To (Rent) Bees or Not to (Rent) Bees?
An Examination of the Farmer's Question

Steven W. Wilcox, David R. Just, C.-Y. Cynthia Lin Lawell,
Miguel I. Gómez, Heather Grab*

November 5, 2023

Abstract

We analyze a farmer's pollination resource decision theoretically using optimal control theory and empirically using farm-level data from US apple farmers, establishing new stylized facts regarding managed pollination. We find that the likelihood of renting honey bees is increasing and concave in farm scale, but that output price and natural open cover have larger positive effects on this choice. Using a shift-share instrument, we estimate managed pollination demand elasticities in the range of -0.78 to -1.24. We apply new methods that use demand shocks to inform and bound demand elasticities, and obtain results consistent with our econometric estimates of the elasticity of managed pollination demand. Finally, we estimate semi-parametric response functions relating yield and profits to honey bee colonies per acre. Our results suggest that the optimal stocking density is around 2 and 4 colonies per acre for Western and Eastern states, respectively. Shape restriction tests are consistent with a concave relationship. Western apple farmers receive a larger return from the marginal colony rented, and yield in Eastern states exhibits a concave relationship with natural forest cover.

Keywords: agriculture, pollination, specialty crops, biodiversity conservation, value-chains
***JEL* Classification:** Q12, Q11, Q18, Q20

*Corresponding author: Steven W. Wilcox (steven.wilcox@usu.edu). Wilcox is affiliated with the Department of Applied Economics at Utah State University. All other authors are affiliated with Cornell University. Just: drj3@cornell.edu. Lin Lawell: clinlawell@cornell.edu. Gómez: mig7@cornell.edu. Grab: hlc66@cornell.edu. We gratefully acknowledge financial support from three awards from the Cornell Atkinson Center for Sustainability: a Sustainable Biodiversity Award, an Atkinson Venture Fund (AVF) Grant, and an AVF-Engaged Cornell grant for undergraduate involvement. We are indebted to Craig Khalke and the Lake Ontario Fruit Program for assistance and very helpful details on apple production in the Northeast. We also appreciate comments from a number of apple growers and consultants in the Northeast, including Debbie Biltonen, Mike Biltonen, and Ian Merwin. Carrie Jones at USDA-ERS and Donnie Fike at USDA-NASS kindly assisted us with many aspects of USDA-ARMS data. Anna Espinoza, Adeline Sutton, Jacob Llodra, and Katie Warner provided excellent research assistance. We thank Todd Gerarden, Ivan Rudik, Bryan Danforth, Maria Van Dyke, Mia Park, Chris Barrett, Kathy Baylis, and Jo Albers for helpful comments; and we thank seminar and conference participants at Cornell University, the Food Studies International Conference, the Association of Environmental and Resource Economists (AERE) Summer Conference, and the Agricultural and Applied Economics Association (AAEA) Annual Meeting. Just, Lin Lawell, and Gómez are Faculty Fellows at the Cornell Atkinson Center for Sustainability. All errors are our own.

1 Introduction

Interest in pollination has grown substantially in recent decades. Many factors have contributed to this trend, including the reporting of colony collapse disorder (CCD) (vanEngelsdorp and Meixner, 2010), evidence of instability in wild pollination stocks (Cameron et al., 2011; Grab et al., 2019), claims that crop production may be pollination-limited (Reilly et al., 2020), and concerns over pesticides and food security, among others. In response to these developments, governments have adopted legislation to protect pollinators with pesticide bans (e.g., see Casert (2018) on EU policy) and adapted incentive programs to encourage conservation of pollinators (e.g., US Farm Bill Conservation Programs). The private sector is even developing drone-based pollination technology (Garfield, 2018).

Farmers of pollination-dependent crops grow much of the world's nutritious and high-value fruits, nuts, and vegetables;[1] and face very complex production decisions (Ridley and Devadoss, 2021). An important decision faced by producers of pollination-dependent crops is whether and how much to use managed pollination services. The most common market transaction between managed pollination service providers and crop producers in the US[2] seems to be the rental of domesticated honey bee colonies during the bloom period.[3] Farmer pollination choices are critical as they impact farm-level outcomes like yield and fruit quality (Roubik, 2002; Garibaldi et al., 2013; Park et al., 2016; Russo et al., 2017; Danforth, Minckley, and Neff, 2019), local pollination resources within and beyond the farm-gate (Kennedy et al., 2013; Park et al., 2015; Grab et al., 2018), and market-level outcomes through shifts in the supply and demand of both pollination resources and agricultural commodities (Rucker, Thurman, and Burgett, 2012; Goodrich, Williams, and Goodhue, 2019b).

[1] Crops that require or benefit greatly from insect pollination include almonds, coffee, apples, avocados, cherries, peaches, blueberries, among many others.

[2] Research on pollination outside of Europe and the United States seems to be fairly rare, with contributions outside of economics being more common. Examples of work outside of the US and Europe include Kasina et al. (2009), Imbach et al. (2017), and Narjes and Lippert (2021).

[3] Additional market-based options for managed pollination services include: the purchase of so-called "buckets of bees" for the bloom period (e.g. single containers of cultivated bumble bees, which do not live past a single season); as well as bloom season rental of mason bees or blue orchard bees (Ward, Whyte, and James, 2010).

In this paper, we make several advances in the study of farmer pollination choice. Our primary objectives are two-fold: (i) develop a more formal rendering of the farmer's pollination resource decision-making problem to establish intuition on factors that impact the use of managed pollination services; and (ii) undertake the first known empirical study of pollination choice using farm-level production data.

From a theoretical perspective, we use optimal control theory to characterize the general nature of the farmer's pollination resource decision-making problem with a specific focus on several factors that affect whether and how much a farmer uses managed pollination services to fulfill pollination resource needs. Our model makes several predictions on how the use of managed pollination services is influenced by output price, managed pollination input price, total factor productivity,[4] farm capital, and local wild pollination stocks. We then apply our theory to farm-level data on apple farmers in the US to empirically explore: (1) the factors that affect the binary choice to use managed pollination; (2) the responsiveness of managed pollination demand to managed pollination price; and (3) the relationship between yield, profit, and managed pollination use.

To examine the determinants of the binary choice to use managed pollination, we estimate a logit fixed effects regression using methods developed by Bergé (2018) for efficiently estimating models with large numbers of fixed effects. We find that the likelihood of renting bees at the block level is increasing and concave in the scale of production (e.g., size of block, number of bearing blocks), but that a unit increase in scale only increases the likelihood of renting bees by around 1-7%. Apple prices and natural open cover[5] have larger positive effects on this choice.

To study managed pollination demand responsiveness to managed pollination price, we undertake the first known empirical estimate of an own-price elasticity of demand for the number of honey bee colonies rented. Here we focus on variation at the random block level. For these efforts we rely on a shift-share instrumental variables strategy that interacts distance from the zip

[4] In our setting, total factor productivity can be captured by measures of the scale of production, such as farm size and trees per acre.

[5] We define natural open cover as the proportion of apple-specific and/or tree-crop specific areas within a county in any of the following cover types: clover, wildflowers, shrubland, herbaceous wetlands, developed open space, and wetlands.

codes where apple farms are located to the approximate center of almond production in Fresno County, California (our share) with the total almond acreage in California (our shift). This instrument accounts for demand for honey bees from almond growers in California (where most of US almond production takes place) and its well-documented effects on the availability and distribution of honey bees (e.g., see Ward, Whyte, and James (2010); Rucker, Thurman, and Burgett (2012); Goodrich, Williams, and Goodhue (2019b)). We find statistically significant point estimates for the demand elasticity in the range of -0.78 to -1.24, which imply that demand for managed pollination services is weakly inelastic to weakly elastic. We also apply new methods from Petterson, Seim, and Shapiro (2023) that use demand shocks to inform and bound demand elasticities, and obtain results consistent with our econometric estimates of the elasticity of managed pollination demand.

For our study of the relationship between production outcomes and use of managed pollination, we employ semi-parametric optimal binscatter developed by Cattaneo et al. (2021) to estimate response functions relating yield and profits to honey bee colonies per acre. We interpret our results as semi-parametric marginal product and marginal value product curves, respectively. We find strong evidence for at least local concavity in these relationships and we use estimated first derivatives to locate local optima for honey bee colony density, which we find is around 2 and 4 colonies rented per acre for Eastern and Western states, respectively. We further find that apple farmers in Western states get a larger return to the marginal honey bee colony rented than apple farmers in Eastern states, and that yield in Eastern states exhibits a concave relationship with natural forest cover. Results of formal hypothesis tests regarding the parametric form of the response function and shape restrictions are consistent with the visual observation of concavity and diminishing returns to managed pollination use.

The remainder of our paper proceeds as follows. Section 2 summarizes related prior literature. Section 3 presents our theoretical model. Section 4 describes our data. Section 5 presents our empirical analysis of the binary choice to use managed pollination. Section 6 presents our empirical analysis of the elasticity of demand for managed pollination. Section 7 presents our empirical analysis of the relationship between yield, profit, and managed pollination use. We discuss and conclude in Section 8.

2 Previous Literature

Since the seminal work by Meade (1952), economic analyses of pollination resources and pollination-dependent sectors include advances in understanding the value provided by pollination resources to society (Penn, Hu, and Penn, 2019; Lippert, Feuerbacher, and Narjes, 2021), the state and nature of pollination service markets (Willett and French, 1991; Rucker, Thurman, and Burgett, 2012; Goodrich, Williams, and Goodhue, 2019b; Fei et al., 2021), the impacts of colony collapse disorder (CCD) on beekeepers and pollination markets (Champetier, Sumner, and Wilen, 2015; Rucker, Thurman, and Burgett, 2019), and the use of beekeeping for poverty alleviation (Albers and Robinson, 2011). In previous work on decision-making by pollination-dependent farmers, Simpson (2018) develops and calibrates a theoretical model to understand the extent to which farmers might set aside land for wild pollinators, and Wu and Atallah (2019) develop a dynamic model to simulate trade-offs between pollination and pest control. The majority of directly related economics literature has focused heavily on beekeepers, almond growers, and the West Coast of the US; and there is a paucity of theoretical and empirical work focused on farmers (Baylis, Lichtenberg, and Lichtenberg, 2021).

Our work complements these prior efforts in several ways. First, our theory model adds more generality by looking at the full suite of input groups within an inherently dynamic production system (e.g., tree crops for fruits and nuts). Second, our empirical application to apple farmers in the US builds on the majority of directly related economics literature, which has focused heavily on beekeepers, almond growers, and the West Coast of the US. Indeed, since the seminal work by Meade (1952), little direct focus appears to have been placed on the setting of apple production and pollination. Third, to the best of our knowledge, there are no directly comparable empirical contributions to ours in the economics literature or related entomology or ecology literatures.[6]

3 Theory Model

Farmers of pollination-dependent crops face very complex production decisions, which in many cases involve long-term investments – especially for long-lived tree fruits and nuts – and production is often labor intensive (Ridley and Devadoss, 2021). In addition to decisions about general production strategies and other input choices, producers of pollination-dependent crops make decisions about whether and how much to use managed pollination services (e.g., renting honey bees); and whether and how much to invest in wild pollination (e.g., setting aside land for planting wildflower strips, or other natural cover (Cohen, 2022)). We discuss pollination choice in more detail in Appendix A.1.

To condense the pollination-dependent farmer's problem into a sufficiently parsimonious model, our discussion in Appendix A.1 and available information highlight a few issues. First, labor and traditional farm capital in the form of farm implements, buildings, and the biological capital of, say, an orchard, are the center of the pollination-dependent farmer's problem. Therefore, we treat labor L as a flow and capital K as a stock, maintained via investment variable I_k, to capture the majority of activity, which, in cases without insect pollination, also comprise pollination inputs in their entirety. Second, outside of some special cases,[7] the most common pollination scenario may be one in which the farmer chooses some amount M of managed pollinators through seasonal arrangement with a pollination service provider (e.g., renting honey bee colonies) but may also benefit from a local wild stock of pollinators W. The presence of wild pollination stock W is likely

[6]For example, although many studies from ecologists have studied various measures of pollinator presence and measures of production in great detail (e.g. Roubik (2002), Park et al. (2016), Blitzer et al. (2016), and Reilly et al. (2020)), no work to our knowledge has measured these variables outside of small-scale experiments, nor have they combined such observations with the realized production behavior of the farmers from whose land they are collecting data.

[7]As an example of such a special case, consider a multi-output pollination-dependent farmer who also produces honey and manages their own stock of honey bee hives for their pollination needs. This is a rare exception to the norm in our empirical setting given that in 2007, 3% of US apple farmers owned their own honey bees and among this unique group some farmers still reported renting honey bees (authors' calculations using the 2007 USDA-ARMS). Another special case might be one in which a farmer can locate next to an apiary.

a function of some local natural resource endowment of habitat, which can be invested in via investment variable I_w.[8] In practice, the effect of I_w on W will likely be limited given stochastic pollinator population dynamics and spillover effects from neighboring land use, though for simplicity we assume these complications are not present. As such, we assume the farmer may choose some managed pollination amount M as a flow, we assume the farmer may benefit from the wild pollination stock W, and we assume the farmer may choose to invest in the wild pollination stock W via wild pollination investment variable I_w. Third, although many pollination-dependent farmers are multi-output firms, many seem to specialize in one particular fruit or nut with smaller amounts of annual or perennial crops. Thus, to further simplify matters, we focus on a single output case, denoted Q.

Profit each period $\pi(K,L,M,W)$ is given by crop revenue minus costs:

$$\pi(K,L,M,W) = p_c Q - \left[p_k I_k + p_l L + p_m M + p_w I_w\right], \tag{1}$$

where we assume costs are linear, and where we assume output Q is given by the following two-level CES production function (Sato, 1967) for generality and to isolate parameters governing input relationships, variable input effectiveness, and total factor productivity:

$$Q = A\left[\gamma_o(\alpha_m M^{-\rho_o} + \alpha_w W^{-\rho_o})^{\frac{\rho}{\rho_o}} + \gamma_{kl}(\alpha_k K^{-\rho_{kl}} + \alpha_l L^{-\rho_{kl}})^{\frac{\rho}{\rho_{kl}}}\right]^{-\frac{1}{\rho}}. \tag{2}$$

The parameter $\rho \in [-1,\infty]$ governs the degree of substitutability between the pollination input group (managed pollination M and wild pollination W) and the traditional input group (capital K and labor L); the higher the ρ, the more the input groups are complements rather than substitutes. Similarly, the parameter ρ_o governs the degree of substitutability between the inputs within the pollination input group (managed pollination M and wild pollination W), while the parameter ρ_{kl} governs the degree of substitutability between the inputs within the traditional input group (capital

[8] In practice, I_w might comprise a bundle of management and land-use practices, including, for example, planting mixes of flowers and diverse vegetation types around farm fields (Grab et al., 2018; Cohen, 2022) – possibly as part of a broader integrated pest management (IPM) plan.

K and labor L). The γ parameters reflect the efficiency of input groups, and the α parameters reflect the efficiency of individual inputs within groups. A is total factor productivity, which in our setting captures measures of the scale of production such as farm size and trees per acre. Prices for Q, I_k, L, M, and I_w respectively are p_c, p_k, p_l, p_m, and p_w. The price p_k for capital investment I_k might be the per land unit area cost. The price p_w for wild pollination investment I_w might be similarly constructed to reflect different forms of wild pollination investment I_w, some of which may be visible in the market (e.g., land values, pollinator habitat enhancement).

The dynamic optimization problem faced by the farmer is to choose an optimal trajectory for labor L, managed pollination M, capital investment I_k, and wild pollination investment I_w to maximize the present discounted value of the stream of profits, $\pi(K,L,M,W)$, over some long horizon, here specified as infinite, subject to equations of motion, non-negativity constraints, and state variable initial conditions. The optimal control problem is given by:

$$
\begin{aligned}
&\max_{\{I_k(t),\ L(t),\ M(t),\ I_w(t)\}} \int_0^\infty \pi(K(t),L(t),M(t),W(t))e^{-rt}dt \\
&\qquad\text{s.t. } \dot{K}(t) = \delta_k K(t) + I_k(t) \ \ : \ \lambda_k(t) \\
&\qquad\dot{O}_w(t) = F(W(t)) + I_w(t) - \delta_{mw}M(t) - \delta_{kw}I_k(t) \ \ : \ \lambda_w(t) \\
&\qquad K(t),\ L(t),\ M(t),\ W(t) \geq 0 \\
&\qquad K(0) = K_o,\ W(0) = W_o.
\end{aligned}
\tag{3}
$$

We hold interest rate r fixed. The equation of motion for capital K is straightforward with a depreciation rate δ_k, and is associated with the capital co-state variable (or current-value shadow price of capital) $\lambda_k(t)$. The equation of motion for wild pollination stock W includes a non-linear biological production function $F(W)$, per-period investment I_w, and the negative impacts of managed pollinators M (e.g., resource competition) and of the expansion of farm capital stock K, and is associated with the wild pollination co-state variable (or current-value shadow price of wild pollination) $\lambda_w(t)$. The δ parameters capture effects within respective equations of motion. The initial values for capital and wild pollination stock at time $t = 0$ are K_o and W_o, respectively.

Full analytical solutions are very difficult owing to the presence of more than one state variable

(Weitzman, 2009). Nevertheless, since our focus here is managed pollination M, which we specify as a flow, we can reasonably invoke a few simplifying assumptions to isolate some of the underlying factors governing the use of managed pollination M. In particular, after using the Maximum Principle, and applying a *ceteris paribus* assumption to inputs, state variables, and other variable parameters at some time t, we can work with the first-order condition for managed pollination M. With these assumptions, we derive expressions and signing regimes for the following comparative statics: $\frac{dM}{dp_c}$, $\frac{dM}{dp_m}$, $\frac{dM}{dA}$, $\frac{d^2M}{dA^2}$, $\frac{dM}{dK}$, and $\frac{dM}{dW}$. We summarize our findings on the factors that affect managed pollination use M in the set of propositions below, where $\eta_{M,j}$ denotes the elasticity of managed pollination M with respect to parameter j. Details and proofs are provided in Appendix A.2.

Proposition 1 (Managed pollination price p_m): Managed pollination use M is decreasing in managed pollination price p_m. The own-price elasticity of demand for managed pollination, $\eta_{M,pm}$, declines in magnitude with managed pollination use M. If production is linear with respect to M (i.e., if there are no diminishing returns to M), then managed pollination use M will be perfectly elastic with respect to managed pollination price p_m. The greater the diminishing returns to M in production, the less elastic M will be with respect to managed pollination price p_m.

Proposition 2 (Output price p_c): Managed pollination use M is increasing in output price p_c.

Proposition 3 (Total factor productivity A): Managed pollination use M is increasing in total factor productivity A, which in our setting captures measures of the scale of production such as farm size and trees per acre. The more production is curved with respect to M (i.e., the greater the diminishing returns to M), the less responsive M is to increases in total factor productivity. Total factor productivity A has a small and concave effect on managed pollination use M when: (a) the magnitude of total factor productivity A exceeds the magnitude of managed pollination use M; and (b) positive amounts of all inputs, especially capital K and labor L, are used; and (c) either input groups are strong complements ($\rho \geq 1$) and pollination inputs are on the spectrum of substitutes ($-1 < \rho_o < 0$), or input groups are strong complements and pollination inputs are weak complements ($\rho > \rho_o > 0$).

Proposition 4 (Capital stock K): As the farm capital stock K increases, so does managed pollination use M.

Proposition 5 (Wild pollination stock W): Managed pollination use M is decreasing in local wild pollination stock W when either (a) pollination inputs are substitutes and input groups are complements ($-1 < \rho_o < 0$ and $\rho > 0$), or (b) input groups are complements and pollination inputs are weak complements ($\rho > \rho_o$).

4 Data

We apply our theory to data on apple production in the US to empirically explore (i) the factors that affect the binary choice to use managed pollination; (ii) the own-price elasticity of demand for managed pollination; and (iii) the relationship between yield, profit, and managed pollination use.

Apples are a useful crop to study farmer pollination behavior. Apples are a widely produced and consumed commodity around the world. Apples are not considered to be a "honey crop", as nectar from apples does not produce palatable honey, and this translates into higher pollination rental fees for apple farmers to mitigate against the fact that beekeepers do not gain forage resources to produce palatable honey from pollinating apples (Rucker, Thurman, and Burgett, 2012). For additional background information on apple production, see Appendix B.1.

For our empirical analysis, we leverage rich, farm-level microdata from the 2007 USDA Agricultural Resource Management Survey (USDA-ARMS). The USDA National Agricultural Statistics Service (USDA-NASS) imposes stringent conditions and restrictions on the use of its USDA-ARMS data, including strict security measures, data confidentiality, and the required use of provided replication weights. We access the USDA-ARMS data via the NORC Data Enclave. The data contain detailed information on production for the 2007 production year (roughly March-November), as well as data on managed pollination use over the years 2006-2007, for apple farmers in seven US states: California (CA), Michigan (MI), New York (NY), North Carolina (NC), Oregon (OR), Pennsylvania (PA), and Washington (WA).[9] There are 1057 farmers who have sufficient responses for our research; Figure B.1 in Appendix B shows their distribution by state.

We merge the 2007 USDA-ARMS data with publicly available data on weather from PRISM (Daly et al., 2008); data on road and Euclidean distances that we construct and calculate; apple prices and California almond production from the USDA-NASS; and remotely sensed measures of land cover from the USDA Cropland Data Layer (CDL) (Boryan et al., 2011). We describe and discuss our data, data sources, and data construction in more detail in Appendix B.2.

At an aggregate level, our data reveal significant structural differences in the apple production sector that are well known to industry veterans. Among the more prominent stylized facts in the data are that there are prominent differences between production strategies and outcomes between apple farmers in West Coast states, and apple farmers in Midwest and East Coast states (which we refer to collectively as the 'Eastern' states) – a fact that reportedly has much to do with the higher volume of production that comes from Washington State, and the higher prevalence of plant diseases that farmers in Eastern states face, which are associated with higher moisture (Kahlke, 2019; Biltonen, 2020). In Tables B.1-B.5 in Appendix B, we present summary statistics of our data to highlight average values for a number of dimensions across all states, West Coast states, and Eastern states; and test for differences in means between Western and Eastern states. The tables show that, on average, Western operations are larger, more recently established, have managers who are slightly better educated, are more likely to rent honey bees, are more intensively farmed (more trees per acre), achieve higher yields and revenues, and are also more profitable ($6,200 per acre versus $3,343 per acre for Eastern states); while Eastern operations face higher honey bee rental costs and use honey bee colonies more intensively (more colonies per acre).

There are instances where farmers rented honey bees one year but not the other. Non-negligible proportions of farmers reported never renting honey bees over 2006-2007 (33% of Eastern farmers, 18% of Western farmers). Never renting honey bees (or never using managed pollination) is a

[9] The USDA-ARMS is designed to be nationally representative as well as representative at the level of a state.

notable strategy as it suggests that farmers are relying on local wild pollination stocks.[10]

To delve deeper into these aggregate differences, we construct a range of nested boxplots that showcase variation in important dimensions of managed pollination use and production by state, and also within states by farmers who did or did not rent honey bees. As seen in Figure 1, most growers rent between 1 to 4 colonies per acre, though the median seems to be closer to 1-2 colonies per acre. Figures B.2 and B.3 in Appendix B provide boxplots by state for farm and orchard characteristics, and for managed pollination measures.

In Figure 2, and in Figure B.4 in Appendix B, we see very interesting variation in production outcomes, costs, revenues, and profits, not only by state, but whether or not a farmer rents honey bees. Growers who rent honey bees have higher median production, and a higher interquartile-range of production. This is particularly the case for total output in bushels per acre and fresh market production, but less so for processed yield.[11] The same observation can be made for revenues and profits per acre. Although this simple observation is provocative, it is only an association as it may well be the case the farmers who rent bees are farming more intensely than farmers who do not rent bees, or some other unobserved factor may account for this structural difference. We examine the extent to which there is genuinely a marginal gain from renting honey bees or not in our study of yields, profits and honey bee use in Section 7.

[10] Although other strategies are possible, such as locating next to an apiary, the more likely scenario is that wild pollinators are the primary source of pollination for these apple farmers.

[11] The distinction between fresh market yield and processed yield is fruit quality and point of sale. Fresh market production is sold for sale in outlets like grocery stores (e.g., a box or bag of whole apples), and generally receives a higher per unit price because the fruit is more evenly shaped and appealing to consumers. In contrast, processed yields are sold to firms that process apples for products like apple juice, apple sauce, etc. and farmers receive lower prices for this output because fruit is of lower quality.

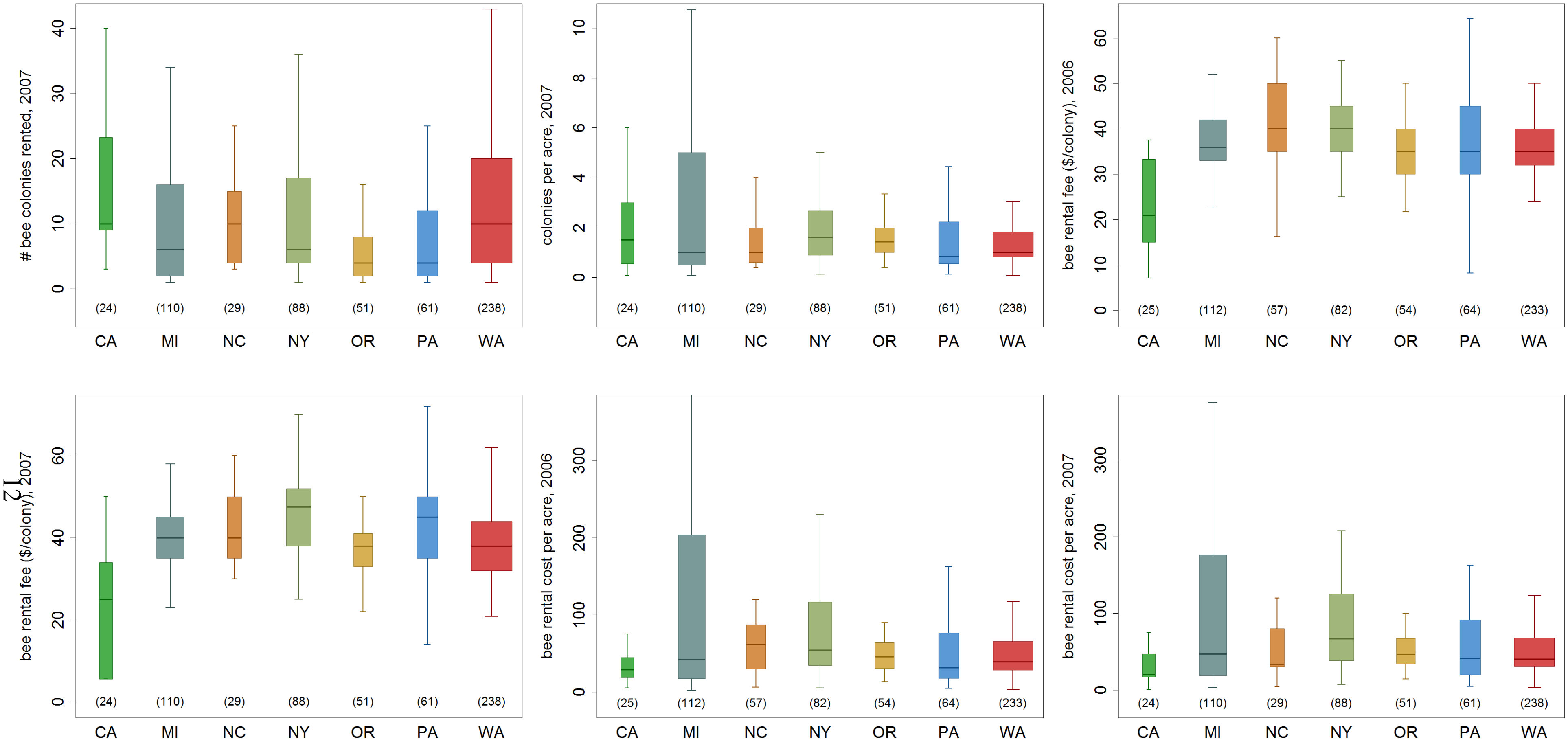

Figure 1: Weighted boxplots by state capturing: the number of honey bee colonies rented in 2007; the number of honey bee colonies rented per acre in 2007; honey bee rental fee ($/colony) in 2006; honey bee rental fee ($/colony) in 2007; honey bee rental costs per acre in 2006; and honey bee rental costs per acre in 2007. All variables are comprised of random block-level variation. Numbers in parentheses next to state abbreviations indicate the sample size per boxplot.

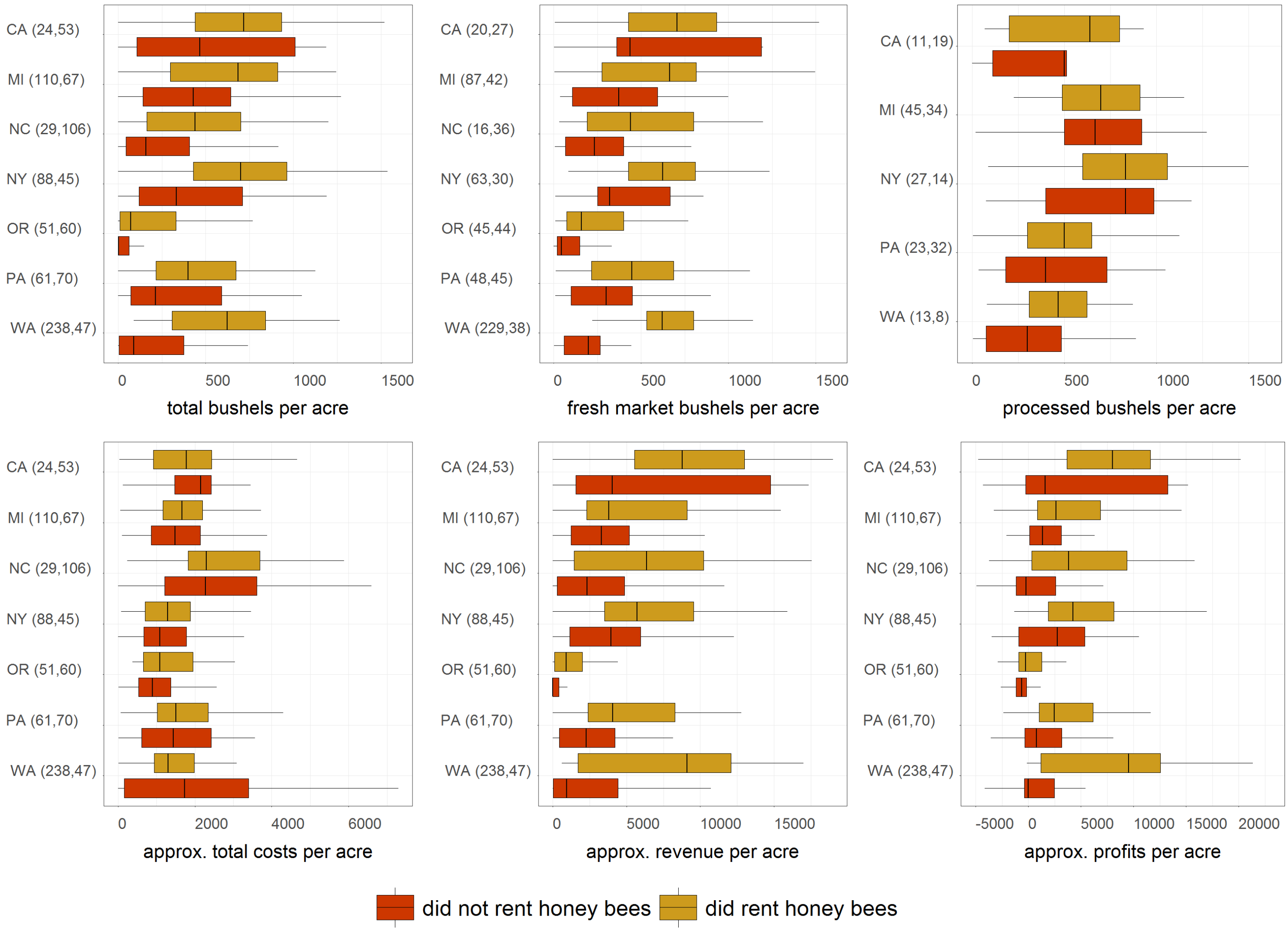

Figure 2: Weighted boxplots by state and if an apple farmer rented honey bees for: total yield in bushels per acre; fresh market yield in bushels per acre; processed yield in bushels per acre; approximate total costs per acre; approximate total revenue per acre; and approximate profits per acre. All variables comprised random block-level variation. Note that since the 2007 USDA-ARMS did not request information on output prices, we use state-level average apple prices to approximate revenue and profits. Numbers in parentheses indicate the sample size per state, and the choice to rent honey bees or not. For example, in the top left plot the notation for the bottom rows, WA (238, 47), indicates of the 286 apple farmers sampled in Washington State, 238 reported renting honey bees, while 47 reported not renting honey bees. Note farmers in some states did not report any processed yields, hence the top right panel only shows processed yield data for a subset of all seven states.

