Does Colony Loss Reduce Honey Yields?

Abstract

While most agricultural goods is marked by increasing productivity, yields of honey measured in pounds produced per colony has fallen 29% since 2000. Throughout this period, both honey bee colony loss rates and pollination service income – the other main source of revenue for beekeepers – have also been high, suggesting a link between poor colony health and the long-distance movement and co-location of colonies in California to service its large and lucrative almond bloom. We use beekeepers survey data from USDA's National Agricultural Statistics Service to estimate the response of honey yields to changes in the beekeeper's colony loss rates and the beekeeper's share of colonies moved to California for almond pollination. We estimate that, on average, a 1 percentage point increase in the loss rate decreases honey yield by .362 pounds, but smaller beekeepers experience a greater yield reduction for a given loss rate than larger beekeepers. We also estimate that a 10 percentage point increase in beekeeper's share of colonies moved to California in the almond pollination season decreases yield by 0.8 pounds.

Peyton Ferrier
USDA Economic Research Service
pferrier@ers.usda.gov
202-694-5224

Disclaimer:
The findings and conclusions in this preliminary presentation have not been formally disseminated by the U. S. Department of Agriculture and should not be construed to represent any Agency determination or policy. This research was supported in part by the intramural research program of the U.S. Department of Agriculture, Economic Research Service.

This manuscript was prepared for presentation as at Agricultural and Applied Economics Association's sponsored sessions at the annual meeting of the Allied Social Sciences Association in Atlanta Georgia, January 5-7 2018.

Keywords: pollinators, honey, productivity, colony loss
JEL Codes: Q1, Q51, Q57

I. Introduction

While most agricultural goods is marked by increasing productivity, yields of honey measured in pounds produced per colony has fallen 29% since 2000. Throughout this period, both honey bee colony loss rates and pollination service income – the other main source of revenue for beekeepers – have also been high, suggesting a link between poor colony health and the long-distance movement and co-location of colonies in California to service its large and lucrative almond bloom. Almond pollination service fees are approximately three times that of other crops and provide one-third of beekeepers revenue (Ferrier et al, 2018). While much media attention has arisen over the possibility that shortfalls of honey bees as crop pollinators will harm agricultural productivity, relatively little attention has been paid to its effect of colony losses on honey yields

Moreover, even if servicing the almond bloom has no effect of colony loss, it may still affect yield through output substitution effects. Beekeepers earn most of their income from honey sales and pollination services fees. Colonies of honey bees (*Apis melifera*) are mobile, being relatively easily moved on trailer trucks in standardized hives. Beekeepers can increase both honey yields and pollination service income by moving colonies between distant apiary sites that offer either large floral blooms for honey production or lucrative colony rental fees. While in most cases, moving colonies may be thought to have a neutral or beneficial effect on yields if it secures better forage land, almonds produce a bitter tasting honey that is not typically sold to consumers or including in production figures. Colony movement in and out of California is substantial and costly in itself. In 2016, the share of all U.S. colonies in California fell from 54 to 26 percent between the 1st and 3rd quarters of 2016 while the share in the Northern Great Plains (Montana, North Dakota, South Dakota, and Minnesota) rose from 12 to 33 percent.

Pollination fees are much higher for almonds and nearly all U.S. almond production occurs in the Central Valley of California. **Figure 1** shows the average pollination fees received by beekeepers in the Pacific Northwest for various crops and the rise in almond pollination service fees between 2004 and 2006, along with the relative stability of fees for other crops. As described in Ferrier et al. (2018), pollination service revenue represented 11 percent of beekeeper revenues in 1988 and 41 percent of revenue in 2016. Almonds comprise 82 percent of all pollination serve revenue, 60 percent of all colony rentals and 52 percent of acres on which rental colonies are placed.

<< Figure 1 – Average Pollination Service Fees Received by Pacific Northwest Beekeepers >>

What makes almond pollination fees so high? Almonds bloom for a few weeks between mid-February and mid-March, and honey bee colonies need to be staged nearby in preparation for the bloom season. Seasonally, the fewest number of colonies are available in the first quarter of the year and the costs of moving colonies potentially thousands of miles to California are high. Moreover, as described by Cheung (1973) and Rucker et al. (2012), colony rental fees are lower for crops that produce larger amounts or higher quantity honey which the beekeeper retains as an in-kind payment. Among crops paying pollination service fees, almond honey is bad-tasting and unmarketable (Nordhaus, 2006). Also, almonds are highly pollinator-dependent with yields estimated to fall over 90 percent in the complete absence of insect pollinators, a circumstance that makes almond producers' demand for colonies very unresponsive to pollination service fee increases. Almonds have also experienced rapid industry growth. **Figure 2** shows that since 1987 almond acreage has grown nearly 139 percent and tree plantings have grown 39 percent

(Ferrier et al., 2018). The United States produces 77 percent of the world's almonds and nearly all those almonds are produced in California.

<< Figure 2 – Growth in U.S. Almonds Acreage and Trees per Acres>>

In recent years, colony health and quality concerns may have raised almond pollination service fees as well. Markets for pollination services are organized through formal and informal contracts specifying the timing, duration and placement of colony rentals. A quality standard that specifies the colony's population may also be included as well. The use of standardized hives allows a colony's population to be visually assessed by the number of removable frames covered with bees and filled with brood. Colony size can differ substantially. Large colonies with more full frames[1] are better at both pollinating and producing honey. Goodrich (2017) explains that while most almond pollination service contracts specify an 8-frame minimum, incentives and enforcement mechanisms vary across grower-beekeeper agreements[2]. Independent inspection may have become more common in recent years if high almond pollination fees have encouraged beekeepers to bring smaller or weaker colonies into pollination service contracts. Coordinating contracts and moving colonies longer distances to reach California increase beekeeper costs and high fees are thought to have initially drawn in colonies from closer states first, resulting in increasingly larger shares of colonies being brought to California in January and February to service the crop's growing acreage.

