

Fresno vs. Merced: Where should I invest in farmland?

Jared Cummings

September 2021

Abstract

An analysis of the sale price-per-acre of small acreage farms (less than 120) in the Fresno and Merced areas. I conclude that there is a statistically significant difference between the median prices in the two cities. Various nonparametric procedures are compared to the conclusions reached by the classical two-sample t-test at the $\alpha = 0.2$ level, a level of significance chosen to align practical significance with statistical significance. On the basis of this study there is a data-driven reason to prefer investment in one location over another.

Contents

| | |
|---|----------|
| Introduction | 2 |
| Data | 2 |
| Results | 2 |
| Conclusions and Recommendations | 4 |
| Appendix 1: Data and Code | 7 |
| Appendix 2: Selection of level of significance | 7 |

Introduction

For high net worth individuals and investment corporations, agricultural land can be an attractive vehicle for generating consistent revenue. To maximize returns, investors seek plots with high output at the lowest price. In California’s central valley some of the most fertile soil for fruit and nut crops can be found. For an investor interested in expanding their holdings of small acreage farms (less than 120 acres) in this region who has already decided to target the cities of Fresno and Merced, should they prioritize one city over the other? Phrased differently, even though the land is considered equally fertile, is there a difference in sale price-per-acre between properties in Fresno and Merced? If so, which area is a better deal for passive-income investing in small-cap farmland?

In this analysis I test the hypothesis that there is no difference between the median price-per-acre of farms in Fresno and Merced using a Mann-Whitney U-test. For this test I determine statistical significance at the 0.2 level in order to align practical importance with the hypothesis decision.

Data

The data used for this analysis is a sample of 20 farmland transactions in each of the cities of Fresno and Merced, California in the years 2000-2020 where less than 120 acres were exchanged. Acreage and sale price (inflation adjusted) are recorded for each sale. We investigate the price-per-acre of the sales to investigate if there is a difference in land value between these two productive farm regions of California’s central valley.

Table 1: Summary statistics of small farm sale price-per-acre in the Fresno and Merced in 2000-2020.

| City | n | Mean | StDev | Median | IQR |
|--------|----|------|-------|--------|------|
| Fresno | 31 | 2.28 | 2.14 | 2.27 | 2.39 |
| Merced | 38 | 2.07 | 3.01 | 1.20 | 1.50 |

Table 1 gives summary statistics for the data, partitioned by city. For the Fresno data, we find that the mean and median are close, and the two measures of spread, the standard deviation and interquartile range (IQR) are also comparable. These similarities are incident to a relatively balanced dispersion of sale prices-per-acre in and around Fresno. For the Merced data, a larger discrepancy exists between the same measures of center and spread, indicating the likely presence of outliers. Attempted inference on the means of these two samples will have issues because of this. Analysis should instead focus on the difference in median price.

Figure 1 displays the data in two boxplots, one for each city. The distribution of price-per-acre in both Fresno and Merced is visibly right skewed by the occasional high values shown as stand-alone points. This skewness more prominent in Merced than Fresno. Because of this data attribute, and because price-per-acre takes only positive values, the most popular method for testing differences between two sample averages, a two sample t-test, will not be able to give reliable results for this analysis. A different method, one that is robust to these quirks of our data and resistant to potential outliers will instead be employed. This decision is detailed more completely in the Methodology section.

Results

Using prepackaged functions from [R Core Team \(2016\)](#), both a two-sample t-test and Mann-Whitney U-test were performed. Table 2 reports the p-values for the tests to three significant digits. For the t-test, the hypotheses tested were:

H_0 : No difference in mean price-per-acre between Fresno and Merced

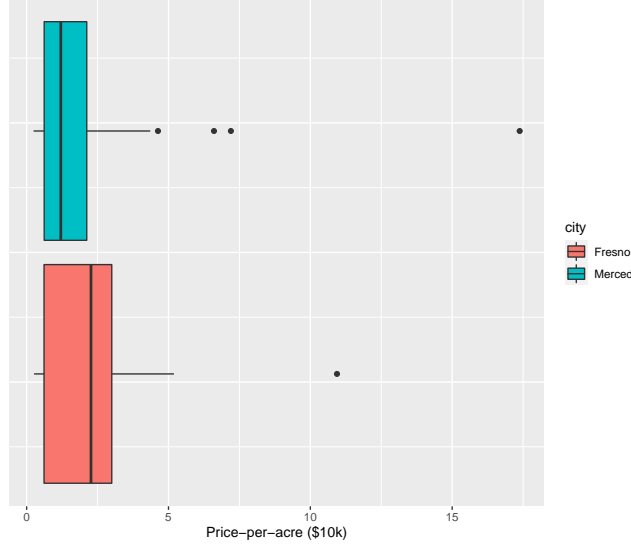


Figure 1: Boxplot comparison of price-per-acre of Fresno and Merced farm sales.

