# The Evolution of the California Blueberry Industry: A Social Network Analysis Approach

Zoe T. Plakias

Department of Agricultural and Resource Economics,
University of California, Davis

plakias@primal.ucdavis.edu

***Selected Paper prepared for presentation at the 2014 AAEA/EAAE/CAES Joint Symposium: Social Networks, Social Media and the Economics of Food, Montreal, Canada, 29-30 May 2014***

# The Evolution of the California Blueberry Industry: A Social Network Analysis Approach

Zoë Plakias

*Ph.D. Candidate, Department of Agricultural and Resource Economics, University of California, Davis*

E-mail: plakias@primal.ucdavis.edu

May 27, 2014

## Abstract

In this paper, I estimate local and industry peer effects related to the adoption of blueberries as a crop among California growers. I employ publicly available data from the California Department of Pesticide Regulation for the years 2001 through 2011. Geographic and inter-industry network analyses complement the econometric estimation and provide greater insight into the patterns of peer effects. Industry peer effects (i.e. those connections to other growers through growing crops other than blueberries) have a positive and statistically significant effect on the probability of adopting blueberries. Local effects play a significant, though seemingly less important, role. The geographic and social network analyses corroborate these results and provide greater depth. I find this type of analysis can offer some insight about crop adoption without the expense necessary for most social network studies.

## 1 Introduction

Between 2005 and 2012, blueberry consumption in the United States tripled to 1.34 pounds per capita (Perez and Plattner 2013). This shift in demand for blueberries is likely associated with the increased demand for foods with attributes believed to be beneficial to human health (blueberries are known to have high levels of antioxidants). Historically, California has not been a blueberry-growing state. The US Department of Agriculture did not even report blueberry production in California until 2005. But recent years have seen rapid growth in the California blueberry industry; between 2005 and 2012, California saw a fourfold increase in blueberry production, to 40.5 million pounds (ERS 2013). In 2012, 7% of the U.S. blueberry crop was produced in California, making it the 7th largest blueberry producing state in the US (ERS 2013).

It comes as no surprise that an outward shift in the demand for blueberries would lead to increased production of blueberries. The increased demand would lead to a higher price for a given quantity, and the increased profits would induce current producers to expand production and/or new growers to enter. However, basic economic theory does not tell us *how* producers learned about blueberries and decided to adopt them, particularly in California, where blueberries are a relatively unknown crop.

The purpose of this work is to better understand the patterns and mechanisms of adoption of blueberries in California. I have two research questions:

1. How do production of the same crops as established blueberry growers (other than blueberries) and geographic proximity to established blueberry growers affect the probability of a California producer adopting blueberries?

2. What are the geographic and crop groupings that define the networks that are formed as a result of blueberry adoption?

To answer these questions, I construct a model of local and industry peer effects. I estimate my empirical model using publicly available data on location and crop choice for growers in California for the years 2001-2011. I also provide visualizations of the networks that developed within the blueberry industry in these years and identify cliques to better understand the specific (other) crops and regions that link producers in the blueberry industry.

I find that the industry peer effect is both positive and significant in all years for potential first adopters. The local effect varies in significance and sign, with significance primarily in the later 2000s, suggesting the local effect is more important during this period. However, even when significant, the sign of the effect varies by year. These results indicate that industry influences are very important, and local influences can also be a factor, with both past successes and past failures potentially driving adoption rates. From the analysis of the geographic networks, it appears there may have been experimentation that occurred in both the far northern and southern regions of California to determine the economically and agronomically optimal growing region for blueberries in the state. This experimentation is suggested by the fact that the growers producing in the northernmost and southernmost regions of the state each produced only in the middle of the study period, starting production later than the first adopters and ceasing production by 2011. The analyses of the inter-industry networks suggest that the blueberry network is tied to the berry, stone fruit, nut and olive, and grape industries through growers' cropping patterns.

I make several contributions with this work. First of all, I consider an industry peer effect using crop choice patterns. Although there is a wealth of peer effects literature relating to technology adoption in agriculture, to my knowledge this is the first use of crop choice to understand industry peer effects. Secondly, I utilize multiple methodologies; I complement the peer effects estimation with geographical and industry network analyses to provide a more nuanced view of the peer effect estimation results. My final contribution is my demonstration that in cases where basic grower data are available, some initial insights can be gained about social networks prior to social network survey data collection. These initial analyses can help inform later data collection activities.

In section 2, I discuss how my work relates to the existing literature. In section 3, I describe my data. In section 4, I present my model. In section 5, I outline the estimation procedure. In section 6, I discuss my estimation results. In section 7, I provide geographic and social network analyses of my data to complement the estimation results. Finally, Section 8 concludes and gives an overview of the next steps in my work.

## 2 Related Literature

Sociologists in the mid 20th century that saw and began to write about the importance of peer effects and social learning in agricultural technology adoption. Work by Ryan and Gross (1943) on hybrid corn adoption in Iowa predated the seminal economics work of Griliches (1957) by more than a decade. Griliches suggested that adoption patterns follow "an adjustment path, moving more or less consistently towards a new equilibrium position." He also suggested that location mattered, due to the need for "availability" of the technology before adoption could occur, and the "acceptance problem"—that is, behavioral differences between producers when it comes to technology adoption. My work in this paper is consistent with an "epidemic" or "learning" model of technology diffusion, first described by Mansfield (1961), although not in those terms. Many studies have followed Griliches and Mansfield, building on their ideas.

My work is most related to three areas of the literature on agricultural technology adoption that followed from the early work of Griliches and Mansfield: social learning, peer effects and spatial effects. Most of the work in this area of the literature has been conducted in developing countries, although the same principles apply in California. The first area, social learning, generally assumes a specific avenue for or structure of the influence of one economic actor on another. Foster and Rosenzweig (1995) use a target input model and assume learning by doing with neighbor effects. Estimating their model using data on the adoption of high yielding varieties of several crops in India, they find that lack of knowledge about how to produce the new variety is a major barrier to adoption and that having an experience neighborhood makes a new adopter considerably more profitable, suggesting the presence and importance of social learning in this setting. Munshi (2004) also studied high yield variety adoption in India and used a similar model, but introduced an additional component of heterogeneity. He showed that for those crops sensitive to minor fluctuations in certain key inputs (such as soil characteristics), heterogeneity across farms may make a neighbor's experience less relevant to a grower and diminish the role of social learning.