---

[1] Full frames are those frames contained honey, pollen or brood over most of their surface area.

[2] Goodrich further notes that colony quality inspections are less important for crops producing good honey because the crop producer's desire for large colonies for better pollination coincides with the beekeepers desire for large colonies producing better honey.

<< Figure 3 – Honey Yields for the United States and Select States >>

As pollination service income rose since 2000, honey yields fell. Although honey bees forage and produce whenever the air temperature exceeds 55 degrees, spring and summer are the most productive seasons. Beekeepers typically collect honey at the end of the peak summer flower bloom, leaving some honey is left in the hive for the colony's subsistence in winter. **Figure 3** shows honey yields since 1990 nationally and for select states. The linkage between pollination service income and honey yields may occur through several routes. First, as emphasized by various authors, honey and pollination services are co-products. An increase in pollination service fees causes both a substitution effect reducing honey production, as beekeepers direct more colonies from primarily honey producing uses to primarily pollination service uses, and an output effect increasing honey production, as beekeepers increase the number of colonies that are now more profitable[3]. Second, higher pollination service fees may induce an income effect causing beekeepers to work less. Beekeeping remains relatively labor intensive and there may be a limited ability to outsource certain crucial managerial functions including securing additional forage area, health monitoring, or colony transportation[4].

Since 2000, colony loss rates have also risen considerably over historic rates and these losses may have contributed to the fall in honey yields. Winter loss rates have been roughly to an average of 28.7 percent between 2006 and 2016, roughly double their historic rate of 15 percent. Ferrier et al. (2018) provides an overview of various emerging factors thought to have contributed to poor honey bee health. Despite these elevated loss rates, total colony numbers remained stable during this time period (2007-16) owing to the ability of beekeepers to split

---

[3] Across these model, co-product production is modeled either sequentially (Lee, Sumner, and Champetier, 2015) or simultaneously (Champetier et al., 2015, Cheung, 1973, Rucker, Thurman and Burgett 2012).

[4] Champetier (2015) also describes how beekeepers reduce honey harvesting in order to increasing colony stocks in response to a permanent increase in pollination fees. This reduced yield represents a transitory adjustment cost lasting only until beekeeper reach a new target colony stock.

colonies as a way to replace lost colonies. Using State-level data on rates of beekeeper colony losses and additions, Ferrier et al. (2018) also showed that colony loss and addition rates are correlated but those loss rates and rates of change in the number of colonies from year-to-year are not.

How might colony losses reduce honey yields? Lost colonies are typically replaced through the splitting of the remaining healthy colonies. To split a colony, beekeepers divide off a portion of a large colony and place that portion within a new hive while installing a new queen in the mother colony. Between bee populations of the same size, two small colonies will produce less honey than a single large colony because each colony must devote a larger share of the colony to reproductive and hive maintenance functions rather than foraging (Farrar, 1968). Split colonies eventually grow to reach a fully mature size, but in the intervening time honey production falls.

This research estimates how much recent decreases in honey yields can be explained by colony loss or colony movement in the pursuit of almond pollination fees. Production and honey yields data are from the USDA's National Agricultural Statistics Service's (NASS) annual survey of beekeepers. Loss rate and movement data are from NASS's quarterly honey colony survey. This survey identifies the States in which the beekeeper's colonies are located at the start of each quarter which we use to compute the share of colonies the beekeeper's moves to California during the almond bloom. Using the beekeeper ID variable common to both datasets, we are able link the two data between the first quarter of 2015 and third quarter of 2017, identifying yields, losses and movement by beekeeper.

To account for productivity differences across beekeepers, we include operation size and lagged yields as controls. To account for regional productivity shocks, we construct an estimate of the beekeeper expected yield based on state of operation in two steps. First, we calculate the average yield for each State in which the beekeeper operates while excluding that individual beekeepers colonies and production. Then, we multiple those State yields by beekeeper's share of colonies in each of those states in the 3rd quarter of the year, the period which most of beekeepers honey is produced. To address for the possibility that omitted variables affect individual beekeeper productivity, we also report the results of a difference based estimator. In this model, the dependent variable yield and the key exogenous variables are stated as the percentage difference from their previous year's value.

Both our dependent variable, yield, and a key dependent variable, the loss rate, are constructed with the number of colonies in the denominator. To address possible bias caused by measurement error, we instrument for the loss rate using a variable distance that equals the average difference of the distance of the beekeeper colonies from California between the 1st and 3rd quarter of the year. The intuition is that the further the beekeeper moves colonies between their main honey producing areas in the 3rd quarter and California in the 1st quarter, the more stress and loss they will incur.

The remainder of the paper is as follows. Section II discuss the datasets used in the estimation. Section III discusses estimation method including tests of the instruments. Section IV presents estimation results and findings. Section V presents a calculation of how much colony loss and movement may have been responsible for the reduction in honey yield since 2000 along with conclusions.

II. Data and Model

Our analysis uses data from two separate surveys of U.S. beekeepers collected by NASS which are used to generate the annual Honey Report and the quarterly Honey Bee Colony report. Individual responses by beekeepers to these surveys is provided under confidentiality restrictions with the author to prevent the revelation of personal information or individual responses. Aside from the identification variable that allows us to link survey respondents across the two datasets, all personally-identifying respondent information, including names, address, and demographics, is withheld. The Honey Production dataset is collected once annually for all beekeepers. The Colony Loss dataset is collected quarterly for beekeepers with 50 or more colonies. The datasets are merged through a beekeeper ID variable common to both.