H_1 : Difference in mean price-per-acre exists between Fresno and Merced

For this application of the Mann-Whitney test, the hypotheses were similar:

H_0 : No difference in median price-per-acre between Fresno and Merced

H_1 : Difference in median price-per-acre exists between Fresno and Merced

Though both tests fail to reject the null hypothesis at the $\alpha = 0.20$ level, there is not agreement between the two tests about the weight of evidence against the null hypotheses. The t-test indicates that there the evidence in favor of the alternate hypothesis is practically non-existent (p-value: 0.746) while the Mann-Whitney test makes it a near borderline decision (p-value: 0.225). If a single observation were added, the t-test would likely continue to fail to reject the null hypothesis, while the Mann-Whitney could very easily be tipped the other way.

Table 2: P-values for tests of equal mean/median

| test | pvalue |
|---------------------|--------|
| 2-sample t-test | 0.746 |
| Mann-Whitney U-test | 0.225 |

In addition to these two procedures, a Monte Carlo permutation test and bootstrap test were performed. Conceptually, these two tests “shuffle” which of the two groups the observed data fall into. Under the null hypothesis of these tests there is no difference in groups, so theoretically the test statistics derived under these tests should be very similar to the value actually observed if the null hypothesis is true. For these tests, a difference in medians was computed as the test statistic. The hypotheses for both tests are the same as for the Mann-Whitney test.

A Monte Carlo permutation test shuffles the data by sampling without replacement from it. This is in contrast to an exact permutation test in which every possible permutation of the data (again, without repeating values) is defined and a test statistic computed. The exact test is prohibitive for data that is “large” in the classical sense, so for this application we apply only the Monte Carlo version. Figure 2 reports the results of the Monte Carlo permutation test with Monte Carlo error assessed on the p-value.

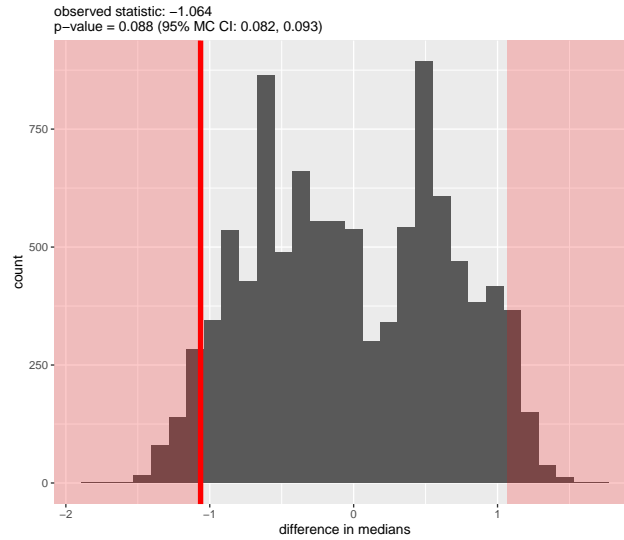


Figure 2: Results of Monto Carlo permutation test on difference in medians between Fresno and Merced small-acreage farm price-per-acre.

The distinction between a permutation test and the bootstrap test is not arbitrary. Unlike a permutation test, a bootstrap test shuffles the data by sampling with replacement. This results in an symmetric and bell-curved distribution that is an always-available tool for evaluating uncertainty in many statistical applications. The nonparametric bootstrap is one of the greatest advances in computer age statistical inference. Figure 3 gives the results of the bootstrap test on this farmland data, with Monte Carlo error assessed on the p-value.

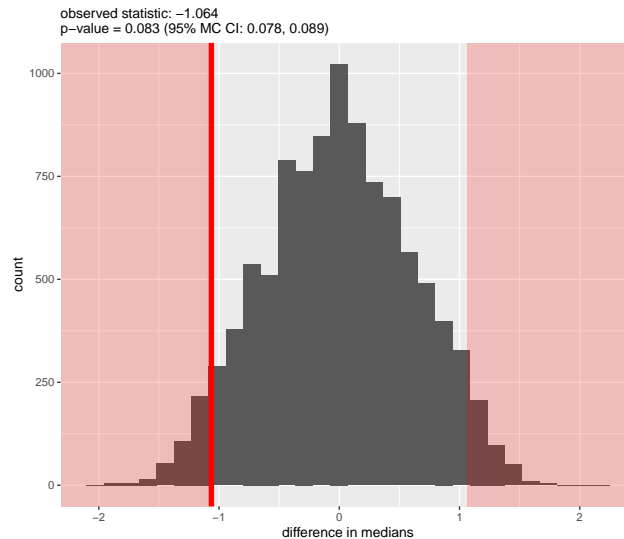


Figure 3: Results of bootstrap test on difference in medians between Fresno and Merced small-acreage farm price-per-acre.