In a more recent paper, Conley and Udry (2010) eschew the assumption of a specific structure for social learning, instead testing a set of hypotheses derived from assumptions about the behavior of Ghanaian pineapple producers regarding fertilizer use. In their work, they discuss the importance of appropriately characterizing and identifying producers' "information neighbors." In this setting, an "information neighbor" is a person whom a producer exchanges information with about agricultural production. In this region of Ghana, an "information neighbor" is likely be a friend or family member, but this idea also applies in a developed setting where producers may interact as acquaintances or colleagues within an industry. However, in a developed region like California, with high literacy rates and internet connectivity, as well as more developed infrastructure of all kinds, an "information neighbor" may well be someone outside a grower's geographic area. Regardless of the differences, this term is useful for helping to frame the technology adoption decision faced by potential adopters of California blueberries.

Although social learning is certainly a type of peer effect, there is also a peer effects literature that does not assume a specific form or structure for learning, but instead simply that the average characteristics of an individual's peer group have an effect on some specific outcome for the individual. Much of this literature is solely econometric and does not include a theoretical component. This approach is ess prevalent in the agricultural technology adoption literature. Bandiera and Rasul (2006), who look at the adoption of sunflowers as a crop in Mozambique, employ a theoretical model of social learning but estimate an empirical model that is quite simple. They ask how many people the farmer knows that are planting the crop, and of these, how many are the farmer's family and friends. They find that farmers are more likely to adopt sunflowers when some family members and friends adopt but less likely to adopt if many family members and friends adopt. Like Conley and Udry (2010), they highlight the importance of choosing the peer group carefully, as the results will be sensitive to the way peer groups are defined in the study. Oster and Thornton (2009) consider how peer effects influence the take-up of a non-agricultural technology – menstrual cups – among women in Nepal. Using data from a randomized control trial, they look at the effect of having peers who are in the treatment group.

Although many of the works already cited have some sort of implicit spatial component, in some cases it may be appropriate to model space explicitly. Case (1992) uses a spatial autoregressive model with a farmer's peers as his neighbors and considers adoption of a particular agricultural technology, a rice harvesting tool in Java, Indonesia. In Java, we expect rice to be a staple crop, so a farmer's neighbors are also his peers in a common industry. In a region with such crop diversity as California, we might expect geographic proximity to play less of a role than participation in a common industry. I examine this hypothesis using a spatially-defined peer adoption probability as an explanatory variable.

## 3 Data

I use data from the Pesticide Use Reporting data set, publicly available from the California Department of Pesticide Regulation (DPR 2013). These data include the field location and treated crop of every pesticide application (organic and conventional) in California between 1990 and 2012. The data set also includes a unique grower identification number for each application. Thus, the data I employ are grower-level, region-level crop mix data for each year. The possible regional levels of analysis include Public Land Survey System (PLSS) townships and counties. The PLSS is a grid system covering the entire United States and serves as the basis for the county road structure and the dividing line between many fields in California. However, at the township level of resolution (36 square miles), the variation in the number blueberry growers is minimal due to the small number of blueberry growers. Therefore I estimate the model at the county level.[1]

**Table 1: Summary Statistics**

| *1. Year* | *2. CA Growers (All Crops)* | *3. No. in Multiple Counties* | *4. Mean No. of Crops* | *5. CA Growers (Blueberries)* | *6. No. in Multiple Counties* | *7. Mean No. of Crops* | *8. New CA Blueberry Growers* |
|---|---|---|---|---|---|---|---|
| 2000 | 21924 | 766 | 1.929 (1.825) | 19 | 1 | 3.947 (4.075) | — |
| 2001 | 21061 | 728 | 1.889 (1.800) | 30 | 3 | 3.600 (3.103) | 17 |
| 2002 | 20842 | 714 | 1.886 (1.779) | 32 | 5 | 3.531 (2.615) | 15 |
| 2003 | 21130 | 727 | 1.884 (1.753) | 33 | 4 | 3.636 (3.200) | 15 |
| 2004 | 20826 | 710 | 1.875 (1.753) | 41 | 3 | 2.585 (1.483) | 19 |
| 2005 | 20649 | 704 | 1.861 (1.729) | 48 | 2 | 3.625 (3.618) | 22 |
| 2006 | 20544 | 660 | 1.866 (1.746) | 52 | 3 | 3.904 (3.942) | 19 |
| 2007 | 19995 | 712 | 1.864 (1.721) | 61 | 2 | 3.443 (3.922) | 22 |
| 2008 | 19525 | 673 | 1.886 (1.745) | 70 | 4 | 4.071 (5.162) | 26 |
| 2009 | 16933 | 528 | 1.867 (1.727) | 68 | 2 | 3.515 (4.237) | 20 |
| 2010 | 16965 | 510 | 1.889 (1.790) | 72 | 5 | 4.153 (5.194) | 21 |
| 2011 | 17084 | 522 | 1.907 (1.805) | 73 | 2 | 3.822 (4.843) | 20 |

Summary statistics are provided in Table 1. Column 2 lists the number of California growers who could potentially grow blueberries. This number includes all California growers with cropland who reported applying pesticides of any kind to cropland in the given year. Column 3 reports the number of the growers in Column 2 who were operating in multiple counties that year. Column

[1]There are 58 counties in California. Over the course of the sample the number of counties with blueberry production ranged from 13 (in 2000) to 22 (in 2007).

4 reports the mean number of crop produced by the growers in Column 2. Column 5 reports the number of blueberry growers in California in the given year. Columns 6 and 7 are analogous to columns 3 and 4. Finally, column 8 reports the number of blueberry growers who were not reported as growing blueberries in the previous year.

# 4 Model

To answer my research questions, I model the choice to adopt blueberries as:

$$y_{i,r,t} = \beta_1 y_{i,r,t-1} + \beta_2 \bar{y}_{-i,r,t-1} + \beta_3 \frac{\Sigma_{p_j \in P_i} \; y_{-i,p_j,t-1}}{\Sigma_{p_j \in P_i} \; n_{-i,p_j,t-1}} + \beta_4 Z_r, \tag{4.1}$$

where $y_{i,r,t}$ is a binary outcome variable indicating the adoption choice of producer $i$ in region $r$ at time $t$, $y_{i,r,t-1}$ is the producer's adoption choice in the previous period, $\bar{y}_{-i,r,t-1}$ is the mean adoption choice of producers in the same region (excluding individual $i$), the second to last term is the mean adoption choice of producers in the industry peer group (again excluding individual $i$), and $Z_i$ is a vector of regional characteristics.

Grower $i$'s regional peers include any growers that grew any crops in the same county (or counties) that grower $i$ grew crops in during year $t-1$. Grower $i$'s industry peers include any growers that grew in $t-1$ any of the crops that grower $i$ grew in $t-1$ and grew blueberries. Hence, peer groups are not necessarily symmetric. For example, if grower $i$ grew lemons in year $t-1$ and grower $j$ grew lemons and avocados in year $t-1$, grower $j$ would be in grower $i$'s peer group (along with all other lemon growers) in year $t$ and grower $i$ would be in grower $j$'s peer group (along with all other lemon growers) in year $t$. However grower $j$'s peer group in year $t$ would also include all growers who produced avocados in year $t-1$.