The Honey Production data covers the four years between 2014 and 2017. The Colony Loss data covers the 14 quarters from the 1st quarter of 2015 to the 2nd quarter of 2018. Beekeepers both enter and leave the sample over time. Both surveys are designed to be nationally representative and comprehensive, capturing large shares of the total number of U.S. beekeepers. While the survey's construction allows beekeepers to specify that they produced honey in multiple States (which can be different from their home State). In the Honey Production data, only about 10 percent of beekeepers reported honey production occurring across multiple States, a rate far lower than that shown in the Honey Colony data. This lack of difference suggests that the exact location at which honey is produced may not be precisely tracked for multi-state operations in the Honey Production data. For this reason, we perform our analysis on each beekeeper's total yield, rather than that beekeepers yield by State. Moreover, we use the Colony Loss Data's location of colonies in the 3rd quarter of the year to identify where honey production occurred and to construct our expected yield variable.

In the Honey Production data, beekeepers report the year's total honey production, all colonies and colonies producing honey. Honey yield is then total production divided by total colonies producing honey. Our dataset covers the period 2014-17 although honey reports surveys themselves extend back far longer. In the colony loss survey, beekeepers report, for each quarter, their total numbers of colony maintained, lost, added or re-queened. Our dataset covers the 10 quarters running from the first quarter of 2015 through the second quarter of 2017, the full range of the available data. The two datasets were linked using an ID variable common to both data sets. **Table 1** provides summary statistics on our key variables.

<< Table 1 – Summary Statistics on Beekeeper Data by Size in 2016 >>

Several modeling concerns inform our estimation method in three ways. First, our dependent variable – yield – and a key explanatory variable, the loss rate, being constructed with the total number of colonies in its denominator. We instrument for the loss rate to avoid possible bias introduced by measurement error bias (Angrist and Krueger, 1999). As our size of operations control variable, we use the beekeeper's lagged rather than the current number of colonies to similarly avoid measurement error bias through that term.

Second, we believe that State level productivity shocks may be uncorrelated across years and States. For instance, high rainfall in North Dakota in 2017 that raising yields is uncorrelated with either rainfall levels in North Dakota in 2016 or Georgia in 2017. To avoid the need to fit numerous fixed effect variables by State and year, we employ our expected yield variable based on the State yield and the locations of beekeeper's colonies in the 3rd quarter the period in which most honey production occurs. Specifically, the beekeeper's expected yield is each beekeepers share of colonies in each State in the 3rd quarters times that State's average yield[5]. The expected yield term also controls for year-to-year productivity shocks that occur at the state-level.

---

[5] While calculating the expected yield of an individual beekeeper, we use state average yields that exclude that beekeeper's individual contribution to that state's production.

Third, both colony numbers and losses are seasonal and, in some cases, unobserved due to managerial reasons[6]. Following the approach of Ferrier et al. (2018), we calculate both two-quarter and four-quarter loss rate as the sum of losses divided by the sum losses of colonies across the previous two or four previous quarters respectively[7]. Since we have only 10 quarters of loss data, we used the two quarter loss rate in estimation to preserving observations and finding little effect on our results in initial testing. In the end, our two-quarter loss rate covers the two quarters preceding the 3rd quarter. In both cases, we pool the data owing to its short-time frame (3 years).

Yield Model

We estimate the effect on honey yields from the beekeeper's colony loss rate and the share of colonies the beekeeper locates in California (a proxy for whether beekeepers pollinate almonds.) Because our data is too short to estimate fixed effects across beekeepers, we instead pool our data and control for potential productivity differences with several control variables. Define the following variables:

y, the beekeeper's honey yield per colony,

size, the beekeeper's size,

lag_y, the beekeeper's lagged yield,

ey, the beekeeper's expected yield,

cali, the share of the beekeeper's colonies in California, and

loss, the beekeeper's loss rate.

---

[6] The act of checking colonies for losses can kill the colonies themselves if temperatures are low. NASS allows beekeepers to leave data fields blank if these values are unavailable.

[7] If losses are not reported, both loss and colonies are treated as zero in that period.

Honey yield (measured in pounds per colony), **y**, is the dependent variable in regression described in equation (1). Three variables are used to control for unobserved differences in beekeeper productivity: **size,** the number of colonies the beekeeper maintained in the previous year; **lag_y**, the beekeeper's yield from the previous year; and **ey**, the beekeeper's expected yield calculated as the sum of the products of the beekeeper's 3rd quarter shares of colonies in each state and that state's average yield. Here, lagged yields captures the effects of the quality of the beekeeper's apiaries and knowledge as well as any productivity shocks that might be correlated across beekeepers at the state level since beekeepers remain at the same apiary sites over time. Size is also likely to capture the beekeeper's knowledge and control for the beekeeper's ability to move colonies between apiary sites with differing levels of productivity. Alternatively, expected yields largely capture the effect of general productivity shocks at the state levels caused by year-to-year variation in weather or habitat.

Colonies can potentially be moved into California shortly before or after the almond bloom which lasts from mid-February to mid-March. To address the potential for late arriving or early departing colonies, ***cali*** is defined as the larger of the beekeeper's share of colonies in California in the either the first and second quarters of the year. The variable ***loss*** is the sum of the first and second quarter losses divided by the sum of beekeeper's colonies in those quarters. In cases where observations are missing for a single quarter, both losses and colonies are counted as zeroes. We also include an interaction variable between the loss rate and the size variable (**loss*size**). Both the **cali** and **size** are multiplied by 100 so that **loss** equaling 23 means that the beekeeper sent 23% of colonies.

The equation to be estimated is then:

$$y = \beta_0 + \beta_{size} size + \beta_{lag\_y} lag_y + \beta_{ey} ey + \beta_{cali} cali + \beta_{loss}\, loss + \beta_{int}\, loss * size \quad (1)$$

While Equation (1) includes controls for productivity differences across beekeepers, regional shocks to productivity within states may not be accounted for with either the lagged or expected yields terms. While sacrificing degrees of freedom through the lag structure, the alternative approach of control for productivity with a difference model places the year-to-year difference in yields as the left-hand side variable and the year-to-year differences in the similarly constructed non-stationary explanatory variables as the right-hand side variables. Specifically, the new variables are:

***dy*,** the year-to-year percentage change in the beekeeper's yield

dey, the year-to-year percentage change in the beekeeper's expected yield,

dcali, the year-to-year change in the beekeeper's share of colonies in California, and

dloss, the year-to-year change in the beekeeper's loss rate.

