Appendix B: Synthetic Control Method Results

We use the SCM to construct synthetic controls for each SWD detected county, separately. This involved 581 runs of the method, yielding 581 synthetic controls; one for each of the 581 counties where SWD is detected. The feasible pool from which the synthetic controls can be constructed from consists of all the counties in US states where at least one SWD detection occurred during our data time period. This includes both counties where SWD is eventually detected (i.e., "ever detected" counties) and counties where SWD has never been detected as of April 2018 (i.e., "never detected" counties).[1] Each synthetic control is weighted to resemble the given treated county under consideration on observable characteristics prior to initial SWD detection.

Mathematically, let J be the number of available control counties and define the $J \times 1$ weighting vector $W = (w_1, \dots, w_J)'$ such that $\sum_{j=1}^{J} w_j = 1$ and $w_j \geq 0$ for all $j = (1, \dots, J)$. Each scalar w_j represents the non-negative weight placed on the j^{th} county in the synthetic control. Let H_0 be a $K \times 1$ vector of K pesticide use variables and other observable characteristics for the treated county prior to SWD detection. Let H_1 be a $K \times J$ matrix of comparable data vectors for each of the J counties in the feasible pool. Following Abadie and Gardeazabal (2003) as referenced in the main text, the vector of weights W^* is chosen such that,

[1] As an additional robustness check, we also ran the SCM using an alternative feasible pool that consisted of only SWD "ever detected" counties. The pool size from this run was significantly reduced, but the results were similar, albeit slightly attenuated, relative to the baseline feasible pool consisting of both "ever" and "never" detected counties.

$$W^* = \underset{W}{\operatorname{argmin}}(H_0 - H_1 W)'V(H_0 - H_1 W)$$

subject to: [B1]

$$w_j \geq 0, \ \sum_{j=1}^{J} w_j = 1 \quad for \ \ j = (1, \ldots, J)$$

where V is a positive-definite importance matrix for the variables in H_0 and H_1. We include in H_0 and H_1 the same set of covariates as were included in eq. (1) in the main text: insecticide or fungicide use and a quadratic function of annual maximum temperature, annual minimum temperature, and annual precipitation.

To compare a given SWD detected county to its synthetic control, we calculate the following DID estimator following Bohn et al. (2014), as cited in the main text,

$$DID_{SWD} = \left(Pesticide_{post}^{SWD} - Pesticide_{post}^{synthSWD}\right) - \left(Pesticide_{pre}^{SWD} - Pesticide_{pre}^{synthSWD}\right) \qquad [B2]$$

where $Pesticide_{post}$ is the mean pesticide use (either insecticides or fungicides) after SWD detection and $Pesticide_{pre}$ is the mean use before SWD detection.

Falsification tests are used to estimate uncertainty on DID_{SWD} (Abadie et al. 2010). That is, the SCM is applied to each unit in the feasible pool as if it had experienced an SWD detection. DID_{SWD} is then compared to the distribution of placebo DID estimates (DID_{PL}) obtained from each falsification test. Following Galiani and Quistorff (2016), a two-sided p-value can be then calculated,

$$\text{p-value} = \Pr(|DID^{PL}| \geq |DID^{SWD}|) = \frac{|DID_j^{PL}| \geq |DID^{SWD}|}{J} \qquad [B3]$$

where DID^{PL} is the distribution of placebo DID estimates averaged over the post-treatment period, DID^{SWD} is the estimated average DID effect for a given SWD detected county from eq. (B2), and DID_j^{PL} is the j^{th} placebo county average DID estimate for $j = (1, \ldots, J)$. In what follows, we weight each DID_j^{PL} by the pre-treatment root mean squared prediction error, as suggested by Abadie et al. (2010). This has the effect of giving more weight to placebo units with good pre-treatment match to the SWD detected county and less weight to those units with poor pre-treatment fit.

Table B1 shows the final results after averaging across all 581 runs of the SCM. We have split the results into insecticides (column 1) and fungicides (column 2). First, note the near zero values in the first row labeled "Average pre-treatment difference." This indicates that the SCM-produced counterfactual is a good fit to the treated during the pre-treatment period. After differencing the first and second rows, we arrive at the DID estimate. On average, SWD detection results in an additional 10.37 kg/mi^2 of insecticides and an additional 9.06 kg/mi^2 of fungicides being used. Both estimates are significant at the 5% level. These estimates are similar in magnitude to those in Table 2 of the main text, where the traditional DID approach was used. This provides additional evidence of an SWD-induced pesticide use story.

Table B1: SCM Results of the Impact of SWD Detection on County-Level Pesticide Use (Averaged Across All Detected Counties)

	(1)	(2)
	Insecticide Use (kg/mi^2)	Fungicide Use (kg/mi^2)
Average pre-treatment difference	0.139	-0.078
Average post-treatment difference	10.51	8.98
DID estimate	10.37	9.06
p-value	0.04	0.04

Notes: each row has been averaged across all 581 runs of the synthetic control method, separately for insecticides and fungicides. DID estimates calculated using eq. (B2). P-values calculated using eq. (B3). SWD detected counties and their idiosyncratic synthetic controls are matched on pre-treatment pesticide use and a quadratic function of annual maximum temperature, annual minimum temperature, and annual precipitation.

Appendix B References:

Galiani, S., & Quistorff, B. (2016). The synth_runner package: Utilities to automate synthetic control estimation using synth. Working Paper, University of Maryland College Park.