5 Empirical Analysis of the Choice to Rent Bees

For the binary choice to use pollination service markets (in our case, renting honey bees), we seek to study associations between the discrete choice to use managed pollination and important farm characteristics, state variables, and parameters, each of which addresses particular elements of Propositions 2-5. To explore these associations, we estimate the following logit fixed effects regression using methods developed by Bergé (2018) for efficiently estimating maximum likelihood models with large numbers of fixed effects:

$$\Pr(y_{isct} = 1) = \mathbf{x}'_{\mathbf{isct}}\beta + \lambda_s + \sigma_t + \varepsilon_{it}, \tag{4}$$

where y_{isct} is a dummy variable for farmer i in state s and county c renting honey bees in year t; $\mathbf{x}'_{\mathbf{isct}}$ is a vector of covariates, including measures of farm production scale (to proxy for total factor productivity), output price, and remotely sensed measures of natural open cover and natural forest cover (to proxy for wild pollinator stocks and landscape heterogeneity); λ_s and σ_t are state and time fixed effects; and ε_{it} is the error term. We describe and discuss our econometric model and the methods developed by Bergé (2018) in more detail in Appendix C.1.

Results for the average partial effects from the logit fixed effects regression in equation (4) are shown in Table 1. Since we employ second-degree polynomials in a number of regressors, Table C.1 in Appendix C provides respective point estimates (including for both linear and quadratic terms). Specifications (1), (3), and (5) use data from all states, the Eastern states, and the Western states, respectively; there are some missing observations for trees per acre and average age of trees, which reduces the respective sample sizes somewhat.[12] Specifications (2), (4), and (6) include the number of farm vehicles and implements (the best available measure for the farm physical capital stock), which unfortunately has many missing observations.

Some of the main takeaways are as follows. First, consistent with Proposition 2, the likelihood

[12]Our panel data set of 1057 farmers has 2114 observations across 2006-2007; 2056 observations are retained after missing observations for trees per acre and average age of trees are dropped. Trees per acre and average age of trees are measures of scale distinct from apple bearing acres and total bearing apple blocks, and therefore may have unique relationships to the binary choice to rent bees that are worth testing. Results are very similar if these variables are dropped and the maximum respective sample sizes are employed.

of renting bees has a significant positive association with our measure of apple output price, the total utilized production price (a weighted average of the fresh market and processed prices).

Second, consistent with Proposition 3, the likelihood of renting bees has a significant positive association with measures of farm scale, particularly block-level apple bearing acres and the total number of operation-level bearing apple blocks. Notably, as shown in Table C.1, there is a statistically significant concave relationship between these measures of scale and the choice to rent honey bees. Nevertheless, the effect of farm scale is small in magnitude: an increase in bearing acreage at the block level increases the likelihood of renting honey bees by around 1% (or less) and an increase in the number of bearing apple blocks increases the likelihood of renting by 1-7%. Output price exerts much greater influence than scale, suggesting marginal product considerations influence this choice more than scale. Proposition 3 indicates that the effect of total factor productivity (proxied here by measures of scale) will be small when there are diminishing returns to the use of managed pollination M – a notion that we study directly in Section 7.

Third, somewhat consistent with Proposition 4, we see that the available measure of physical capital stocks (number of farm vehicles and farm implements) has a positive (but insignificant) association with renting bees for both the full sample of all states and the Eastern states subsample.

Fourth, the choice to rent bees has a significant positive and concave relationship with natural open cover. The effect of natural open cover is comparable to that of output price. Although areas with more natural open cover (and hence less vegetative structure) will not necessarily have lower wild pollinator stocks suitable for apple pollination, and therefore will not necessarily require managed pollination, if wild pollinator stocks are indeed less prevalent in areas with more natural open cover, this result would be consistent with Proposition 5. For natural forest cover, results are insignificant and of mixed sign. There is some empirical work that indicates that natural forest cover near apple orchards can be source areas for wild pollination stocks that can enhance production (Park et al., 2015; Kammerer et al., 2016).

Finally, we see that farmers were slightly less likely to rent bees in 2007 and that farmers who deliberately scout for pests and have recently attended a pest management training are more likely to rent bees.

In Tables C.2 and C.3 in Appendix C we show average partial effects and point estimates for analogous logit fixed effects regressions of the binary choice to *never* rent honey bees during 2006-2007. Consistent with results shown in Table 1, results in Tables C.2 and C.3 have the opposite sign and are of similar or greater statistical significance. Notably, our measure of physical capital, number of farm vehicles and implements, is negative and statistically significant. This is consistent with Proposition 4 as it indicates that the likelihood of never renting honey bees over 2006-2007 decreases with a farm's physical capital stock.

Table 1: Average partial effects from weighted logit regression of the binary choice to rent honey bees.

	Dependent variable is probability of renting honey bees					
	(1)	(2)	(3)	(4)	(5)	(6)
apple bearing acres	0.006***	0.007***	0.002	0.002	0.009*	0.010**
	(0.0023)	(0.0024)	(0.0024)	(0.0024)	(0.0049)	(0.0049)
total bearing apple blocks	0.07***	0.02***	0.01***	0.01***	0.03**	0.04**
	(0.004)	(0.005)	(0.003)	(0.002)	(0.012)	(0.017)
trees per acre	0.00017	0.00030	0.00049	0.00106***	0.00017	0.00008
	(0.000264)	(0.000317)	(0.000555)	(0.000212)	(0.000251)	(0.000180)
average age of trees	0.0001	0.0010	0.0033	0.0052	0.0002	0.0026
	(0.00345)	(0.00425)	(0.00610)	(0.00846)	(0.00126)	(0.00424)
total utilized production price ($/lb)	0.98***	0.98**	2.60***	3.17***	0.71***	0.45***
	(0.314)	(0.443)	(0.428)	(0.596)	(0.194)	(0.171)
natural forest cover (county proportion)	0.080	0.108	0.006	−0.301	0.127	0.451
	(0.1804)	(0.2021)	(0.1953)	(0.2867)	(0.2714)	(0.3082)
natural open cover (county proportion)	0.95	1.17*	1.72*	0.98	0.60	1.19*
	(0.588)	(0.705)	(0.913)	(1.248)	(0.604)	(0.645)
number farm vehicles and implements		0.002		0.005		−0.001
		(0.0029)		(0.0041)		(0.0021)
recent pest training (dummy)	0.09	0.09**	0.20***	0.18***	−0.04	−0.02
	(0.061)	(0.045)	(0.019)	(0.036)	(0.052)	(0.061)
deliberate pest scouting (dummy)	0.119***	0.102**	0.002	−0.053	0.158***	0.132***
	(0.0395)	(0.0434)	(0.0426)	(0.1142)	(0.0604)	(0.0383)
year 2007 (dummy)	−0.04**	−0.04	−0.06***	−0.07***	−0.03***	−0.01***
	(0.015)	(0.026)	(0.010)	(0.022)	(0.003)	(0.004)
State FE	Y	Y	Y	Y	Y	Y
Sample	All	All	East	East	West	West
Standard Errors	C,S	C,S	C,S	C,S	C,S	C,S
Pseudo R^2	0.33	0.37	0.20	0.30	0.50	0.54
# Observations	2056	1514	1136	858	920	656

Notes: Table presents average partial effects from weighted logit regression of the binary choice to rent honey bees on block- and operation-level characteristics, output price, physical capital, and land cover measures. For land cover measures, we use remotely sensed measures of natural open cover and natural forest cover proportions at the county level. Standard errors are clustered at both the state (S) and county (C) levels, and are in parentheses. Significance codes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

6 Elasticity of Demand for Managed Pollination

6.1 Estimating Honey Bee Demand

To estimate the own-price demand elasticity for managed pollination use, we estimate the demand for honey bees from apple farmers using an instrumental variables strategy to address the endogeneity of price in models of demand (Manski, 1995; Goldberger, 1991; Angrist, Graddy, and Imbens, 2000; Lin, 2011). Our first-stage equation is given by:

$$p_{m,isct} = \delta_1 Z_{sct} + \mathbf{X'}_{\mathbf{isct}}\alpha + \mathscr{D}_l + \gamma_t + \nu_{isct}, \tag{5}$$

where $p_{m,isct}$ is the price of managed pollination services (here the honey bee rental fee per colony) faced by farm i, in state s, county c, and year t; Z_{sct} is our shift-share instrument for price (Goldsmith-Pinkham, Sorkin, and Swift, 2020; Borusyak, Hull, and Jaravel, 2022), which interacts almond acreage in California with distance measures to California; $\mathbf{X}'_{\mathbf{isct}}$ is comprised of farm and orchard characteristics; $\mathscr{D}_l$ and γ_t are dummies for location and time, respectively;[13] and ν_{isct} is the first-stage error term.

Our second-stage managed pollination demand equation is given by:

$$M_{isct} = \beta_1 \widehat{p}_{m,isct} + \mathbf{X'}_{\mathbf{isct}}\theta + \mathscr{D}_l + \gamma_t + \varepsilon_{isct}, \tag{6}$$

where M_{isct} is the number of bee colonies demanded by farm i in state s, county c, and time t; $\widehat{p}_{m,isct}$ is the predicted price from the first stage; and ε_{isct} is the second-stage error term. We describe and discuss shift-share instruments and our econometric model of honey bee demand in more detail in Appendix D.1.

The own-price elasticity of demand for managed pollination, η_{M,p_m}, is given by:

[13] For our two-period panel, including a dummy for 2007 is equivalent to including year fixed effects. As discussed in more detail in Appendix D.1, we employ a subset of state-specific dummies rather than the full set of state fixed effects because state fixed effects effectively eliminate the strength of our instrument; the remaining variation in distance to Fresno, CA (the share in our shift-share instrument), after inclusion of state fixed effects, is only weakly correlated with price. Analogous challenges have been encountered by Bruno and Jessoe (2021), who find that year fixed effects remove excessive variation when estimating groundwater demand elasticities in California. In Section 6.2, we use new methods developed by Petterson, Seim, and Shapiro (2023) to assess the bounds and robustness of our results.

$$\eta_{M,p_m} = \frac{dM}{dp_m}\frac{p_m}{M} = \beta_1 \frac{p_m}{M}. \tag{7}$$

To identify β_1, the coefficient on managed pollination price in the second-stage managed pollination demand equation (6), we use a shift-share instrument $Z_{sct} = d_{sc} * s_t$ for price that interacts distance from the zip codes where apple farms are located to the approximate center of almond production in Fresno County, California (our time-invariant "share" d_{sc}) with the total almond acreage in California (our time-varying "shift" s_t).

The main insight we leverage for our identification strategy is the reality that California almonds have exerted tremendous influence on pollination service markets in the US for decades, but especially since 2004. This is because California is essentially the only place in the US where almond production takes place with any real scale,[14] and almonds require imported pollination services to produce marketable almonds: during the almond bloom period, almonds use an estimated 80% of US colonies, some coming from places as distant as Florida (Rucker, Thurman, and Burgett, 2012; Goodrich, 2017; Goodrich, Williams, and Goodhue, 2019a). After the almond pollination season is done in early spring (e.g., March), many beekeepers migrate to different parts of the country chasing bloom periods for various crops (Rucker, Thurman, and Burgett, 2012; Goodrich, 2017; Goodrich, Williams, and Goodhue, 2019a).

For the exogenous, time-varying "shift" s_t in our shift-share instrument Z_{sct}, we use total almond acres in California in period t. Figure 3 shows the time-series of total almond acreage in California since 1995, broken out with total bearing and non-bearing acreage. An increase in acreage and supply of almonds in California has a well-documented effect of increasing demand for pollination services nationally (e.g., see Ward, Whyte, and James (2010); Rucker, Thurman, and Burgett (2012); Goodrich, Williams, and Goodhue (2019b)), which in turn affects the equilibrium price to rent bees. On the other hand, since the apple farmers in our sample were not almond

[14]According to USDA-NASS Quick Stats, California is the only state where production is reported for almonds. This can be easily verified by a query of the USDA-NASS Quick Stats portal by searching for 'AREA BEARING & NON-BEARING' for almonds: `https://quickstats.nass.usda.gov`.

growers,[15] for our sample of apple farmers, the option to grow almonds, and the acreage and supply of almonds in California, should have no bearing on the number of honey bee colonies they rent for a random block of apples. As a consequence, almond acreage in California is correlated with the price of managed pollination, but does not affect the demand for managed pollination from apple growers except through its price.

Since pollination services require transportation (generally by semi-trucks), distances and other transportation costs become important determinants of prices and exposure to "shifts" from places like California. Thus, for the exogeneous, time-invariant "share" d_{sc} in our shift-share instrument Z_{sct}, we study Euclidean and road distances from zip code, county, and state centroids to the centroid for Fresno County, California (a credible center for California almond production). Our preferred instrument interacts total almond acres in California in period t with the Euclidean distance from zip code centroids to the centroid for Fresno County, California.

Table 2 presents the honey bee demand IV estimation results (weighted) from estimating the second-stage regression in equation (6). Table D.1 in Appendix D presents the respective first-stage results. The own-price demand elasticity for managed pollination use, as calculated using equation (7), is evaluated at the mean price and quantity in the data for the respective sample of data.

Although the 2007 USDA-ARMS collected data on the binary choice to rent honey bees, and the costs to rent honey bees per colony over 2006-2007, data on the quantity of honey bees demanded in 2006 is not available except in the instance a farmer reported not renting bees (in which case we know quantity rented is zero). Specifications (1) and (2) use data from 2007 only. Specifications (3) and (4) employ an unbalanced panel over 2006-2007 that includes all observations from 2007, as well as growers who reported not renting bees in 2006, for whom we know the number of colonies rented in 2006 is zero (thereby eliminating the need for quantity imputation).

[15] Only one of the 1057 farmers in the 2007 USDA-ARMS who have sufficient responses for our research reported growing both almonds and apples; this sole grower is in California. There were four others growers in the 2007 USDA-ARMS who also reported growing both almonds and apples, all in California, but none of these other four growers filled out the Phase II survey, which suggests that their apples may be non-bearing (see Appendix B.2 for more information on the USDA-ARMS Phase II survey).

Specification (5) is a balanced panel that includes all growers in the data for both 2006 and 2007: if a grower rented in both years, we impute the number of colonies rented in 2007 to be the number of colonies rented in 2006; if a grower rented bees in 2006 but not in 2007, we impute the quantity rented in 2006 by multiplying the acreage of the selected apple block with the state level average honey bee stocking density in 2007. Our preferred specifications are specifications (2), (4), and (5), which include a larger set of controls. We describe and discuss our specifications for honey bee demand in more detail in Appendix D.1.

Our main finding from our IV results in Table 2 is that the demand for managed pollination is weakly inelastic to weakly elastic. Across our preferred specifications (specifications (2), (4), and (5)), the estimated elasticity of demand for managed pollination ranges from -0.78 to -1.24. In contrast, the analogous OLS estimates, reported in Table D.2 in Appendix D, are strongly inelastic and statistically insignificant. Kleibergen and Papp first-stage F-statistics, F_{kp}, are well-above conventional levels indicating sufficient instrument strength (F_{kp} is equal to the robust first-stage F-statistics suggested by Montiel Olea and Pflueger (2013) when the model is just-identified (Andrews, Stock, and Sun, 2019)). Regression-based Durban-Wu-Hausman (DWH) statistics (Wooldridge, 2010) strongly reject exogeneity of the endogenous variable. Model diagnostic statics also offer support for our identification strategy.

We also see that, consistent with Proposition 3, honey bee demand is increasing and concave in measures of the scale of production, as captured by apple bearing acres at the block level, though here with magnitudes that are more meaningful economically than those in our empirical results for the binary choice to rent honey bees.

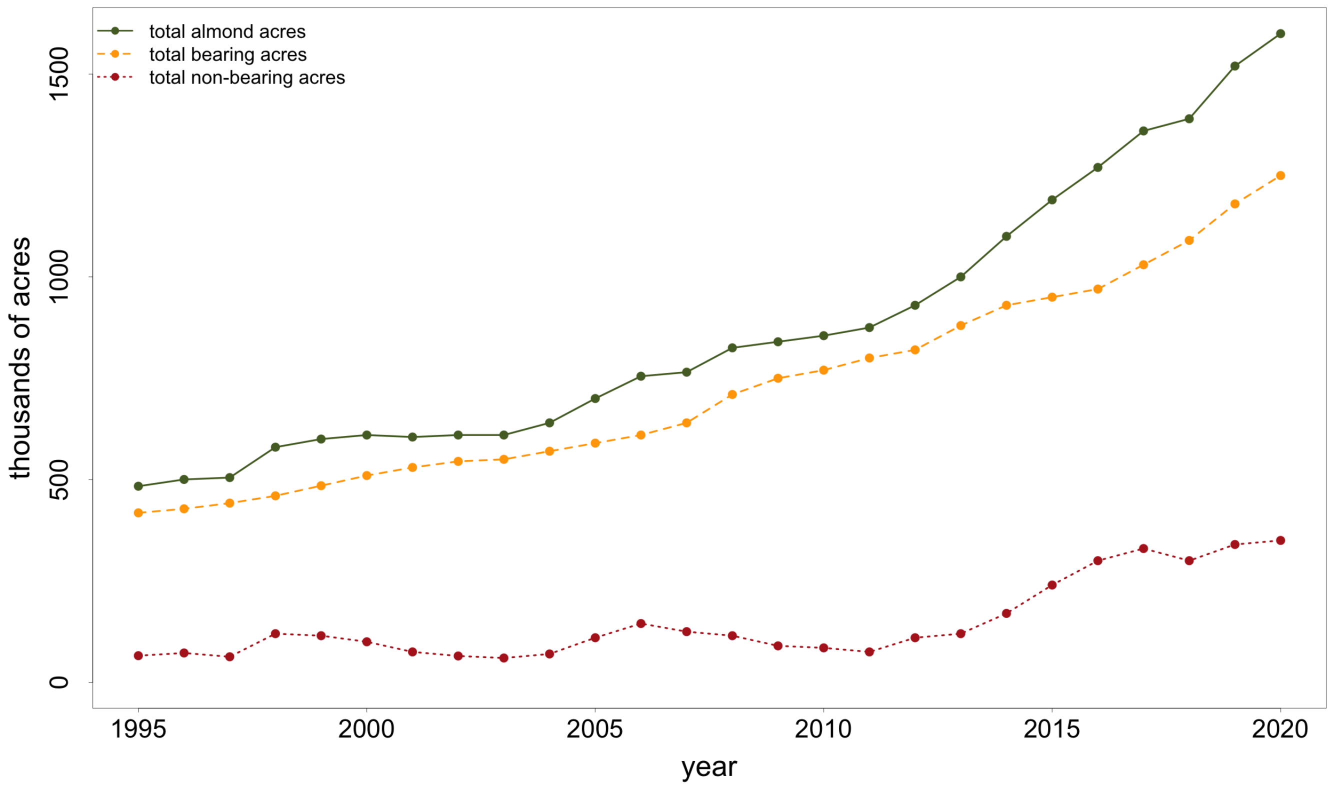

Figure 3: Almond acreage in California, 1995-2020
Data source: USDA-NASS 2020 California Almond Acreage Report.

Table 2: Honey bee demand own-price elasticity estimation, IV results (weighted).

Dependent variable is the number of honey bee colonies rented					
	(1)	(2)	(3)	(4)	(5)
honey bee rental fee ($/colony)	-0.44*	-0.25**	-0.51*	-0.32**	-0.27***
	(0.254)	(0.107)	(0.260)	(0.992)	(0.069)
apple bearing acres	1.13***	1.13***	1.00***	0.99***	1.14***
	(0.156)	(0.208)	(0.173)	(0.242)	(0.203)
apple bearing acres, squared	-0.0008**	-0.0008	-0.0005	-0.0005	-0.0009
	(0.00035)	(0.00045)	(0.00040)	(0.00053)	(0.00044)
deliberate pest scouting (dummy)		3.12***		2.07**	3.17***
		(0.996)		(0.751)	(0.690)
CA (dummy)	-12.38**	-8.20***	-12.37**	-8.90***	-8.60***
	(5.685)	(1.486)	(5.865)	(1.693)	(1.203)
MI (dummy)	1.40	1.04**	0.89	0.32	1.35***
	(2.857)	(0.388)	(2.416)	(0.304)	(0.325)
PA (dummy)		-4.82***		-3.96***	-5.01***
		(0.813)		(0.507)	(0.540)
year 2007 (dummy)			10.53***	9.04***	0.98**
			(1.367)	(1.166)	(0.287)
Constant	19.85*	9.23**	13.57	5.92	9.41**
	(10.212)	(3.893)	(9.656)	(3.940)	(3.232)
Elasticity at mean	-1.39*	-0.78**	-1.97*	-1.24***	-0.81***
Data included in sample:					
All observations from 2007	Y	Y	Y	Y	Y
Growers who did not rent in 2006	N	N	Y	Y	Y
Growers who rented bees in 2006	N	N	N	N	Y
Standard Errors	C	C,S	C	C,S	C,S
First-stage F-statistic, F_{kp}	22.42	118.03	21.41	87.04	128.39
DWH	1.80	2.08	2.24	3.16	3.19
Adjusted R^2	0.49	0.52	0.44	0.50	0.53
# Observations	1057	1057	1487	1487	2114

Notes: Table presents IV results for honey bee demand (weighted). Although the 2007 USDA-ARMS collected data on the binary choice to rent honey bees, and the costs to rent honey bees per colony over 2006-2007, data on the quantity of honey bees demanded in 2006 is not available except in the instance a farmer reported not renting bees (in which case we know quantity rented is zero). Specifications (1) and (2) use data from 2007 only. For growers who rented honey bees in 2007, we use the grower's rental fee for the price. If a grower reported not renting bees in 2007, they did not report a bee rental fee; to deal with this we use the state average rental fee in 2007 for the price. Specifications (3) and (4) employ an unbalanced panel over 2006-2007 that includes all observations from 2007, as well as growers who reported not renting bees in 2006, for whom we know the number of colonies rented in 2006 is zero (thereby eliminating the need for quantity imputation), and for whom we use the state average rental fee in 2006 as the price in 2006. Specification (5) is a balanced panel that includes all growers in the data for both 2006 and 2007: if the grower rented bees in 2007 and 2006, we impute the number of colonies rented in 2007 to be the number of colonies rented in 2006; if the grower rented bees in 2006 but not in 2007, we impute the quantity rented in 2006 by multiplying the acreage of the selected apple block with the state level average honey bee stocking density in 2007. We use the state average rental fee in 2006 as the price in 2006. For specifications (1) and (2), the instrument Z_{sct} is the Euclidean distance from the centroids of zip code units of farm locations to the centroid of Fresno, County California. For specifications (3), (4), and (5) the instrument Z_{sct} is the interaction between the distance from zip code centroids where farms are located to the centroid of Fresno, County California and the total almond acres in California in year t. Elasticity is evaluated at the mean price and quantity in the data for the respective sample of data. Standard errors are clustered at the county (C) and/or state (S) level, and are in parentheses. Significance codes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

6.2 Bounding the Elasticity of Demand

As a further step to study honey bee demand, we employ new methods developed by Petterson, Seim, and Shapiro (2023) for studying bounds on elasticities. Petterson, Seim, and Shapiro (2023) show that economic intuitions about the plausible size of demand shocks can be informative about and help bound the elasticity of demand. In particular, Petterson, Seim, and Shapiro (2023) develop methods for determining demand elasticities that are consistent with a given bound B on the plausible size of demand shocks.

We use two alternative demand shocks for our demand shock bound B. One demand shock we use are shocks to demand in the year 2007. Estimates of demand shocks from 2007 are obtained from point estimates on the dummy for 2007 in specifications (3), (4), and (5) of our econometric model of honey bee demand in Table 2. Shocks in 2007 are potentially informative as this was a year of diesel price shocks and widespread onset of colony collapse disorder (CCD) (vanEngelsdorp and Meixner, 2010). For the second demand shock, following Appendix C of Petterson, Seim, and Shapiro (2023), we use the absolute value of the (smoothed) differenced quantity[16] for observations for which the differenced price is equal to zero. We describe our application of the bounding methods developed by Petterson, Seim, and Shapiro (2023) in more detail in Appendix D.2.

We first present the results from using shocks to demand in the year 2007 as our demand shock, and from using the balanced panel of all 1057 farmers that we use in specification (5) in Table 2 as our subsample of data. Following the approach taken in Figure 3 of Petterson, Seim, and Shapiro (2023), the shaded region in Figure 4 depicts the set of all demand functions consistent with a demand shock bound B of twice the maximum estimated shock in 2007, where the maximum estimated shock in 2007 is the coefficient on the year 2007 dummy in specification (3) in Table 2. As explained in more detail in Appendix D.2, a demand function consistent with a demand shock bound B is a downward sloping line that passes through the origin as well as through

[16] As described in more detail in Appendix D.2, to address data limitation challenges and comply with USDA NASS conditions and restrictions on data confidentiality, we apply data smoothing methods to the first differenced quantity to average out outliers.

all of the dotted intervals. A demand shock bound B of twice the maximum estimated shock in 2007 implies a bound on the demand elasticity (when evaluated at mean price and quantity) of -1.3. Then, following the approach taken in Figure 4 of Petterson, Seim, and Shapiro (2023), in Figure 5 we plot the range of honey bee demand elasticities that are consistent with bounds on the plausible size of shocks to demand in the year 2007 ranging from the minimum estimated shock in 2007 to twice the maximum estimated shock in 2007. The implied elasticities consistent with bounds on the plausible size of shocks to demand in the year 2007 range from -0.18 (for the minimum estimated shock in 2007) to -1.3 (for twice the maximum estimated shock in 2007). These results are consistent with our econometric estimates of the demand elasticity from our preferred specifications (specifications (2), (4), and (5)) of the IV regressions in Table 2, all of which lie within this range.

We next present the results from using as our demand shock the absolute value of the differenced quantity for observations for which the differenced price is equal to zero. We once again use the balanced panel of all 1057 farmers that we use in specification (5) in Table 2 as our subsample of data. As seen in Figure D.1 in Appendix D, a demand shock bound B of twice the maximum absolute value of the differenced quantity for observations for which the differenced price is equal to zero implies a bound on the demand elasticity of -1.56. Figure 6 plots the range of honey bee demand elasticities that are consistent with bounds on the plausible size of shocks to demand ranging from the mean absolute value of the differenced quantity for observations for which the differenced price is equal to zero, to twice the maximum absolute value of the differenced quantity for observations for which the differenced price is equal to zero. The implied elasticities (when evaluated at mean price and quantity) consistent with bounds on the plausible size of demand shocks range from -0.33 (for the mean absolute value of the differenced quantity for observations for which the differenced price is equal to zero) to -1.56 (for twice the maximum absolute value of the differenced quantity for observations for which the differenced price is equal to zero). These results are again consistent with our econometric estimates of the demand elasticity from our preferred specifications (specifications (2), (4), and (5)) in Table 2, all of which lie within this range.

Finally, we present the results from using shocks to demand in the year 2007 as our demand shock, and from using the subset of 430 farmers from the sample used in specifications (3)-(4) in Table 2 who have data in both 2006 and 2007 as our subsample of data. The advantage of this sample is that it does not require any quantity imputation; the disadvantage is that this sample does not have much variation in the absolute value of the differenced quantity for observations for which the differenced price is equal to zero, and therefore is not amenable to using our second demand shock. As seen in Figure D.2 in Appendix D, a demand shock bound B of twice the maximum estimated shock in 2007 implies a bound on the demand elasticity of -1.42. As seen in Figure D.3 in Appendix D, the elasticities that are consistent with a range of bounds on the plausible size of shocks to demand in the year 2007 are once again consistent with our econometric estimates of the elasticity of demand.

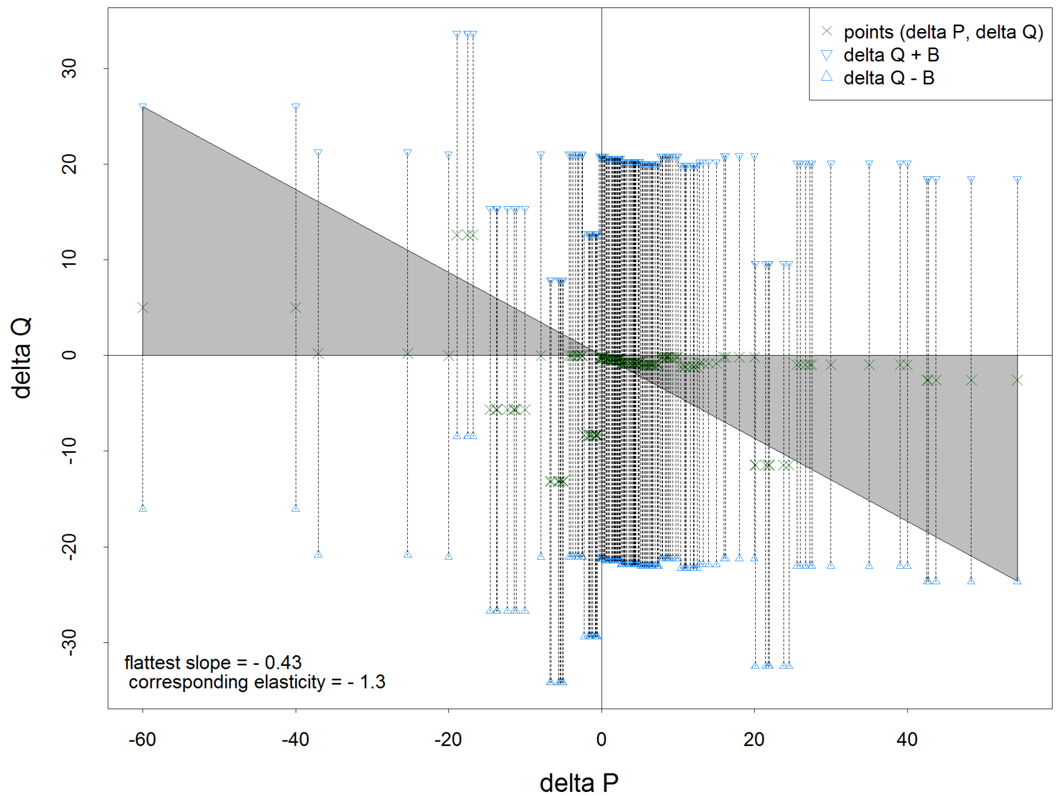

Figure 4: Figure illustrates the construction of bounds on the honey bee demand elasticity from using shocks to demand in the year 2007 as our demand shock and a demand shock bound B of 21.05, which is twice the maximum estimated shock in 2007. The subsample is the balanced panel of all 1057 farmers that we use in specification (5) in Table 2. The cross-hatches depict a scatterplot of the first differenced price on the x-axis and smoothed first differenced quantity on the y-axis. The dotted interval around each cross-hatch has radius of $B = 21.05$. The shaded region depicts all demand functions consistent with an upper bound of $B = 21.05$ on the maximum absolute value of the demand shock. These are the downward-sloping lines that pass through the origin and through all of the dotted intervals. The implied bound on the slope is -0.43 and the corresponding bound on demand elasticity (when evaluated at mean price and quantity) is -1.3.

