Examining Vaccinations Among Students in California's School System

Ben Harwood

September 26, 2020

Abstract

As COVID-19 continues to throw the country out of whack, there is an increased risk from disinformation around vaccinations, especially with the potential vaccine in development. Vaccines in general are under attack by anti-vaxxers and those claiming religious exemptions, and our children (and indeed the population in general) run the risk of being exposed to and contracting diseases that we have the ability to prevent. This report examines vaccination rates in California schools in hopes of determining how best to allocate funds for vaccination awareness and completeness.

Introduction

Books have been written about the importance of the announcement in 1955 that Jonas Salk had developed a vaccine for poliomyelitis (see [1], for example). As well they should be, polio had been ravaging the US and Europe for many years, and this vaccination was set to change the landscape for children everywhere. The same can be said for measles, indeed John Franklin Enders shared the Nobel Prize in Physiology or Medicine in 1954 for his development of the measles vaccine. Yet despite the mountains of scientific evidence, many people believe vaccines are dangerous (the so-called "anti-vaxxers").

Scepticism and protests around vaccinations are not new. For example a major smallpox epidemic broke out in Stockholm, Sweden after a campaign motivated by religious objections and concerns about effectiveness and individual rights dropped the vaccination rate to 40% (see [2]). Vaccine protest didn't occur in the U.S. until 2005, when a measles outbreak (a disease that had been declared eliminated in the country) was attributed to parents in Indiana who did not have their children vaccinated. There are numerous reasons motivating vaccine hesitancy, including religious objection, personal liberty, and myths connecting vaccines to serious conditions such as autism and Guillain-Barré syndrome. With misinformation aplenty through social media and word of mouth, vaccination awareness is of utmost importance, especially now that COVID-19 is running rampant.

The goal of this report is to determine what factors influence vaccination rates and develop guidance for allocating financial assistance to the school districts in the state to improve vaccination rates (and reporting compliance).

1 Analysis and Models

Many questions were of interest in this study, both on a macro level and more focused. Specifically:

1. How have vaccination rates at the national level varied over time?
2. Which vaccinations have the highest and lowest rate of vaccination, and which vaccination rate was the most volatile?
3. Is there a credible difference in reporting proportions between private and public schools?
4. How do vaccination rates in California compare to national vaccination rates?
5. How are individual vaccine rates related among districts, i.e. if a student is missing a single vaccine are they missing others?
6. Can whether each school in a district reported vaccination rates be predicted, and if so how?
7. Can the percentage of all enrolled students that are up-to-date on all of their vaccinations be predicted, and if so how?
8. Can the percentage of all enrolled students claiming religious exemptions to vaccination be predicted, and if so how?

Combined, these questions will paint the "big picture" and inform the question posed at the end of the introduction. To answer them, numerous statistical tests and methods were utilized. This section will address those techniques and the data utilized to perform them.

1.1 Data

Given the breadth of the questions just posed, multiple datasets were required. Here is a brief description of each.

1. Time series data from the World Health Organization reporting U.S. vaccination rates for five common vaccines from 1980 to 2013: DTP1 (first dose of diptheria/tetanus/pertussis); HepB_BD (birth dose of hepatits B); Pol3 (third dose of polio); Hib3 (influenze third dose); MCV1 (first dose of measles).

2. A list of 7381 California kindergartens (public and private) and whether they reported their vaccination data to the state in 2013.
3. A 13 variable sample of 700 California school districts including the following information:
   (a) DistrictName : District name
   (b) WithoutDTP : Percentage of students without the DTP vaccine
   (c) WithoutPolio : Percentage of students without the Polio vaccine
   (d) WithoutMMR : Percentage of students without the MMR vaccine
   (e) WithoutHepB : Percentage of students without the Hepatitis B vaccine
   (f) PctUpToDate : Percentage of all enrolled students with completely up-to-date vaccines
   (g) DistrictComplete : Boolean indicating whether or not the district's reporting was complete
   (h) PctBeliefExempt : Percentage of all enrolled students with belief exceptions
   (i) PctChildPoverty (X_1) : Percentage of children in the district living below the poverty line
   (j) PctFreeMeal (X_2) : Percentage of children in the district eligible for free student meals
   (k) PctFamilyPoverty (X_3) : Percentage of familes in the district living below the poverty line
   (l) Enrolled (X_4) : Total number of enrolled students in the district
   (m) TotalSchools (X_5): Total number of different schools in the district

   Many of these variables are highly correlated, which is to be expected.