I use a peer effects framework similar to that employed by De Giorgi et al. (2010). For identification, they exploited the fact that peer groups were not fully overlapping for all individuals. In their case, the authors looked at students who attended the same classes. Because groups of students do not all attend the same classes, the peer groups for each student varied, mitigating the "reflection" problem identified by Manski (1993). The "reflection" problem is the idea that a peer effect is the aggregation of individuals' effects, so the peer effect may simply be a "reflection" of the individual effects and not actually a peer effect. In addition, I employ lagged peer effects in my analysis, which also helps to mitigate the reflection problem, as we would not expect a producer's adoption choice in time $t$ to influence others' choices in time $t-1$.

Another problem common to this type of analysis was described by Moffitt (2001). In particular, he highlighted that with most peer effects models it is not possible to disentangle the effects of social interactions from the effects of other unobserved variables when both are correlated with group outcomes. As with many peer effects studies, that is the case for the model I have presented. Social interactions are far from the only factors at play here. Other factors include agronomic, climatic, economic and behavioral factors. Nonetheless, this common problem does not eliminate the value of the analysis. We can learn a great deal about which groups of factors might be the most important from observing the patterns of these effects over time in conjunction with geographic and industry network analyses.

# 5 Estimation

I employ a logit framework to estimate the model. Although the data are constructed as an unbalanced panel, there is insufficient within variation in the dependent variable to use panel estimation techniques. I thus estimate eleven annual logit models for the years 2001-2011. To construct the industry peer group, I dropped all non-agricultural pesticide applications in the data while retaining all other agriculture, including livestock and research crops. This broad definition of crop allows peer effects to work through a variety of channels.

If a grower operated in multiple counties, the grower's local peer group is constructed by taking the sum of all adoption decisions of growers in all counties in which the grower operated during that year and dividing it by the total number of growers in those counties. Thus, if grower $i$ and grower $j$ both operates in County A, but grower $i$ also operates in a County B, the local peer effects will be different for growers $i$ and $j$ and will reflect the influence of peers in all counties in which they operate. By constructing the local peer group this way, I am treating a grower-county as a unique individual and I am counting that grower's adoption decision multiple times (once for each county in which the grower is operating in each year in which he/she is operating there). The effect of constructing the data this way is to weight the adoption decisions of growers who operate in multiple counties more heavily.

The construction of the industry peer variable is similar. In this case, I use grower-crop years. Thus, a producer is only counted once for each crop in each year, regardless of the number of counties in which she produces that crop. Again by treating a grower-crop as a unique individual, I am counting a grower's decision multiple times (once for each crop the grower produces in each year). This weights the adoption decisions of producers with a larger crop mix more heavily.

For both the local and industry peer variables, the effect of the variable construction is that peers who overlap with a grower both by county and by crop, or in multiple counties or multiple crops, are counted more times. In social network analysis parlance, this overlap between growers can be interpreted as "tie strength." My construction gives more weight to peers with stronger ties to the grower.

The inclusion of lagged, rather than contemporaneous, peer effects is logical in the case of agriculture because producers make their growing decisions well in advance of the start of the growing season and there is ample evidence that many producers wait for other producers to test out new crops to learn from their successes and failures. However, since the blueberry industry in California is still rather nascent, current blueberry producers could still be considered early adopters. If this is true, a contemporaneous peer variable may be more logical.

In my estimation, I also employ one grower characteristic – whether or not the grower's home county is on the Central Coast.[2] This indicator variable serves as a proxy for blueberry price, which is not available at the county level for most counties in the sample. Through conversations with University of California Cooperative Extension (UCCE) specialists, I learned that Central Coast blueberries are harvested earlier in the season than the rest of the blueberries in California and earn a significantly higher price. What limited data are available corroborate this finding.[3] The variable is equal to 1 when a grower's home county is Santa Cruz, Monterey, San Luis Obispo,

[2]To ensure the dependent variable was unique for each grower-year, I identified a home county for each grower. Although the home county was listed for most growers, some home counties were missing. If no home county was listed but a producer grew in one county only for the entire 12 years, I assumed this county was the home county. However, if a grower grew in multiple counties and no home county was listed in the data, I dropped the observation (fewer than 1,000 observations were dropped for this reason over the entire sample period).

[3]The Agricultural Market Service Market News database reports blueberry price for Fresno only. Furthermore, for most counties in California, blueberries are not a valuable crop relative to others in the county to warrant inclusion in a separate line of the County Agricultural Commissioner's report; blueberries are lumped in with 'other fruit.' Recently, as the blueberry market has grown, some counties have begun reporting prices. These numbers support the idea that the Coastal producers get a considerably higher price for their blueberries than the Central Valley producers

Santa Barbara or Ventura.

Along with price, another factor we might expect to affect a producer's adoption choice is the expected profitability of outside options in time $t$. Given the many crops that California producers grow, it would be difficult to obtain data to control for this factor. However, because I use lagged rather than contemporaneous peer effect variables, omitted variable bias owing to the omission is only a concern if we think the expected profitability of the outside option in time $t$ to is correlated with producers' adoption choices in time $t-1$. Given that the adoption choice in time $t-1$ is made using information from time $t-2$ and expectations about time $t-1$, this is unlikely to be a major problem. However, to the extent that one thinks relative profitability of crops remains fairly constant across time periods, omitted variable bias could indeed be a concern.

## 6 Estimation Results

I estimate the model twice, once including all growers in all years, and once excluding growers who grew blueberries in the previous year. The results of the empirical estimation with all growers in all years are provided in Table 1, and the results of the estimation excluding growers who grew blueberries in the previous year are provided in Table 2.

In Table 2, the values in Column 1 indicate the difference in the marginal effect between those growers who produced blueberries in the previous year and those who did not on the probability of producing blueberries in the year indicated. The marginal effects are large in magnitude, positive and significant in all years. The values in Column 2 of Table 2 indicate the average marginal effect of an increase of one percentage point in the average probability of growers in one's local peer group growing blueberries in the previous period. These effects are statistically significant and positive, although small in magnitude, for several years in the early 2000s. The values in Column 3 indicate the average marginal effect of an increase of one percentage point in the average probability of growers in one's industry peer group growing blueberries in the previous period. These effects are statistically significant and positive (except for one year) in the later 2000s, and once in the early