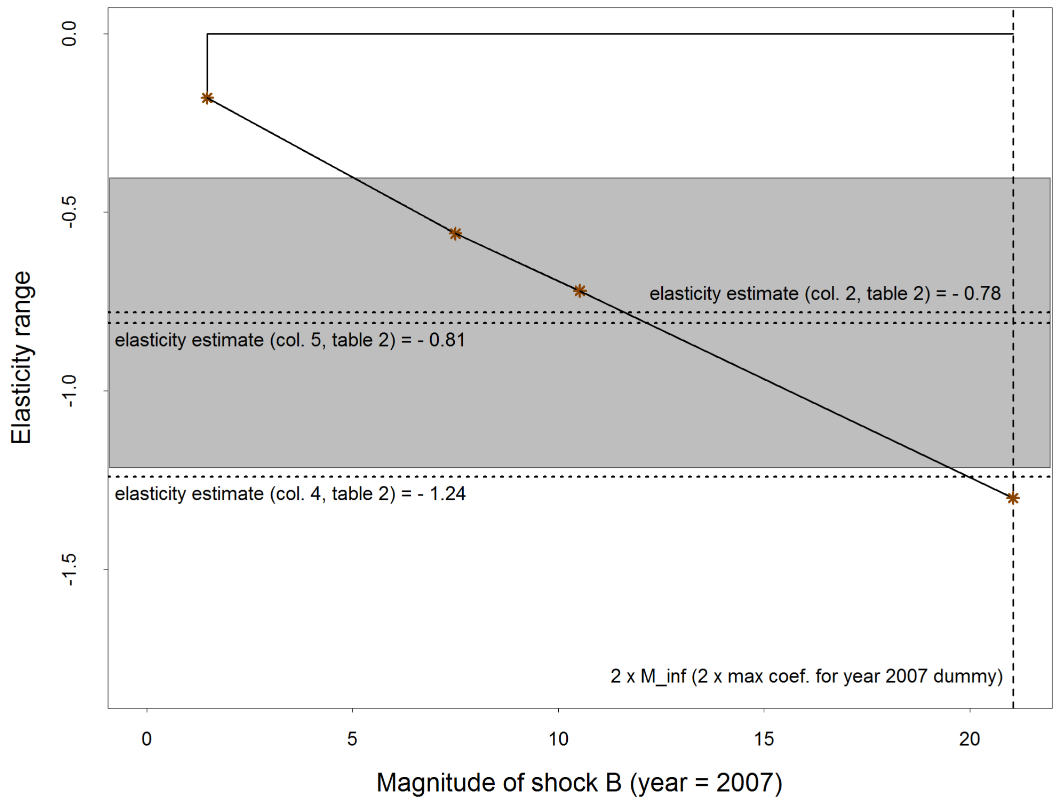

Figure 5: Figure plots the range of honey bee demand elasticities (when evaluated at mean price and quantity) that are consistent with bounds on the plausible size of shocks to demand in the year 2007 ranging from the minimum estimated shock in 2007 to twice the maximum estimated shock in 2007. The subsample is the balanced panel of all 1057 farmers that we use in specification (5) in Table 2. Estimates of shocks from 2007 are obtained from point estimates on the dummy for 2007 included in specifications (3), (4), and (5) in Table 2. The dashed vertical line is at twice the maximum estimated shock in 2007. The horizontal dotted lines depict the point estimates for the demand elasticity from specifications (2), (4), and (5) in Table 2, and the shaded region depicts the associated 95% confidence interval for the estimated demand elasticity from specification (5) in Table 2.

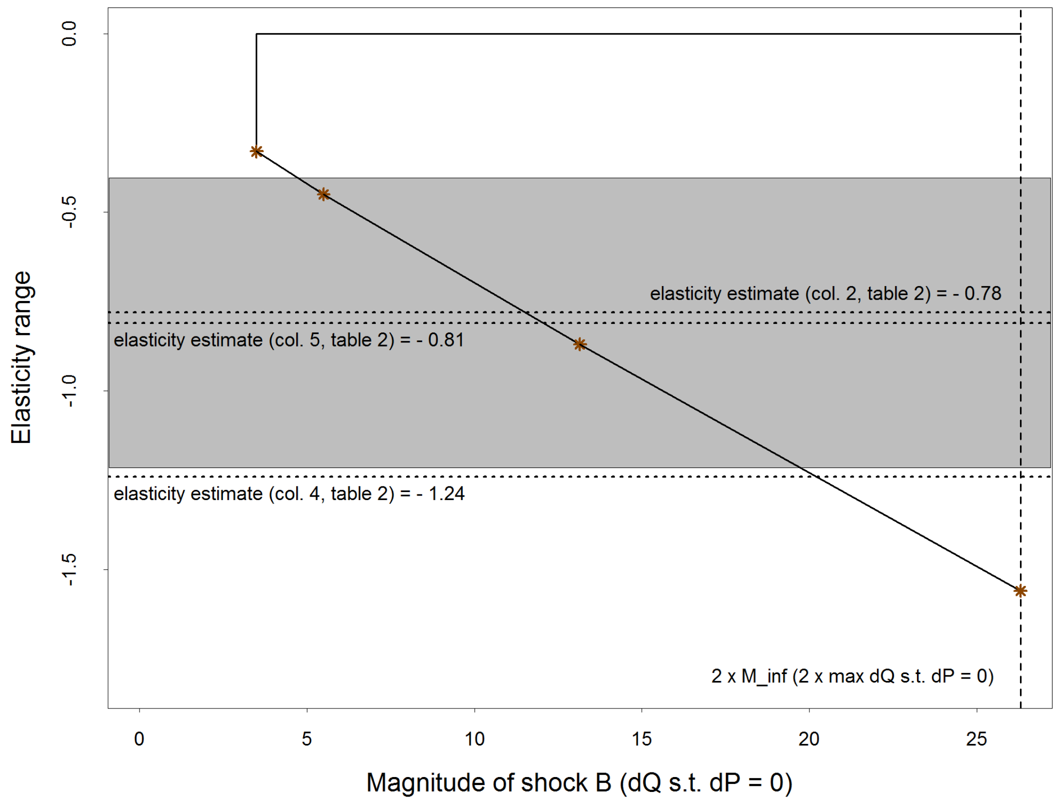

Figure 6: Figure plots the range of honey bee elasticities (when evaluated at mean price and quantity) that are consistent with bounds on the plausible size of shocks to demand ranging from the mean absolute value of the smoothed differenced quantity for observations for which the differenced price is equal to zero, to twice the maximum absolute value of the smoothed differenced quantity for observations for which the smoothed differenced price is equal to zero. The subsample is the balanced panel of all 1057 farmers that we use in specification (5) in Table 2. The dashed vertical line is at twice the maximum absolute value of the smoothed differenced quantity for observations for which the differenced price is equal to zero. The horizontal dotted lines depict the point estimates for the demand elasticity from specifications (2), (4), and (5) in Table 2, and the shaded region depicts the associated 95% confidence interval for the estimated demand elasticity from specification (5) in Table 2.

7 Relationship Between Yield, Profits, and Honey Bee Use

A key component of our theoretical propositions has to do with the shape of the response function between output and managed pollination use (see Lemma 1 in Appendix A). In this section, we address this notion empirically, and also estimate optimal honey bee stocking densities with respect to yield and profits. These relationships are potentially very meaningful economically, particularly since our descriptive analysis in Figure 2 suggests significant differences in production outcomes between farmers who do and do not rent honey bees, particularly for yield and profits.

To study how yield and profits vary with managed pollination use, we rely on optimal binscatter estimators from Cattaneo et al. (2021) to estimate the following semi-parametric function:

$$y_{isct} = \mu(x_{isct}) + \mathbf{w}'_{\mathbf{isct}}\gamma + \varepsilon_{isct}, \tag{8}$$

where y_{isct} is either block-level profits or yield per acre for farmer i in state s, county c in year t; x_{isct} is the number of honey bee colonies per acre employed for pollination at the random apple block level; and $\mathbf{w}_{\mathbf{isct}}$ is a vector of covariates, which in respective estimations include average monthly precipitation and temperature over January-September (the months leading into the main harvest period), trees per acre, apple bearing acres, average age of trees, labor hours for pruning/thinning, harvesting, machine and land preparation work, full time and seasonal labor, and county-level natural forest cover and natural open cover. The goal is to recover the unknown function $\mu(x_{isct})$, which is the functional relationship between outcome (profits or yield) y_{isct} and honey bee colonies per acre x_{isct}.[17]

We rely on semi-parametric methods developed by Cattaneo et al. (2021) in order to estimate response functions $\mu(x_{isct})$, plot them with confidence bands, and estimate the respective first and second derivatives. With estimation of the first derivative we can identify global and local maxima of yield and profits where the first derivative is zero and the second derivative is negative. Finally, we apply formal t-tests developed by Cattaneo et al. (2021) for the parametric form of the response

[17] As discussed in more detail in Appendix E.1, honey bee colonies rented per acre is arguably exogenous to yield and profit since honey bees are rented during the bloom period, several months before yield and profits are realized; and since farmers are unable to precisely control insect pollination.

function and for shape restrictions on the first and second derivatives (i.e., monotonicity, concavity, and convexity). For further robustness checks, we estimate linear fixed effects models and employ second-order polynomials in honey bee colonies per acre and other covariates. We describe and discuss the Cattaneo et al. (2021) methods and how we apply them in more detail in Appendix E.1.

Figures 7 and 8 show estimated response curves for block-level yield (bushels/acre) and profits ($/acre), respectively, as semi-parametric functions of honey bee colonies per acre, from applying optimal binscatter with quantile-spaced bins. Each figure presents results for the pooled sample of all states, and for the Eastern and Western states subsamples, and range from including no covariate adjustment, to covariate adjustment, to covariate-adjustment with state dummies. Therefore, column 3 in each figure reflects the greatest degree of controls and fixed effects. Each figure also includes a trimmed scatter between y_{isct} (yield or profits) and x_{isct}, honey bees colonies rented per acre, which excludes the 99th centile as it can make a tremendous difference in the legibility of the figure. Results for parametric tests and shape restriction tests focused on yield are provided in Tables 3 and 4 (results are similar when profits are the outcome of interest). Supplementary results are provided in Appendix E for linear fixed effects specifications (Tables E.1-E.3), alternative versions of Figures 7 and 8 that use equally spaced bins (Figures E.1 and E.2), and parametric and shape restriction tests focused on yield for Eastern and Western states subsamples (Tables E.4-E.7).

The main takeaway from these empirical results are as follows. First, as is apparent in results using either quantile-spaced bins (Figures 7 and 8), equally spaced bins (Figures E.1 and E.2), or linear fixed effects models (Tables E.1-E.3), and consistent with Lemma 1 in Appendix A, yield and profits are concave in honey bee colonies per acre. For both yield and profits, the optimal number of honey bee colonies per acre is approximately 3 to 4 for the pooled sample of all states, around 2 honey bee colonies per acre for Eastern states, and around 4 honey bee colonies per acre for Western states. We also see that the marginal returns for Western states for 1 additional colony per acre tends to be larger than for Eastern states.

Second, formal parametric tests shown in Table 3 for the pooled sample reject hypotheses that the response function for yield is constant, linear, or cubic in honey bee colonies per acre, but, consistent with Lemma 1 in Appendix A, generally do not reject concavity. Third, in Table 4 we see that for the pooled sample, a monotonically increasing function for yield is rejected, a monotonically decreasing function is not soundly rejected, convexity is rejected for the second derivative, but we see again see that, consistent with Lemma 1 in Appendix A, concavity is not rejected. Respective tests focused on the subsamples of Western and Eastern states in Appendix E yield qualitatively similar results.

A variety of additional noteworthy findings are apparent from respective linear fixed effects regressions of yield and profits (Tables E.1-E.3), which we summarize briefly here and direct the reader to Appendix E for further details. First, results from linear fixed effects regressions show that, consistent with Lemma 1 in Appendix A, yield and profits are concave in honey bee colonies, and that these relationships that are highly statistically significant and economically meaningful.

Second, for the relationship between profits and measures of production scale, we find that, consistent with Proposition 3, the relationship with block size exhibits concavity and significance, and the relationship with age of trees seems to exhibit concavity with much greater economic significance than trees per acre (Table E.3). Third, for the relationship between yield and measures of production scale, we find that the association with block size is negative and marginally significant, and the relationship with age of trees exhibits concavity and significance, more so than that with trees per acre (Table E.1).

Fourth, it is apparent that labor hours put into production make a positive difference for yield (Table E.1), but can also have negative impacts on profits as one would expect (Table E.3).

Fifth, we observe some significant relationships with land cover. Particularly notable is the relationship between yield and natural forest cover, which becomes concave and significant in Eastern states (Table E.1). This is noteworthy given existing evidence for natural forests in some contexts being sources of wild pollinator stocks which may enhance apple yield and fruit quality (Park et al., 2015; Kammerer et al., 2016). To explore the durability of this finding, we run additional linear fixed effects models for the Eastern states with alternative measures of natural forest cover that reflect buffers of 1000 and 3000 meters around apple production areas within counties observed in the 2007 USDA-ARMS, and find that the concave relationship holds (Table E.2). We also estimate optimal binscatter curves focused on yield and natural forest cover (Figure E.3 in Appendix E), which further demonstrate that yield is concave in natural forest cover for Eastern states.

Finally, weather covariates exhibit a variety of logical relationships with profits and yield that are worthy of deeper analysis in subsequent research. For example, a wet May is negative and significant for yield, which may reflect conditions that result in poor fruit set[18] (i.e., adequate pollination inducing fruit set is difficult in rainy conditions) as is a warm January (which can prematurely bring trees out of dormancy). Nevertheless, since we only have a cross-section to work with for yield and profits, we expect our weather-related findings are not capturing the full range of relationships with weather shocks.

[18]Fruit set is the biological process in which flowers become fruit and potential fruit size is determined (Mid Valley Agricultural Services, 2006). When seed formation is complete and well-distributed, the fruit is considered to be more appealing (e.g., consistent shape and fruit quantity/quality), which generally means a higher price is received by the farmer.

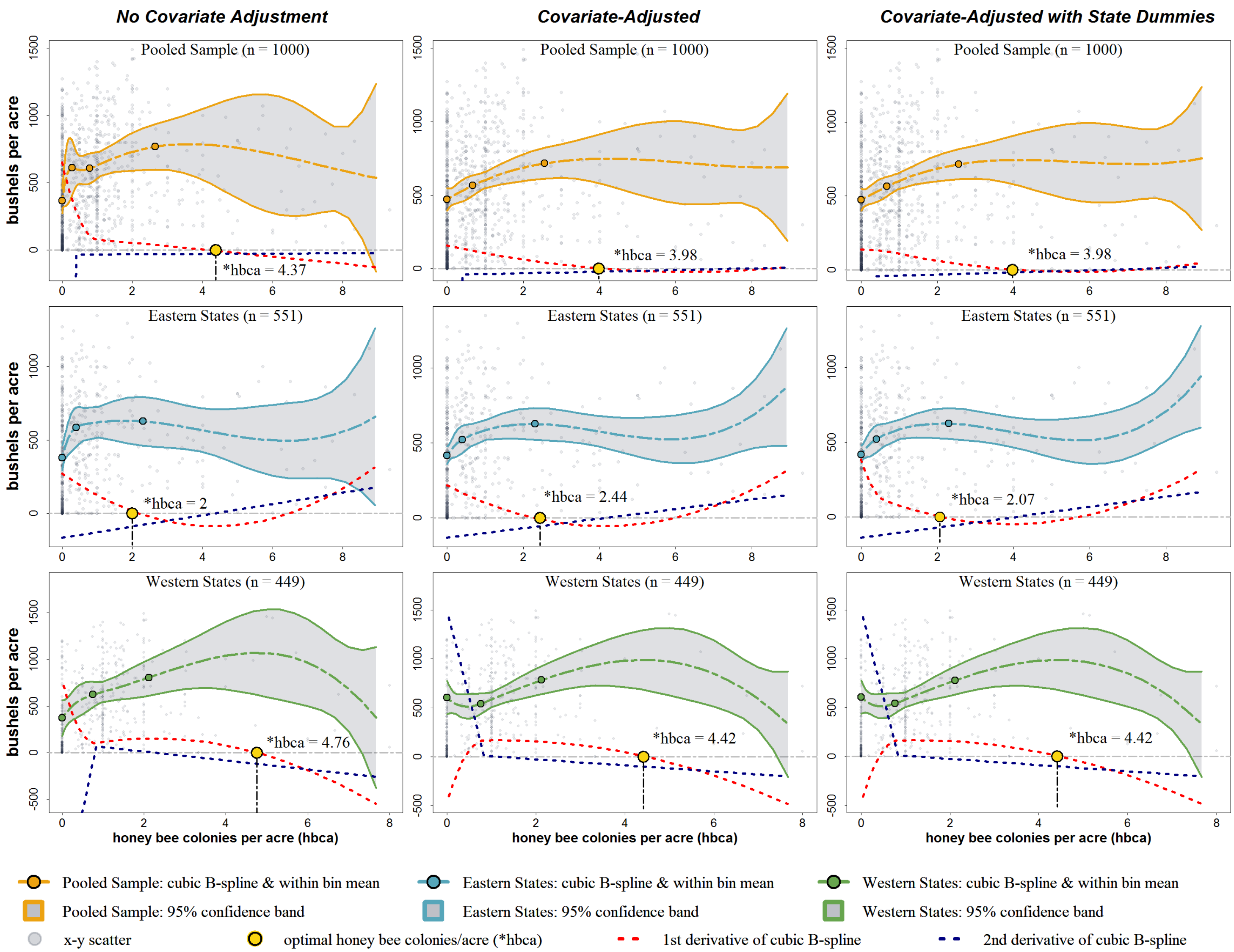

Figure 7: Optimal binscatter (following Cattaneo et al. (2021)) of *yield* in bushels per acre on the semi-parametric function $\mu(x)$, where x is honey bee colonies per acre, which is defined as the number of honey bee colonies rented divided by selected block size in acres. Each panel trims the 99% centile of the outcome variable and honey bee colonies per acre to reduce the influence of extreme outliers that can dramatically affect the readability of the figure. Column 1 is the optimal binscatter of yield on honey bee colonies per acre. Column 2 includes covariate-adjustment using the same covariates employed in the linear models in Tables E.1-E.3 of Appendix E, with the exception of the polynomial versions of some of these variables. Column 3 employs the same model in Column 2 but includes state dummies. These estimations employ *quantile-spaced*, data-driven rule of thumb bin selection, and cubic B-splines within and between bins. Confidence bands are bootstrapped with n draws. Optimal honey bee colonies per acre are plotted where the estimated first derivative (in red) of the response function equals zero and the response function is at a global (or local) maximum. Second derivatives are also plotted in dark blue.

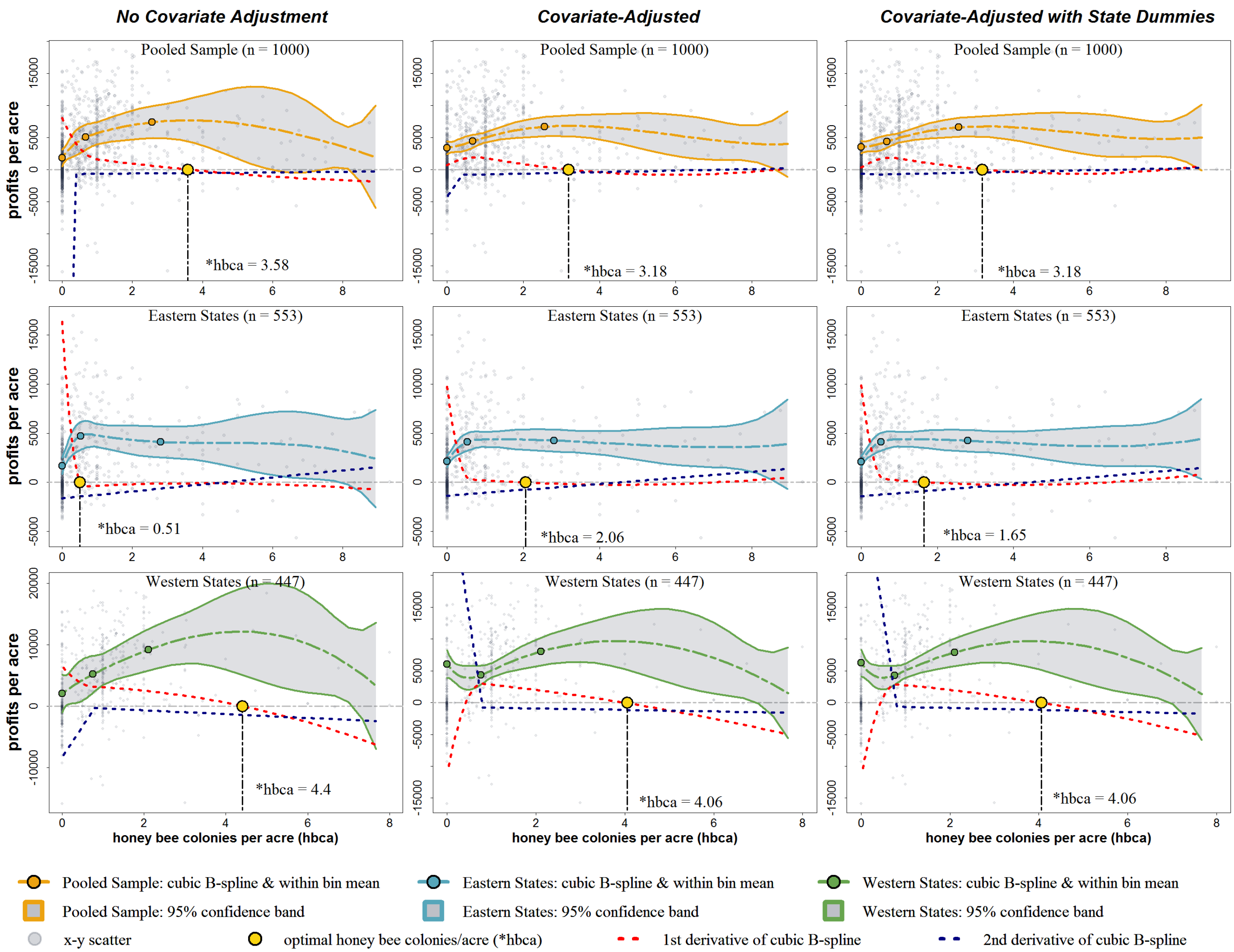

Figure 8: Optimal binscatter (following Cattaneo et al. (2021)) of *profits* in dollars per acre on the semi-parametric function $\mu(x)$, where x is honey bee colonies per acre, which is defined as the number of honey bee colonies rented divided by selected block size in acres. Each panel trims the 99% centile of the outcome variable and honey bee colonies per acre to reduce the influence of extreme outliers that can dramatically affect the readability of the figure. Column 1 is the optimal binscatter of yield on honey bee colonies per acre. Column 2 includes covariate-adjustment using the same covariates employed in the linear models in Tables E.1-E.3 of Appendix E, with the exception of the polynomial versions of some of these variables. Column 3 employs the same model in Column 2 but includes state dummies. These estimations employ *quantile-spaced*, data-driven rule of thumb bin selection, and cubic B-splines within and between bins. Confidence bands are bootstrapped with n draws. Optimal honey bee colonies per acre are plotted where the estimated first derivative (in red) of the response function equals zero and the response function is at a global (or local) maximum. Second derivatives are also plotted in dark blue.

Table 3: Parametric tests of response function $\mu(x)$ for yield.

	no covariate adjustment	covariate-adjusted	covariate-adjusted with state dummies
constant	6.369	3.694	3.637
	(0.000)	(0.001)	(0.000)
linear	4.061	1.840	1.687
	(0.000)	(0.122)	(0.128)
quadratic	1.763	0.628	0.752
	(0.450)	(0.763)	(0.608)
cubic	1.373	0.008	0.086
	(0.697)	(1.000)	(0.999)
# Bins	3.000	3.000	3.000
# Observations	1000	1000	1000
# Distinct values	265	265	265

Notes: Table presents t-statistics (p-values in parentheses) from parametric tests of the response function $\mu(x)$ for yield for specifications using observations from all states ('Pooled Sample'). Yield is in bushels per acre; x is honey bee colonies per acre defined as the number of honey bee colonies rented divided by selected block size in acres. Tests employ rule of thumb approach for selection of the number of bins (Cattaneo et al. 2021), quantile-spaced bins, and sample weights. Significance codes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 4: Shape restriction tests of response function $\mu(x)$ for yield.

	no covariate adjustment	covariate-adjusted	covariate-adjusted with state dummies
non-positive	6.073	3.329	3.322
	(0.000)	(0.004)	(0.004)
non-negative	−1.522	−0.639	−0.441
	(0.385)	(0.878)	(0.935)
concave	−0.195	0.075	0.075
	(1.000)	(0.983)	(0.983)
convex	−3.186	−1.808	−1.808
	(0.002)	(0.170)	(0.170)
# Bins	2	2	2
# Observations	1000	1000	1000
# Distinct values	265	265	265

Notes: Table presents t-statistics (p-values in parentheses) from shape restriction tests of the response function $\mu(x)$ for yield for specifications using observations from all states ('Pooled Sample'). Yield is in bushels per acre; x is honey bee colonies per acre defined as the number of honey bee colonies rented divided by selected block size in acres. Monotonicity tests are applied to the first derivative of respective optimal binscatter curves for the models represented in each column. Tests for concavity and convexity are applied to the respective second derivatives. Tests employ data-driven rule of thumb approach for selection of the number of bins (Cattaneo et al. 2021), quantile-spaced bins, and sample weights. Significance codes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

8 Discussion and Conclusion

Pollination-dependent farmers are important actors in the pollination and food production space, and their needs and choices need to be better understood in order to maintain the production of nutritious specialty crops that are critical to food security and welfare around the world, as well as the value chains that these production sectors underpin (Schmit et al., 2018). In particular, pollination choices of pollination-dependent farmers have important implications for production outcomes, food supplies, and pollination resources within and beyond the farm gate.

In this paper, we develop new theory surrounding the decision to use managed pollination services. We derive several propositions regarding the impact of several factors on these choices, including output price, pollination input prices, total factor productivity, physical capital, and wild pollination resource stocks; as well as related theory on how the shape of the response function between output and managed pollination use relates to relationships between these factors and managed pollination use. We take this theory to data, undertaking, to our knowledge, the first empirical study of pollination behavior in conjunction with farm-level production data.

Our empirical findings are broadly congruent with our theoretical propositions. Consistent with Proposition 1, we find that demand for managed pollination declines in its own price, and is weakly inelastic to weakly elastic (Sections 6.1 and 6.2). Employing a novel instrumental variables strategy that exploits the well-documented supply shifts caused by high early season demand for managed pollination services from California's almond sector, we estimate that the managed pollination demand elasticity ranges from -0.78 to -1.24; these results are consistent with bounds on the elasticity we calculate using new methods from Petterson, Seim, and Shapiro (2023). Proposition 1 suggests that weakly inelastic to weakly elastic demand is likely to arise when returns are diminishing in managed pollination use – a condition for which we find strong evidence.

Consistent with Proposition 2, we find that the likelihood of renting honey bees increases with output price (Section 5).

Consistent with Proposition 3, which suggests that farm scale may have a small and concave effect on managed pollination use, we see evidence that managed pollination use is increasing and concave in measures of the scale of production. While the marginal effect for the binary choice to rent bees is small (Section 5), magnitudes become economically significant for quantity demanded (Section 6.1).

Consistent with Proposition 4, we find that the likelihood of never renting honey bees over 2006-2007 decreases with the farm's physical capital stock (Section 5).

We see notable relationships with land cover measures that proxy for landscape heterogeneity and wild pollination stocks (Section 5), some of which may be suggestive of Proposition 5. We find that the likelihood of renting honey bees increases with natural open cover. For Eastern states, we find that yield is concave in natural forest cover.

Consistent with Lemma 1 in Appendix A, we find that yield and profits are concave in managed pollination use (Section 7). Our finding of diminishing returns to the use of managed pollination is not only noteworthy by itself, but also helps to explain why demand for managed pollination is not perfectly elastic (Proposition 1) and why the scale of production does not have greater economic importance for the binary choice of whether to use managed pollination (Proposition 3).

The methods that we employ from Cattaneo et al. (2021) also permit us to estimate first and second derivatives of underlying response functions, which in turn allow us to estimate optimal honey bee stocking densities for yield and profits. For farmers in Western states, the optimal stocking density is around 4 colonies per acre; for farmers in Eastern states, the optimal stocking density is closer to 2 colonies per acre. Estimated response curves also suggest that farmers in Western states experience a greater return to the marginal honey bee colony than farmers in Eastern states. On average, descriptive data suggests Eastern farmers are closer to being at optimal levels of managed pollination use than farmers in Western states.

To our knowledge, many of these empirical measurements are the first of their kind, which by itself emphasizes how little is known about these production sectors and associated pollination demand around the world. For example, Rucker, Thurman, and Burgett (2012) and Ferrier et al. (2018) give some limited attention to farmer pollination demand through the lens of stocking densities, and use theory and aggregate cost data to argue that honey bee stocking densities are likely to be held more or less fixed by farmers, suggesting that demand is either perfectly inelastic or at least strongly inelastic.[19] There are reports that some apple growers in the Northeast consider honey bee rental an inconsequential cost analogous to "cheap insurance" (Biltonen, 2020; Kahlke, 2019), suggesting that demand may be highly inelastic. In contrast, our empirical analyses of the elasticity of demand suggest that, while not perfectly elastic, managed pollination demand may be more elastic than prior work has assumed and perhaps conventional wisdom sometimes suggests. To our knowledge, no prior research has attempted to empirically estimate an own-price demand elasticity for managed pollination.

Similarly, our finding of significant regional differences in optimal managed pollination use levels, and differing marginal returns per marginal colony, raise questions about the underlying factors in these agro-ecological systems that produce these divergent scenarios. Indeed, it is plausible that sources of wild pollination stocks from forested regions, particularly in the Eastern states, may be providing a significant pollination subsidy to apple farmers in these states, and this may in part explain why apple farmers in Eastern states do not see larger marginal returns for the marginal honey bee colony. This notion is in fact consistent with our finding that yield is concave in natural forest cover for Eastern states. Indeed, if this hypothesis could be more rigorously tested the implications may be significant. Credible ways to measure the state of wild pollination stocks at the farm level, perhaps through combinations of remote sensing and traditional field methods from entomology, could greatly enhance such research endeavors and provide a much clearer picture as to whether pollination dependent sectors are over- or under-supplied from a pollination perspective.

Our work also offers a useful setting to consider the ideas of Zilberman, Lu, and Reardon (2019), who study innovations in bioeconomy supply chains and the choices firms make in the process of introducing innovations[20] to resolve resource demands either internally or externally via service providers. This lens is instructive for pondering how pollination choice has evolved in the past, and how it may evolve in the future, and what this may mean for the state of both wild

[19] Rucker, Thurman, and Burgett (2012) assume fixed bee stocking densities to sign comparative statics.

[20] Lee, Sumner, and Champetier (2019) address related ideas with their simulations of almond sector innovations, such as self-pollination almond varieties, to assess impacts to pollination service providers and honey output.

and domesticated pollination resource stocks.

Moving forward, we suggest that fruitful research endeavors abound to replicate the kinds of empirical work we have accomplished in this paper with more recent data and in other pollination-dependent sectors around the world. This type of work seems to us, critical for tracking relationships between pollination sectors and agriculture, and placing policy-making on better footing. Future refinements to our contributions might be made for advising farmers on optimal pollination strategies; uncovering the complexities of where wild pollination stocks do or do not subsidize production effectively and why; and analyzing the spatial externalities that arise from wild pollination resources. Gaining a better understanding of optimality from a pollination perspective would better allow pollination-dependent farmers to manage pollination resources sustainability and to find innovative ways to resolve pollination resource needs within bioeconomy value chains (Zilberman, Lu, and Reardon, 2019).

References

Albers, H., and E. Robinson. 2011. "The Trees and the Bees: Using Enforcement and Income Projects to Protect Forests and Rural Livelihoods Through Spatial Joint Production." *Agricultural and Resource Economics Review* 40:424–438.

Andrews, I., J.H. Stock, and L. Sun. 2019. "Weak instruments in instrumental variables regression: Theory and practice." *Annual Review of Economics* 11:727–753.

Angrist, J.D., K. Graddy, and G.W. Imbens. 2000. "The Interpretation of Instrumental Variables Estimators in Simultaneous Equations Models with an Application to the Demand for Fish." *Review of Economic Studies* 67:499–527.