The difference model is then estimated in Equation (2) as.

$$dy = \beta_0 + \beta_{size} * size + \beta_{dey} dey + \beta_{dcali} dcali + \beta_{dloss}\, dloss + \beta_{int}\, dloss * size \quad (2)$$

The **dy** and **dey** yields are percentage changes in **y** and **ey** from their initial values and multiplied by 100 (i.e. **dy** equals 10 indicates a 10 percent increase in yields). Because the underlying **cali** and **loss** variables already expressed as percentages, the **dcali** and **dloss** variables are just simple year-to-year difference in **cali** and **loss** variables. **Size,** again acts as a control variable.

As with the yield model, we instrument for the **dloss** variable with **dist**, a yearly dummy for 2017, and our set of exogenous variables. We also include an interaction terms for the **dloss** variable times by the **size** variable. Although the size variable changes slightly from year-to-year, its incorporation to our estimation is to allow for the effects of losses to vary across producers. In the next section, we discuss how we address various estimation issues involving how the **size** variable may effect regression weighting.

III. Estimation

Yield is measured as average production per colony. For large beekeepers, this average is based on a larger number of observed colonies. As we are mainly interested in market effects, we weighted our estimator by the number of colonies the beekeeper maintains in the current year. This had a large impact on the estimation results and suggested our later incorporation of the inclusion of the interaction effects (i.e. **loss*size, dloss*size**).

As previously mentioned, we also instrument for the **dloss** variable in Equation (2) using the **move** variable and yearly dummies. **Table 2** provides the outcome for regression of our three instrumental variables - **move** and two yearly dummies - on the yield level, **y**.

<< Table 2 – Regression Results of our three Instrumental Variables on Yield (*y*) >>

Table 2 shows that the **move** variable is of the expected sign and significant, indicating that moving colonies a larger distance between the 1st and 3rd quarter increases losses. In the next section, we report estimates of equations (1) and (2) based on ordinary least squares and iterated two stage least squares (IT2SLS) using move, the two yearly dummies and all our other explanatory variables as instruments. In each case, we weight the regression outcome for the number of colonies.

IV. Findings

Tables 3 and 4 show the results of our yields level and yield difference. ***Table 3*** shows that the p-value of .1891 for the Hausman test indicates that OLS and IV estimates are not significantly different. However, this results may reflect a weakness of our instruments so we continue to focus our discussion on the IV results. Both **Tables 3 and 4** show a general pattern of showing that both the share of colonies moved to California (**cali**) and the loss rate (**loss**) reducing honey yields. Moreover, the effect of losses dissipates with the size of the beekeeper. In **Table 3**, lagged yield (**lag_y**) and expected yield (**ey**) are strongly significant and in line with our intuition that the productivity is correlated for the same beekeeper across years and different beekeepers within state. While we attribute the expected yield variable controlling for state effects to weather, the relationship may also arise from land-use policies such as requirements on pesticide spray notification or apiary registration laws[8]. Size is estimated to reduce yield, but this effect is partially offset through its interaction effect with the colony loss rate (**loss*size**, which is positive). All our estimates indicate that the colony loss reduced yield much more for smaller beekeepers than larger ones. **Table 4** showed a similar direction and magnitude of the effects of **Table 3**, but were often not significant, a difference we attribute to weak instruments.

<< Table 3 - Estimation of the Yield Difference Model in Equation (1) >>

<< Table 4 - Estimation of the Yield Difference Model in Equation (2) >>

Both **Tables 3 and 4** show that yields fall with the share of colonies moved to California. The IV estimator in **Table 3** indicates that a 10 percentage point increase in the beekeeper's share of colonies brought to California reduces yields by .77 pounds (or 1.5 percent based on an average yield of 52.7 pounds). Alternatively, **Table 4** indicates that the same change reduces yields by 2.4 percent.

Owing to the **loss*size** interaction term , the marginal effect of **loss** in **Tables 3 and 4** can only be interpreted in the context of beekeeper size of operations. **Table 5** provides various

---

[8] For instance, Montana and South Dakota allow registered apiaries to restrict other beekeepers from placing colonies in the same area.

effects based on different sizes. The “Threshold Size” is the beekeeper size after which the estimated marginal effect of colony loss switches from negative to positive and exceeds 23,000 colonies in each case. Marginal effects are also provided for operations with 1,000, 5,000, 10,000 and 20,000 colonies and shows that large beekeepers have lower marginal effects. In **Table 5** our IV estimator for the yield model indicates that a 1 percentage point increase in the two-quarter loss rate decreases yields by 2.63 pounds for beekeepers with only 1,000 colonies, 0.45 pounds for beekeepers with 20,000 colonies, and 0.36 for all beekeepers on average. Based on an average yield of 52.7 pounds, this represents a 0.69 percent reduction in yield. For comparison, the yield difference model estimates that average yield falls 0.66 percent.

<< **Table 5- Marginal effect of a one percentage point increase in the 2-quarter loss rate** >>

The calculation method for the loss rate matters in its interpretation. Our 2-quarter loss rate is a quarterly average rate of loss measured as the sum of colonies loss in each quarter divided by the sum of colonies maintained in each quarter. Another commonly reported potential measure of colony loss is the total number of colonies lost in a year divided by the average number of colonies maintained across periods of measurement, a rate of total loss. If the beekeeper’s number of colonies loss is constant throughout the year, then our 2-quarter loss rate is ¼ of this rate of total loss.