**Table 2: Average marginal effects (all growers)**

| *Year* | *1. Grew Last Year* | *2. Industry Peers* | *3. Local Peers* | *4. Central Coast* | *N* |
|---|---|---|---|---|---|
| 2001 | 0.77603*** | 0.00004*** | 0.00044** | 0.00015 | 16764 |
| 2002 | 0.72565*** | 0.00001 | -0.00079 | 0.00004 | 16297 |
| 2003 | 0.63815*** | 0.00002 | 0.00150** | -0.00087*** | 16350 |
| 2004 | 0.67028*** | 0.00018*** | -0.00109 | 0.00156 | 16529 |
| 2005 | 0.70474*** | 0.00001 | -0.00004 | -0.00083 | 16309 |
| 2006 | 0.74098*** | 0.00004 | 0.00079** | 0.00029 | 16332 |
| 2007 | 0.72197*** | 0.00003 | 0.00121*** | -0.00054 | 16138 |
| 2008 | 0.78413*** | 0.00001 | 0.00150*** | 0.00204*** | 15529 |
| 2009 | 0.84015*** | 0.00003 | -0.00112* | -0.00141 | 13631 |
| 2010 | 0.86877*** | 0.00004 | 0.00104 | 0.00417** | 13568 |
| 2011 | 0.80271*** | 0.00002 | -0.00006 | -0.00016 | 13682 |

* $p < 0.1$    ** $p < 0.05$    *** $p < 0.01$

2000s, and larger in magnitude than the industry peer effects. Finally, the values in column 4 indicate the average marginal effect of growing in one of the counties on the Central Coast. This effect is significant and negative for one year in the early 2000s and significant and positive for two non-consecutive years in the later 2000s.

The values in Column 1 of Table 3 indicate the average marginal effect of an increase of of one percentage point in the average probability of growers in one's local peer group growing blueberries in the previous period. The marginal effects are positive and significant in all years and appear to be larger in magnitude than the same column in Table 2 for most years. The values in Column 2 indicate the average marginal effect of an increase of one percentage point in the average probability of growers in one's industry peer group growing blueberries in the previous period. Similar to the results in Table 2, these effects are statistically significant and (positive in two years and negative in two years) in the later 2000s, statistically significant and positive once in the early 2000s, and range in magnitude.

**Table 3: Average marginal effects (potential first adopters)**

| *Year* | *1. Industry Peers* | *2. Local Peers* | *3. Central Coast* | *N* |
|---|---|---|---|---|
| 2001 | 0.00011*** | 0.00044** | 0.00045 | 16749 |
| 2002 | 0.00019** | -0.00011 | -0.00008 | 16274 |
| 2003 | 0.00006* | 0.00031 | — | 16325 |
| 2004 | 0.00013** | -0.00084 | 0.00184 | 16501 |
| 2005 | 0.00009 | -0.00052 | 0.00149 | 16273 |
| 2006 | 0.00009** | 0.00048 | 0.00098 | 16292 |
| 2007 | 0.00017*** | 0.00094*** | 0.00059 | 16091 |
| 2008 | 0.00032*** | 0.00084** | 0.00198** | 15477 |
| 2009 | 0.00050*** | -0.00118*** | 0.00084** | 13574 |
| 2010 | 0.00016** | 0.00031 | 0.00535*** | 13513 |
| 2011 | 0.00025** | -0.00152** | 0.01573 | 13620 |

* $p < 0.1$    ** $p < 0.05$    *** $p < 0.01$

Combined, these results suggest that many growers who tried blueberries were successful (demonstrated by the fact they continued to produce blueberries). In addition, the success or failures (indicated by the negative peer effect) of nearby growers may have been a factor. Although we might be concerned that results related to a local peer effect are the result of the "reflection" problem highlighted by Manski (1993), the similar results related to the local peer effect in Table 3 should quell these fears. Growers in Table 3 did not produce blueberries the previous year and therefore any changes in their probability of growing blueberries could not have been the result of their own experience with blueberries. Finally, comparing the industry effects in Tables 2 and 3, we find that the industry peer effect may well be important, particularly for first adopters. This could be due to similarities in production techniques, marketing channels and/or needed agronomic conditions among certain industries. I consider these industry networks in more detail in the next section.

## 6.1 Caveats

Although I have estimated a peer effects model, it is important to note that the interpretation of my results should not be limited to a peer effects story. Growers in the same location or growing the same other crops might choose to grow blueberries for many reasons without ever interacting. For instance, two growers in the same region are likely facing a similar crop choice set based on the agronomic conditions in their region. This could lead them to optimize the same way, even if they never interacted. Similarly, two growers growing the same other crops may face the same economic conditions, including: profitability of blueberries relative to their other crops and potential benefits from economies of scale or scope in production or marketing. Although there is still a compelling story to be told without controlling for these factors—since the industry and local peer effects suggest the increased relative importance of different factors at different times—I will attempt to explore these possible factors more in future iterations of this paper.

# 7 Peer Networks

The results of the econometric estimation suggest the industry peer effect and, to a lesser extent, the local peer effect, are important factors in blueberry adoption. Given the small number of blueberry growers, the geographic patterns are perhaps best understood using maps. The crop patterns, on the other hand, can be understood using social network analysis. In this section I present maps of California throughout the sample period to understand the evolution of geographic patterns of adoption and visualizations and cluster analyses of crop networks to better understand the inter-industry relationships among blueberry growers.