Baylis, K., E.M. Lichtenberg, and E. Lichtenberg. 2021. "Economics of Pollination." *Annual Review of Resource Economics* 13:335–354.

Bergé, L. 2018. "Efficient estimation of maximum likelihood models with multiple fixed-effects: the R package FENmlm." Working paper.

Biltonen, M. 2020. personal communication.

Blitzer, E.J., J. Gibbs, M.G. Park, and B.N. Danforth. 2016. "Pollination services for apple are dependent on diverse wild bee communities." *Agriculture, Ecosystems and Environment* 221:1–7.

Borusyak, K., P. Hull, and X. Jaravel. 2022. "Quasi-experimental shift-share research designs." *Review of Economic Studies* 89:181–213.

Boryan, C., Z. Yang, R. Mueller, and M. Craig. 2011. "Monitoring US agriculture: the US Department of Agriculture, National Agricultural Statistics Service, Cropland Data Layer Program." *Geocarto International* 26:341–358.

Bruno, E.M., and K. Jessoe. 2021. "Missing markets: Evidence on agricultural groundwater demand from volumetric pricing." *Journal of Public Economics* 196:104374.

Cameron, S.A., J.D. Lozier, J.P. Strange, J.B. Koch, N. Cordes, L.F. Solter, and T.L. Griswold. 2011. "Patterns of widespread decline in North American bumble bees." *Proceedings of the National Academy of Sciences* 108:662–667.

Casert, R. 2018. "EU Moves to Full Ban on Pesticides That Harm Bees." *The Associated Press* 27 April 2018, URL: `https://apnews.com/article/3df5af693d494f34b6d4a9df00bfbed2`.

Cattaneo, M.D., R.K. Crump, M.H. Farrell, and Y. Feng. 2021. "On binscatter."

Champetier, A., D.A. Sumner, and J.E. Wilen. 2015. "The bioeconomics of honey bees and pollination." *Environmental and Resource Economics* 60:143–164.

Cohen, H. 2022. "Planting mixes of flowers around farm fields helps keep bees healthy." *Morning Ag Clips* 6 April 2022, URL: `https://www.morningagclips.com/planting-mixes-of-flowers-around-farm-fields-helps-keep-bees-healthy/`.

Daly, C., M. Halbleib, J.I. Smith, W.P. Gibson, M.K. Doggett, G.H. Taylor, J. Curtis, and P.P. Pasteris. 2008. "Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States." *International Journal of Climatology* 28:2031–2064.

Danforth, B.N., R.L. Minckley, and J.L. Neff. 2019. "Solitary bees and agricultural pollination." In *The Solitary Bees: Biology, Evolution, Conservation*. Princeton University Press, chap. 13.

Fei, C.J., K.M. Williamson, R.T. Woodward, B.A. McCarl, and J. Rangel. 2021. "Honey bees, almonds, and colony mortality: An economic simulation of the US pollination market." *Land Economics* 97:688–703.

Ferrier, P.M., R.R. Rucker, W.N. Thurman, and M. Burgett. 2018. "Economic Effects and Responses to Changes in Honey Bee Health." Working paper.

Garfield, L. 2018. "Walmart has hinted that it's building crop-pollinating robot bees." *Business Insider* 14 March 2018, URL: `http://www.businessinsider.com/walmart-robot-bees-farming-patent-2018-3`.

Garibaldi, L.A., I. Steffan-Dewenter, R. Winfree, M.A. Aizen, R. Bommarco, S.A. Cunningham, C. Kremen, L.G. Carvalheiro, L.D. Harder, O. Afik, et al. 2013. "Wild pollinators enhance fruit set of crops regardless of honey bee abundance." *Science* 339:1608–1611.

Goldberger, A. 1991. *A Course in Econometrics*. Cambridge, MA: Harvard University Press.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift. 2020. "Bartik instruments: What, when, why, and how." *American Economic Review* 110:2586–2624.

Goodrich, B., J. Williams, and R. Goodhue. 2019a. "The great bee migration: Supply analysis of honey bee colony shipments into California for almond pollination services." Working paper.

Goodrich, B.K. 2017. "The California Almond Pollination Market: Contracts and Honey Bee Colony Health." PhD dissertation, University of California, Davis.

Goodrich, B.K., J.C. Williams, and R.E. Goodhue. 2019b. "The great bee migration: Supply analysis of honey bee colony shipments into California for almond pollination services." *American Journal of Agricultural Economics* 101:1353–1372.

Grab, H., M.G. Branstetter, N. Amon, K.R. Urban-Mead, M.G. Park, J. Gibbs, E.J. Blitzer, K. Poveda, G. Loeb, and B.N. Danforth. 2019. "Agriculturally dominated landscapes reduce bee phylogenetic diversity and pollination services." *Science* 363:282–284.

Grab, H., K. Poveda, B. Danforth, and G. Loeb. 2018. "Landscape context shifts the balance of costs and benefits from wildflower borders on multiple ecosystem services." *Proceedings of the Royal Society B: Biological Sciences* 285:20181102.

Imbach, P., E. Fung, L. Hannah, C.E. Navarro-Racines, D.W. Roubik, T.H. Ricketts, C.A. Harvey, C.I. Donatti, P. Läderach, B. Locatelli, et al. 2017. "Coupling of pollination services and coffee suitability under climate change." *Proceedings of the National Academy of Sciences* 114:10438–10442.

Kahlke, C. 2019. personal communication.

Kammerer, M.A., D.J. Biddinger, N.K. Joshi, E.G. Rajotte, and D.A. Mortensen. 2016. "Modeling local spatial patterns of wild bee diversity in Pennsylvania apple orchards." *Landscape Ecology* 31:2459–2469.

Kasina, J., J. Mburu, M. Kraemer, and K. Holm-Mueller. 2009. "Economic benefit of crop pollination by bees: a case of Kakamega small-holder farming in western Kenya." *Journal of Economic Entomology* 102:467–473.

Kennedy, C.M., E. Lonsdorf, M.C. Neel, N.M. Williams, T.H. Ricketts, R. Winfree, R. Bommarco, C. Brittain, A.L. Burley, D. Cariveau, et al. 2013. "A global quantitative synthesis of local and landscape effects on wild bee pollinators in agroecosystems." *Ecology Letters* 16:584–599.

Lee, H., D.A. Sumner, and A. Champetier. 2019. "Pollination markets and the coupled futures of almonds and honey bees: simulating impacts of shifts in demands and costs." *American Journal of Agricultural Economics* 101:230–249.

Lin, C.Y.C. 2011. "Estimating supply and demand in the world oil market." *Journal of Energy and Development* 34:1–32.

Lippert, C., A. Feuerbacher, and M. Narjes. 2021. "Revisiting the economic valuation of agricultural losses due to large-scale changes in pollinator populations." *Ecological Economics* 180:106860.

Manski, C.F. 1995. *Identification Problems in the Social Sciences*. Cambridge, MA: Harvard University Press.

Meade, J.E. 1952. "External economies and diseconomies in a competitive situation." *Economic Journal* 62:54–67.

Mid Valley Agricultural Services. 2006. "Fruit Set Factors." *Mid Valley Agricultural Services Newsletter* October 2006.

Montiel Olea, J.L., and C. Pflueger. 2013. "A robust test for weak instruments." *Journal of Business and Economic Statistics* 31:358–369.

Narjes, M.E., and C. Lippert. 2021. "Regional differences in farmers' preferences for a native bee conservation policy: The case of farming communities in Northern and Eastern Thailand." *PloS ONE* 16:e0251206.

Park, M.G., E. Blitzer, J. Gibbs, J.E. Losey, and B.N. Danforth. 2015. "Negative effects of pesticides on wild bee communities can be buffered by landscape context." *Proceedings of the Royal Society B: Biological Sciences* 282:20150299.

Park, M.G., R.A. Raguso, J.E. Losey, and B.N. Danforth. 2016. "Per-visit pollinator performance and regional importance of wild Bombus and Andrena (Melandrena) compared to the managed honey bee in New York apple orchards." *Apidologie* 47:145–160.

Penn, J., W. Hu, and H.J. Penn. 2019. "Support for solitary bee conservation among the public versus beekeepers." *American Journal of Agricultural Economics* 101:1386–1400.

Petterson, M.S., D.G. Seim, and J.M. Shapiro. 2023. "Bounds on a slope from size restrictions on economic shocks." *American Economic Journal: Microeconomics* 15:552–572.

Reilly, J., D. Artz, D. Biddinger, K. Bobiwash, N. Boyle, C. Brittain, J. Brokaw, J. Campbell, J. Daniels, E. Elle, et al. 2020. "Crop production in the USA is frequently limited by a lack of pollinators." *Proceedings of the Royal Society B* 287:20200922.

Ridley, W., and S. Devadoss. 2021. "Challenges for the US Fruit Industry: Trends in Production, Consolidation, and Competition." *Choices* 36.

Roubik, D.W. 2002. "Tropical agriculture: the value of bees to the coffee harvest." *Nature* 417:708.

Rucker, R.R., W.N. Thurman, and M. Burgett. 2019. "Colony collapse and the consequences of bee disease: Market adaptation to environmental change." *Journal of the Association of Environmental and Resource Economists* 6:927–960.

—. 2012. "Honey bee pollination markets and the internalization of reciprocal benefits." *American Journal of Agricultural Economics* 94:956–977.

Russo, L., M. Park, E. Blitzer, and B. Danforth. 2017. "Flower handling behavior and abundance determine the relative contribution of pollinators to seed set in apple orchards." *Agriculture, Ecosystems and Environment* 246:102–108.

Sato, K. 1967. "A two-level constant-elasticity-of-substitution production function." *Review of Economic Studies* 34:201–218.

Schmit, T.M., R.M. Severson, J. Strzok, and J. Barros. 2018. "Economic contributions of the apple industry supply chain in New York State." Extension Bulletin EB 2018-03, Charles H. Dyson School of Applied Economics and Management.

Simpson, R.D. 2018. "Conservation incentives from an ecosystem service: How much farmland might be devoted to mative pollinators?" *Environmental and Resource Economics*, pp. 1–18.

vanEngelsdorp, D., and M.D. Meixner. 2010. "A historical review of managed honey bee populations in Europe and the United States and the factors that may affect them." *Journal of Invertebrate Pathology* 103:S80–S95.

Ward, R., A. Whyte, and R. James. 2010. "A tale of two bees: Looking at pollination fees for almonds and sweet cherries." *American Entomologist* 56:170–177.

Weitzman, M.L. 2009. *Income, Wealth, and the Maximum Principle*. Cambridge, MA: Harvard University Press.

Willett, L.S., and B.C. French. 1991. "An econometric model of the US beekeeping industry." *American Journal of Agricultural Economics* 73:40–54.

Wooldridge, J.M. 2010. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

Wu, L., and S.S. Atallah. 2019. "Farm and landscape-level trade-offs between pest control and pollination service provision: the case of neonicotinoid insecticides."

Zilberman, D., L. Lu, and T. Reardon. 2019. "Innovation-induced food supply chain design." *Food Policy* 83:289–297.

Online Appendix for:

To (Rent) Bees or Not to (Rent) Bees? An Examination of the Farmer's Question

Steven W. Wilcox, David R. Just, C.-Y. Cynthia Lin Lawell, Miguel I. Gómez, Heather Grab*

November 5, 2023

*Corresponding author: Steven W. Wilcox (steven.wilcox@usu.edu). Wilcox is affiliated with the Department of Applied Economics at Utah State University. All other authors are affiliated with Cornell University. Just: drj3@cornell.edu. Lin Lawell: clinlawell@cornell.edu. Gómez: mig7@cornell.edu. Grab: hlc66@cornell.edu. We gratefully acknowledge financial support from three awards from the Cornell Atkinson Center for Sustainability: a Sustainable Biodiversity Award, an Atkinson Venture Fund (AVF) Grant, and an AVF-Engaged Cornell grant for undergraduate involvement. We are indebted to Craig Khalke and the Lake Ontario Fruit Program for assistance and very helpful details on apple production in the Northeast. We also appreciate comments from a number of apple growers and consultants in the Northeast, including Debbie Biltonen, Mike Biltonen, and Ian Merwin. Carrie Jones at USDA-ERS and Donnie Fike at USDA-NASS kindly assisted us with many aspects of USDA-ARMS data. Anna Espinoza, Adeline Sutton, Jacob Llodra, and Katie Warner provided excellent research assistance. We thank Todd Gerarden, Ivan Rudik, Bryan Danforth, Maria Van Dyke, Mia Park, Chris Barrett, Kathy Baylis, and Jo Albers for helpful comments; and we thank seminar and conference participants at Cornell University, the Food Studies International Conference, the Association of Environmental and Resource Economists (AERE) Summer Conference, and the Agricultural and Applied Economics Association (AAEA) Annual Meeting. Just, Lin Lawell, and Gómez are Faculty Fellows at the Cornell Atkinson Center for Sustainability. All errors are our own.

A Theory Model

A.1 Pollination Choice

Farmers of pollination-dependent crops grow much of the world's nutritious and high-value fruits, nuts, and vegetables. They face very complex production decisions, which in many cases involve long-term investments – especially for long-lived tree fruits and nuts – and production is often labor intensive and dependent on migrant farmworkers (Ridley and Devadoss, 2021). Agronomic trends are also evolving away from traditional, low-density, very long-lived plants to high intensity, university-driven plant materials that are selected to tolerate greater density of fruit into smaller area on smaller trees, which are easier to harvest (Robinson et al., 2013).

Pollination choice, and general production strategies that affect pollination resource demand (e.g., planting density), are very important. Although insect pollination is not necessary for all crops, many crops require or benefit greatly from pollination from insects or other organisms.[1] In addition to decisions about general production strategies and other input choices, producers of pollination-dependent crops make decisions about whether and how much to use managed pollination services (e.g., renting honey bees); and whether and how much to invest in wild pollination (e.g., setting aside land for planting wildflower strips, or other natural cover (Cohen, 2022)).[2] Available research suggests that farmer pollination choices are critical as they impact farm-level outcomes like yield and fruit quality (Roubik, 2002; Garibaldi et al., 2013; Park et al., 2016; Russo et al., 2017; Danforth, Minckley, and Neff, 2019), local pollination resources within and beyond the

[1] In many cases crops are sufficiently self-pollinating (e.g., grasses), and in some special cases pollination is carried out with a combination of human labor and motorized or non-motorized farm implements (e.g., paint brushes, compressed air, and possibly drone-based technology in the future). Researchers are also developing self-pollinating plant varieties (Lee, Sumner, and Champetier, 2019). Widely consumed crops that require or greatly benefit from insect pollination include almonds, coffee, apples, avocados, cherries, peaches, blueberries, among many others.

[2] Although reliable global data on variation in pollination practices is not known to exist, available information suggests that farmers of pollination-dependent crops vary widely in the form of pollination they depend upon and in how critical pollination is viewed and valued as a resource. For example, contrast the well-trodden tail of US almond growers' seeming absolute dependence on imported honey bee colonies with reports that for many Northeast US apple growers pollination is almost an afterthought compared to concerns surrounding labor and traditional farm capital (Biltonen, 2020; Kahlke, 2019). Other qualitative variation is documented by Narjes and Lippert (2019) and Narjes and Lippert (2021) who document relationships between beekeepers and longan fruit farmers in Thailand. There are also special cases like vanilla in Madagascar, which is entirely dependent on labor for pollination (Kaila and Boone, 2020).

farm-gate (Kennedy et al., 2013; Park et al., 2015; Grab et al., 2018), and market-level outcomes through shifts in the supply and demand of both pollination resources and agricultural commodities (Rucker, Thurman, and Burgett, 2012; Goodrich, Williams, and Goodhue, 2019).

A.2 Proofs of Propositions 1 through 5

We take our start from the optimal control theory model in the main text:

$$\begin{aligned}
\max_{\{I_k(t),\ L(t),\ M(t),\ I_w(t)\}} \ & \int_0^\infty \pi(K(t),L(t),M(t),W(t))e^{-rt}dt \\
\text{s.t.} \quad & \dot{K}(t) = \delta_k K(t) + I_k(t) \ \ : \ \lambda_k(t) \\
& \dot{O}_w(t) = F(W(t)) + I_w(t) - \delta_{mw}M(t) - \delta_{kw}I_k(t) \ \ : \ \lambda_w(t) \\
& K(t),\ L(t),\ M(t),\ W(t) \geq 0 \\
& K(0) = K_o,\ W(0) = W_o,
\end{aligned} \tag{A.1}$$

and its current-value Hamiltonian:

$$\begin{aligned}
H_c &= \pi(K,\ L,\ M,\ W) \ + \ \lambda_k\left[\delta_k K + I_K\right] + \lambda_w\left[F(W) + I_w - \delta_{mw}M - \delta_{kw}I_k\right] \\
&= p_c\Big(A\left[\gamma_o(\alpha_m M^{-\rho_o} + \alpha_w W^{-\rho_o})^{\frac{\rho}{\rho_o}} + \gamma_{kl}(\alpha_k K^{-\rho_{kl}} + \alpha_l L^{-\rho_{kl}})^{\frac{\rho}{\rho_{kl}}}\right]^{\frac{-1}{\rho}}\Big) - \\
&\quad \Big(p_k I_k + p_l L + p_m M + p_w I_w\Big) + \\
&\quad \lambda_k\left[\delta_k K + I_K\right] + \lambda_w\left[F(W) + I_w - \delta_{mw}M - \delta_{kw}I_k\right].
\end{aligned} \tag{A.2}$$

After applying the Maximum Principle and simplifying expressions we arrive at the following optimality conditions:

[#1L]: $p_c(t)\frac{\partial Q}{\partial L} = p_l(t)$

[#1M]: $p_c(t)\frac{\partial Q}{\partial M} = p_m(t) + p_w(t)\delta_{mw}$

[#1w]: $p_w(t) = \lambda_w(t)$

[#1k]: $p_k(t) = \lambda_k(t) - p_w(t)\delta_{kw}$

[#2w]: $\dot{p}_w = p_w(t)[r - \frac{\partial F(W)}{\partial W}] - p_c(t)\frac{\partial Q}{\partial W}$

[#2k]: $\dot{p}_k = p_k(t)(r - \delta_k) + p_w(t)\delta_{kw}(\frac{\partial F(W)}{\partial W} - \delta_k) + p_c(t)(\delta_{kw}\frac{\partial Q}{\partial W} - \frac{\partial Q}{\partial k})$

[#3w]: $\lim_{t\to\infty} p_w(t)W(t)e^{-rt} = 0$

[#3k]: $\lim_{t\to\infty}(p_k(t) + p_w(t)\delta_{kw})K(t)e^{-rt} = 0$

Per our assumptions stated in the main text, we focus on the respective first-order condition for M, denoted [#1M].

We start with the following Lemma:

Lemma 1:

First, we show that, for most reasonable values of the parameters, the production function is weakly concave in M: $\frac{\partial^2 Q}{\partial M^2} \leq 0$. Since a more explicit expression for this derivative is useful for our results of interest, we can also establish the negativity result analytically.

Given the stated structure of output Q as given by a 2-level CES production (Sato, 1967), we use the following expressions to simplify notation:

$$\text{Let } B = \alpha_m M^{-\rho_o} + \alpha_w W^{-\rho_o}$$

$$\text{Let } C = \alpha_k K^{-\rho_{kl}} + \alpha_l L^{-\rho_{kl}}$$

$$\text{Let } D = \gamma_o(\alpha_m M^{-\rho_o} + \alpha_w W^{-\rho_o})^{\frac{\rho}{\rho_o}} + \gamma_{kl}(\alpha_k K^{-\rho_{kl}} + \alpha_l L^{-\rho_{kl}})^{\frac{\rho}{\rho_{kl}}}$$

$$\Rightarrow Q = A[\gamma_o(B)^{\frac{\rho}{\rho_o}} + \gamma_{kl}(C)^{\frac{\rho}{\rho_{kl}}}]^{\frac{-1}{\rho}} = A[D]^{\frac{-1}{\rho}}$$

Taking this simplified notation as our starting point we see the marginal product with respect to M is as below and that the sign is positive.

$$\frac{\partial Q}{\partial M} = \frac{Q\,(B(M,W))^{\frac{\rho-\rho_o}{\rho_o}}\ \alpha_m\,\gamma_o}{D\,M^{(\rho_o+1)}}$$

$$\Rightarrow\ \text{sign}\left[\frac{\partial Q}{\partial M}\right] \geq 0\ \text{ since all terms are} \geq 0.$$

Taking the above equation as the starting point for the second derivative we have the following expression for the second derivative with respect to M, with many simplifying steps omitted.

$$\frac{\partial^2 Q}{\partial M^2} = \frac{\partial}{\partial M}\left(\frac{QB(M,W)^{\frac{\rho-\rho_o}{\rho_o}}\alpha_m\gamma_o}{DM^{\rho_o+1}}\right)$$

$$= \frac{\alpha_m\gamma_o QB(M,W)^{\frac{\rho-\rho_o}{\rho_o}}}{\left(DM^{\rho_o+1}\right)^2}\left[\alpha_m\gamma_o B(M,W)^{\frac{\rho-\rho_o}{\rho_o}}(1+\rho) - \left(\gamma_o(B(M,W))^{\frac{\rho-\rho_o}{\rho_o}} + \frac{\gamma_{kl}(C(K,L))^{\frac{\rho}{\rho_{kl}}}}{B(M,W)}\right)\left(\alpha_m(\rho-\rho_o)\right) - D(1+\rho_o)M^{\rho_o}\right]$$

Since we know for sure that $\frac{\alpha_m\gamma_o B(M,W)^{\frac{\rho-\rho_o}{\rho_o}}Q}{\left(DM^{\rho_o+1}\right)^2} > 0$, we can resolve the sign of $\frac{\partial^2 Q}{\partial M^2}$ analytically as shown below:

$$\text{sign}\left[\frac{\partial^2 Q}{\partial M^2}\right] =$$
$$\text{sign}\left[\alpha_m\gamma_o B(M,W)^{\frac{\rho-\rho_o}{\rho_o}}(1+\rho) - \left(\gamma_o(B(M,W))^{\frac{\rho-\rho_o}{\rho_o}} + \frac{\gamma_{kl}(C(K,L))^{\frac{\rho}{\rho_{kl}}}}{B(M,W)}\right)\left(\alpha_m(\rho-\rho_o)\right) - D(1+\rho_o)M^{\rho_o}\right]$$

$$\Rightarrow \frac{\partial^2 Q}{\partial M^2} \leq 0 \iff$$
$$\left(\gamma_o(B(M,W))^{\frac{\rho-\rho_o}{\rho_o}} + \frac{\gamma_{kl}(C(K,L))^{\frac{\rho}{\rho_{kl}}}}{B(M,W)}\right)\left(\alpha_m(\rho-\rho_o)\right) + D(1+\rho_o)M^{\rho_o} \geq \alpha_m\gamma_o B(M,W)^{\frac{\rho-\rho_o}{\rho_o}}(1+\rho)$$

From the above we can see that a negative sign is likely analytically. This is because under realistic assumptions, the above inequality is likely to hold given that D, for example, is the majority of what comprises Q. Hence it should hold that $D \geq B(M,W)^{\frac{\rho-\rho_o}{\rho_o}}$.

By this additional work, we assume for the remaining expressions that $\frac{\partial Q}{\partial M} \geq 0$ and $\frac{\partial^2 Q}{\partial M^2} \leq 0$.

Proof of Proposition 1:

Taking [#1M] as the starting point, the respective total derivative and expression for $\frac{dM}{dp_m}$ are as follows:

$$-dp_m + p_c \frac{\partial^2 Q}{\partial M^2} dM = 0$$

$$\Rightarrow \frac{dM}{dp_m} = \frac{1}{p_c \frac{\partial^2 Q}{\partial M^2}}$$

The elasticity η_{M,p_m} of M with respect to managed pollination price p_m is then given by:

$$\eta_{M,p_m} = (\frac{p_m}{M})\frac{dM}{dp_m} = \frac{p_m}{Mp_c \frac{\partial^2 Q}{\partial M^2}}$$

With the intermediate results from Lemma 1, we obtain:

$$\text{sign}\left[\frac{dM}{dp_m}\right] = \text{sign}\left[\frac{1}{p_c \frac{\partial^2 Q}{\partial M^2}}\right] = \frac{(+)}{(+)(-)} \leq 0$$

The remainder of our claims under Proposition 1 follow immediately from the preceding work. Specifically, we see that, *ceteris paribus*, managed pollination use M is decreasing in managed pollination price p_m (since $\frac{dM}{dp_m} \leq 0$), and the own-price elasticity η_{M,p_m} declines in magnitude with with managed pollination use M and with output price p_c.

Under the assumptions from Lemma 1 that Q is concave in M, output Q will exhibit diminishing returns to M. The more production is curved with respect to M (i.e., the greater the diminishing returns to M), the less elastic M will be with respect to managed pollination price (i.e., the less responsive M will be to increases in price).

If production is linear with respect to M (no diminishing returns, hence $\frac{\partial^2 Q}{\partial M^2} = 0$), then demand will be perfectly elastic.

QED.

Proof of Proposition 2:

Taking [#1M] as the starting point, the respective total derivative and expression for $\frac{dM}{dp_c}$ are as follows:

$$\frac{\partial Q}{\partial M} dp_c + p_c \frac{\partial^2 Q}{\partial M^2} dM = 0$$

$$\Rightarrow \frac{dM}{dp_c} = \frac{\frac{-\partial Q}{\partial M}}{p_c \frac{\partial^2 Q}{\partial M^2}}$$

The elasticity η_{M,p_c} of M with respect to output price p_c is then given by:

$$\eta_{M,p_c} = (\frac{p_c}{M}) \frac{dM}{dp_c} = \frac{\frac{-\partial Q}{\partial M}}{M \frac{\partial^2 Q}{\partial M^2}}$$

With the intermediate results in Lemma 1, we obtain:

$$\text{sign}\left[\frac{dM}{dp_c}\right] = \text{sign}\left[\frac{\frac{-\partial Q}{\partial M}}{p_c \frac{\partial^2 Q}{\partial M^2}}\right] = \frac{(-)(+)}{(+)(-)} \geq 0$$

The remainder of our claims under Proposition 2 follow immediately from the preceding work. Specifically, we see that, *ceteris paribus*, managed pollination use M is increasing in output price

p_c (since $\frac{dM}{dp_c} \geq 0$), while the elasticity of M with respect to output price p_c does not depend on p_c.

QED.

Proof of Proposition 3:

We establish Proposition 3 in a few steps. First, we establish conditions where $\frac{dM}{dA} > 0$ is likely to hold. Second, we show conditions where $\frac{dM}{dA}$ is likely to be small in magnitude. Third, we show conditions where $\frac{d^2M}{dA^2} < 0$. Finally, we show when all three conditions hold.

Using [#1M] as our starting point, we arrive at the following expression for $\frac{dM}{dA}$,

$$p_c \frac{\partial^2 Q}{\partial A \partial M} dA + p_c \frac{\partial^2 Q}{\partial M^2} dM$$

$$\Rightarrow \frac{dM}{dA} = \frac{-\frac{\partial^2 Q}{\partial A \partial M}}{\frac{\partial^2 Q}{\partial M^2}}$$

By Lemma 1, we established the claim that $\frac{\partial^2 Q}{\partial M^2} \leq 0$. Therefore, we must show that $\frac{\partial^2 Q}{\partial A \partial M} > 0$. Using the earlier expression for $\frac{\partial Q}{\partial M}$ as our starting point we have the following expression, which is positive assuming it is always optimal to use some amount of inputs:

$$\frac{\partial^2 Q}{\partial A \partial M} = \frac{\alpha_m \; \gamma_o (B)^{\frac{\rho - \rho_o}{\rho_o}}}{D^{\frac{\rho+1}{\rho}} \; M^{(\rho_o + 1)}} > 0$$

Thus, $\frac{dM}{dA} > 0$: the use of managed pollination M is increasing in total factor productivity and/or size of farms A.

Moreover, the more production is curved with respect to M (i.e., the greater the diminishing returns to M), the less responsive M is to increases in total factor productivity.

After putting the full result together and simplifying we get the next expression:

$$\frac{dM}{dA} = \frac{-DM^{\rho_o+1}}{A\left[B(M,W)^{\frac{\rho-\rho_o}{\rho_o}}\alpha_m\gamma_o(1+\rho) - D\left(\frac{\alpha_m(\rho-\rho_o)+(\rho_o+1)M^{\rho_o}B(M,W)}{B(M,W)}\right)\right]}.$$

By inspection of the last expression, we can establish conditions where $\frac{dM}{dA}$ is likely to be small in magnitude. We see the denominator term in square brackets is the same term that was used to establish Lemma 1; we showed in our proof of Lemma 1 that this term is likely to be negative.

Combined with conditions such that $\rho > 0$ (i.e., input groups complements) and $-1 < \rho_o < 0$ (i.e., pollination inputs substitutes), or $\rho > \rho_o > 0$ (i.e., input groups being strong complements and pollination inputs being weak complements), the above inequalities ensure the term in square brackets will be negative. These conditions also suggest that the entire term in square brackets may be a comparatively small magnitude.

If the term in square brackets is small in magnitude, then A and M^{ρ_o+1} are likely to play important roles in the magnitude of the effect. This further suggests that for farmers with high total factor productivity A, the change in M with respect to A will be small in magnitude, particularly if $A \geq M^{\rho_o+1}$.

The line of logic above suggests that, although we expect use of managed pollination M to be increasing in farm size, the magnitude of the effect may be comparatively small. This means that other factors may be playing a larger role in determining a farmers' managed pollination use.

Now we need to establish conditions where $\frac{d^2M}{dA^2} < 0$. Using the preceding expression as a starting point, we arrive the following result for this second derivative.

$$\frac{d^2M}{dA^2} = \; = \frac{DM^{\rho_o+1}}{A^2\left[B(M,W)^{\frac{\rho-\rho_o}{\rho_o}}\alpha_m\gamma_o(1+\rho) - D\left(\frac{\alpha_m(\rho-\rho_o)+(\rho_o+1)M^{\rho_o}B(M,W)}{B(M,W)}\right)\right]} \leq 0$$

The last inequality will hold because, once again, we see in the denominator term in square brackets is the same term that was used to establish Lemma 1; we showed in our proof of Lemma

1 that this term is likely to be negative. The relationship between A and M^{ρ_o+1} is also similar to that found for $\frac{dM}{dA}$, but now the rate of diminishing returns in M will increase with unit increases in A because A^2 is present.

From the preceding results and Lemma 1, it is apparent that when the conditions below hold, use of M will increase with A; $\frac{dM}{dA}$ will be small in magnitude; and M will be concave in A.

(1) Positive amounts of all inputs are used, especially capital and labor (this will ensure that $D > B^{\frac{\rho-\rho_o}{\rho_o}}$).

(2) Total factor productivity exceeds use of managed pollination: $A > M$.

(3) Input groups are strong complements (i.e., $\rho \geq 1$) and pollination inputs are on the spectrum of substitutes (i.e., $-1 < \rho_o < 0$); or, $\rho > \rho_o > 0$ (i.e. input groups are strong complements and pollination inputs are weak complements). (Either condition will ensure that $\rho - \rho_o > 0$ and $\rho_o + 1 > 0$, which ensures that the term in square brackets will be negative.)

QED.

Proof of Proposition 4:

We establish that $\frac{dM}{dK} > 0$.

Starting again from condition [#1M], we arrive at the following expression for $\frac{dM}{dK}$.

$$p_c \frac{\partial^2 Q}{\partial K \partial M} dK + p_c \frac{\partial^2 Q}{\partial M^2} = 0$$

$$\Rightarrow \frac{dM}{dK} = \frac{-\frac{\partial^2 Q}{\partial K \partial M}}{\frac{\partial^2 Q}{\partial M^2}}$$

Given Lemma 1, it suffices to show $\frac{\partial^2 Q}{\partial K \partial M} > 0$.

Using the same expression for the marginal product with respect to M as the starting point, we find the following expression for the second order derivative in the numerator of $\frac{dM}{dK}$.

$$\begin{aligned}\frac{\partial^2 Q}{\partial K \partial M} &= \frac{\partial}{\partial K}\left(\frac{QB(M,W)^{\frac{\rho-\rho_o}{\rho_o}}\alpha_m\gamma_o}{DM^{\rho_o+1}}\right) \\ &= \frac{\partial Q}{\partial M}\left(\frac{C(K,L)^{\frac{\rho-\rho_{kl}}{\rho_{kl}}}\gamma_{kl}\alpha_k(\rho+1)}{DK^{\rho_{kl}+1}}\right) > 0\end{aligned}$$

All the terms in this expression are positive, regardless of the sign of the ρ parameters. Therefore, we can conclude that $\frac{dM}{dK} > 0$.