V. Conclusion

Much of the focus of policy debate on the consequences of pollinator health problems has focused on the potential for shortfalls in pollination service markets leading to reduced crop production. As Ferrier et al. (2018) show, increases in pollination service fees have been largely constrained to almonds, strongly suggesting that pollination shortfalls are unlikely for most crops. However, the rapid growth of the almonds industry has had several knock on effects for the beekeeping industry. While raising incomes through higher pollination service fees, it may also raise loss in a manner that reduces honey yields and thus reclaim some of the fee-driven income gain.

Using instrumental variables to control for potential measurement error bias, we estimate that a one percent increase in our 2-quarter loss rate reduces yield by 0.99 percent. The overwinter loss rates of colonies is thought to have roughly doubled from 15 to 30 percent around 2006 when high loss rates led to the systematic data being collected, a change that would represent a 7.5 percent increase in our 2-quarter loss rate and cause a yield reduction of 2.7 pounds or 5.2 percent.

Similarly, increasing plantings of almonds that led to the run-up in almond pollination fees around 2004-06 may have plausibly raised the share of colonies moved to California by 30 percentage points. Our estimates suggest that a 30 percent increase in the share of colonies sent to California would reduce yields by 2.32 pounds or 4.4 percent. Collectively, these two factors (elevated colony loss and movement in service of almond pollination) reduce honey yields 9.6 percent. Collectively, the increased loss rates and movement of colonies in the last two decades explain roughly 33 percent of the total reduction in yields (=9.6/29.3). This back of the envelope calculation is, of course, subject to some caveats. First, since loss rates have a heterogeneous effect across producers of different size, our estimates will under estimate actual yield reductions if loss rates are higher for small beekeepers. **Table 2** suggests this may be the case. Second, as data is not available before 2015, the share of colonies moved to California may have risen by more or less than our simulated 30 percent.

Honey yields are notoriously variable from year-to-year, but surprisingly little research has examined their response to their response to health and productivity shocks in market settings. Unlike fields crops where yield is often assumed to be exogenous to market forces once the crops are planted, yield may respond to market shocks dynamically since splitting colonies less frequently or leaving honey with the colony may aid its growth and health. We control for many of these effects, but note that other exogenous potential supply changes – loss of forage and habitat, changes in climate, sub-lethal colony health problems – may also have reduced yields since 2000 as well. If yields continue to decline further, the role of these factors should be scrutinized as well.

References

Angrist, J.D., and A.B. Krueger (1999) Chapter 23 Empirical strategies in labor economics, ed. Orley Ashenfelter and David Card, vol. Volume 3, Part 1, Elsevier, pp. 1277-1366.

Champetier, A., D.A. Sumner, and J.E. Wilen. 2015. "The bioeconomics of honey bees and pollination." *Environmental and Resource Economics* 60(1):143-164.

Cheung, S.N. 1973. " The Fable of the Bees: An Economic Investigation" *Journal of Law and Economics* 16:11.

Daberkow, S., R. Rucker, W. Thurman, and M. Burgett. 2009. "US honey markets: Recent changes and historical perspective." *American Bee Journal.*

Economic Research Service, 2017, Sugar and Sweeteners Yearbook, United States Department of Agriculture, Washington, DC

Farrar, C.L. 1968. "Productive management of honey-bee colonies." *American Bee Journal.*

Ferrier, P., R.R. Rucker, W.N. Thurman, and M. Burgett. 2018. "Economic Effects and Responses to Changing Honey Bee Health " *USDA ERS Economic Research Report #246.*

Goodrich, B.K. 2017. "Are All Colonies Created Equal? The Role of Honey Bee Colony Strength in Almond Pollination Contracts." *SCC-76 Conference in Pensacola Beach, FL, April 1, 2017.*

Lee, H., D.A. Sumner, and A. Champetier. 2018. "Pollination Markets and the Coupled Futures of Almonds and Honey Bees: Simulating Impacts of Shifts in Demands and Costs." *American Journal of Agricultural Economics.*

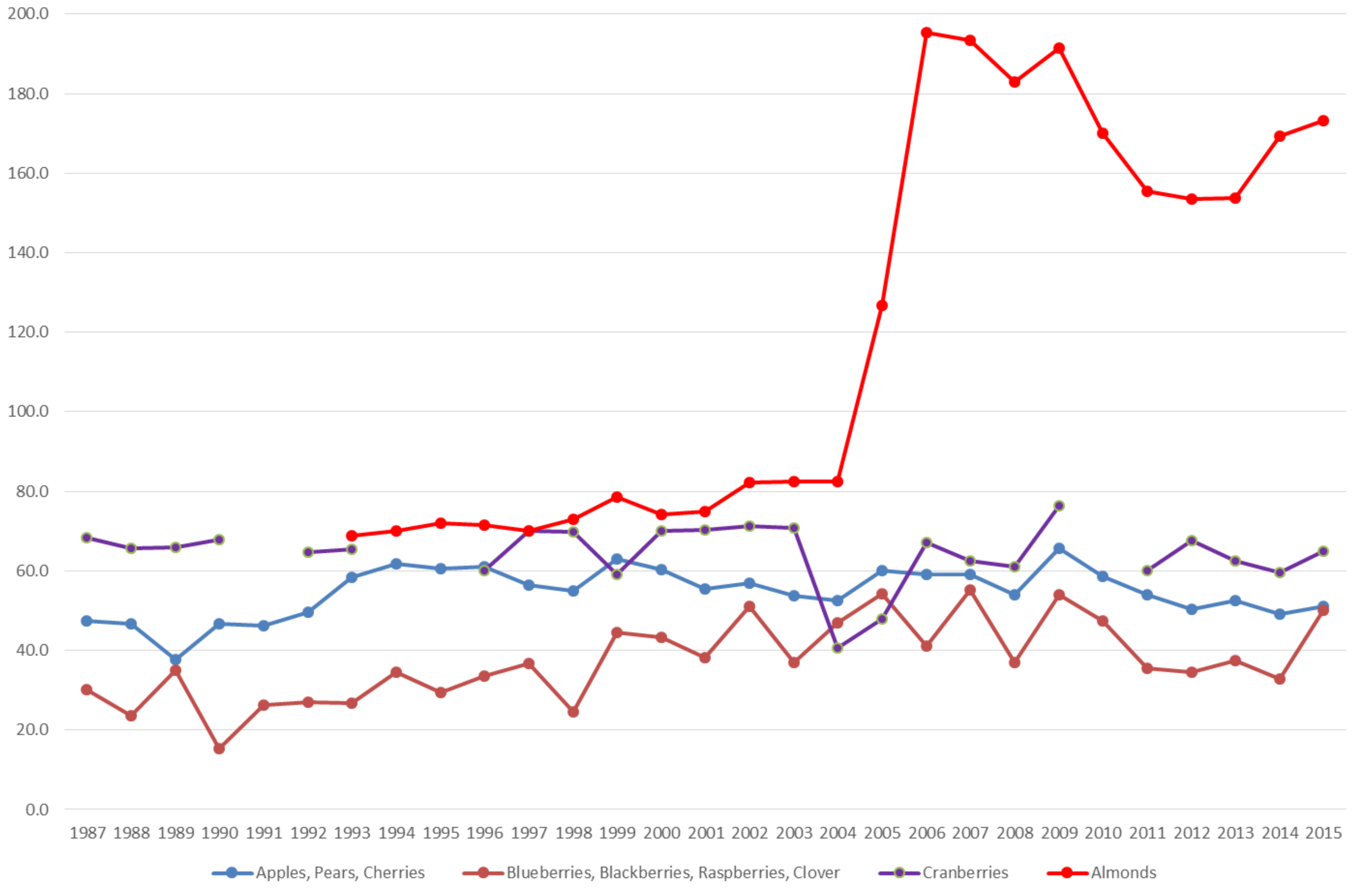
National Agricultural Statistics Service. 2017, California Almond Objective Measurement Report. Sacramento, CA.