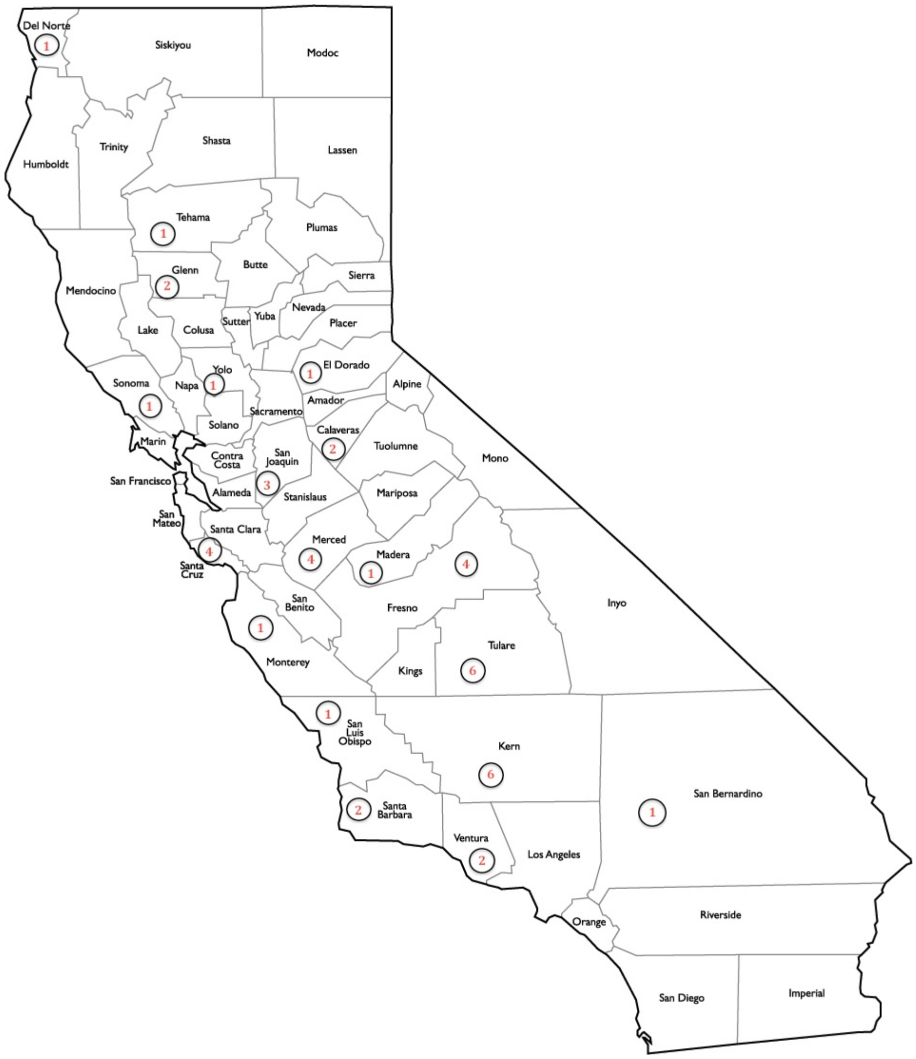
## 7.1 Geographic Networks

Figure 1 shows the counties where blueberry production occurred in the year 2000 and the number of growers in each county. In 2000, early adopters of blueberries were scattered mostly in the Central Coast region and the southern Central Valley.

Figure 2 provides the same information for 2004. By this time, the Central Coast and southern Central Valley regions had more adopters and appear to be establishing themselves as the primary growing regions in the state. However, blueberries are now seen further north in the Central Valley, and near the Oregon border. These results suggest favorable economic conditions warranting experimentation with blueberries, increased knowledge of blueberries as a crop, and growers in the north witnessing the economic success of blueberries further south. Favorable economic conditions could include high prices for blueberries, decreases in profitability of substitutes in production, the realization of economies of scope based on the complementary seasonality of crops, increased market access/decreased transaction costs.

**Figure 1. Blueberry Growers in 2000**

**Figure 2. Blueberry Growers in 2004**

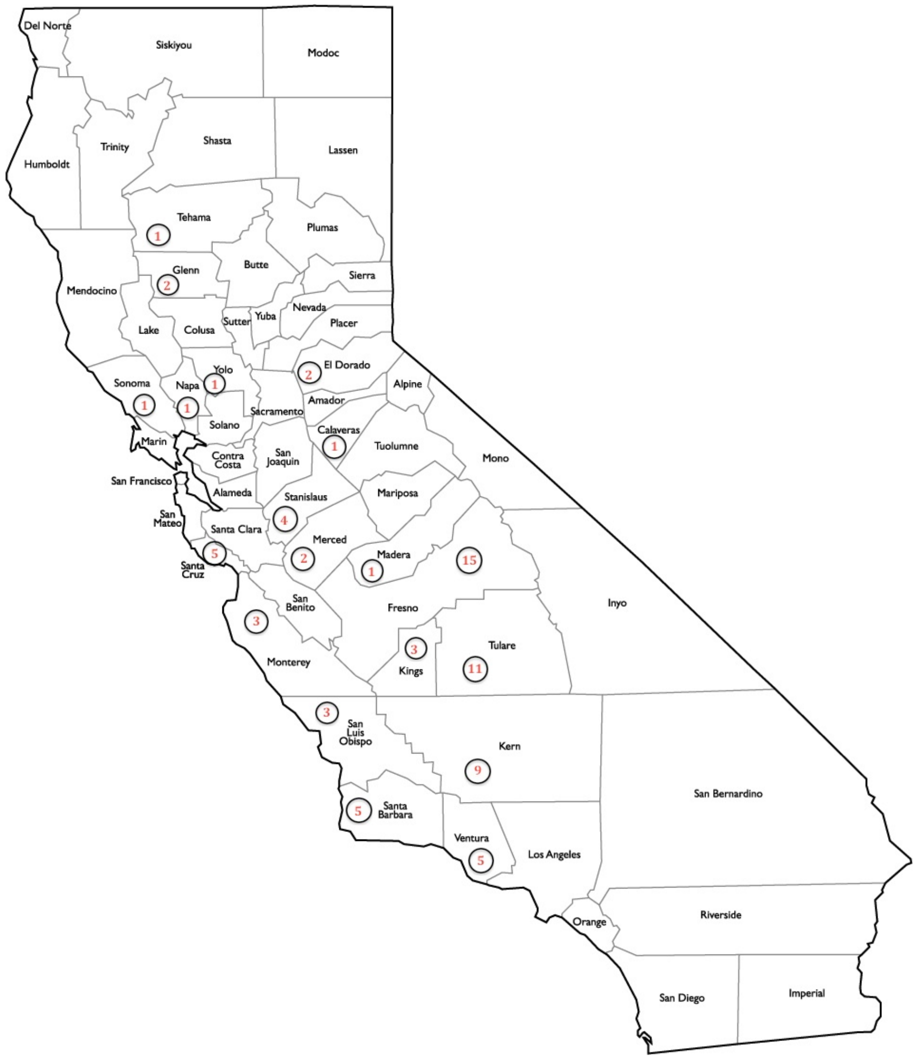



The year 2008 is presented in Figure 3. In 2008, we see more growers in the Delta Region of the Central Valley. Furthermore, no blueberries are produced in Del Norte or Humboldt counties in the north, but now southern California growers in Riverside and Orange Counties appear to be experimenting with blueberries.

**Figure 3. Blueberry Growers in 2008**


Finally, grower adoption in 2011, the last year of my sample, is presented in Figure 4. Here it appears that what may have been experiments at both the far northern and far southern ends of California (in 2006, there were two growers producing as far south as San Diego County) have ceased. The numbers of producers have slowly increased over the course of the sample, and all production is occurring within the Central Coast and Central Valley regions, with the largest number of growers producing in the southern Central Valley.

**Figure 4. Blueberry Growers in 2011**




It seems clear from the number of growers in the primary growing regions that geography played an important role in adoption. Geography may have played a role in a number of ways. First of all, University of California Cooperative Extension (UCCE) researchers found these regions to be well suited agronomically and climatically to the production of blueberries (Jimenez 2005). This would lead to higher yields than in other regions, making blueberries more profitable to these growers relative to growers in other regions. From the geographic patterns alone, we cannot identify how information might have been shared among growers. Doing so would require survey data. Nevertheless, it's clear that other growers in these successful regions learned about blueberries somehow and it's likely they observed and learned from nearby growers' successes and failures. Overall, the pattern of adoption suggests early adopters producing in the regions best suited to blueberries (economically and agronomically), a period of experimentation to determine the geographic limits, north and south, of the most suitable production region, and a subsequent geographic contraction of production into the most productive regions. No doubt more growers will experiment with blueberries in the future. These early successes and failures, as well as trials and reports by state agricultural specialists, have likely helped to provide a more rich information environment for potential adopters going forward.