QED.

Proof of Proposition 5:

Finally, we derive the expression for $\frac{dM}{dW}$ and establish plausible scenarios where $\frac{dM}{dW} < 0$.

$$p_c \frac{\partial^2 Q}{\partial W \partial M} dW + p_c \frac{\partial^2 Q}{\partial M^2} dM$$

$$\Rightarrow \frac{dM}{dW} = \frac{-\frac{\partial^2 Q}{\partial W \partial M}}{\frac{\partial^2 Q}{\partial M^2}}$$

To identify possible signing regimes, we solve and simplify the expression for $\frac{\partial^2 Q}{\partial W \partial M}$.

$$\frac{\partial^2 Q}{\partial W \partial M} = \frac{\partial}{\partial W}\left(\frac{QB(M,W)^{\frac{\rho-\rho_o}{\rho}}\alpha_m\gamma_o}{DM^{\rho_o+1}}\right)$$

$$= \left[\frac{-A\alpha_m\alpha_w\gamma_o B(M,W)^{\frac{\rho-2\rho_o}{\rho_o}}}{O_w^{\rho_o+1}\, D^{\frac{2\rho+1}{\rho}}\, M^{\rho_o+1}}\right]\left[(\rho-\rho_o)D \;-\; \gamma_o(\rho+1)B(M,W)^{\frac{\rho}{\rho_o}}\right]$$

Now putting the pieces together, we arrive at a signable expression for $\frac{dM}{dW}$.

$$\frac{dM}{dW} = \frac{-\left(\frac{\partial^2 Q}{\partial W \partial M}\right)}{\frac{\partial^2 Q}{\partial M^2}} = -\frac{\left[\frac{-A\alpha_m\alpha_w\gamma_o B(M,W)^{\frac{\rho-2\rho_o}{\rho_o}}}{O_w^{\rho_o+1}\, D^{\frac{2\rho+1}{\rho}}\, M^{\rho_o+1}}\right]\left[(\rho-\rho_o)D \;-\; \gamma_o(\rho+1)B(M,W)^{\frac{\rho}{\rho_o}}\right]}{\frac{\alpha_m\gamma_o B(M,W)^{\frac{\rho-\rho_o}{\rho_o}}Q}{\left(DM^{\rho_o+1}\right)^2}\left[B(M,W)^{\frac{\rho-\rho_o}{\rho_o}}\alpha_m\gamma_o(1+\rho) - D\left(\frac{\alpha_m(\rho-\rho_o)+(\rho_o+1)M^{\rho_o}B(M,W)}{B(M,W)}\right)\right]}$$

$$= \left[\frac{\alpha_w M^{\rho_o+1}}{B(M,W)O_w^{\rho_o+1}}\right]\frac{\left[(\rho-\rho_o)D \;-\; \gamma_o(\rho+1)B(M,W)^{\frac{\rho}{\rho_o}}\right]}{\left[B(M,W)^{\frac{\rho-\rho_o}{\rho_o}}\alpha_m\gamma_o(1+\rho) - D\left(\frac{\alpha_m(\rho-\rho_o)+(\rho_o+1)I_m^{\rho_o}B(M,W)}{B(M,W)}\right)\right]}$$

$$\Rightarrow \text{ sign}\left[\frac{dM}{dW}\right] = \text{sign}\left[\left[\frac{\alpha_w M^{\rho_o+1}}{B(M,W)O_w^{\rho_o+1}}\right]\frac{\left[(\rho-\rho_o)D \;-\; \gamma_o(\rho+1)B(M,W)^{\frac{\rho}{\rho_o}}\right]}{\left[B(M,W)^{\frac{\rho-\rho_o}{\rho_o}}\alpha_m\gamma_o(1+\rho) - D\left(\frac{\alpha_m(\rho-\rho_o)+(\rho_o+1)M^{\rho_o}B(M,W)}{B(M,W)}\right)\right]}\right]$$

$$= (+)\frac{(?)}{(-)} \quad \text{...by observation, Lemma 1, and preceding work.}$$

$$\Rightarrow \frac{dM}{dW} \leq 0 \iff (\rho-\rho_o)D \;-\; \gamma_o(\rho+1)B(M,W)^{\frac{\rho}{\rho_o}} \geq 0$$
$$\iff (\rho-\rho_o)D \geq \gamma_o(\rho+1)B(M,W)^{\frac{\rho}{\rho_o}}$$

These conditions are likely to hold when:

- $-1 < \rho_o < 0$ and $\rho > 0$ (pollination inputs are substitutes and input groups are complements); or
- $\rho > \rho_o$ (input groups being strong complements and pollination inputs being weak complements).

QED.

B Data

B.1 Background on Apple Production

Apples are a useful crop to study farmer pollination behavior. Apples are a widely produced and consumed commodity around the world[3] with high cultural value. From a pollination perspective, apples are also unique in the sense that wild pollinators have been shown to be much more effective at inducing fruit set[4] than honey bees are (Blitzer et al., 2016; Russo et al., 2017), with important implications for fruit quality and price received. This may be particularly important for farmers as high quality fruit receives a much better price on average compared to lower quality fruit which is often sold for processing (e.g., apple sauce and other products). A complexity in mapping pollination efficacy to yield, at least with modern approaches to apple production, is that farmers commonly engage in thinning (typically with a chemical agent) immediately after fruit set to encourage the plant to drop poorly pollinated fruit early and thus increase investment in remaining fruit. Another interesting aspect of pollination with apples is that apples are not considered to be a "honey crop" as nector from apples does not produce palatable honey, and this translates into higher pollination rental fees for apple farmers to mitigate against the fact that beekeepers do not gain forage resources to produce palatable honey from pollinating apples (Rucker, Thurman, and Burgett, 2012).

From a production perspective, apples have traditionally been grown in orchards with tall (6-8 meters), widely spaced (80-100 trees per hectare), and very long-lived trees (30-50 years or more). In recent decades, production strategies have started shifting towards more modern approaches where apples are grown in high density plantings on trellis systems, with shorter trees and very small spacings between rows and individual trees (Robinson et al., 2007, 2013). These high density systems bear little resemblance to the orchards of the past, with hopes of increasing yields and

[3]Today, China leads the world in global apple production with the US a fairly distant second (authors' calculations, FAOSTAT). Among states in the US, apple production is highest in Washington followed by New York.

[4]Fruit set is the biological process in which flowers become fruit and potential fruit size is determined (Mid Valley Agricultural Services, 2006). When seed formation is complete and well-distributed, the fruit is considered to be more appealing (e.g., consistent shape and fruit quantity/quality), which generally means a higher price is received by the farmer.

lowering labor costs. Some recommendations put optimal tree height at around 3-4 meters, orchard rows at 3-4 meters apart, and trees spaced within rows at as little as 0.7 meters, resulting in tree densities of 2-3,000 trees per hectare or more at the high end (Robinson et al., 2013).

B.2 Data and Data Sources

For our empirical analysis, we leverage rich, farm-level microdata from the 2007 USDA Agricultural Resource Management Survey (USDA-ARMS), which is designed to be nationally representative as well as representative at the level of a state. The USDA National Agricultural Statistics Service (USDA-NASS) imposes stringent conditions and restrictions on the use of its USDA-ARMS data, including strict security measures, data confidentiality, and the required use of provided replication weights. Qualified researchers at US universities or Government agencies can submit a formal request to the USDA Economic Research Service (ERS) and USDA-NASS to have access granted to USDA-ARMS data for specific research projects (USDA Economic Research Service (ERS), 2022). We access the USDA-ARMS data via the NORC Data Enclave.

The 2007 USDA-ARMS provides rich farm-level data from apple farmers in seven US states: California (CA), Michigan (MI), New York (NY), North Carolina (NC), Oregon (OR), Pennsylvania (PA), and Washington (WA). Useful data comes from the Phase III and Phase II surveys. Phase III covers operation-level data on land, production, and financial information. Phase II provides rich production data for a random operation and a random block of apples within the selected operation. Data at the random apple block level covers all the main aspects of production, including input use, costs and yield, for the 2007 production year, as well as honey bee rental data for the years 2006-2007. Although data on costs and on the binary choice to rent bees are available for 2006-2007, the quantity of honey bee colonies rented is only available for 2007. There are 1057 farmers who have sufficient responses for our research, which comprises the vast majority of the farmers sampled. In Figure B.1, we provide a barplot showing the distribution by state for the responses that comprise our base sample. Our observations span 7 states, 207 counties, and 466 zip codes.

To operationalize and enhance our research objectives, we also merge a variety of other data with the 2007 ARMS. We use data on almond production in California from USDA-NASS.

The 2007 USDA-ARMS did not request information on output prices. Thus, for apple output price, we use the state-level total utilized production price from USDA-NASS, which is a weighted average of fresh market and processed prices. We use the state-level apple price to derive revenue estimates and approximate profits at the random apple block level.

For distance measures, we compute Euclidean distances using R and we also employ the Google Distance Matrix API to derive road distances as alternative "share" variables in our instrument construction.

To derive relevant data on weather covariates that might affect yield, and collect credible proxy measures for landscape influence and local pollinator habitat (the closest proxy available for wild pollinator stocks), we use the closest[5] and most reliable coverage year from the USDA Cropland Data Layer (CDL) (Boryan et al. (2011)) for each state to construct a county-level mask of apple and tree-crop producing regions within each county. Using the resulting boundaries within each county for apple-specific and/or tree-crop-specific regions, as well as the county boundaries themselves, we further use the CDL to construct a variety of variables to characterize land cover heterogeneity, and also credible measures of pollinator habitat quality (Martins, Gonzalez, and Lechowicz, 2015; Park et al., 2015), including the proportion of land area in natural forest cover and the the proportion of land area in natural open cover. We define natural open cover as the proportion of apple-specific and/or tree-crop-specific areas within a county in any of the following cover

[5]Apples are difficult to identify with high accuracy, as are tree crops, therefore classification error in annual CDL layers induce potential for measurement error. Since tree crops are long-lived, there are unlikely to be large year-to-year changes in cover. Therefore we adopted the following rule to construct apple- and tree-crop specific spatial masks and gather other land cover information within county domains, and county-specific apple and tree-crop spatial domains: use the CDL crop mask data for the timepoint closest to 2007 as possible, but if the closest year to 2007 had low cover for apples and tree-crops, use the next closest year of the CDL that had substantially higher cover for apples and/or tree crops. The logic here is that if ARMS data imply that apple growers are present within a county, yet the CDL does not pick up apples or tree-crops, the closest year to 2007 that shows at least some spatial footprint for these crops is likely a more accurate spatial mapping of this agricultural activity than another year that might be closer to 2007. Since we cannot resolve sampled farm locations in space, these boundaries are designed to reflect that average conditions that apple growers face in their respective counties. Crops that are included in our tree crop definition include: apples, cherries, peaches, other tree crops, pears, prunes, plums, nectarines, and apricots; citrus and nut crops were excluded.

types: clover, wildflowers, shrubland, herbaceous wetlands, developed open space, and wetlands. We also employ the tree-crop-specific regions and county boundaries to gather monthly precipitation and temperature data from PRISM spanning January-November of the 2007 production (Daly et al., 2008).

The West Coast states in our data set are California (CA), Oregon (OR), and Washington (WA). The Midwest and East Coast states in our data set (which we refer to collectively as the 'Eastern' states) are Michigan (MI), New York (NY), North Carolina (NC), and Pennsylvania (PA).

B.3 Supplementary Tables and Figures

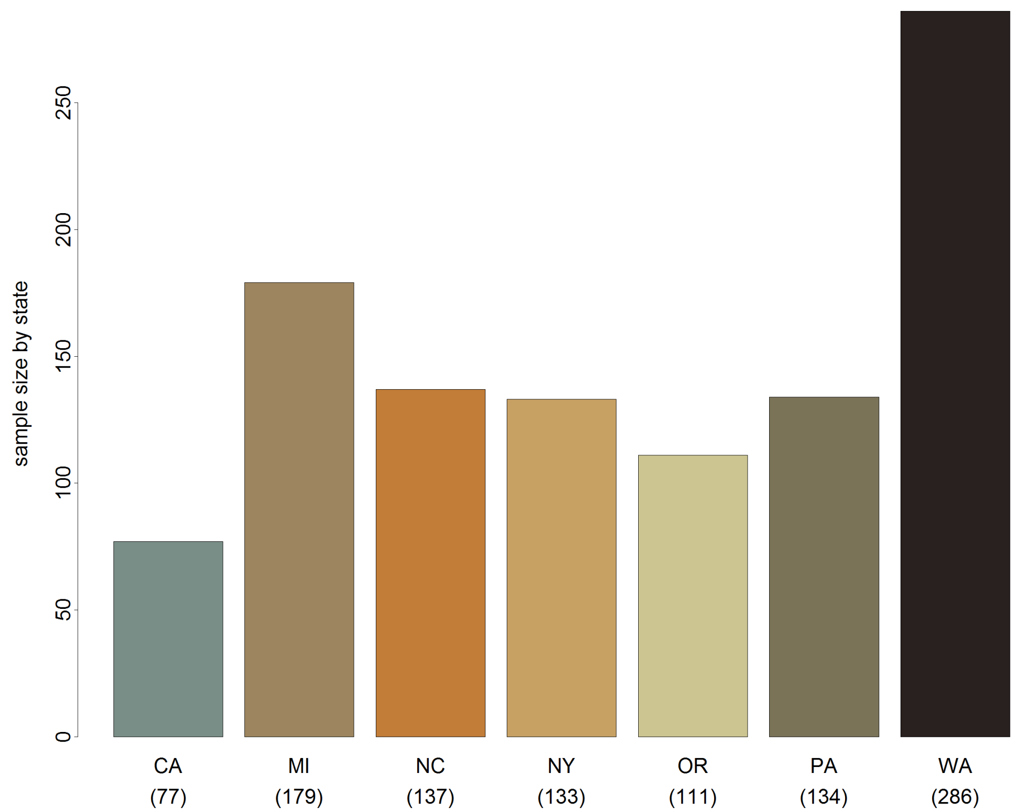

Figure B.1: Distribution of the sample apple farmers from 2007 USDA-ARMS that we employ in our analysis.

Table B.1: Weighted operation-level summary statistics.

Variable	Weighted Means All	West	East	Difference in Mean (West - East)
year apple operation was established	1981.46	1983.88	1978.39	5.49***
	(14.61)	(14.46)	(14.25)	(0.9)
	[1057]	[474]	[583]	
operator has some college (dummy)	0.72	0.76	0.67	0.1***
	(0.45)	(0.43)	(0.47)	(0.03)
	[779]	[337]	[442]	
total cropland acres	330.44	351.34	303.95	47.39
	(777.08)	(832.99)	(699.94)	(50.42)
	[1057]	[474]	[583]	
total apple acres	156.91	177.01	131.44	45.57*
	(397.39)	(511.24)	(161.48)	(24.37)
	[1057]	[474]	[583]	
total bearing apple acres	148.43	169.42	121.84	47.57**
	(393.78)	(509.13)	(149.06)	(24.18)
	[1057]	[474]	[583]	
total non-bearing apple acres	8.47	7.59	9.59	-2
	(19.6)	(19.52)	(19.67)	(1.22)
	[1057]	[474]	[583]	
total organic apple acres	1.01	1.79	0.01	1.79**
	(14.88)	(19.87)	(0.15)	(0.87)
	[1057]	[474]	[583]	
total number apple blocks	15.08	12.13	18.83	-6.7***
	(20.55)	(18.72)	(22.11)	(1.27)
	[1057]	[474]	[583]	
number of farm vehicles and implements	11.62	10.57	12.85	-2.28***
	(10.87)	(11.13)	(10.45)	(0.77)
	[772]	[333]	[439]	
difficulty with pollination (dummy)	0.2	0.2	0.19	0.01
	(0.4)	(0.4)	(0.39)	(0.03)
	[778]	[336]	[442]	
difficulty with labor (dummy)	0.65	0.68	0.61	0.08**
	(0.48)	(0.47)	(0.49)	(0.03)
	[778]	[336]	[442]	
own honey bees (dummy)	0.03	0.01	0.04	-0.03**
	(0.17)	(0.12)	(0.21)	(0.01)
	[779]	[337]	[442]	

Notes: Summary statistics are at the operation level for the selected apple operation. From left to right, columns are as follows: variable; mean for observations from all states (‘All’); mean for West Coast states (‘West’); mean for Midwest and East Coast states (‘East’); mean for West Coast minus mean for Midwest and East Coast states (‘West - East’). Below the means for each variable, standard deviations are in parentheses and sample sizes are in square brackets. Sample sizes may differ from respective full sample sizes because a farmer did not answer the question, or the question was not applicable. Below the differences in mean between West and East, bootstrapped standard errors from two-sample t-tests for the mean of the West Coast minus the mean for East Coast and Midwest states are in parentheses. Significance codes for two-sample t-tests: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table B.2: Weighted random block-level summary statistics focused on pollination.

Variable	Weighted Means All	West	East	Difference in Mean (West - East)
rented bees in 2007 (dummy)	0.74	0.81	0.64	0.17***
	(0.44)	(0.39)	(0.48)	(0.03)
	[1057]	[474]	[583]	
rented bees in 2006 (dummy)	0.74	0.8	0.66	0.14***
	(0.44)	(0.4)	(0.48)	(0.03)
	[1057]	[474]	[583]	
did not rent bees in 2006-2007 (dummy)	0.25	0.18	0.33	-0.14***
	(0.43)	(0.39)	(0.47)	(0.03)
	[1057]	[474]	[583]	
number of bee colonies rented in 2007	17.36	18.39	15.72	2.67
	(30.13)	(30.24)	(29.94)	(2.43)
	[601]	[313]	[288]	
number of bee colonies per acre in 2007	1.91	1.68	2.28	-0.6***
	(2.25)	(1.79)	(2.80)	(0.18)
	[601]	[313]	[288]	
bee rental fee in 2007 ($/colony)	40.54	37.58	45.26	-7.67***
	(13.4)	(10.73)	(15.71)	(1.10)
	[601]	[313]	[288]	
bee rental fee in 2006 ($/colony)	37.33	35.47	40.2	-4.73***
	(12.01)	(10.14)	(13.98)	(0.97)
	[627]	[312]	[315]	
bee rental cost in 2007 ($/acre)	76.92	63.63	98.15	-34.52***
	(102)	(85.17)	(121.43)	(8.38)
	[601]	[313]	[288]	
bee rental cost in 2006 ($/acre)	73.14	62.69	89.27	-26.58***
	(100.73)	(91.75)	(111.49)	(8.04)
	[627]	[312]	[315]	
bee rental proportion of total costs 2007	0.07	0.05	0.1	-0.05***
	(0.11)	(0.08)	(0.15)	(0.01)
	[601]	[313]	[288]	
bee rental proportion of total costs 2006	0.07	0.05	0.09	-0.04***
	(0.11)	(0.08)	(0.15)	(0.01)
	[627]	[312]	[315]	

Notes: Summary statistics are at the block level for the selected block. From left to right, columns are as follows: variable; mean for observations from all states (‘All’); mean for West Coast states (‘West’); mean for Midwest and East Coast states (‘East’); mean for West Coast minus mean for Midwest and East Coast states (‘West - East’). Below the means for each variable, standard deviations are in parentheses and sample sizes are in square brackets. Sample sizes may differ from respective full sample sizes because a farmer did not answer the question, or the question was not applicable. Below the differences in mean between West and East, bootstrapped standard errors from two-sample t-tests for the mean of the West Coast minus the mean for East Coast and Midwest states are in parentheses. Significance codes for two-sample t-tests: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table B.3: Weighted random block-level summary statistics focused on production.

Variable	Weighted Means All	West	East	Difference in Mean (West - East)
year block started production	1989.13	1990.72	1987.15	3.57***
	(11.96)	(11.41)	(12.43)	(0.74)
	[1037]	[467]	[570]	
number of apple trees	3512.68	5027.92	1584.66	3443.26***
	(12248.46)	(16008.70)	(2927.93)	(712.03)
	[1039]	[467]	[572]	
average age of trees	18.94	17.21	21.12	-3.92***
	(12.71)	(12.36)	(12.82)	(0.76)
	[1042]	[466]	[576]	
for fresh market (dummy)	0.84	0.93	0.71	0.22***
	(0.37)	(0.25)	(0.45)	(0.02)
	[1057]	[474]	[582]	
has federal crop insurance in 2007 (dummy)	0.62	0.58	0.68	-0.11***
	(0.48)	(0.49)	(0.47)	(0.03)
	[1057]	[474]	[583]	
yield (bushels/acre)	589.78	650.47	512.87	137.59***
	(422.30)	(455.71)	(361.90)	(26.69)
	[1057]	[474]	[583]	
approximate apple revenue ($) per acre	7225.63	9090.93	4861.83	4229.10***
	(6130.79)	(6845.39)	(3988.85)	(360.30)
	[1057]	[474]	[583]	
approximate apple revenue ($) per tree	56.40	69.86	39.28	30.58**
	(239.54)	(316.04)	(54.08)	(14.44)
	[1039]	[467]	[572]	
approximate apple cost ($) per acre	2285.37	2890.75	1518.19	1372.56***
	(3072.22)	(3893.56)	(1075.89)	(184.08)
	[1057]	[477]	[583]	
approximate apple cost ($) per tree	22	29.03	13.06	15.97*
	(148.10)	(197.31)	(14.89)	(8.98)
	[1039]	[467]	[572]	
approximate profit ($) per acre	4940.26	6200.18	3343.63	2856.54***
	(6416.43)	(7582.36)	(4002.39)	(383.93)
	1057	474	583	
approximate profit ($) per tree	34.40	40.83	26.22	14.61**
	(104.75)	(131.47)	(53.38)	(6.32)
	[1039]	[467]	[572]	

Notes: Summary statistics are at the block level for the selected block. From left to right, columns are as follows: variable; mean for observations from all states (‘All’); mean for West Coast states (‘West’); mean for Midwest and East Coast states (‘East’); mean for West Coast minus mean for Midwest and East Coast states (‘West - East’). Below the means for each variable, standard deviations are in parentheses and sample sizes are in square brackets. Sample sizes may differ from respective full sample sizes because a farmer did not answer the question, or the question was not applicable. Below the differences in mean between West and East, bootstrapped standard errors from two-sample t-tests for the mean of the West Coast minus the mean for East Coast and Midwest states are in parentheses. Revenues and profits are approximate and employ state-level average output prices obtained from USDA-NASS. Significance codes for two-sample t-tests: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table B.4: Summary statistics for apple prices, instruments, and land cover variables.

Variable	Weighted Means All	Weighted Means West	Weighted Means East	Difference in Mean (West - East)
Apple output prices				
total utilized production price ($/lb) in 2007	0.23	0.30	0.16	0.14**
	(0.09)	(0.04)	(0.05)	(0.034)
	[7]	[3]	[4]	
total utilized production price ($/lb) in 2006-2007	0.20	0.26	0.16	0.10***
	(0.07)	(0.05)	(0.04)	(0.025)
	[14]	[6]	[8]	
Instrumental variables				
zip code distance (Euclidean) to Fresno County, CA (km)	2490.48	876.92	3478.72	-2601.80***
	(1313.74)	(383.12)	(341.78)	(35.121)
	[466]	[177]	[289]	
zip code distance (Euclidean) to Fresno County, CA (km) X	1.893E9	0.666E9	2.644E9	-1.977E9***
total almond acres in CA	(0.998E9)	(0.291E9)	(0.260E9)	(0.189E9)
	[932]	[354]	[578]	
Land cover variables				
natural forest cover (county proportion)	0.50	0.44	0.53	-0.09**
	(0.22)	(0.26)	(0.19)	(0.035)
	[207]	[71]	[136]	
natural open cover (county proportion)	0.17	0.31	0.10	0.21***
	(0.16)	(0.19)	(0.06)	(0.023)
	[207]	[71]	[136]	

Notes: From left to right, columns are as follows: variable; mean for observations from all states ('All'); mean for West Coast states ('West'); mean for Midwest and East Coast states ('East'); mean for West Coast minus mean for Midwest and East Coast states ('West - East'). Below the means for each variable, standard deviations are in parentheses and sample sizes are in square brackets. Below the differences in mean between West and East are standard errors in parentheses from two-sample t-tests for the mean of the West Coast minus the mean for East Coast and Midwest states. Apple prices are at the state level, instrumental variables are at the zip code level, and land cover variables are at the county level. There are 7 states, 207 counties, and 466 zip codes. Significance codes for two-sample t-tests: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table B.5: Summary statistics for weather variables.

Variable	Weighted Means All	West	East	Difference in Mean (West - East)
mean temperature (C), winter 2006-2007	1.86	5.37	0.04	5.33***
	(4.64)	(3.27)	(4.18)	(0.374)
	[414]	[142]	[272]	
mean precipitation (mm), winter 2006-2007	2.84	3.62	2.44	1.18***
	(1.98)	(3.14)	(0.62)	(0.266)
	[414]	[142]	[272]	
mean temperature (C), spring 2006-2007	13.47	13.10	13.66	-0.56***
	(2.51)	(2.84)	(2.31)	(0.276)
	[414]	[142]	[272]	
mean precipitation (mm), spring 2006-2007	2.31	1.62	2.67	-1.05***
	(0.97)	(1.09)	(0.67)	(0.10)
	[414]	[142]	[272]	
mean temperature (C), summer 2006-2007	20.83	20.05	21.23	-1.18***
	(2.59)	(3.20)	(2.09)	(0.297)
	[414]	[142]	[272]	
mean precipitation (mm), summer 2006-2007	2.41	0.36	3.48	-3.12***
	(1.81)	(0.43)	(1.23)	(0.083)
	[414]	[142]	[272]	
mean temperature (C), fall 2006-2007	10.07	10.01	10.10	-0.09
	(2.77)	(3.31)	(2.45)	(0.315)
	[414]	[142]	[272]	
mean precipitation (mm), fall 2006-2007	3.10	2.87	3.23	-0.36
	(1.89)	(2.95)	(0.94)	(0.254)
	[414]	[142]	[272]	

Notes: From left to right, columns are as follows: variable; mean for observations from all states ('All'); mean for West Coast states ('West'); mean for Midwest and East Coast states ('East'); mean for West Coast minus mean for Midwest and East Coast states ('West - East'). Below the means for each variable, standard deviations are in parentheses and sample sizes are in square brackets. Below the differences in mean between West and East are standard errors in parentheses from two-sample t-tests for the mean of the West Coast minus the mean for East Coast and Midwest states. Weather variables are at the county level. There are 207 unique counties observed. Significance codes for two-sample t-tests: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

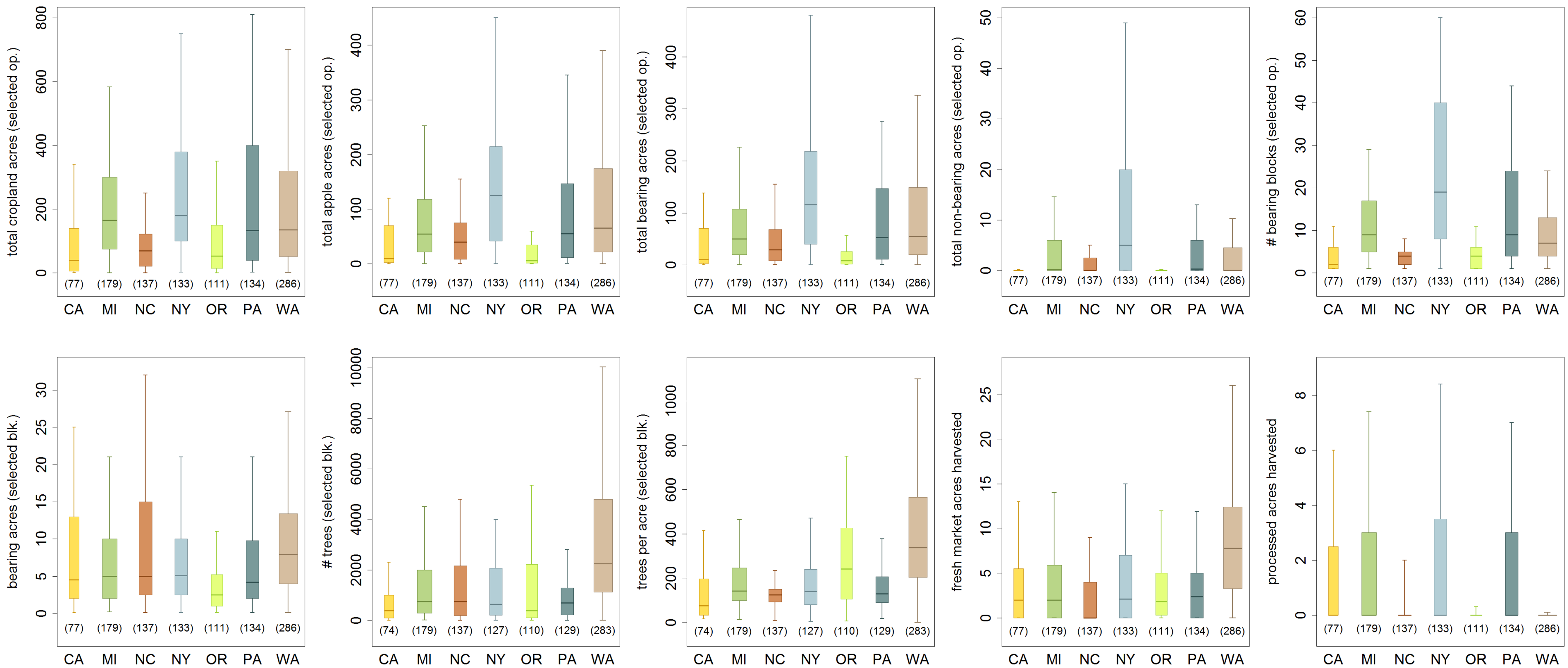

Figure B.2: Weighted boxplots by state for a suite of farm and orchard characteristics. Numbers in parentheses next to state abbreviations indicate the respective sample size per boxplot.