Nordhaus, H. (2011) The Beekeeper's Lament: How one man and half a billion bees help feed America, New York: Harper/Perennial.

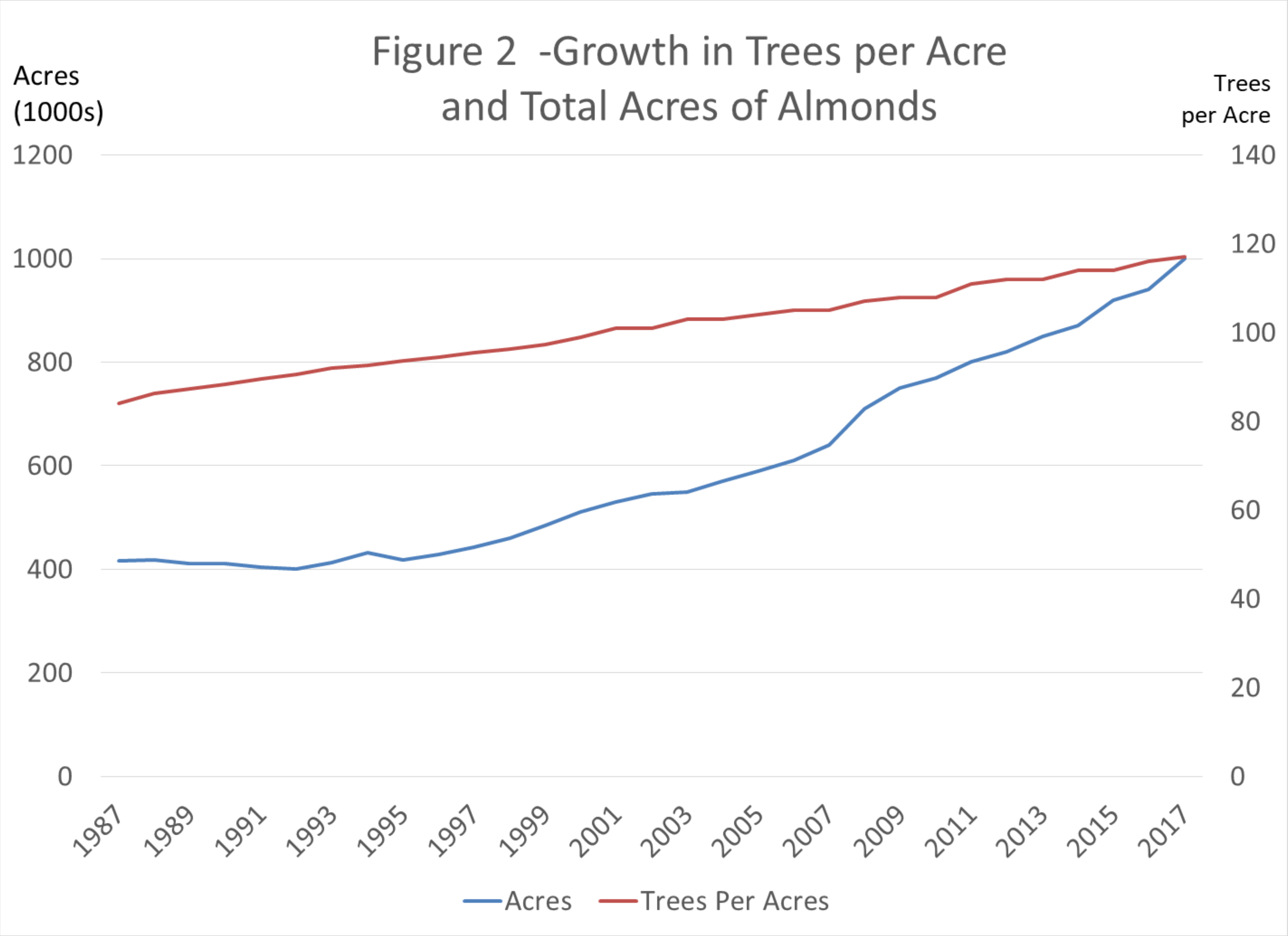
Rucker, R.R., W.N. Thurman, and M. Burgett. 2012. "Honey bee pollination markets and the internalization of reciprocal benefits." *American Journal of Agricultural Economics* 94(4):956-977.