### 7.2 Inter-Industry Networks

The results of the econometric estimation suggested that industry connections outside of blueberries played an important role in adoption. Figure 5 shows the industry connections in 2000. Each circle represents a county, and each square represents a category of crops. An edge or tie between a grower and a crop category indicates that a grower in the county indicated grew a crop in that category in the labeled year.[4] The size of the circle corresponds to the number of blueberry growers in the county, and the width of the line connecting the county to the crop corresponds to the number of growers in that county growing a crop in the indicated crop category. These representations are thus blueberry egonets of a much larger industry affiliation network (with counties as nodes and crop categories grown as events). Since by definition all growers represented in the figure produce blueberries, I have not shown these edges (or the event of blueberries) for ease of viewing. Thus, any counties who are isolates (i.e. not connected to crops) are counties in which blueberry producers are producing only blueberries.[5] Any crop categories that are isolates are not produced by growers who also grow blueberries in that year.

[4]It is important to note that these are not the same peer groups as those used in the econometric estimation. In the estimation, the peer group included non-adopters and the crop choices for those peers were made in the previous period. The networks presented here indicate contemporaneous crop choice for only those growers who produced blueberries in a given year.

[5]Isolate counties are indicated by circles along the side of the networks that are not connected by an edge to crop categories. Larger circles here indicate multiple growers in the county, all producing only blueberries. However, it's not likely that many growers produce blueberries as their only crop. These isolates may exist because in cases where grower ID was not available, a pesticide application license ID variable was used that may have connected to only one of a grower's crops. Thus, the growers in the counties which are isolates may be producing other crops. Nevertheless, definite patterns are discernible with the data we do have.