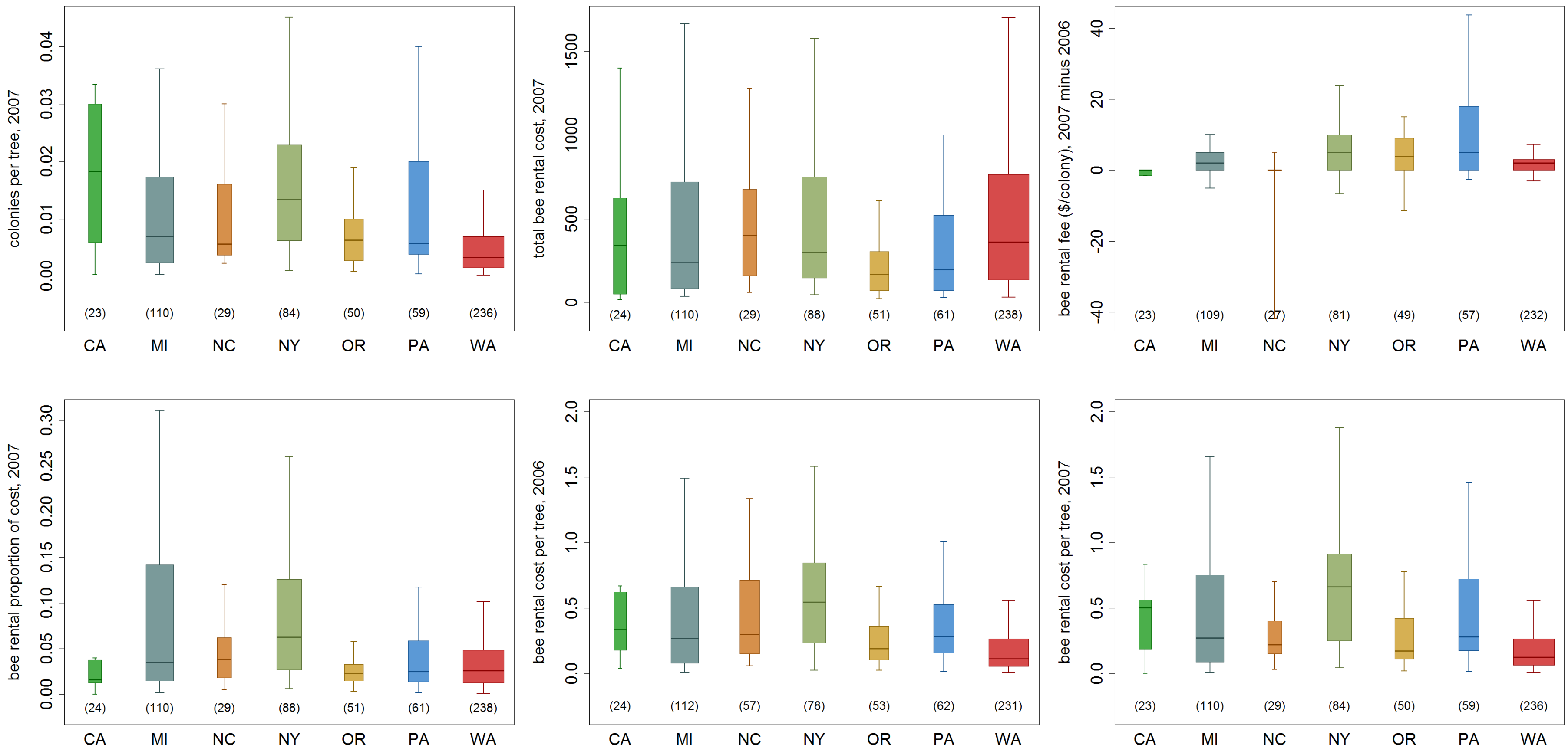

Figure B.3: Weighted boxplots by state capturing: the number of colonies rented per tree in 2007; total honey bee colonies rented per tree; total bee rental costs in 2007; dollars per colony rental fee in 2007 minus dollars per colony rental fee in 2006; bee rental proportion of block level costs in 2007; bee rental costs per tree in 2006; bee rental costs per tree in 2007. All variables comprised random block-level variation. Numbers in parentheses indicate the sample size per state.

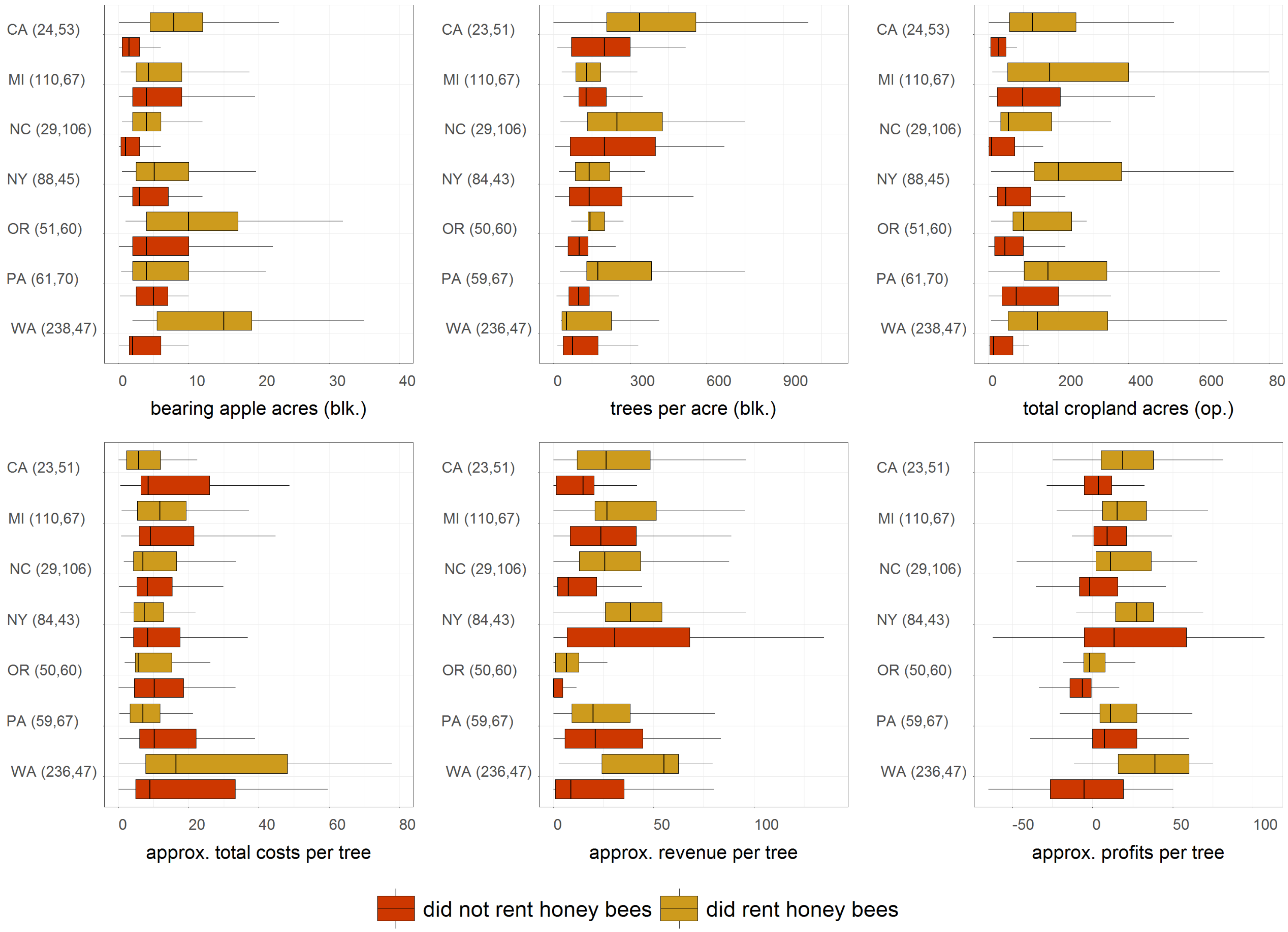

Figure B.4: Weighted boxplots by state and if an apple farmer rented honey bees for: bearing apple acres (block-level); trees per acre (block level); total cropland acreage (operation level); and approximate block level production costs per tree; revenue per tree; and profits per tree. Numbers in parentheses indicate the sample size per state, and the choice to rent honey bees or not. For example, in the top left plot the notation for the bottom rows, WA (238, 47), indicates of the 286 apple farmers sampled in Washington State, 238 reported renting honey bees, while 47 reported not renting honey bees.

C Empirical Analysis of the Choice to Rent Bees

C.1 Econometric Model

For the binary choice to use pollination service markets (in our case, renting honey bees), we seek to study associations between the discrete choice to use managed pollination and important farm characteristics, state variables, and parameters, each of which addresses particular elements of Propositions 2-5. To explore these associations, we estimate the following logit fixed effects regression using methods developed by Bergé (2018) for efficiently estimating maximum likelihood models with large numbers of fixed effects:

$$\Pr(y_{isct} = 1) = \mathbf{x'_{isct}}\beta + \lambda_s + \sigma_t + \varepsilon_{it}, \tag{C.1}$$

where y_{isct} is a dummy variable for farmer i in state s and county c renting honey bees in year t; $\mathbf{x'_{isct}}$ is a vector of covariates, including measures of farm production scale (to proxy for total factor productivity), output price, and remotely sensed measures of natural open cover and natural forest cover (to proxy for wild pollinator stocks and landscape heterogeneity); λ_s and σ_t are state and time fixed effects; and ε_{it} is the error term.

Bergé (2018) employs fixed-point algorithms in combination with the concentrated likelihood function to efficiently estimate maximum likelihood models with large numbers of fixed effects. The concentrated log-likelihood function, $g(\beta) = l(y_{isct}, \mathbf{x'_{isct}}\beta + \lambda_s(\beta) + \sigma_t(\beta))$, treats fixed effects as functions of the parameters of interest and leverages the observation that each fixed effect represents a partition of the data to improve efficiency.[6]

To address Propositions 2-5, we employ several strategies. For Proposition 2, we use state-level total utilized production price (a weighted average of the fresh market and processed prices). For

[6] We also tried using the bias-corrected discrete choice estimator from Stammann, Heiss, and McFadden (2016) and Czarnowske and Stammann (2019), which corrects for the incidental parameter bias in discrete choice fixed effects estimation. Unfortunately, the available code for these bias-corrected discrete choice estimators currently does not permit the use of sampling weights, and USDA-NASS does not permit us to retrieve results from the NORC Data Enclave unless their provided replication weights are applied. We therefore cannot make use of these estimators. Nevertheless, our qualitative results using the bias-corrected discrete choice estimator without sampling weights are similar to the results we report using the fixed effects discrete choice estimator from Bergé (2018) with the provided replication weights.

Proposition 3, we study second-degree polynomials in the apple bearing acreage of the selected block, the total number of bearing apple blocks (in the selected operation), as well as trees per acre and age of apple trees for the selected apple block. For Proposition 4, we use the number of farm vehicles and farm implements available for use on the selected block. For Proposition 5, we study second-degree polynomials in county-level remotely sensed measures of natural open cover and natural forest cover to proxy for wild pollinator stocks and landscape heterogeneity.[7] We also include measures of management (including pest training and scouting effort), state level fixed effects, and a dummy for 2007 (equivalent to year fixed effects in this two-period panel).

In all specifications, we use multi-way clustered standard errors that are clustered at both the state and county levels. Clustering at the county level alone is problematic as there are counties with only one farmer sampled, and this will not address potential spatial correlation within states. Clustering at the state level alone is also problematic as there are too few clusters with only seven states (Cameron and Miller, 2015). Clustering at both the state and county levels therefore offers a conservative approach to standard error estimation.

We also estimate analogous logit fixed effects regressions of the binary choice to *never* rent honey bees during 2006-2007. For these regressions, output price (which is at the state level) and the dummy for 2007 are omitted as there is no temporal variation in our dependent variable.

[7] In our exploration of CDL-based land cover measures, we study measures summarized at different buffer sizes of 500 meters, 1000 meters, 3000 meters, and at the county level. We apply buffers to apple-specific polygons and tree-crop-specific polygons within counties where apple farmers are sampled, and also summarize these cover types at the county level (without buffers). Similar qualitative results are generally apparent using these alternative measures.

C.2 Supplementary Tables and Figures

Table C.1: Point estimates from weighted logit regression of the binary choice to rent honey bees.

	Dependent variable is probability of renting honey bees					
	(1)	(2)	(3)	(4)	(5)	(6)
apple bearing acres	0.040**	0.049**	0.009	0.009	0.086***	0.101***
	(0.0176)	(0.0193)	(0.0162)	(0.0208)	(0.0123)	(0.0131)
apple bearing acres, squared	−0.00008**	−0.00010***	0.000124685	0.00020	−0.00017***	−0.00020***
	(0.000037)	(0.000038)	(0.000155)	(0.000333)	(0.0000297)	(0.000030)
total bearing apple blocks	0.12***	0.12***	0.09***	0.07***	0.27***	0.36***
	(0.027)	(0.040)	(0.020)	(0.025)	(0.021)	(0.043)
total bearing apple blocks, squared	−0.0010***	−0.0008***	−0.0008***	−0.0004***	−0.0021***	−0.0031***
	(0.00020)	(0.00032)	(0.00011)	(0.00012)	(0.00031)	(0.00031)
trees per acre	0.00006	0.00106	0.00317	0.00805***	0.00074	−0.00137
	(0.001922)	(0.002263)	(0.003391)	(0.002013)	(0.001588)	(0.001608)
trees per acre, squared	0.000003	0.000003	−0.000002	−0.000006*	0.000002	0.000004
	(0.0000023)	(0.0000030)	(0.0000044)	(0.0000031)	(0.0000026)	(0.0000039)
average age of trees	0.02	0.07	0.05	0.10	0.06	0.13
	(0.041)	(0.050)	(0.060)	(0.090)	(0.064)	(0.122)
average age of trees, squared	−0.0006	−0.0016**	−0.0009	−0.0018*	−0.0017	−0.0031
	(0.00054)	(0.00071)	(0.00074)	(0.00111)	(0.00200)	(0.00332)
total utilized production price ($/pound)	6.22***	6.70**	13.72***	18.61***	6.38**	4.35
	(2.394)	(3.400)	(2.113)	(2.969)	(2.751)	(2.706)
natural forest cover	−0.50	2.77	−0.07	−2.03	0.01	1.20
	(3.865)	(3.328)	(4.147)	(8.538)	(5.078)	(2.854)
natural forest cover, squared	1.03	−2.07	0.10	0.26	1.20	3.19
	(3.388)	(3.463)	(3.168)	(6.869)	(5.2817)	(4.656)
natural open cover	8.62*	11.27*	21.74*	17.83	11.72	245.0***
	(4.532)	(6.006)	(12.448)	(17.126)	(7.449)	(7.901)
natural open cover, squared	−7.37**	−9.01*	−64.26	−60.0	−9.96**	−20.61***
	(3.139)	(5.029)	(40.250)	(55.187)	(3.985)	(4.228)
number farm vehicles and implements		0.01		0.03		−0.01
		(0.02)		(0.03)		(0.02)
deliberate pest scouting (dummy)	0.76**	0.70*	0.01	−0.31	1.43***	1.27***
	(0.339)	(0.373)	(0.226)	(0.680)	(0.057)	(0.135)
recent pest training (dummy)	0.55	0.60**	1.06***	1.07***	−0.39	−0.20
	(0.335)	(0.306)	(0.179)	(0.204)	(0.427)	(0.526)
year 2007 (dummy)	−0.23*	−0.27	−0.32***	−0.41***	−0.26***	−0.14
	(0.119)	(0.193)	(0.079)	(0.113)	(0.078)	(0.084)
State FE	Y	Y	Y	Y	Y	Y
Sample	All	All	East	East	West	West
Standard Errors	C,S	C,S	C,S	C,S	C,S	C,S
Pseudo R^2	0.33	0.37	0.20	0.30	0.50	0.54
# Observations	2056	1514	1136	858	920	656

Notes: Table presents point estimates from weighted logit regression of the binary choice to rent honey bees on block- and operation-level characteristics, output prices, physical capital, and land cover measures. For land cover measures, we use remotely sensed measures of natural open cover and natural forest cover proportions at the county level. Standard errors are clustered at both the state (S) and county (C) levels, and are in parentheses. Significance codes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table C.2: Average partial effects from weighted logit regression of the binary choice to never rent honey bees.

	Dependent variable is probability of never renting honey bees					
	(1)	(2)	(3)	(4)	(5)	(6)
apple bearing acres	-0.007^{***}	-0.008^{***}	-0.003^{*}	-0.004^{*}	-0.008^{*}	-0.008^{**}
	(0.0024)	(0.0027)	(0.0018)	(0.0020)	(0.0042)	(0.0033)
total bearing apple blocks	-0.02^{***}	-0.02^{***}	-0.02^{***}	-0.01	-0.03^{***}	-0.05^{***}
	(0.005)	(0.006)	(0.006)	(0.007)	(0.007)	(0.009)
trees per acre	-0.0001	-0.0002	-0.0004	-0.0010^{***}	-0.0002	-0.0001
	(0.00029)	(0.00030)	(0.00063)	(0.00024)	(0.00020)	(0.00016)
average age of trees	0.001	0.001	-0.002	-0.002	0.003	0.0002
	(0.0035)	(0.0040)	(0.0058)	(0.0081)	(0.0016)	(0.0036)
natural forest cover	-0.13	-0.16	-0.09	0.19	-0.18	-0.47^{*}
	(0.175)	(0.206)	(0.163)	(0.285)	(0.260)	(0.246)
natural open cover	-1.11^{**}	-1.19^{*}	-1.79^{***}	-1.00	-0.74	-1.20^{**}
	(0.563)	(0.623)	(0.623)	(1.008)	(0.587)	(0.530)
number farm vehicles and implements		-0.006^{*}		-0.012^{*}		-0.001
		(0.0038)		(0.0056)		(0.0044)
deliberate pest scouting (dummy)	-0.12^{***}	-0.10^{**}	-0.01	0.05	-0.15^{***}	-0.12^{***}
	(0.037)	(0.045)	(0.033)	(0.122)	(0.042)	(0.046)
recent pest training (dummy)	-0.10^{*}	-0.10^{**}	-0.22^{***}	-0.19^{***}	0.02	0.01
	(0.055)	(0.041)	(0.048)	(0.029)	(0.035)	(0.045)
State FE	Y	Y	Y	Y	Y	Y
Sample	All	All	East	East	West	West
Standard Errors	C,S	C,S	C,S	C,S	C,S	C,S
Pseudo R^2	0.35	0.41	0.21	0.32	0.57	0.62
# Observations	1028	757	1136	858	920	656

Notes: Table presents average partial effects from weighted logit regression of the binary choice to never rent honey bees on block- and operation-level characteristics, physical capital, and land cover measures. For land cover measures, we use remotely sensed measures of natural open cover and natural forest cover proportions at the county level. Standard errors are clustered at both the state (S) and county (C) levels, and are in parentheses. Significance codes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table C.3: Point estimates from weighted logit regression of the binary choice to never rent honey bees.

	Dependent variable is probability of never renting honey bees					
	(1)	(2)	(3)	(4)	(5)	(6)
apple bearing acres	-0.05^{**}	-0.06^{***}	-0.02	-0.02	-0.09^{***}	-0.09^{***}
	(0.019)	(0.023)	(0.012)	(0.016)	(0.021)	(0.021)
apple bearing acres, squared	0.0001^{***}	0.0001^{***}	-0.0001	-0.0003	0.0002^{***}	0.0002^{***}
	(0.00004)	(0.00004)	(0.00013)	(0.00024)	(0.00005)	(0.00005)
total bearing apple blocks	-0.13^{***}	-0.14^{**}	-0.10^{***}	-0.05	-0.53	-0.53
	(0.032)	(0.056)	(0.020)	(0.043)	(0.337)	(0.337)
total bearing apple blocks, squared	0.0011^{***}	0.0008^{*}	0.0009^{***}	-0.0005	0.0019	0.0019
	(0.00024)	(0.00050)	(0.00013)	(0.00091)	(0.02175)	(0.02175)
trees per acre	0.0003	-0.0003	-0.0020	-0.0069^{***}	0.0009	0.0009
	(0.00227)	(0.00236)	(0.00425)	(0.00152)	(0.00195)	(0.00195)
trees per acre, squared	-0.000003	-0.000004	0.000001	0.000004^{*}	-0.000003	-0.000003
	(0.0000028)	(0.0000031)	(0.0000055)	(0.0000024)	(0.0000041)	(0.0000041)
average age of trees	-0.006	-0.040	-0.045	-0.081	-0.059	-0.059
	(0.0465)	(0.0525)	(0.0572)	(0.0870)	(0.1746)	(0.1746)
average age of trees, squared	0.0004	0.0013^{*}	0.0008	0.0017	0.0018	0.0018
	(0.00064)	(0.00077)	(0.00075)	(0.00113)	(0.00449)	(0.00449)
natural forest cover	3.57	-0.06	0.84	2.58	2.78	2.78
	(3.406)	(3.100)	(4.821)	(10.110)	(2.153)	(2.153)
natural forest cover, squared	-4.43	-1.08	-1.32	-1.48	-8.04^{**}	-8.04^{**}
	(3.048)	(3.131)	(3.884)	(8.731)	(3.401)	(3.401)
natural open cover	-10.52^{**}	-12.19^{**}	-22.91^{**}	-17.42	-32.23^{***}	-32.23^{***}
	(5.210)	(6.175)	(11.068)	(14.939)	(9.867)	(9.867)
natural open cover, squared	9.50^{**}	10.29^{*}	68.48^{*}	57.43	28.92^{***}	28.92^{***}
	(4.022)	(5.464)	(35.324)	(47.673)	(8.126)	(8.126)
number farm vehicles and implements		-0.05^{**}		-0.06^{**}	-0.01	-0.01
		(0.022)		(0.031)	(0.046)	(0.046)
deliberate pest scouting (dummy)	-0.79^{**}	-0.69^{*}	-0.06	0.32	-1.33^{***}	-1.33^{***}
	(0.353)	(0.409)	(0.172)	(0.692)	(0.331)	(0.331)
recent pest training (dummy)	-0.66^{**}	-0.72^{***}	-1.14^{***}	-1.11^{***}	0.09	0.09
	(0.281)	(0.276)	(0.096)	(0.189)	(0.484)	(0.484)
State FE	Y	Y	Y	Y	Y	Y
Sample	All	All	East	East	West	West
Standard Errors	C,S	C,S	C,S	C,S	C,S	C,S
Pseudo R^2	0.35	0.41	0.21	0.32	0.62	0.62
# Observations	1028	757	1136	858	656	656

Notes: Table presents point estimates from weighted logit regression of the binary choice to never rent honey bees on block- and operation-level characteristics, physical capital, and land cover measures. For land cover measures, we use remotely sensed measures of natural open cover and natural forest cover proportions at the county level. Standard errors are clustered at both the state (S) and county (C) levels, and are in parentheses. Significance codes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

D Elasticity of Demand for Managed Pollination

D.1 Econometric Model of Honey Bee Demand

To estimate the own-price demand elasticity for managed pollination use, we estimate the honey bee demand function for apple farmers using an instrumental variables strategy to address the endogeneity problem that arises because observed equilibrium prices and quantities are simultaneously determined in the supply-and-demand system (Manski, 1995; Goldberger, 1991; Angrist, Graddy, and Imbens, 2000; Lin, 2011). Our first-stage equation is given by:

$$p_{m,isct} = \delta_1 Z_{sct} + \mathbf{X'}_{\mathbf{isct}} \alpha + \mathscr{D}_l + \gamma_t + \nu_{isct}, \tag{D.1}$$

where $p_{m,isct}$ is the price of managed pollination services (here the honey bee rental fee per colony) faced by farm i, in state s, county c, and year t; Z_{sct} is our shift-share instrument for price (Goldsmith-Pinkham, Sorkin, and Swift, 2020; Borusyak, Hull, and Jaravel, 2022), which interacts almond acreage in California with distance measures to California; $\mathbf{X}'_{\mathbf{isct}}$ is comprised of farm and orchard characteristics; $\mathscr{D}_l$ and γ_t are dummies for location and time, respectively; and ν_{isct} is the first-stage error term.

Our second-stage managed pollination demand equation is given by:

$$M_{isct} = \beta_1 \widehat{p}_{m,isct} + \mathbf{X'}_{\mathbf{isct}} \theta + \mathscr{D}_l + \gamma_t + \varepsilon_{isct}, \tag{D.2}$$

where M_{isct} is the number of bee colonies demanded by farm i in state s, county c, and time t; $\widehat{p}_{m,isct}$ is the predicted price from the first stage; and ε_{isct} is the second-stage error term.

We use a linear demand specification (e.g., rather than a log-log specification) for several reasons. First, since we find in Proposition 1 that elasticity η_{p_m} is not constant, but instead depends on managed pollination price p_m and managed pollination M, we do not want a functional form (such as a log-log specification) that would assume that the elasticity is constant. Second, we expect the demand function to intersect the price axis and therefore that there is a price above which farmers will not rent bees. Indeed, we observe farmer-years in the data for which quantity rented is zero, which is only possible if the demand function intersects the price axis.[8] Third, we also expect the demand function to intersect the quantity axis and therefore that farmers would rent a finite number of bee colonies even if renting bees were free.

To identify β_1, the coefficient on managed pollination price in the second-stage managed pollination demand equation (D.2), we use a shift-share instrument Z_{sct} for price.[9] As a shift-share instrument, Z_{sct} is an interaction term, $Z_{sct} = d_{sc} * s_t$, of something that is ideally an exogenous, time-invariant "share" d_{sc} with an exogenous, time-varying "shift" s_t, which in combination predict the endogenous variable of interest while not violating the exclusion restriction (Goldsmith-Pinkham, Sorkin, and Swift, 2020; Borusyak, Hull, and Jaravel, 2022). In particular, our shift-share instrument Z_{sct} for price interacts distance from the zip codes where apple farms are located to the approximate center of almond production in Fresno County, California (our share d_{sc}) with the total almond acreage in California (our shift s_t). As explained in detail in Section 6.1 of the paper, this instrument accounts for demand for honey bees from almond growers in California (where most of US almond production takes place) and its well-documented effects on the availability and distribution of honey bees (e.g., see Ward, Whyte, and James (2010); Rucker, Thurman, and Burgett (2012); Goodrich, Williams, and Goodhue (2019)).

Although the 2007 USDA-ARMS collected data on the binary choice to rent honey bees, and the costs to rent honey bees per colony over 2006-2007, data on the quantity of honey bees demanded in 2006 is not available except in the instance a farmer reported not renting bees (in which case we know quantity rented is zero). We therefore use three different subsamples of data to estimate honey bee demand.

The first subsample uses data from 2007 only. For growers who rented honey bees in 2007, we use the grower's rental fee for the price. If a grower reported not renting bees in 2007 (i.e., quantity rented is zero), they did not report a bee rental fee; to deal with this we use the state average rental

[8]The inverse hyperbolic sine transformation was also studied as an alternative approach that permits observations of zero in price or quantity, but this is not our preferred approach as model diagnostics (e.g., lower adjusted R^2) were worse than the linear specifications and respective elasticity point estimates are implausibly large (ranging between -4 and -2).

[9]The indices reflect that we cannot resolve farm locations precisely and are restricted to zip code, county, and state level locations

fee in 2007 for the price. For specifications that use data from 2007 only, the instrument Z_{sct} is the Euclidean distance from the centroids of zip code units of farm locations to the centroid of Fresno, County California. As a consequence, the variation in the instrument Z_{sct} in specifications (1) and (2) comes from variation in the Euclidean distance from the centroids of zip codes where farms are located to the centroid of almond production in Fresno, County California.

The second subsample uses an unbalanced panel over 2006-2007 that includes all observations from 2007, as well as growers who reported not renting bees in 2006, for whom we know the number of colonies rented in 2006 is zero (thereby eliminating the need for quantity imputation), and for whom we use the state average rental fee in 2006 as the price in 2006. For this second subsample, the instrument is the full shift-share version of Z_{sct}, and is therefore the interaction between the distance from zip code centroids where farms are located to the centroid of Fresno, County California and the total almond acres in California in year t.

The third subsample uses a balanced panel that includes all growers in the data for both 2006 and 2007. If a grower rented in both years, we impute the number of colonies rented in 2007 to be the number of colonies rented in 2006. If a grower rented bees in 2006 but not in 2007, we impute the quantity rented in 2006 by multiplying the acreage of the selected apple block with the state level average honey bee stocking density in 2007. We use the state average rental fee in 2006 as the price in 2006. For farmers who never rented bees over 2006-2007, their quantity rented is zero in both years and we use the state average rental fee in the respective year for the price. For this subsample, the instrument is once again the full shift-share version of Z_{sct}, and is therefore the interaction between the distance from zip code centroids where farms are located to the centroid of Fresno, County California and the total almond acres in California in year t.

There are several challenges to estimating demand elasticities in this setting, which stem primarily from having limited temporal and cross-sectional variation (particularly for the number of honey bee colonies rented) and aggregate-level variation in our distance measures (the share in our instrument). Particular limitations come with state and county fixed effects, which eliminate too much information to provide a useful model, as well as high sensitivity to even including a moderate number of covariates. We can, however, effectively use time fixed effects in our specifications that use data from both 2006 and 2007 (in this two-period panel, including a dummy for 2007 is equivalent to including year fixed effects). The issue with state fixed effects is that they remove too much variation and effectively eliminate the strength of our instrument; the remaining variation in distance to Fresno, CA (the share in our shift-share instrument) after inclusion of state fixed effects is only weakly correlated with price. Analogous challenges with the limitations of fixed effects in demand estimation have been encountered by Bruno and Jessoe (2021), who find that year fixed effects remove excessive variation when estimating groundwater demand elasticities in California. In Section 6.2 of our paper, we use methods from Petterson, Seim, and Shapiro (2023) to assess the bounds and robustness of our results. Over a longer panel with an ability to resolve farm-specific locations, we anticipate that our instrument would retain sufficient strength to withstand location- and even farm-level fixed effects. For this reason, we focus on specifications with a limited set of covariates and state level dummies that do not overload the model. As such, our specifications limit us to testing Propositions 1 and 3 in our models. On these scores, our findings are consistent with these propositions.

Standard error estimation comes with the same aforementioned challenges as with our binary choice models. For this reason, standard errors are either clustered at the county level, or multi-way clustered at state and county levels.

D.2 Methods for Bounding the Elasticity of Demand

As a further step to study honey bee demand, we employ new methods developed by Petterson, Seim, and Shapiro (2023) for studying bounds on elasticities. Petterson, Seim, and Shapiro (2023) show that economic intuitions about the plausible size of demand shocks can be informative about and help bound the elasticity of demand. In particular, Petterson, Seim, and Shapiro (2023) develop methods for determining demand elasticities that are consistent with a given bound on the plausible size of demand shocks.

Following the approach taken in Figures 3 and 4 of Petterson, Seim, and Shapiro (2023), we determine the range of demand elasticities consistent with a range of demand shocks. In particular, following the approach taken in Figure 3 of Petterson, Seim, and Shapiro (2023), for each given bound B on the plausible size of shocks to demand, we make a plot with first differenced price on the x-axis and first differenced quantity on the y-axis. Around each point, we construct a dotted interval of radius the length of the given demand shock bound B, thus indicating, for each first differenced price, the plausible range of first differenced quantity given positive or negative demand shocks of magnitude up to the given bound B. In first differences, a demand function is a line through the origin with with non-positive slope. A demand function consistent with a bound B on the maximum absolute value of the demand shock is a downward sloping line that passes through the origin as well as through all of the dotted intervals. We determine the set of all demand functions consistent with a bound B, depict the set on the plot with a shaded region, and use the set to calculate the corresponding bound on demand elasticities (when evaluated at mean price and quantity) consistent with the demand shock bound B. We repeat the approach taken in Figure 3 of Petterson, Seim, and Shapiro (2023) for different demand shock bounds B, obtain the range of demand elasticities consistent with each demand shock bound B, and then, following the approach taken in Figure 4 of Petterson, Seim, and Shapiro (2023), we plot the range of demand elasticities consistent with the range of demand shocks.

We use two alternative demand shocks for our demand shock bound B. One demand shock we use are shocks to demand in the year 2007. Estimates of demand shocks from 2007 are obtained from point estimates on the dummy for 2007 in specifications (3), (4), and (5) of our econometric model of honey bee demand in Table 2 of our paper. Shocks in 2007 are potentially informative as this was a year of diesel price shocks and widespread onset of colony collapse disorder (CCD) (vanEngelsdorp and Meixner, 2010). For the second demand shock, following Appendix C of Petterson, Seim, and Shapiro (2023), we use the absolute value of the differenced quantity for observations for which the differenced price is equal to zero.

As the new methods developed by Petterson, Seim, and Shapiro (2023) involve taking first differences in price and quantity, we apply them to subsamples of our data that have observations in both 2006 and 2007. One such subsample is the subset of 430 farmers from the sample used in specifications (3)-(4) in Table 2 of our paper who have data in both 2006 and 2007. The advantage of this sample is that it does not require any quantity imputation; the disadvantage is that this sample does not have much variation in the absolute value of the differenced quantity for observations for which the differenced price is equal to zero, and therefore is not amenable to using our second demand shock. Another subsample of our data that has observations in both 2006 and 2007 is the balanced panel of all 1057 farmers that we use in specification (5) in Table 2 of our paper.

Large magnitude changes in quantity among those who rent bees in one year but not the other in our data, and the lack of changes in quantity for growers who rent in 2006 resulting from our quantity imputation in specification (5), present challenges for cleanly applying results from Petterson, Seim, and Shapiro (2023), however. To address make the effort more informative, and also to comply with USDA NASS conditions and restrictions on data confidentiality, we apply data smoothing methods to the first differenced quantity to average out outliers, and view our respective analysis as a simulation-focused approach on a pseudo-version of the actual data, which can be used to study the plausibility of our findings. To smooth the first differenced quantity, we average the first differenced quantity over observations that share similar values of first differenced price.[10]

D.3 Supplementary Tables and Figures

[10]We smooth the first differenced quantity when calculating the second demand shock, the absolute value of the differenced quantity for observations for which the differenced price is equal to zero, as well. For the results we report, we discretize the first differenced price using the empirical distribution of first differenced price, and then average the first differenced quantity for each bin of first differenced price. Based on the empirical distribution of first differenced price, we use the following bounds for the bins: -61, -40, -25, -20, -15, -10, -7.5, -5, -2.5, -0.5, 0.5, 2.5, 5, 7.5, 10, 12.5, 15, 20, 25, 40, 60. Our results are fairly robust to different smoothing approaches. For example, results are very similar when we apply smoothing approaches that re-assign each grower's first differenced quantity with the mean the first difference quantity over observations whose first difference price is within a given neighborhood of that grower's first difference price.