Conclusions and Recommendations

On the basis of this analysis, I advise favoring purchases in the Merced market over the Fresno market. Comparative pricing information is available, and it appears that similarly strong investment opportunities can be found in both locales, but there is evidence to suggest that the median price-per-acre in Fresno

is sufficiently different than the median price-per-acre in Merced to be of practical concern. Though the rank-based Mann-Whitney test does not detect significance (p-value: 0.225), it is a boundary line case. The weight of evidence brought by the permutation and bootstrap tests, both of which consider the real values of the observed data, not just their relative ranking, both disclose a significant difference in the median price-per-acre (p-values of 0.092 and 0.080 respectively).

In stating this conclusion, it must be recognized that binary hypothesis testing may not adequately capture all aspects of real and perceived risk. Additional study on this question could be done using Bayesian methods that allow the incorporation of expert opinions and an expanded vocabulary of risk. To further confirm the findings of the study done in this report, I advise a simulation study to confirm that the power of the Mann-Whitney test is sufficient for the designated level of practical importance outlined in Appendix 2.

In how this study was conducted, the benefit of using a nonparametric test cannot be overstated. As discussed in the Data section, there were clear reasons to be entirely skeptical of the results that could be gained through a classical t-test. The dramatic difference in the p-values that the parametric and nonparametric methods gave, as discussed in the Results section, is an indicator of the kinds of disparities that can result in the misapplication of statistical methods. Nonparametric tests, like the Mann-Whitney test, are a responsible option for fitting to a wider variety of data and should be adopted in more applications. The stark difference between the findings of the classical t-test and the two resampling procedures is also a strong indicator of the importance of applying appropriate methods to get specific and actionable insights.

- R Core Team. 2016. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Randles, Ronald H, and Douglas A Wolfe. 1979. "Introduction to the Theory of Nonparametric Statistics." John Wiley.
- Wickham, Hadley. 2011. "Ggplot2." *Wiley Interdisciplinary Reviews: Computational Statistics* 3 (2): 180–85.

Appendix 1: Data and Code

Data for this analysis is available at this [link](#). The data was obtained via personal communication with an industry analyst and the source is not publicly disclosable. Sale prices are given in increments of \$10k.

The code for this analysis can be found [here](#) and is supplemented by code for the resampling tests, which can be found [here](#).

All figures are made with the ggplot2, first introduced in [Wickham \(2011\)](#).

Appendix 2: Selection of level of significance

In the context of agriculture investing, the risk associated with incorrectly concluding that there is a preferred price location when no material difference in per-acre price exists (type I error) is entirely realized in an unnecessary limiting of perspective properties. There is not a measurable financial cost to making this mistake. The consequence of incorrectly concluding that there is no difference in per-acre price (type II error) is much more substantial. If this mistake is made there could be substantial, unnecessary overpayment without a commensurate increase in returns. In classical hypothesis testing, the level of significance, α , represents the probability of making a type I error. To appropriately offset risk, we adjusted α to control the type II error probability of making a “practically significant” over payment of at least 20-25% of per-acre price at roughly 0.05. In making this adjustment we align practical significance with statistical significance.

Figure 4 shows the probability of type II error at this practically significant threshold as a function of the level of significance, α , for a two-sample t-test with equal variance. As mentioned earlier, using this parametric statistical procedure is not appropriate for this data. However, [Randles and Wolfe \(1979\)](#), indicate that in simulation studies performed by the authors, when the non-parametric equivalent of the two-sample t-test is applied to strictly positive and right skewed data the power of the test exceeds $1 - \alpha$. By extrapolating their results to our sample sizes, it is safe to conclude that the type II error probability for the nonparametric test is bounded above by the two-sample t-test’s type II error probability. In future work, a simulation study could be conducted to empirically confirm this behavior if more certainty in the method is desired.

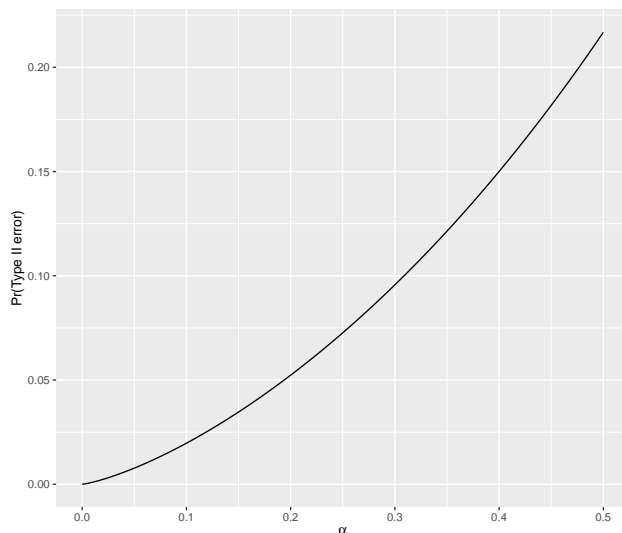


Figure 4: Type II error probability of two-sample t-test as a function of significance level if true difference between Fresno and Merced price-per-acre is 20 to 25 percent