**Figure 5. Inter-Industry Networks Among Blueberry Growers in 2000**

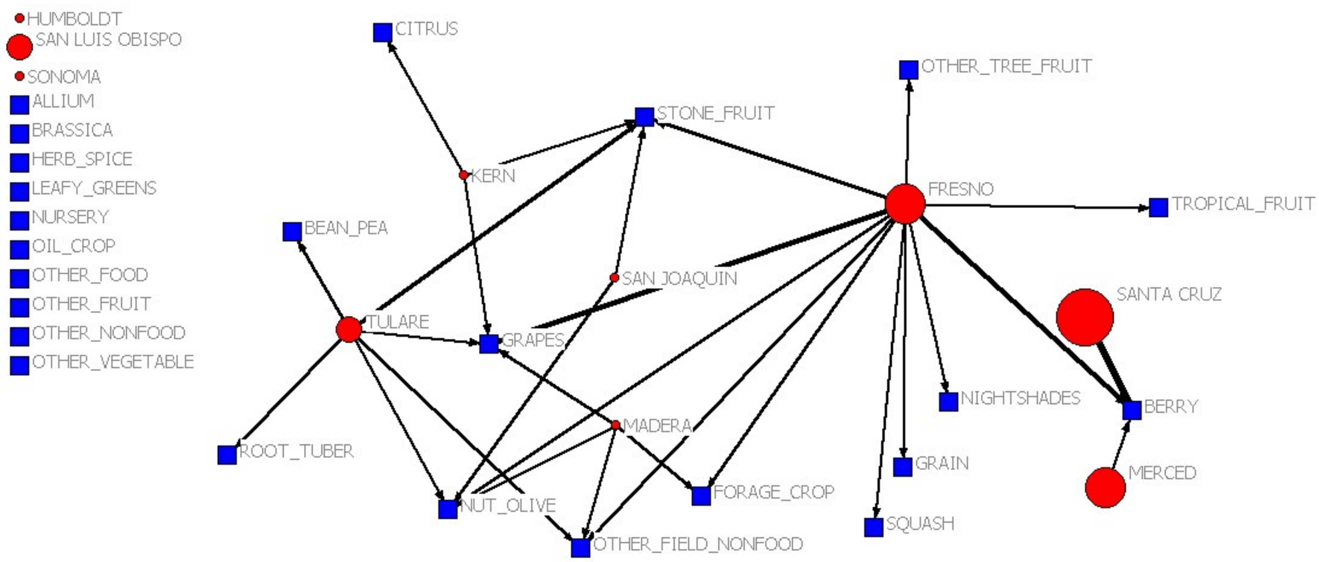



**Figure 6. Inter-Industry Networks Among Blueberry Growers in 2004**

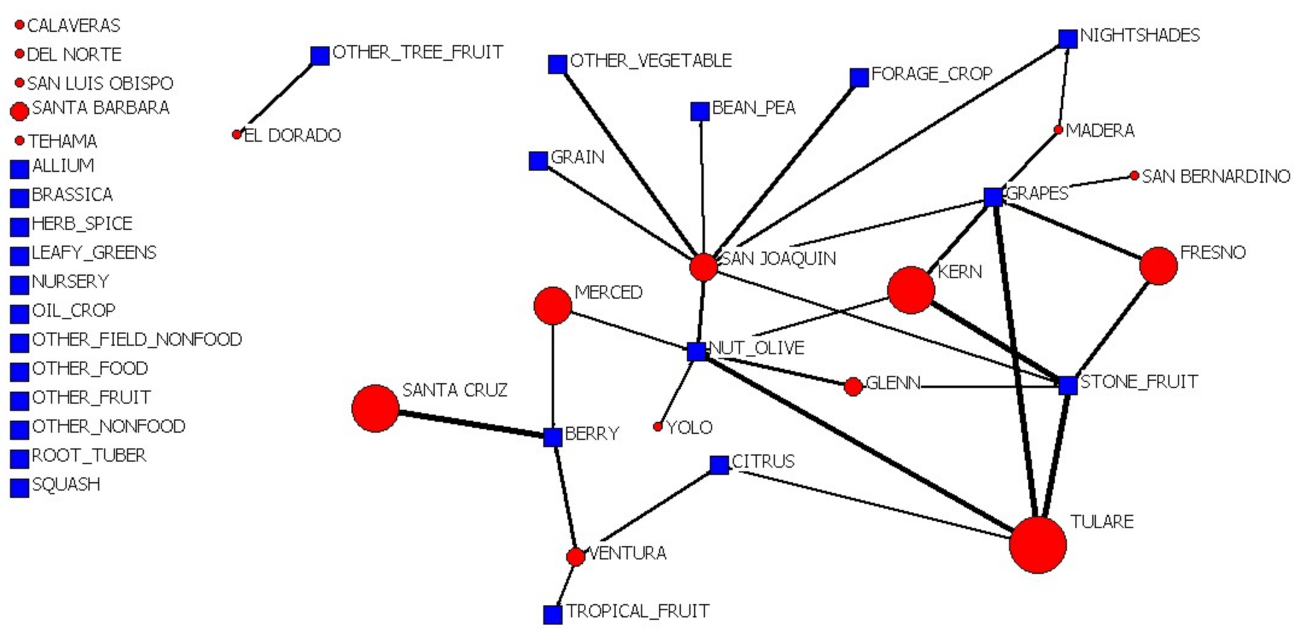


**Figure 7. Inter-Industry Networks Among Blueberry Growers in 2008**

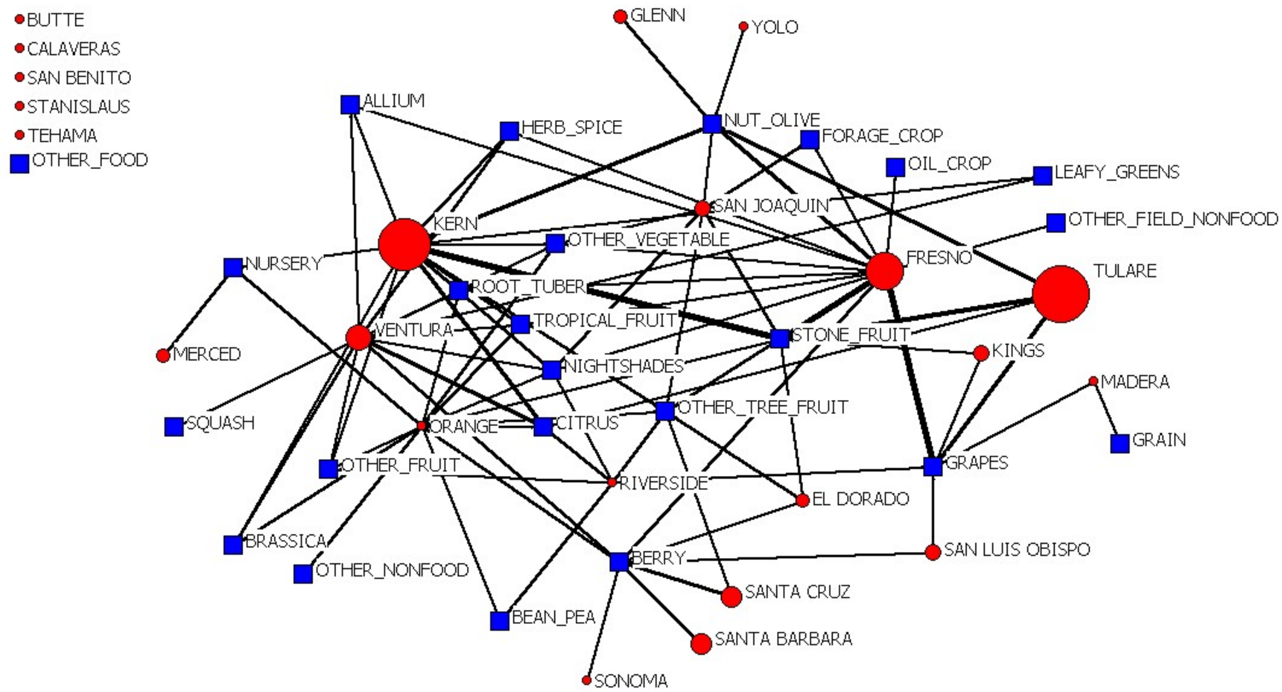
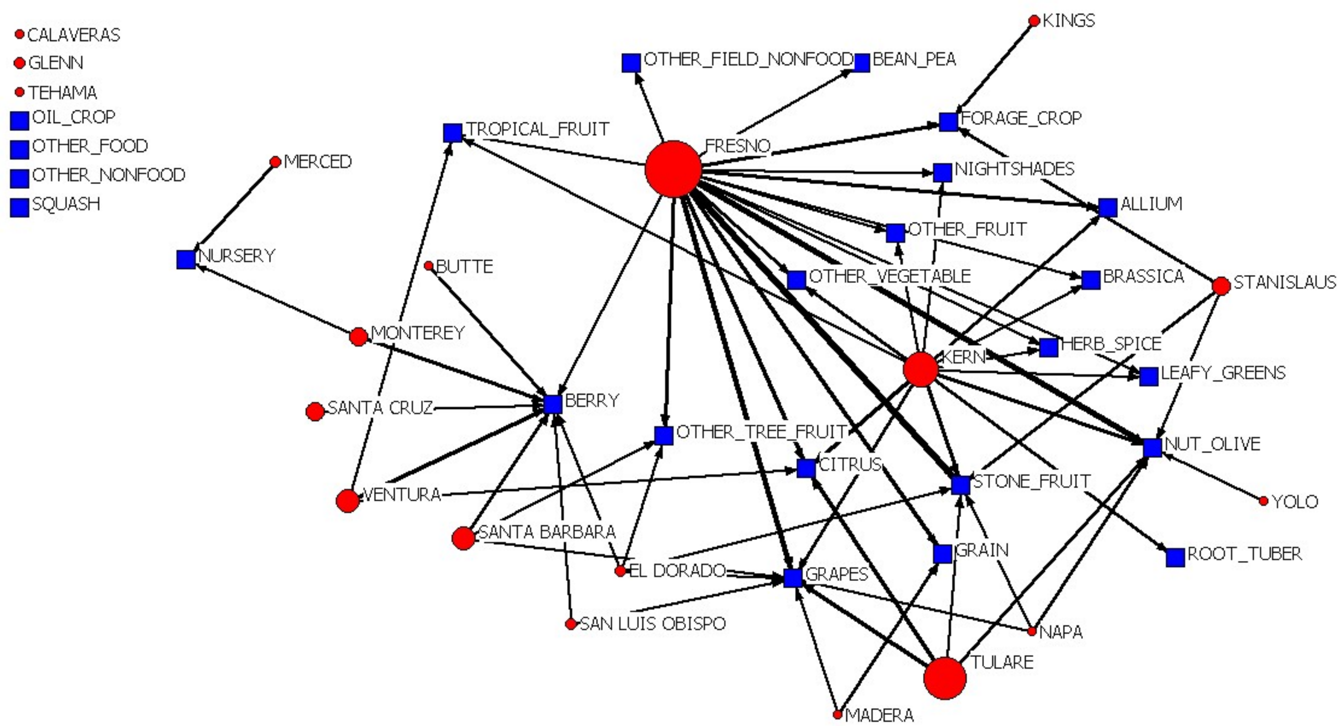