Sagili, R., and D. Caron. 2016. "Honey Bee Pollination in the Pacific Northwest." *American Bee Journal* 156(7).

Figure 1 - Average Real Pollination Service Fees received by Pacific Northwest Beekeepers for Various Crops in dollars per colony

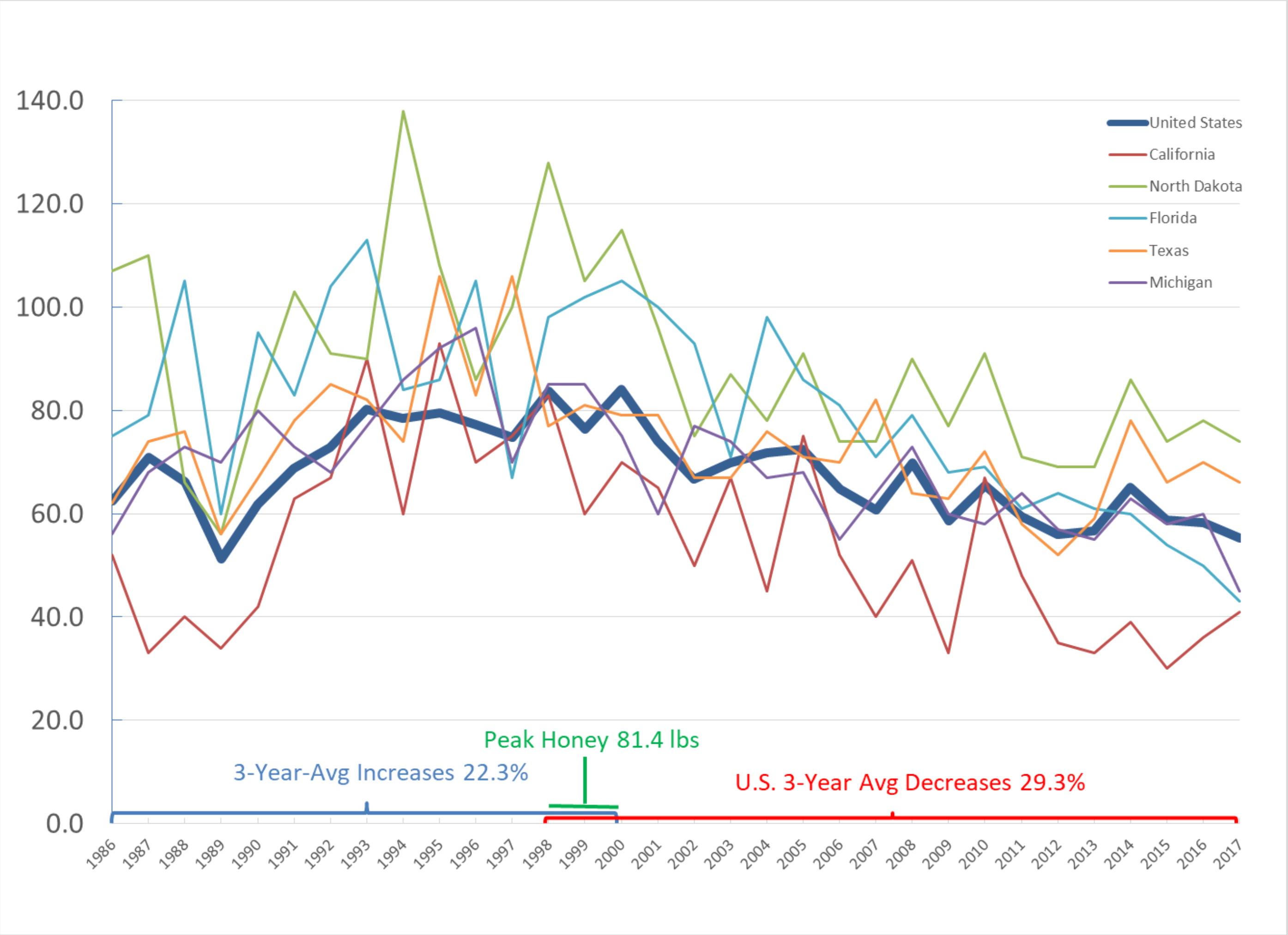


Source: Ferrier et al. (2018) based on Daberkow et al. (2009) and Sagili and Caron (2016)



Source: USDA NASS California Objective Almond Report (2017)

Figure 3 – Honey Yields in the United States and selected states between 1986 and 2017



Source: USDA Economic Research Service Sugar and Sweeteners Yearbook (2018)