Table D.1: First-stage results for honey bee demand elasticity estimation (weighted).

	Dependent variable is the honey bee rental fee per colony				
	(1)	(2)	(3)	(4)	(5)
IV: zip code distance to Fresno County, CA (km)	0.000003***	0.000004***			
	(0.000001)	(0.0000003)			
IV: zip code distance to Fresno County, CA (km) X total almond acres in CA			0.000000004***	0.000000005***	0.000000004***
			(0.0000000008)	(0.0000000005)	(0.0000000004)
apple bearing acres	0.002	-0.009	-0.02	-0.02	0.008
	(0.0290)	(0.0326)	(0.0289)	(0.0337)	(0.019)
apple bearing acres, squared	0.00004	0.00005	0.00007	0.00008	0.0000097
	(0.00006)	(0.00008)	(0.000064)	(0.000079)	(0.0000463)
deliberate pest scouting (dummy)		2.64***		1.65*	1.98
		(0.470)		(0.532)	(0.5346)
CA (dummy)	-13.03***	-11.76***	-13.46***	-12.54***	-12.93***
	(4.940)	(0.407)	(4.314)	(0.503)	(0.408)
MI (dummy)	-4.09***	-5.07***	-2.86***	-4.26***	-3.83***
	(1.460)	(0.541)	(1.306)	(0.568)	(0.396)
PA (dummy)		-4.11***		-5.01***	-5.54***
		(0.813)		(0.828)	(0.610)
year 2007 (dummy)			4.89***	4.31**	3.59***
			(0.880)	(1.069)	(0.854)
Constant	35.15***	32.31***	31.01***	29.44***	30.27***
	(1.533)	(0.601)	(1.393)	(0.759)	(0.678)
Data included in sample:					
All observations from 2007	Y	Y	Y	Y	Y
Growers who did not rent in 2006	N	N	Y	Y	Y
Growers who rented bees in 2006	N	N	N	N	Y
Standard Errors	C	C,S	C	C,S	C,S
First-stage F-statistic, F_{kp}	22.42	118.03	21.41	87.04	128.39
DWH	1.80	2.08	2.24	3.16	3.19
Adjusted R^2	0.25	0.27	0.30	0.31	0.27
# Observations	1057	1057	1487	1487	2114

Notes: Table presents the first-stage results for the honey bee demand IV estimation. Specifications (1) and (2) use data from 2007 only. For growers who rented honey bees in 2007, we use the grower's rental fee for the price. If a grower reported not renting bees in 2007, they did not report a bee rental fee; to deal with this we use the state average rental fee in 2007 for the price. Specifications (3) and (4) employ an unbalanced panel over 2006-2007 that includes all observations from 2007, as well as growers who reported not renting bees in 2006, for whom we use the state average rental fee in 2006 as the price in 2006. For specification (5), which is a balanced panel that includes all growers in the data for both 2006 and 2007, we use the state average rental fee in 2006 as the price in 2006. For specifications (1) and (2), the instrument Z_{sct} is the Euclidean distance from the centroids of zip code units of farm locations to the centroid of Fresno, County California. For specifications (3), (4), and (5) the instrument Z_{sct} is the interaction between the distance from zip code centroids where farms are located to the centroid of Fresno, County California and the total almond acres in California in year t. Standard errors are clustered at the county (C) and/or state (S) level, and are in parentheses. Significance codes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table D.2: Honey bee demand own-price elasticity estimation, OLS results (weighted).

Dependent variable is the number of honey bee colonies rented	(1')	(2')	(3')	(4')	(5')
honey bee rental fee ($/colony)	-0.009	-0.008	-0.007	-0.007	-0.009
	(0.0652)	(0.0697)	(0.060)	(0.071)	(0.074)
apple bearing acres	1.15***	1.13***	1.00***	0.99***	1.15***
	(0.153)	(0.202)	(0.170)	(0.235)	(0.199)
apple bearing acres, squared	-0.0009**	-0.0008	-0.0005	-0.0005	-0.0009*
	(0.00035)	(0.00044)	(0.00039)	(0.00052)	(0.00043)
deliberate pest scouting (dummy)		2.97***		1.48**	2.63***
		(0.807)		(0.632)	(0.679)
CA (dummy)	-4.14	-3.92**	-2.61	-3.01**	-3.84**
	(2.961)	(1.199)	(2.372)	(1.002)	(1.320)
MI (dummy)	2.03	1.31**	1.35	0.57	1.47**
	(2.616)	(0.537)	(2.137)	(0.725)	(0.479)
PA (dummy)		-5.41***		-4.23***	-5.09***
		(0.457)		(0.702)	(0.295)
year 2007 (dummy)			9.58***	7.87***	0.018
			(0.992)	(1.681)	(0.445)
Constant	1.58	-0.23	-5.83	-5.70	-0.12
	(3.600)	(4.140)	(2.878)	(3.303)	(3.954)
Elasticity at mean	-0.03	-0.03	-0.027	-0.027	-0.10
Data included in sample:					
All observations from 2007	Y	Y	Y	Y	Y
Growers who did not rent in 2006	N	N	Y	Y	Y
Growers who rented bees in 2006	N	N	N	N	Y
Standard Errors	C	C,S	C	C,S	C,S
Adjusted R^2	0.53	0.54	0.52	0.52	0.54
# Observations	1057	1057	1487	1487	2114

Notes: Table presents OLS results for honey bee demand (weighted). Although the 2007 USDA-ARMS collected data on the binary choice to rent honey bees, and the costs to rent honey bees per colony over 2006-2007, data on the quantity of honey bees demanded in 2006 is not available except in the instance a farmer reported not renting bees (in which case we know quantity rented is zero). Specifications (1) and (2) use data from 2007 only. For growers who rented honey bees in 2007, we use the grower's rental fee for the price. If a grower reported not renting bees in 2007, they did not report a bee rental fee; to deal with this we use the state average rental fee in 2007 for the price in 2007. Specifications (3) and (4) employ an unbalanced panel over 2006-2007 that includes all observations from 2007, as well as growers who reported not renting bees in 2006, for whom we know the number of colonies rented in 2006 is zero (thereby eliminating the need for quantity imputation), and for whom we use the state average rental fee in 2006 as the price in 2006. Specification (5) is a balanced panel that includes all growers in the data for both 2006 and 2007: if the grower rented bees in 2007 and 2006, we impute the number of colonies rented in 2007 to be the number of colonies rented in 2006; if the grower rented bees in 2006 but not in 2007, we impute the quantity rented in 2006 by multiplying the acreage of the selected apple block with the state level average honey bee stocking density in 2007. We use the state average rental fee in 2006 as the price in 2006. Elasticity is evaluated at the mean price and quantity in the data for the respective sample of data. Standard errors are clustered at the county (C) and/or state (S) level, and are in parentheses. Significance codes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

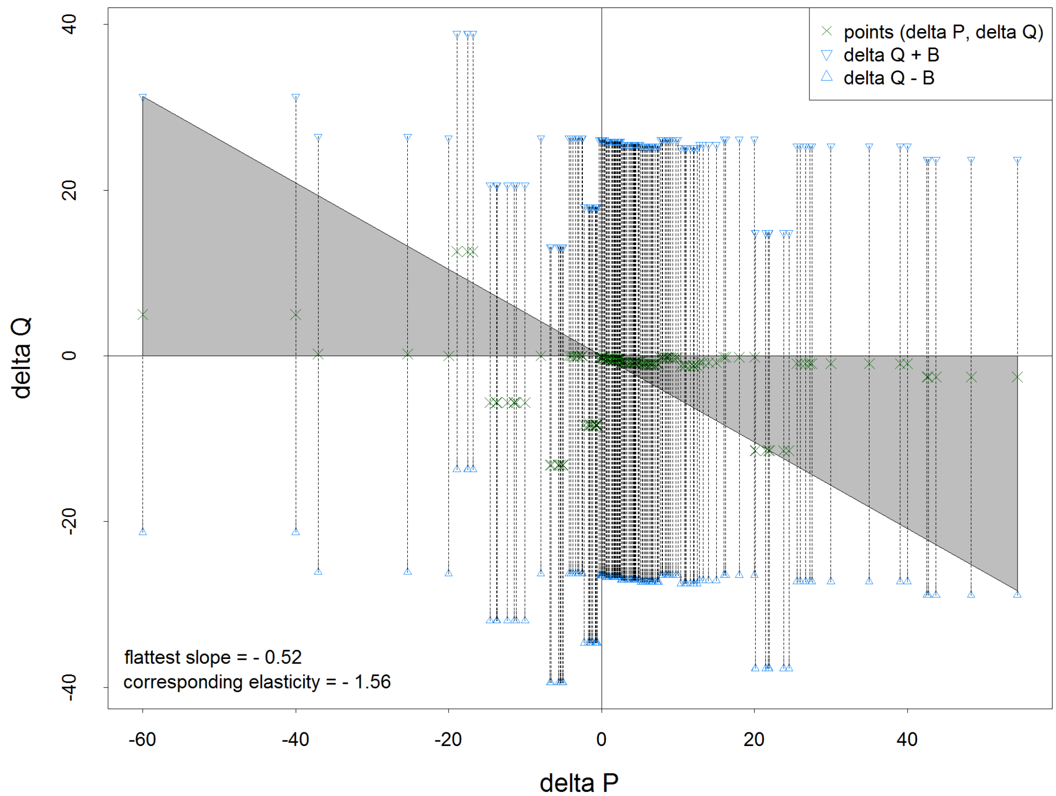

Figure D.1: Figure illustrates the construction of bounds on the honey bee demand elasticity from using the absolute value of the smoothed differenced quantity for observations for which the differenced price is equal to zero as our demand shock and a demand shock bound B of 26.3, which is twice the maximum absolute value of the smoothed differenced quantity for observations for which the differenced price is equal to zero. The subsample is the balanced panel of all 1057 farmers that we use in specification (5) in Table 2. The cross-hatches depict a scatterplot of the first differenced price on the x-axis and smoothed first differenced quantity on the y-axis. The dotted interval around each cross-hatch as radius of $B = 26.3$. The shaded region depicts all demand functions consistent with an upper bound of $B = 26.3$ on the maximum absolute value of the demand shock. These are the downward-sloping lines that pass through the origin and through all of the dotted intervals. The implied bound on the slope is -0.52 and the corresponding bound on demand elasticity (when evaluated at mean price and quantity) is -1.56.

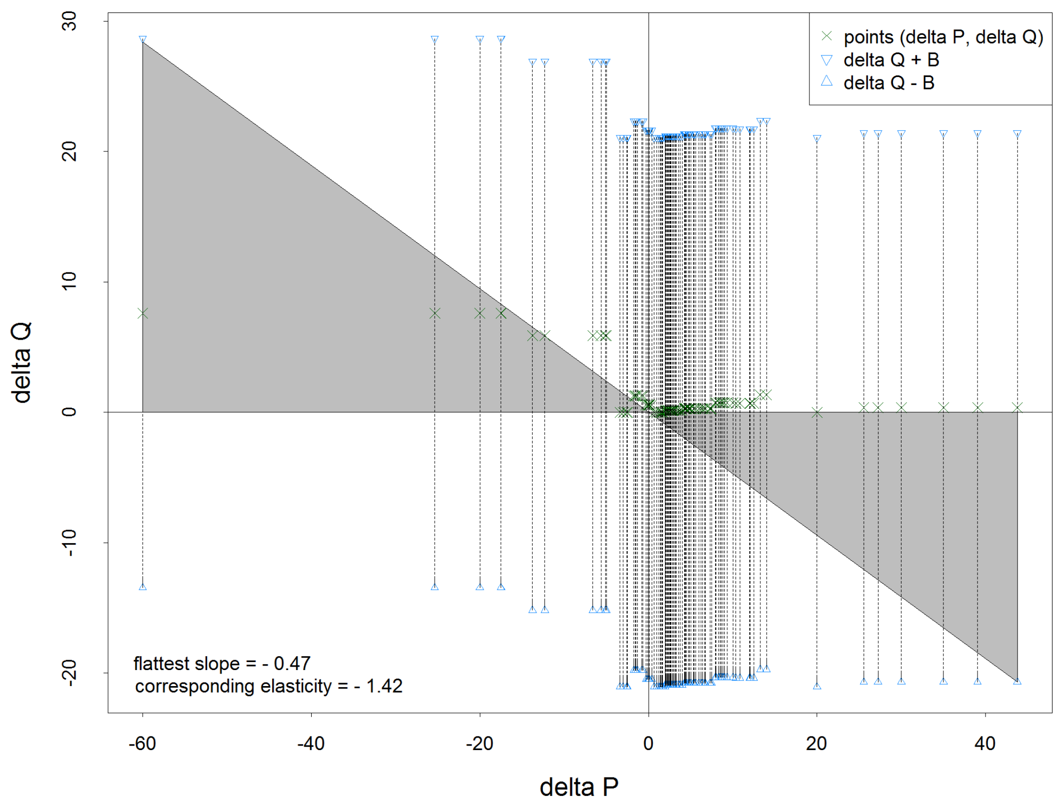

Figure D.2: Figure illustrates the construction of bounds on the honey bee demand elasticity (when evaluated at mean price and quantity) from using shocks to demand in the year 2007 as our demand shock and a demand shock bound B of 21.05, which is twice the maximum estimated shock in 2007. The subsample is the the subset of 430 farmers from the sample used in specifications (3)-(4) in Table 2 who have data in both 2006 and 2007. The cross-hatches depict a scatterplot of the first differenced price on the x-axis and smoothed first differenced quantity on the y-axis. The dotted interval around each cross-hatch as radius of $B = 21.05$. The shaded region depicts all demand functions consistent with an upper bound of $B = 21.05$ on the maximum absolute value of the demand shock. These are the downward-sloping lines that pass through the origin and through all of the dotted intervals. The implied bound on the slope is -0.47 and the corresponding bound on demand elasticity (when evaluated at mean price and quantity) is -1.42.

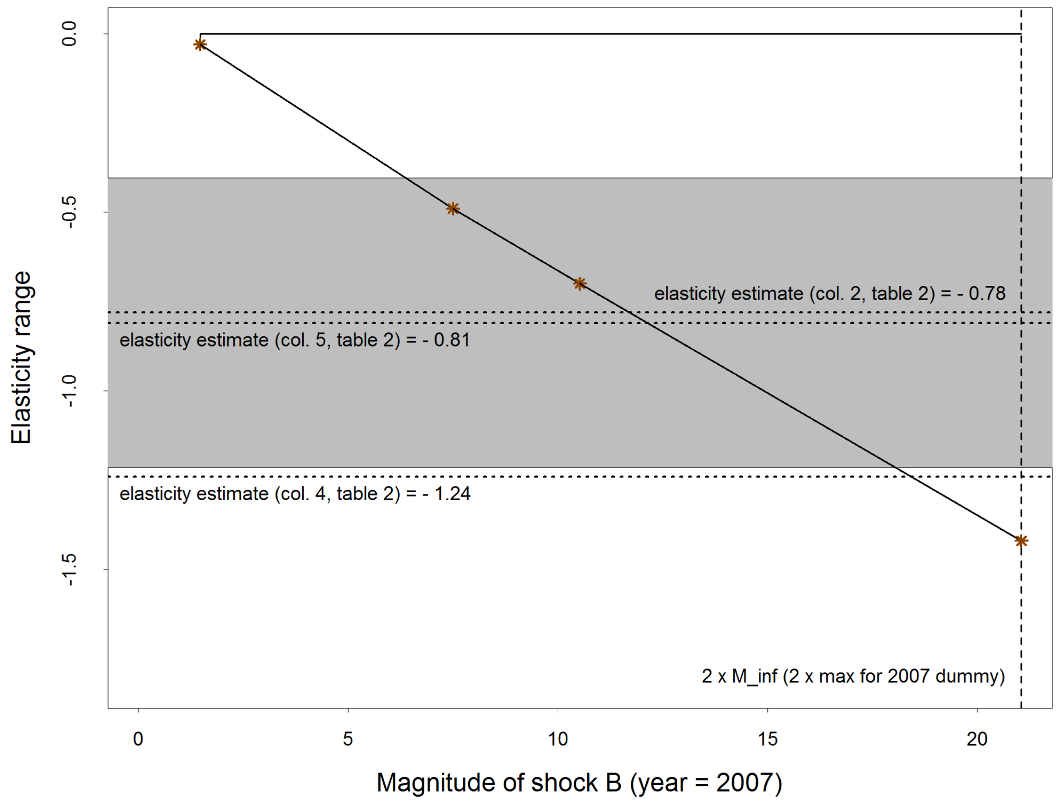

Figure D.3: Figure plots the range of honey bee demand elasticities (when evaluated at mean price and quantity) that are consistent with bounds on the plausible size of shocks to demand in the year 2007 ranging from the minimum estimated shock in 2007 to twice the maximum estimated shock in 2007. The subsample is the subset of 430 farmers from the sample used in specifications (3)-(4) in Table 2 who have data in both 2006 and 2007. Estimates of shocks from 2007 are obtained from point estimates on the dummy for 2007 included in specifications (3), (4), and (5) in Table 2. The dashed vertical line is at twice the maximum estimated shock in 2007. The horizontal dotted lines depict the point estimates for the demand elasticity from specifications (2), (4), and (5) in Table 2, and the shaded region depicts the associated 95% confidence interval for the estimated demand elasticity from specification (5) in Table 2.

E Relationship Between Yield, Profits, and Honey Bee Use

E.1 Econometric Model

To study how yield and profits vary with managed pollination use, we rely on optimal binscatter estimators from Cattaneo et al. (2021) to estimate the following semi-parametric function:

$$y_{isct} = \mu(x_{isct}) + \mathbf{w}'_{\mathbf{isct}}\gamma + \varepsilon_{isct}, \tag{E.1}$$

using the following p-th order polynomial, q-times continuously differentiable, covariate-adjusted least-squares extended binscatter estimator:

$$\hat{\mu}^{(v)}(x_{isct}) = \hat{\mathbf{b}}_q^{(v)}(x_{isct})'\hat{\beta}, \quad \begin{bmatrix} \hat{\beta} \\ \hat{\gamma} \end{bmatrix} = \arg\min_{\beta,\gamma} \sum_{i=1}^{n} (y_{isct} - \hat{\mathbf{b}}_q^{(v)}(x_{isct})'\beta - \mathbf{w}_{\mathbf{isct}}{}'\gamma)^2, \;\; 0 \leq v, \;\; q \leq p, \tag{E.2}$$

where y_{isct} is either block-level profits or yield per acre for farmer i in state s, county c in year t; x_{isct} is the number of honey bee colonies per acre employed for pollination at the random apple block level for farmer i in state s, county c in year t; $\mu(x_{isct})$ is some unknown function of x_{isct}; p is the polynomial order used; v is the desired derivative to be approximated; q is the level of smoothness imposed across bins; $\mathbf{w}_{\mathbf{isct}}$ is a vector of covariates; and where the model can be expanded to include dummies for fixed effects. The condition $q \leq p$ requires that a least squares p-th order polynomial is constructed in each bin, v refers to the derivative of interest, and $\hat{\mathbf{b}}_q^{(v)}$ reflects a spline basis to approximate $\mu^{(v)}(\cdot)$ (B-splines are employed). The goal is to recover the unknown function $\mu(x_{isct})$, which in our case is the functional relationship between outcome (profits or yield) y_{isct} and honey bee colonies per acre x_{isct}.

We use the optimal binscatter estimators from Cattaneo et al. (2021). Cattaneo et al. (2021) are the first to formalize the ad-hoc binscatter approaches that have long been in use and they offer a number of innovations to improve upon prior practice,[11] including: formalization within the frame-

[11] Common approaches have involved binning the regressor of interest to some ad-hoc number of bins, and then plotting the mean of the response variable of interest within bins, while also applying residualization in the Frisch-Waugh sense.

work of semi-linear least squares approximations (since $\mu(x)$ is non-linear, Frisch-Waugh logic is not applicable); data-driven selection of the number and placement of bins (using a data-driven rule of thumb approach, or integrated mean squared error (IMSE) criteria); smoothness restrictions using splines (between bins); smooth confidence bands; estimation of response function derivatives; and parametric (e.g., concavity) and shape restriction (e.g., monotonicity) t-tests.[12]

We use data-driven rule of thumb bin selection and provide results for both quantile-spaced and equally spaced bins. Compelling arguments can be made for quantile-spaced bins perhaps being preferable. In our setting, equally spaced bins produced more distinct value per bin, and therefore have some appealing properties, as one element of the methods of Cattaneo et al. (2021) that drives the data-driven approach is having sufficient degrees of freedom in regards to the number of distinct values within bins. Integrated mean squared error (IMSE)-based bin selection rules also have appealing properties over rule of thumb data-driven selection as potentially being more adaptable, but they also have a higher bar for the number of distinct values within bins. For these reasons, we opt for the data-driven rule of thumb approach and provide results for both quantile- and equally spaced bins for robustness.

At present, these methods are not adapted to address endogeneity beyond controlling for fixed effects and covariate adjustment. Hence estimation of $\mu(\cdot)$ remains vulnerable to endogeneity concerns. In our setting, while honey bees are an input, and hence potentially endogenous, honey bee colonies rented per acre is arguably exogenous to yield and profit, since honey bees are rented during the bloom period, several months before yield and profits are realized. Honey bee colonies per acre is also likely exogenous to yield and profits because farmers are unable to precisely control insect pollination. Beyond placement of pallets of colonies around orchards, little can be done to ensure honey bees pollinate crops as desired. Indeed, any effort to directly engage with bees can result in significant harm from bee stings (a reality clearly demonstrated by beekeepers' extensive use of protective clothing and implements to prevent injury when working with their colonies). Moreover, rented honey bees may end up foraging outside of the locations they are brought to

[12] Implementation software for R is known as *bingsreg*.

pollinate crops. For example, findings from McArt et al. (2017) suggest that honey bees utilized for apple pollination in New York may conduct a significant amount of foraging in non-crop areas. Since farmers have limited control over these potentially dangerous insects with minds of their own, they are unable to precisely control insect pollination. There are therefore good reasons to view insect pollination as an exogenous process on some level, hence omitted variable bias may be the larger issue and we employ a highly relevant set of controls in $\mathbf{w_{isct}}$ to mitigate this concern (including monthly weather variation, farm labor, landscape cover measures, and farm scale measures). Although some amount of bias may be present in our estimations, in our view, the novelty of the opportunity to make *any* estimation of the functional relationship between realized farm-level production outcomes and managed pollination outweighs concerns of bias – particularly given the innovative tools provided by Cattaneo et al. (2021). To assess the stability of the relationships between colonies per acre and production outcomes, we also assess these relationships using standard linear fixed effects models to assess the stability of parametric point estimates.[13]

For further robustness checks, we estimate linear fixed effects models and employ second-order polynomials in honey bee colonies per acre and other covariates. For the linear fixed effects regressions, standard error estimation comes with the same aforementioned challenges as with our binary choice models. For this reason, we use either Huber-White heteroskedasticity-robust standard errors, or standard errors that are multi-way clustered at the county and state levels.

E.2 Supplementary Tables and Figures

[13]We have explored a two-stage least squares set-up, using prior year pollination prices and/or the aforementioned shift-share instrument, as plausible instruments for honey bee colonies per acre. Unfortunately, none of these instruments predict colonies per acre with sufficient strength to make associated tests useful.

Table E.1: Weighted linear fixed effects regressions of yield.

	Dependent variable is block-level apple yield (bushels/acre)			
	(1)	(2)	(3)	(4)
Honey bee colonies per acre				
honey bee colonies per acre	110.031***	110.031***	157.534**	67.836*
	(19.494)	(23.817)	(36.068)	(22.308)
honey bee colonies per acre, squared	-10.996***	-10.996**	-15.143	-5.787
	(3.136)	(3.365)	(8.264)	(3.818)
Measures of production scale				
apple bearing acres	-1.037	-1.037	-1.818**	-1.377
	(1.024)	(0.744)	(0.199)	(3.434)
apple bearing acres, squared	0.0003	0.0003	0.003*	-0.005
	(0.003)	(0.002)	(0.001)	(0.052)
trees per acre	0.158	0.158	-0.169	0.682
	(0.143)	(0.313)	(0.280)	(0.307)
trees per acre, squared	-0.0002**	-0.0002	-0.00002	-0.001
	(0.0001)	(0.0002)	(0.0002)	(0.0004)
average age of trees	10.219***	10.219***	8.658**	15.724*
	(2.279)	(2.726)	(1.380)	(6.451)
average age of trees, squared	-0.117***	-0.117***	-0.084*	-0.154*
	(0.033)	(0.021)	(0.022)	(0.057)
Labor input variables				
pruning/thinning hours	-0.059***	-0.059***	-0.039	-0.116
	(0.018)	(0.012)	(0.016)	(0.050)
harvesting hours	0.073***	0.073***	0.044**	0.117*
	(0.015)	(0.010)	(0.010)	(0.043)
land prep and machine hours	0.283**	0.283***	0.287**	0.645**
	(0.110)	(0.053)	(0.037)	(0.166)
pest scouting hours	0.127***	0.127***	0.137**	0.651*
	(0.043)	(0.014)	(0.022)	(0.215)
part time and seasonal hours	0.007	0.007*	0.0003	0.006
	(0.012)	(0.003)	(0.003)	(0.027)
full time hours	0.075***	0.075**	0.066	0.065**
	(0.024)	(0.028)	(0.036)	(0.014)
Land cover variables				
natural forest cover	-219.613	-219.613	-281.935**	759.463***
	(186.591)	(220.246)	(62.627)	(112.357)
natural forest cover, squared	185.792	185.792	300.656	-1,020.779***
	(249.471)	(403.285)	(142.217)	(22.319)
natural open cover	-160.877	-160.877	-1,258.030	-3,550.023*
	(525.951)	(683.298)	(1,141.463)	(1,280.448)
natural open cover, squared	-567.918	-567.918	408.736	8,171.569
	(595.006)	(759.403)	(1,298.848)	(4,152.272)
Weather variables				
Jan. average precipitation (mm)	-42.780	-42.780	6.025	-33.643
	(29.738)	(36.270)	(77.009)	(38.332)
Jan. average temperature (C)	-114.840***	-114.840*	-51.497	-217.290

	(33.833)	(50.809)	(62.456)	(101.180)
Feb. average precipitation (mm)	3.904	3.904	-72.013*	171.598***
	(14.490)	(21.730)	(17.522)	(15.796)
Feb. average temperature (C)	61.916	61.916	-138.821	177.358
	(49.780)	(85.217)	(59.304)	(112.731)
Mar. average precipitation (mm)	19.598	19.598	55.934	-198.468*
	(28.516)	(41.088)	(45.298)	(80.035)
Mar. average temperature (C)	21.888	21.888	79.640	104.990
	(51.375)	(86.584)	(148.471)	(56.589)
Apr. average precipitation (mm)	28.696	28.696	176.996**	37.360*
	(29.602)	(35.735)	(20.137)	(11.821)
Apr. average temperature (C)	37.609	37.609	-167.630	-110.854*
	(63.338)	(55.664)	(166.129)	(45.685)
May average precipitation (mm)	-108.142**	-108.142*	-359.718	-73.980*
	(46.097)	(50.813)	(164.899)	(29.323)
May average temperature (C)	-20.035	-20.035	-12.161	159.515
	(59.615)	(71.919)	(24.359)	(94.450)
Jun. average precipitation (mm)	-56.635	-56.635**	283.116	9.808
	(35.165)	(20.462)	(187.178)	(34.478)
Jun. average temperature (C)	-11.240	-11.240	123.762	-27.402
	(55.419)	(60.573)	(187.206)	(91.193)
Jul. average precipitation (mm)	-94.401***	-94.401**	-230.351	-102.319**
	(35.549)	(35.346)	(206.623)	(22.429)
Jul. average temperature (C)	65.610	65.610*	-240.338	65.218
	(54.926)	(32.966)	(96.740)	(56.350)
Aug. average precipitation (mm)	22.829	22.829	-754.282	3.241
	(19.669)	(30.622)	(310.293)	(25.175)
Aug. average temperature (C)	-66.040	-66.040	29.549	-134.824
	(58.443)	(51.631)	(73.414)	(63.034)
Sep. average precipitation (mm)	119.168***	119.168	-55.096	76.542
	(44.303)	(67.851)	(35.449)	(69.675)
Sep. average temperature (C)	24.369	24.369	308.891*	-4.208
	(63.923)	(63.645)	(96.817)	(33.437)
State FE	Y	Y	Y	Y
Sample	All	All	West	East
Standard Errors	HW	S,C	S,C	S,C
Adjusted R^2	0.31	0.31	0.32	0.33
# Observations	1000	1000	449	551

Notes: Table presents results from weighted linear fixed effects regressions of block-level yield in bushels per acre regressed on honey bee colonies per acre, block characteristics, labor inputs, monthly average temperature and precipitation (Jan-Sept), and remotely sensed land cover measures to proxy for wild bee habitat and landscape heterogeneity. Standard errors are either Huber-White robust standard errors (HW), or multi-way clustered at the county (C) and state (S) levels, and are in parentheses. Significance codes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table E.2: Weighted linear fixed effects regressions of yield for Eastern states using alternative measures of land cover.

	Dependent variable is block-level apple yield (bushels/acre)		
	(1)	(2)	(3)
Honey bee colonies per acre			
honey bee colonies per acre	67.836*	67.727*	68.685*
	(22.308)	(23.488)	(22.966)
honey bee colonies per acre, squared	-5.787	-6.275	-6.449
	(3.818)	(3.922)	(3.905)
Land cover variables			
natural forest cover (county)	759.463***		
	(112.357)		
natural forest cover (county), squared	-1,020.779***		
	(22.319)		
natural forest cover (3000 m. buffer)		1,157.634**	
		(317.524)	
natural forest cover (3000 m. buffer), squared		-1,567.622*	
		(545.776)	
natural forest cover (1000 m. buffer)			903.257*
			(308.524)
natural forest cover (1000 m. buffer), squared			-1,123.942
			(502.324)
natural open cover (county)	-3,550.023*		
	(1,280.448)		
natural open cover (county), squared	8,171.569		
	(4,152.272)		
natural open cover (3000 m. buffer)		-1,467.648	
		(1,297.936)	
natural open cover (3000 m. buffer), squared		1,253.466	
		(2,403.695)	
natural open cover (1000 m. buffer)			-1,331.240
			(1,522.495)
natural open cover (1000 m. buffer), squared			900.154
			(2,854.113)
Measures of production scale	Y	Y	Y
Labor input variables	Y	Y	Y
Weather variables	Y	Y	Y
State FE	Y	Y	Y
Sample	East	East	East
Standard Errors	HW	S,C	S,C
Adjusted R^2	0.33	0.34	0.34
# Observations	551	551	551

Notes: Table presents results for weighted linear regressions of block-level yield in bushels per acre regressed on honey bee colonies per acre, block characteristics, labor inputs, monthly average temperature and precipitation (Jan-Sept), and alternative wild bee habitat proxies for the Eastern states. Remotely sensed habitat measures include county-level proportions in specification (1), and the proportion within a buffer of 3000 and 1000 meters around apple producing areas in specifications (2) and (3). A caveat for specifications (2) and (3) is that apple-specific areas are not necessarily identifiable within the USDA-CDL within a given county. To deal with this, we use proportions within buffers around tree crop producing areas, or the county proportion (if tree crop areas are not identifiable). Of the 551 observations in these regressions, 431 farms are in counties where apple specific areas can be identified, 6 farms are in counties where tree crop areas can be identified, and the remaining 114 farms are in counties where natural forest and open cover can only be summarized at the county level. Standard errors are either Huber-White robust standard errors (HW), or multi-way clustered at the county (C) and state (S) levels, and are in parentheses. Significance codes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table E.3: Weighted linear fixed effects regressions of profits.