**Figure 8. Inter-Industry Networks Among Blueberry Growers in 2011**



These network visualizations indicate that growers in the counties with the most blueberry producers grow a diverse array of crops, indicating the importance of geography in predicting adoption patterns. However, it's also apparent that crop type and geography interacts to play an important role. Stone fruits, grapes, berries, nuts and olives are all grown by blueberry growers in multiple counties, emphasizing the importance of the connections across industries, and not just space. To complement the information provided by these visualized networks, I also measure network degree centralization and density. Network degree centralization tells us the variation in the number of connections between growers. Network degree centralization of 100% would indicate that $n-1$ growers were producing blueberries and growing one other crop produced by the remaining producer but by no other growers. This network would look like a star with a single grower at the center. Alternatively, zero network degree centralization is achieved when all growers have the same number of connections with other growers (for instance, if they are in a circle). In this context, lower network degree centralization would perhaps indicate a uniform peer effect across individuals. Density is a related term that tells us the ratio of the number of ties between growers out of the total possible ties. Thus, density would is zero for a grower if he or she only grows blueberries. A grower's density would be 1 if he or she grew a crop in all of the crop categories produced by other growers. In this context, network density tells us the diversification of blueberry growers' crop mixes. Network degree centralization and average density are reported below in Table 4.

Table 4 shows no strong patterns in density or network centralization over time, although there does appear to be a noisy and subtle downward trend in both. This would indicate that blueberry growers are becoming more homogeneous in terms of the number of crops categories grow and they are growing in fewer categories. Growing in fewer categories makes sense because even if a grower is able to grow many crops successfully on his land from an agronomic standpoint, there are likely to be economies of scale in both production and marketing that can be achieved through the production of similar crops. Overall, these results suggest that blueberry producers may be moving from an experimentation phase and incorporating blueberries into a more stable and established crop mix.

**Table 4: Network Characteristics**

| *Year* | *Network Degree Centralization (%)* | *Average Density* |
|---|---|---|
| 2000 | 24.183 | 0.357 |
| 2001 | 12.956 | 0.257 |
| 2002 | 15.108 | 0.304 |
| 2003 | 17.288 | 0.350 |
| 2004 | 26.571 | 0.220 |
| 2005 | 15.319 | 0.266 |
| 2006 | 10.753 | 0.245 |
| 2007 | 11.375 | 0.192 |
| 2008 | 10.290 | 0.229 |
| 2009 | 13.921 | 0.220 |
| 2010 | 15.252 | 0.258 |
| 2011 | 18.267 | 0.209 |

## 8 Conclusion and Next Steps

The results of these analyses suggest that relationships between growers across industries and space play an important role in determining adoption of blueberries. In particular, the industry peer effect is positive and significant in all but one years for potential first adopters. The network analyses corroborate these results, suggesting a core set of industries that have important relationships with blueberries: stone fruit, grapes, berries, nuts and olives. The majority of these crops are perennials and berries may grow in bush form (depending on the type), which suggests that growers may be more likely to adopt if they already grow a crop that requires a related production practice and/or planning horizon. Furthermore, it leaves open the possibility that growers may have shared information about blueberries among other growers in these core industries. The local peer effect is less consistent, although it is significant in most of the later periods of the sample. The sign of the effect changes by year, suggesting that successes or failures of nearby growers, either due to economic or agronomic factors, may well have influenced potential adopters. Combining this result with geographic analysis, it appears that an experimentation phase occurred in the middle of the 2000s, with growers at the extreme ends of the state trying their hand at blueberries. However, this did not last long, and the 2011 map shows production in two main regions: the southern Central Valley and the Central Coast. Overall, this analysis provides valuable insights into the blueberry industry and hints at the mechanisms that may have led to adoption of blueberries among California growers.

My goal with this paper is to lay the foundation for a social network survey and more in-depth analysis of the social networks in the blueberry industry in California. Along with revising and strengthening the core analyses of this work, my next step is to work on developing this survey. As with any surveys, social network surveys are time-consuming and costly. In addition, social network data are rarely collected by sources other than researchers studying social networks, so primary data collection is usually necessary for these types of studies. My data and methodology in this paper have allowed me to take a first step in understanding blueberry adoption in California without having to collect social network data. When I do collect survey data, the results I have obtained in this initial analysis will be used to help me better understand specific mechanisms of learning and complement the work I have already done which relates to patterns but is not specific enough to uncover mechanisms.

## References


Bandiera, O., and I. Rasul. 2006. "Social Networks and Technology Adoption in Northern Mozambique." *The Economic Journal* 116:869–902.

Case, A. 1992. "Neighborhood Influence and Technological Change." *Regional Science and Urban Economics* 22:491–508.

Conley, T.G., and C.R. Udry. 2010. "Learning about a New Technology: Pineapple in Ghana." *American Economic Review* 100:35–69.

De Giorgi, G., M. Pellizzari, and S. Redaelli. 2010. "Identification of Social Interactions through Partially Overlapping Peer Groups." *American Economic Journal: Applied Economics* 2:241–275.

[DPR 2013] California Department of Pesticide Regulation. 2013. "California Pesticide Information Portal." `http://calpip.cdpr.ca.gov/main.cfm`.

[ERS 2013] Economic Research Service, United States Department of Agriculture. 2013. "U.S. Blueberry Industry." `http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1765`.

Foster, A.D., and M.R. Rosenzweig. 1995. "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture." *Journal of Political Economy* 103:1176–1209.

Griliches, Z. 1957. "Hybrid Corn: An Exploration in the Economics of Technological Change." *Econometrica* 25:501–522.

Jimenez, M., F. Carpenter, R.H. Molinar, K. Wright, and K.R. Day. 2005. "Blueberry Research Launches Exciting New California Crop." *California Agriculture* 59:65–69.

Mansfield, E. 1961. "Technical Change and the Rate of Imitation." *Econometrica* 29:741–766.

Manski, C.F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *The Review of Economic Studies* 60:531–542.

Moffitt, R.A. 2001. *Social Dynamics*, Cambridge, MA: MIT Press, chap. Policy Interventions, Low Level Equilibria, and Social Interactions. pp. 45–82.

Munshi, K. 2004. "Social Learning in a Heterogeneous Population: Technology Diffusion in the Indian Green Revolution." *Journal of Development Economics* 73:185–213.

Oster, E., and R. Thornton. 2009. "Determinants of Technology Adoption: Private Value and Peer Effects in Menstrual Cup Take-Up." Working Paper No. 14828, National Bureau of Economic Research.

Perez, A., and K. Plattner. 2013. "Fruit and Tree Nuts Outlook." Situation and Outlook No. FTS-356, United States Department of Agriculture Economic Research Service.

Ryan, B., and N.C. Gross. 1943. "The Diffusion of Hybrid Seed Corn in Two Iowa Communities." *Rural Sociology* 8:15–24.