Table 1 – Summary Statistics on Beekeeper Data by Size in 2016

		Small <300		Mid 300 to 1000		Large 1000 to 10000		Very Large Greater than 10000		All	
Label	Description	MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD
N	Observations	354		298		714		136		1502	
y	Yield	49.839	28.135	49.546	31.101	54.031	30.051	60.092	28.567	52.7	29.83
size	All colonies	139.28	70.02	576.9	180.66	3640.87	2208.15	23893.9	16225.69	4041.53	8231.81
	Honey producing colonies	96.11	64.53	378.19	214.23	2241.76	1629.5	12541.3	11730.03	2298.91	5003.01
ey	expected yield	52.748	17.109	53.433	15.14	55.435	16.153	61.609	12.412	54.96	16.05
loss	2 quarter loss rate (100%=100)	18.736	17.968	16.466	15.092	11.819	11.13	12.501	13.887	14.31	14.24
	4 quarter loss rate (100%=100)	16.737	14.383	17.932	16.802	14.495	11.948	15.226	13.243	15.77	13.79
loss*size	Loss * Colonies	2121.47	4115.55	6037.8	6869	25746.1	32318.42	136010.67	208488.82	28064.15	79944.51
cali	CA Colony Share (max Q1, Q2)	9.072	25.955	25.466	38.793	43.614	39.708	51.304	40.774	33.67	39.99
move	Avg Colony Mov. (Q1 to Q3)	77.92	399.98	324.18	584.17	533.94	661.39	697.89	786.45	429.09	651.75
prod	Honey Production in Pounds	5302.27	5842.67	20392.7	19788.79	122433.7	118559.17	790954.79	902053.65	135114.21	354376.31
Difference Variables											
diff_y	Diff. in yield (100%=100)	14.95	81.31	98.57	941.74	13.86	92.21	5.85	46.13	30.2	427.12
dloss	Diff. in Loss Rate Across Years	-6.11	24.686	-0.773	22.243	-1.479	15.603	3.199	15.966	-1.594	18.791
dloss*size		-425.49	4755.95	286.75	7220.82	-1960.79	40598.3	37956.89	191888.06	3014.81	70898.43
diff_ac	Diff. in all colonies	-19.13	93.43	-63.98	386.71	-499.11	2328.33	3858.7	12858.59	94.93	4351.69
diff_cali	Diff. of CA Colonies Share	0.733	17.636	-3.429	28.463	-3.743	31.005	-4.285	26.597	-3.07	28.46
diff_ey	Difference of Expected Yield	-1.698	23.382	-2.471	18.663	-2.128	21.615	3.592	19.907	-1.58	21.4
	Total Colonies	49,305		171,916		2,599,581		3,249,570		6,070,378	
	Share of Colonies	0.8%		2.8%		42.8%		53.5%		100.0%	

Table 2 – Regression Results of our three Instrumental Variables on Yield (y)

Parameter	**Estimate**	**Est. St.d Err**	**t-value**	**P-value**
Intercept	8.6485	0.5229	16.5381	0.0000
2016 Dummy	3.2728	0.6144	5.3272	0.0000
2017 Dummy	0.5281	0.8034	0.6573	0.5111
Move	0.0030	0.0004	7.3646	0.0000
DF Model	DF Error	MSE	Root MSE	R-Square
4	2175	897762.3381	947.5032	0.0410

Table 3 - Estimation of the Yield Difference Model in Equation (1)

Parameter		OLS	IV (IT2SLS)
Intercept	Estimate	2.2434	34.4808
	t-value	0.9690	1.5282
	p-value	0.3327	0.1266
Lagged Col. (size in 1000s)	Estimate	-0.058	-1.352
	t-value	-1.1530	-2.0221
	p-value	0.2490	0.0433
Lagged Yield (y_lag)	Estimate	0.4627	0.4365
	t-value	30.8274	12.2101
	p-value	0.0000	0.0000
Expected Yield (ey)	Estimate	0.5410	0.5471
	t-value	14.7807	10.6636
	p-value	0.0000	0.0000
Share moved to CA (Cali)	Estimate	-0.0425	-0.0772
	t-value	-3.4951	-3.7882
	p-value	0.0005	0.0002
Loss rate (loss)	Estimate	-0.2169	-2.7440
	t-value	-3.1547	-1.6562
	p-value	0.0016	0.0978
Loss*size (size in 1000s)	Estimate	0.0076	0.1145
	t-value	2.3777	2.0143
	p-value	0.0175	0.0441
	DF Model	7	7
	DF Error	2169	2169
	SSE	5478954654	8933502806
	MSE	2526027.96	4118719.6
	Root MSE	1589.3483	2029.4629
	R-Square	0.4721	
	Adj R-Sq	0.4706	
Hansen's Test of Instruments between OLS to IT2SLS Estimations			
	d.f	6	
	m-stat	8.7334	
	P-value	0.1891	

Table 4 - Estimation of the Yield Difference Model in Equation (2)

Parameter		OLS	IV (IT2SLS)
Intercept	Estimate	-7.7366	-9.5269
	t-value	-4.4296	-1.4933
	p-value	0.0000	0.1356
Lagged Col. (size in 1000s)	Estimate	0.5972	0.5132
	t-value	7.3591	4.7484
	p-value	0.0000	0.0000
Diff Exp Yield (dey)	Estimate	0.0141	-0.0240
	t-value	0.3103	-0.1704
	p-value	0.7564	0.8647
Diff Cali Share (dcali)	Estimate	0.3962	0.2962
	t-value	6.0888	0.8390
	p-value	0.0000	0.4016
Diff in Loss Rate (sloss)	Estimate	-0.5307	-2.9548
	t-value	-3.5442	-0.3523
	p-value	0.0004	0.7247
Dloss*size (size in 1000s)	Estimate	0.01702	0.11041
	t-value	2.6165	0.3873
	p-value	0.0090	0.6986
	DF Model	6	6
	DF Error	1232	1232
	MSE	9856955	11953693
	Root MSE	3139.57	3457.41
	R-Square	0.0798	
	Adj R-Sq	0.07607	

Hansen's Test of Instruments between OLS to IT2SLS Estimations

	d.f	6
	m-stat	1.8870
	P-value	0.9297

Table 5- Marginal effect of a one percentage point increase in the 2-quarter loss rate

Yield Level Model in Equation 1							
	Colony Counts					Weighted Average	
	Threshold	1,000	5,000	10,000	20,000	As Level	*As Percent*
OLS	28,590	-0.209	-0.179	-0.141	-0.065	-0.059	*-0.11%*
IV	23,957	-2.629	-2.171	-1.599	-0.453	-0.362	*-0.69%*
Yield Difference Model in Equation 2							
	Colony Counts					Weighted Average	
	Threshold	1,000	5,000	10,000	20,000	*As Level*	As Percent
OLS	31,179	-0.51%	-0.45%	-0.36%	-0.19%	*-0.093*	-0.18%
IV	26,763	-2.84%	-2.40%	-1.85%	-0.75%	*-0.347*	-0.66%