	Dependent variable is block-level apple profits ($/acre)			
	(1)	(2)	(3)	(4)
Honey bee colonies per acre				
honey bee colonies per acre	1,549.111***	1,549.111**	2,075.516	1,006.975*
	(279.126)	(419.982)	(954.052)	(327.935)
honey bee colonies per acre, squared	-171.891***	-171.891*	-203.250	-117.006
	(44.908)	(71.225)	(200.751)	(54.302)
Measures of production scale				
apple bearing acres	56.507***	56.507***	32.510***	66.840**
	(14.660)	(13.141)	(1.560)	(20.797)
apple bearing acres, squared	-0.028	-0.028	0.019	-0.601**
	(0.041)	(0.034)	(0.007)	(0.145)
trees per acre	0.945	0.945	-2.066	8.091
	(2.041)	(5.724)	(3.433)	(4.684)
trees per acre, squared	-0.003	-0.003	-0.001	-0.004
	(0.002)	(0.004)	(0.003)	(0.003)
average age of trees	79.385**	79.385**	93.205***	42.677
	(32.630)	(24.938)	(8.587)	(26.753)
average age of trees, squared	-1.449***	-1.449***	-0.834**	-0.839**
	(0.474)	(0.382)	(0.131)	(0.145)
Labor input variables				
pruning/thinning hours	-1.427***	-1.427***	-1.647***	-0.804
	(0.257)	(0.073)	(0.111)	(0.557)
harvesting hours	-0.042	-0.042	0.671*	-0.516*
	(0.214)	(0.189)	(0.160)	(0.175)
land prep and machine hours	2.355	2.355	1.910	6.088**
	(1.573)	(1.298)	(1.579)	(1.367)
pest scouting hours	-1.257**	-1.257***	-0.717*	7.222
	(0.616)	(0.291)	(0.189)	(4.059)
part time and seasonal hours	0.096	0.096	0.023	0.345
	(0.175)	(0.111)	(0.073)	(0.189)
full time hours	0.265	0.265***	0.218	0.644
	(0.344)	(0.052)	(0.254)	(0.756)
Land cover variables				
natural forest cover	-5,984.983**	-5,984.983	5,945.144	-514.450
	(2,671.699)	(4,518.220)	(2,644.068)	(6,370.526)
natural forest cover, squared	5,059.542	5,059.542	-12,532.420	-1,758.821
	(3,572.048)	(6,085.677)	(5,807.915)	(3,844.049)
natural open cover	10,508.770	10,508.770	-26,537.320	-46,329.760***
	(7,530.827)	(6,983.854)	(14,737.680)	(5,403.654)
natural open cover, squared	-26,533.220***	-26,533.220***	15,685.780	182,183.100**
	(8,519.587)	(5,901.124)	(19,905.380)	(31,886.730)
Weather variables				
Jan. average precipitation (mm)	-1,145.566***	-1,145.566*	-1,621.610	-1,180.238**
	(425.806)	(473.491)	(1,103.637)	(235.590)
Jan. average temperature (C)	-2,243.378***	-2,243.378**	-2,798.264*	-621.450

	(484.444)	(724.466)	(866.428)	(538.706)
Feb. average precipitation (mm)	807.156***	807.156*	438.406	2,475.269**
	(207.472)	(397.865)	(490.717)	(546.171)
Feb. average temperature (C)	-492.970	-492.970	-7,305.001	702.282
	(712.774)	(1,473.341)	(3,164.172)	(880.198)
Mar. average precipitation (mm)	79.614	79.614	4,784.945**	-2,021.727**
	(408.308)	(625.294)	(710.380)	(488.922)
Mar. average temperature (C)	2,153.403***	2,153.403	7,509.021	760.360
	(735.609)	(1,687.360)	(2,662.791)	(848.077)
Apr. average precipitation (mm)	87.671	87.671	-1,339.856	317.139*
	(423.852)	(686.790)	(3,750.771)	(129.491)
Apr. average temperature (C)	1,283.824	1,283.824	5,123.948	-1,493.162**
	(906.906)	(1,473.880)	(2,539.635)	(290.012)
May average precipitation (mm)	-1,060.300	-1,060.300	-6,940.986*	-155.317
	(660.042)	(1,021.956)	(1,813.886)	(560.079)
May average temperature (C)	1,671.445*	1,671.445*	-2,281.300**	1,108.217
	(853.603)	(700.738)	(453.605)	(632.785)
Jun. average precipitation (mm)	-1,512.475***	-1,512.475*	-2,662.714	539.282
	(503.516)	(744.122)	(1,133.519)	(819.057)
Jun. average temperature (C)	-1,910.459**	-1,910.459	-1,009.436	679.472
	(793.513)	(1,397.355)	(2,575.067)	(1,118.294)
Jul. average precipitation (mm)	192.324	192.324	-4,150.616**	-857.996
	(509.013)	(741.224)	(439.469)	(460.047)
Jul. average temperature (C)	904.911	904.911	-5,751.975	126.626
	(786.463)	(1,246.463)	(2,259.378)	(1,057.315)
Aug. average precipitation (mm)	-638.195**	-638.195	764.923	302.561
	(281.627)	(613.630)	(654.339)	(444.199)
Aug. average temperature (C)	-1,341.222	-1,341.222	5,083.671*	1,113.124
	(836.812)	(1,714.507)	(1,237.662)	(1,148.521)
Sep. average precipitation (mm)	1,281.286**	1,281.286	-2,224.227	665.372
	(634.348)	(825.776)	(1,299.754)	(466.498)
Sep. average temperature (C)	63.660	63.660	-1,098.149	-3,121.417
	(915.286)	(857.229)	(2,596.546)	(1,343.887)
State FE	Y	Y	Y	Y
Sample	All	All	West	East
Standard Errors	HW	S,C	S,C	S,C
Adjusted R^2	0.427	0.427	0.516	0.318
# Observations	1,000	1,000	449	551

Notes: Table presents results from weighted linear fixed effects regressions of block-level profits per acre regressed on honey bee colonies per acre, block characteristics, labor inputs, monthly average temperature and precipitation (Jan-Sept), and remotely sensed land cover measures to proxy for wild bee habitat and landscape heterogeneity. Standard errors are either Huber-White robust standard errors (HW), or multi-way clustered at the county (C) and state (S) levels, and are in parentheses. Significance codes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

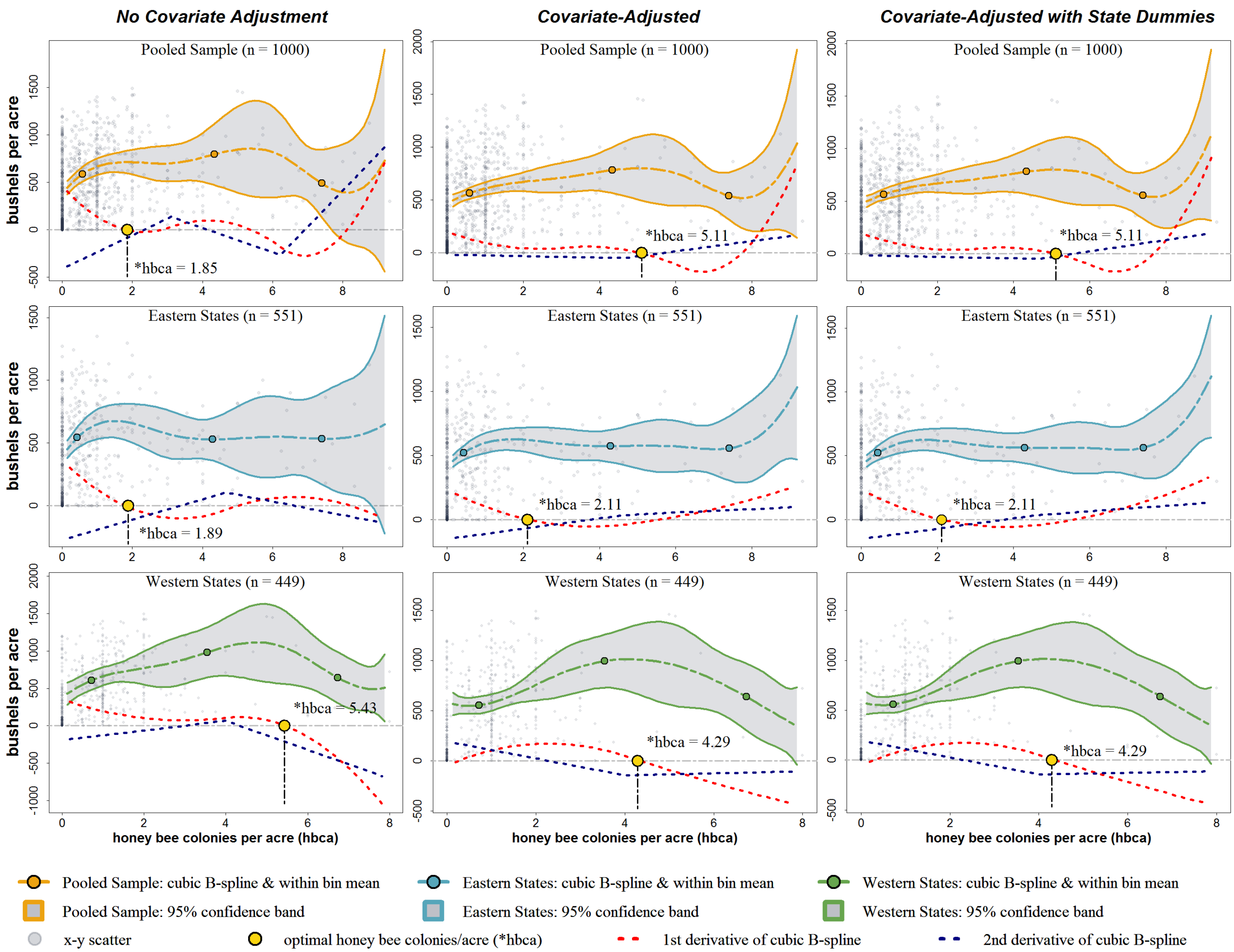

Figure E.1: Optimal binscatter (following Cattaneo et al. (2021)) of *yield* in bushels per acre on the semi-parametric function $\mu(x)$, where x is honey bee colonies per acre, which is defined as the number of honey bee colonies rented divided by selected block size in acres. Each panel trims the 99% centile of the outcome variable and honey bee colonies per acre to reduce the influence of extreme outliers that can dramatically affect the readability of the figure. Column 1 is the optimal binscatter of yield on honey bee colonies per acre. Column 2 includes covariate-adjustment using the same covariates employed in the linear models in Tables E.1-E.5, with the exception of the polynomial versions of some of these variables. Column 3 employs the same model in Column 2 but includes state dummies. These estimations employ *equally spaced*, data-driven rule of thumb bin selection, and cubic B-splines within and between bins. Confidence bands are bootstrapped with n draws. Optimal honey bee colonies per acre are plotted where the estimated first derivative (in red) of the response function equals zero and the response function is at a global (or local) maximum. Second derivatives are also plotted in dark blue.

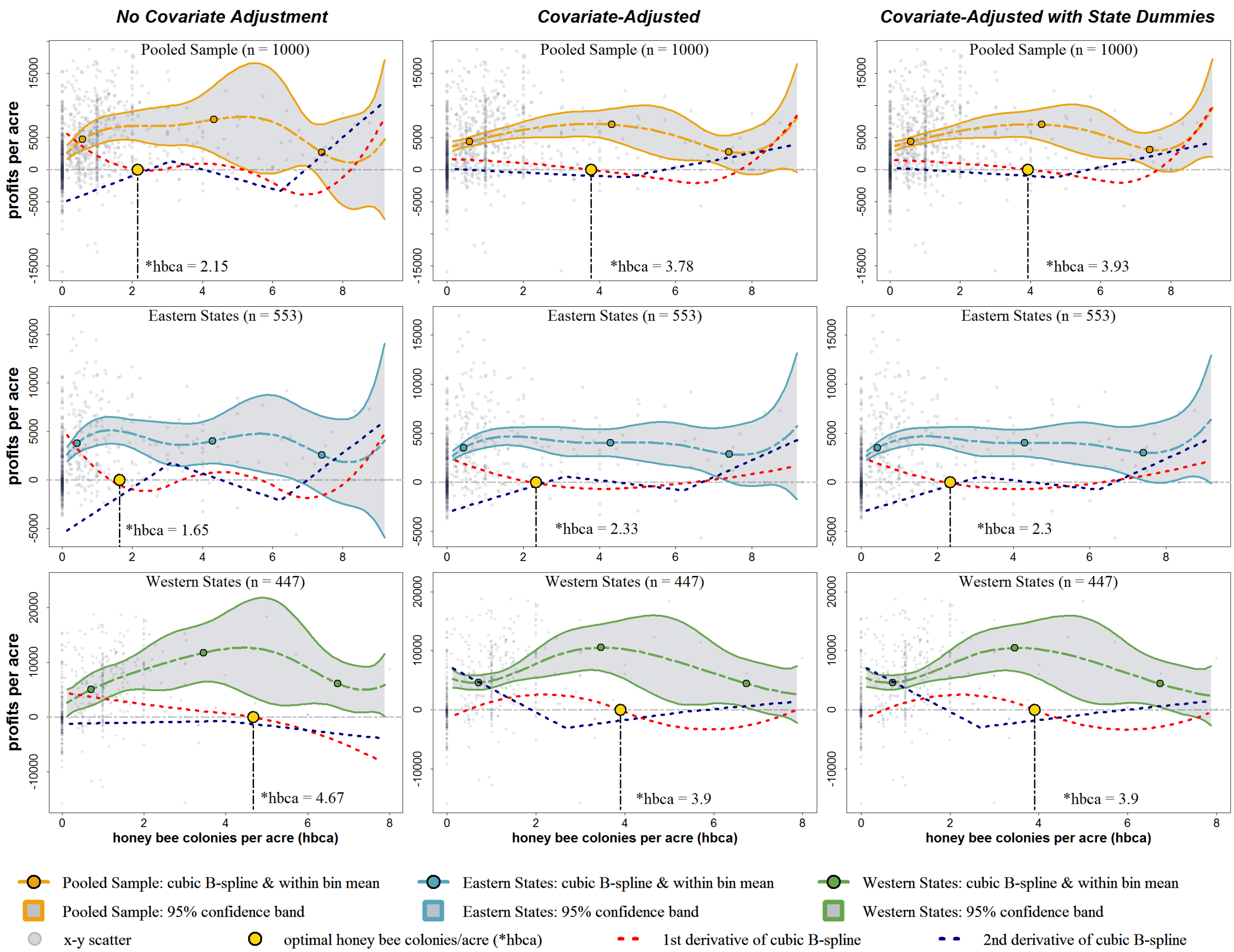

Figure E.2: Optimal binscatter (following Cattaneo et al. (2021)) of *profits* in dollars per acre on the semi-parametric function $\mu(x)$, where x is honey bee colonies per acre, which is defined as the number of honey bee colonies rented divided by selected block size in acres. Each panel trims the 99% centile of the outcome variable and honey bee colonies per acre to reduce the influence of extreme outliers that can dramatically affect the readability of the figure. Column 1 is the optimal binscatter of yield on honey bee colonies per acre. Column 2 includes covariate-adjustment using the same covariates employed in the linear models in Tables E.1-E.5, with the exception of the polynomial versions of some of these variables. Column 3 employs the same model in Column 2 but includes state dummies. These estimations employ *equally spaced*, data-driven rule of thumb bin selection, and cubic B-splines within and between bins. Confidence bands are bootstrapped with n draws. Optimal honey bee colonies per acre are plotted where the estimated first derivative (in red) of the response function equals zero and the response function is at a global (or local) maximum. Second derivatives are also plotted in dark blue.

Table E.4: Parametric tests of response function $\mu(x)$ for yield for Eastern states.

	no covariate adjustment	covariate-adjusted	covariate-adjusted with state dummies
constant	4.509	3.479	3.453
	(0.000)	(0.000)	(0.000)
linear	3.491	2.306	2.374
	(0.004)	(0.029)	(0.019)
quadratic	2.633	1.797	2.281
	(0.071)	(0.086)	(0.029)
cubic	1.635	0.486	0.420
	(0.525)	(0.686)	(0.718)
# Bins	3	3	3
# Observations	551	551	551
# Distinct values	159	159	159

Notes: Table presents t-statistics (p-values in parentheses) from parametric tests of the response function $\mu(x)$ for yield for specifications using the Eastern states subsample. Yield is in bushels per acre; x is honey bee colonies per acre defined as the number of honey bee colonies rented divided by selected block size in acres. Tests employ rule of thumb approach for selection of the number of bins (Cattaneo et al. 2021), quantile-spaced bins, and sample weights. Significance codes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table E.5: Parametric tests of response function $\mu(x)$ for yield for Western states.

	no covariate adjustment	covariate-adjusted	covariate-adjusted with state dummies
constant	4.006	3.181	3.130
	(0.000)	(0.003)	(0.003)
linear	4.086	3.586	3.545
	(0.000)	(0.001)	(0.001)
quadratic	1.101	1.987	2.001
	(0.856)	(0.077)	(0.070)
cubic	0.843	0.891	0.891
	(0.938)	(0.526)	(0.526)
# Bins	3	3	3
# Observations	449	449	449
# Distinct values	136	136	136

Notes: Table presents t-statistics (p-values in parentheses) from parametric tests of the response function $\mu(x)$ for yield for specifications using the Western states subsample. Yield is in bushels per acre; x is honey bee colonies per acre defined as the number of honey bee colonies rented divided by selected block size in acres. Tests employ rule of thumb approach for selection of the number of bins (Cattaneo et al. 2021), quantile-spaced bins, and sample weights. Significance codes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table E.6: Shape restriction tests of response function $\mu(x)$ for yield for Eastern states.

	no covariate adjustment	covariate-adjusted	covariate-adjusted with state dummies
non-positive	4.970	3.750	3.857
	(0.000)	(0.000)	(0.000)
non-negative	−1.139	−1.190	−1.371
	(0.625)	(0.599)	(0.496)
concave	2.402	2.566	2.983
	(0.033)	(0.020)	(0.003)
convex	−3.758	−3.126	−3.297
	(0.000)	(0.001)	(0.000)
# Bins	2	2	2
# Observations	551	551	551
# Distinct values	159	159	159

Notes: Table presents t-statistics (p-values in parentheses) from shape restriction tests of the response function $\mu(x)$ for yield for specifications using the Eastern states subsample. Yield is in bushels per acre; x is honey bee colonies per acre defined as the number of honey bee colonies rented divided by selected block size in acres. Monotonicity tests are applied to the first derivative of respective optimal binscatter curves for the models represented in each column. Tests for concavity and convexity are applied to the respective second derivatives. Tests employ data-driven rule of thumb approach for selection of the number of bins (Cattaneo et al. 2021), quantile-spaced bins, and sample weights. Significance codes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table E.7: Shape restriction tests of response function $\mu(x)$ for yield for Western states.

	no covariate adjustment	covariate-adjusted	covariate-adjusted with state dummies
non-positive	3.464	3.344	3.311
	(0.003)	(0.003)	(0.003)
non-negative	−1.990	−2.282	−2.261
	(0.157)	(0.090)	(0.092)
concave	0.565	1.505	1.534
	(0.834)	(0.306)	(0.290)
convex	−2.614	−2.909	−2.886
	(0.030)	(0.013)	(0.014)
# Bins	2	2	2
# Observations	449	449	449
# Distinct values	136	136	136

Notes: Table presents t-statistics (p-values in parentheses) from shape restriction tests of the response function $\mu(x)$ for yield for specifications using the Western states subsample. Yield is in bushels per acre; x is honey bee colonies per acre defined as the number of honey bee colonies rented divided by selected block size in acres. Monotonicity tests are applied to the first derivative of respective optimal binscatter curves for the models represented in each column. Tests for concavity and convexity are applied to the respective second derivatives. Tests employ data-driven rule of thumb approach for selection of the number of bins (Cattaneo et al. 2021), quantile-spaced bins, and sample weights. Significance codes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

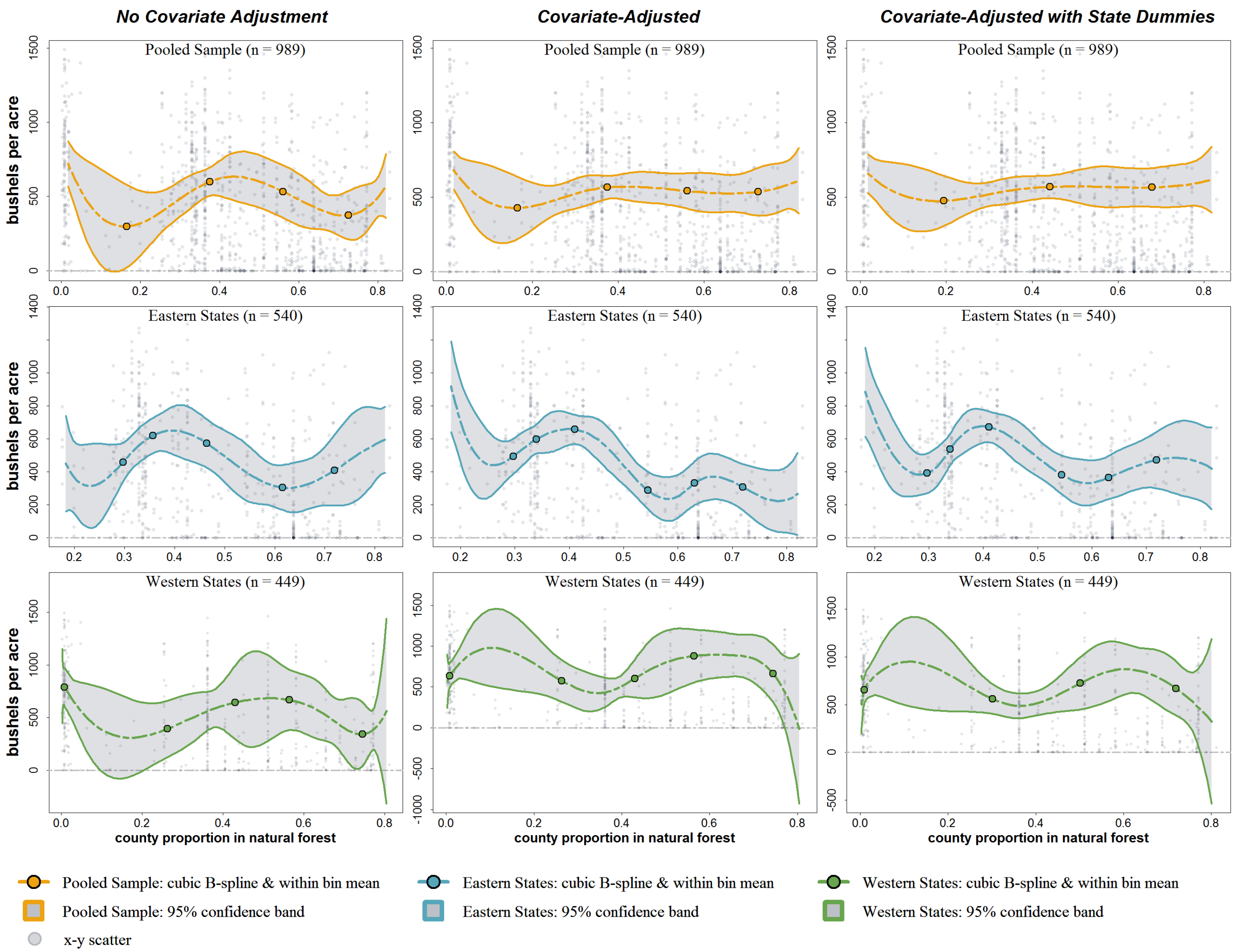

Figure E.3: Optimal binscatter (following Cattaneo et al. (2021)) of *yield* in bushels per acre on the semi-parametric function $\mu(x)$, where x is the county-level proportion in natural forest cover. Each panel trims the 99% centile of the outcome variable and honey bee colonies per acre to reduce the influence of extreme outliers that can dramatically affect the readability of the figure. Column 1 is the optimal binscatter of yield on natural forest cover. Column 2 includes covariate-adjustment using the same covariates employed in the linear models in Tables E.1-E.5, with the exception of the polynomial versions of some of these variables. Column 3 employs the same model in Column 2 but includes state dummies. These estimations employ *quantile-spaced*, data-driven rule of thumb bin selection, and cubic B-splines within and between bins. Confidence bands are bootstrapped with n draws.

References

Angrist, J.D., K. Graddy, and G.W. Imbens. 2000. "The Interpretation of Instrumental Variables Estimators in Simultaneous Equations Models with an Application to the Demand for Fish." *Review of Economic Studies* 67:499–527.

Bergé, L. 2018. "Efficient estimation of maximum likelihood models with multiple fixed-effects: the R package FENmlm." Working paper.

Biltonen, M. 2020. personal communication.

Blitzer, E.J., J. Gibbs, M.G. Park, and B.N. Danforth. 2016. "Pollination services for apple are dependent on diverse wild bee communities." *Agriculture, Ecosystems and Environment* 221:1–7.

Borusyak, K., P. Hull, and X. Jaravel. 2022. "Quasi-experimental shift-share research designs." *Review of Economic Studies* 89:181–213.

Boryan, C., Z. Yang, R. Mueller, and M. Craig. 2011. "Monitoring US agriculture: the US Department of Agriculture, National Agricultural Statistics Service, Cropland Data Layer Program." *Geocarto International* 26:341–358.

Bruno, E.M., and K. Jessoe. 2021. "Missing markets: Evidence on agricultural groundwater demand from volumetric pricing." *Journal of Public Economics* 196:104374.

Cameron, A.C., and D.L. Miller. 2015. "A practitioner's guide to cluster-robust inference." *Journal of Human Resources* 50:317–372.

Cattaneo, M.D., R.K. Crump, M.H. Farrell, and Y. Feng. 2021. "On binscatter."

Cohen, H. 2022. "Planting mixes of flowers around farm fields helps keep bees healthy." *Morning Ag Clips* 6 April 2022, URL: `https://www.morningagclips.com/planting-mixes-of-flowers-around-farm-fields-helps-keep-bees-healthy/`.

Czarnowske, D., and A. Stammann. 2019. "Fixed effects binary choice models: Estimation and inference with long panels."

Daly, C., M. Halbleib, J.I. Smith, W.P. Gibson, M.K. Doggett, G.H. Taylor, J. Curtis, and P.P. Pasteris. 2008. "Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States." *International Journal of Climatology* 28:2031–2064.

Danforth, B.N., R.L. Minckley, and J.L. Neff. 2019. "Solitary bees and agricultural pollination." In *The Solitary Bees: Biology, Evolution, Conservation*. Princeton University Press, chap. 13.

Garibaldi, L.A., I. Steffan-Dewenter, R. Winfree, M.A. Aizen, R. Bommarco, S.A. Cunningham, C. Kremen, L.G. Carvalheiro, L.D. Harder, O. Afik, et al. 2013. "Wild pollinators enhance fruit set of crops regardless of honey bee abundance." *Science* 339:1608–1611.

Goldberger, A. 1991. *A Course in Econometrics*. Cambridge, MA: Harvard University Press.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift. 2020. "Bartik instruments: What, when, why, and how." *American Economic Review* 110:2586–2624.

Goodrich, B.K., J.C. Williams, and R.E. Goodhue. 2019. "The great bee migration: Supply analysis of honey bee colony shipments into California for almond pollination services." *American Journal of Agricultural Economics* 101:1353–1372.

Grab, H., K. Poveda, B. Danforth, and G. Loeb. 2018. "Landscape context shifts the balance of costs and benefits from wildflower borders on multiple ecosystem services." *Proceedings of the Royal Society B: Biological Sciences* 285:20181102.

Kahlke, C. 2019. personal communication.

Kaila, H.K., and C. Boone. 2020. "The Agricultural Resource Curse? The Welfare Effects of the Vanilla Price Surge." Working paper.

Kennedy, C.M., E. Lonsdorf, M.C. Neel, N.M. Williams, T.H. Ricketts, R. Winfree, R. Bommarco, C. Brittain, A.L. Burley, D. Cariveau, et al. 2013. "A global quantitative synthesis of local and landscape effects on wild bee pollinators in agroecosystems." *Ecology Letters* 16:584–599.

Lee, H., D.A. Sumner, and A. Champetier. 2019. "Pollination markets and the coupled futures of almonds and honey bees: simulating impacts of shifts in demands and costs." *American Journal of Agricultural Economics* 101:230–249.

Lin, C.Y.C. 2011. "Estimating supply and demand in the world oil market." *Journal of Energy and Development* 34:1–32.

Manski, C.F. 1995. *Identification Problems in the Social Sciences*. Cambridge, MA: Harvard University Press.

Martins, K.T., A. Gonzalez, and M.J. Lechowicz. 2015. "Pollination services are mediated by bee functional diversity and landscape context." *Agriculture, Ecosystems and Environment* 200:12–20.

McArt, S.H., A.A. Fersch, N.J. Milano, L.L. Truitt, and K. Böröczky. 2017. "High pesticide risk to honey bees despite low focal crop pollen collection during pollination of a mass blooming crop." *Scientific Reports* 7:1–10.

Mid Valley Agricultural Services. 2006. "Fruit Set Factors." *Mid Valley Agricultural Services Newsletter* October 2006.

Narjes, M.E., and C. Lippert. 2019. "The optimal supply of crop pollination and honey from wild and managed bees: An analytical framework for diverse socio-economic and ecological settings." *Ecological Economics* 157:278–290.

—. 2021. "Regional differences in farmers' preferences for a native bee conservation policy: The case of farming communities in Northern and Eastern Thailand." *PloS ONE* 16:e0251206.

Park, M.G., E. Blitzer, J. Gibbs, J.E. Losey, and B.N. Danforth. 2015. "Negative effects of pesticides on wild bee communities can be buffered by landscape context." *Proceedings of the Royal Society B: Biological Sciences* 282:20150299.

Park, M.G., R.A. Raguso, J.E. Losey, and B.N. Danforth. 2016. "Per-visit pollinator performance and regional importance of wild Bombus and Andrena (Melandrena) compared to the managed honey bee in New York apple orchards." *Apidologie* 47:145–160.

Petterson, M.S., D.G. Seim, and J.M. Shapiro. 2023. "Bounds on a slope from size restrictions on economic shocks." *American Economic Journal: Microeconomics* 15:552–572.

Ridley, W., and S. Devadoss. 2021. "Challenges for the US Fruit Industry: Trends in Production, Consolidation, and Competition." *Choices* 36.

Robinson, T., S. Hoying, M.M. Sazo, A. DeMarree, and L. Dominguez. 2013. "A vision for apple orchard systems of the future." *New York Fruit Quarterly* 21:11–16.

Robinson, T., S.A. Hoying, A. DeMaree, K. Iungerman, and M. Fargione. 2007. "The evolution towards more competitive apple orchard systems in New York." *New York Fruit Quarterly* 15:3–9.

Roubik, D.W. 2002. "Tropical agriculture: the value of bees to the coffee harvest." *Nature* 417:708.

Rucker, R.R., W.N. Thurman, and M. Burgett. 2012. "Honey bee pollination markets and the internalization of reciprocal benefits." *American Journal of Agricultural Economics* 94:956–977.

Russo, L., M. Park, E. Blitzer, and B. Danforth. 2017. "Flower handling behavior and abundance determine the relative contribution of pollinators to seed set in apple orchards." *Agriculture, Ecosystems and Environment* 246:102–108.

Sato, K. 1967. "A two-level constant-elasticity-of-substitution production function." *Review of Economic Studies* 34:201–218.

Stammann, A., F. Heiss, and D. McFadden. 2016. "Estimating fixed effects logit models with large panel data."

USDA Economic Research Service (ERS). 2022. "ARMS Farm Financial and Crop Production Practices: Contact Us." URL: `https://www.ers.usda.gov/data-products/arms-farm-financial-and-crop-production-practices/contact-us/`.

vanEngelsdorp, D., and M.D. Meixner. 2010. "A historical review of managed honey bee populations in Europe and the United States and the factors that may affect them." *Journal of Invertebrate Pathology* 103:S80–S95.

Ward, R., A. Whyte, and R. James. 2010. "A tale of two bees: Looking at pollination fees for almonds and sweet cherries." *American Entomologist* 56:170–177.