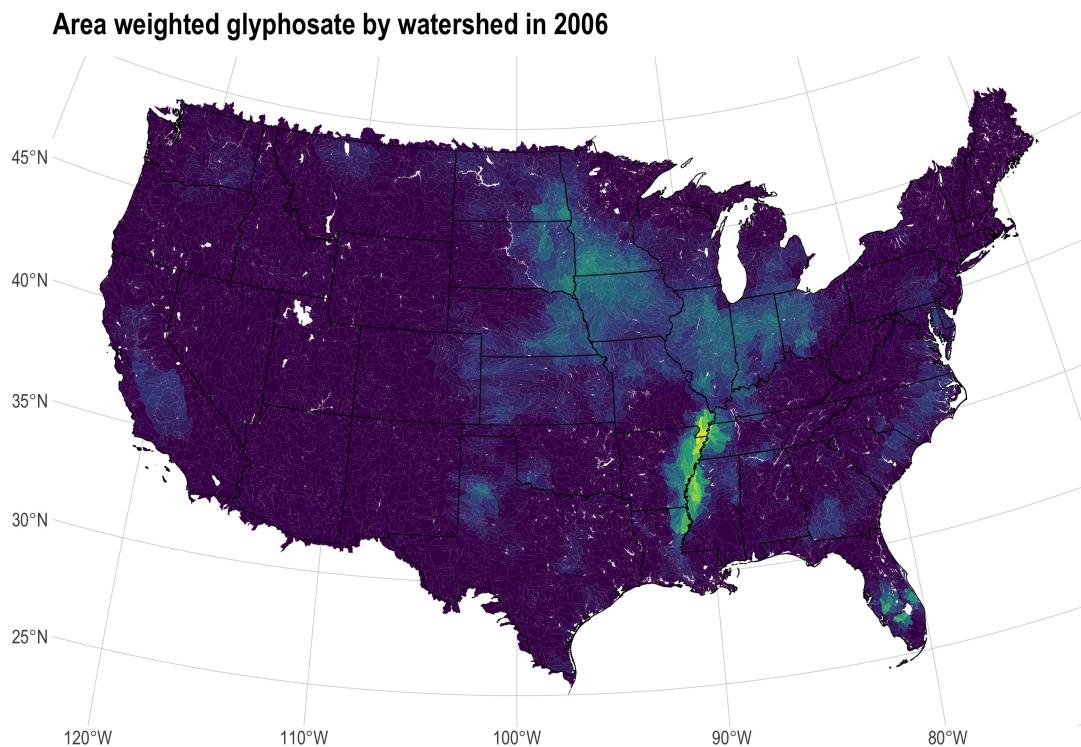
**Figure A1:** Attainable yield for Soy, Corn, and Cotton by county.



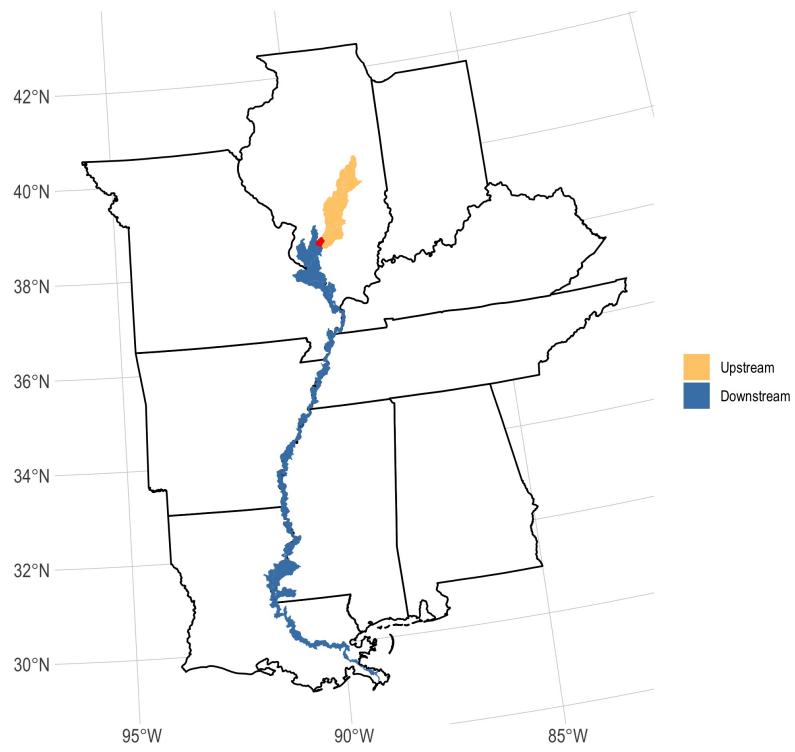
This map shows the difference between attainable yield for the high and low input scenarios for each of soy, corn, and cotton. The difference for each crop has been standardized.



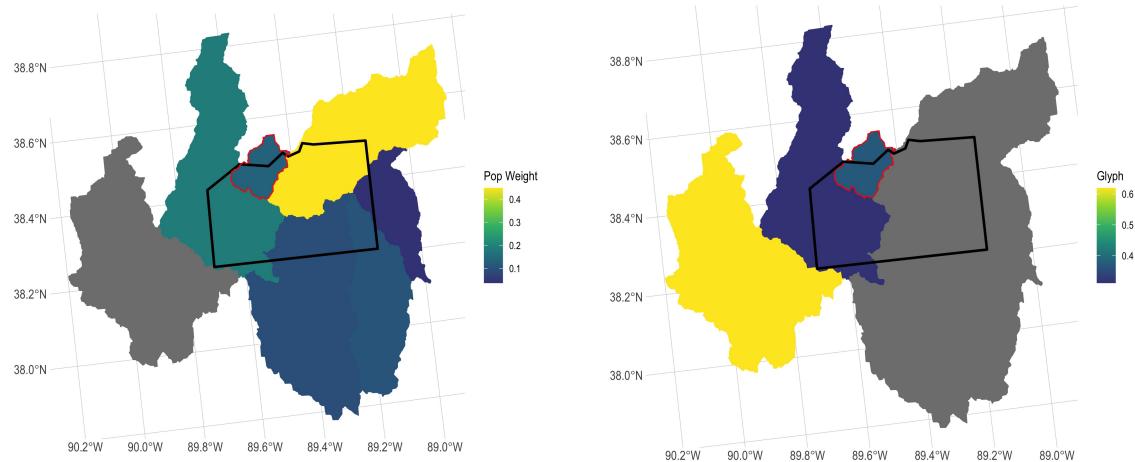
**Figure A2:** Distribution of the difference in attainable yield for high relative to low input scenarios. Dashed lines represent options to split high and low suitability for each crop.



**Figure A3:** Area weighted glyphosate per square kilometer by watershed in 2006.



**Figure A4:** Watersheds upstream and downstream of the example watershed, which is highlighted in red.



**Figure A5:** Right shows upstream glyphosate for each watershed touching Washington County. Left shows the population weights used to take the weighted sum of upstream glyphosate.