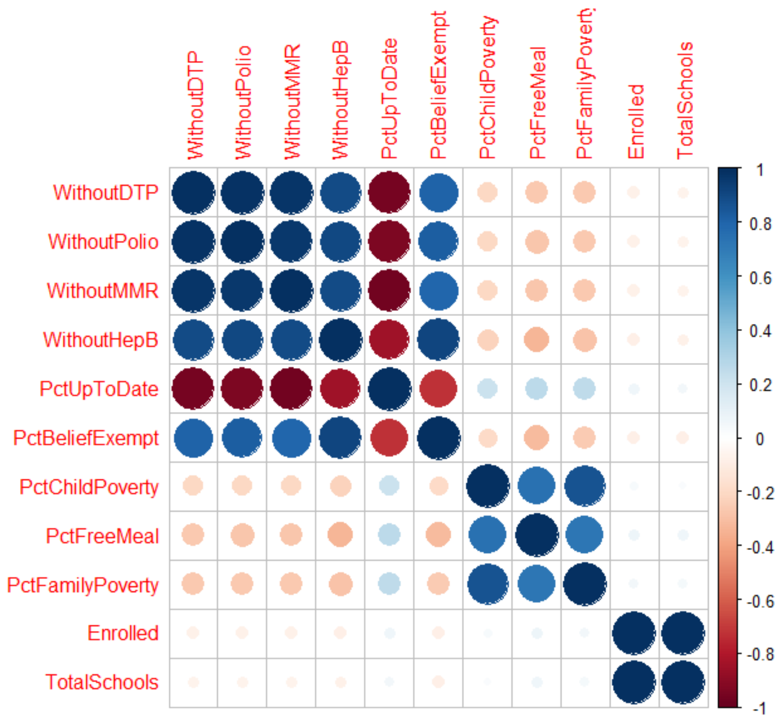


Figure 1: District variable correlation

As mentioned, none of these correlations are especially surprising. Skewness was calculated for each numeric variable to test for univariate normality. Each of the variables, except for PctChildPoverty and PctFreeMeal, required transformation to reduce skew. The following table shows the original skew measure and the transformed skew measure (with applied transformation indicated)

	Base Skew	Transformation	Transformed Skew
WithoutDTP	2.186	$\sqrt{X}$	0.594
WithoutPolio	2.292	$\sqrt{X}$	0.652
WithoutMMR	2.134	$\sqrt{X}$	0.611
WithoutHepB	2.834	$\sqrt{X}$	0.827
PctUpToDate	-2.15	$\sqrt{(100-X)}$	0.597
PctBeliefExempt	3.231	$\sqrt{X}$	0.987
PctChildPoverty	0.824	NA	0.827
PctFreeMeal	-0.132	NA	-0.129
PctFamilyPoverty	1.231	$\sqrt{X}$	0.211
Enrolled	20.608	$\log X$	-0.072
TotalSchools	20.264	$\log X$	0.521

Table 1: Variable Skewness

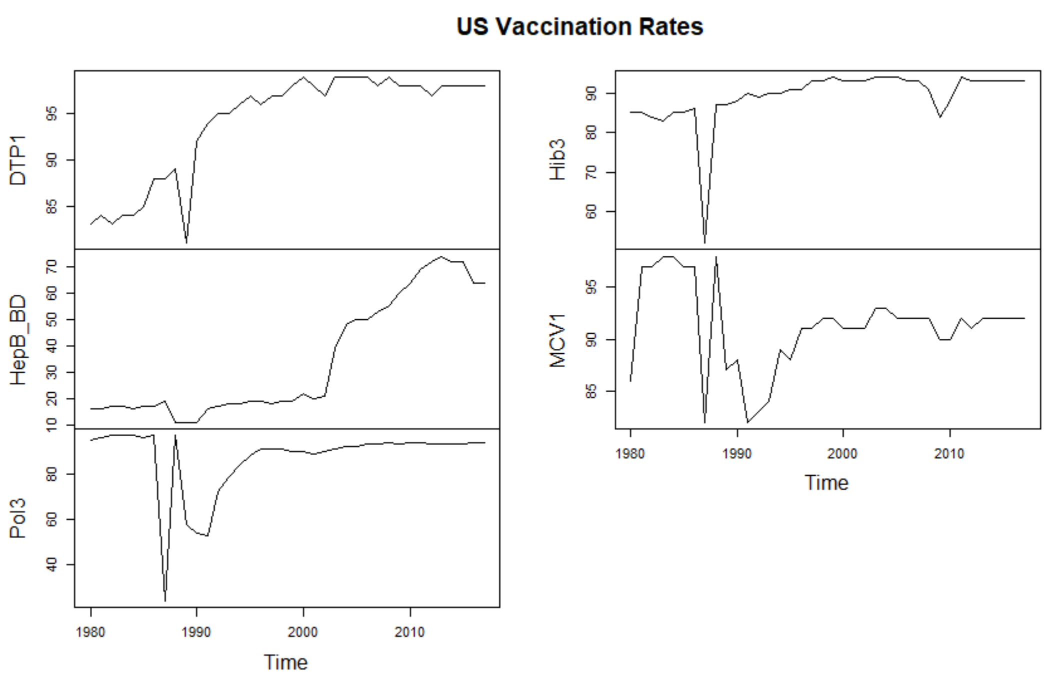
The observant reader will notice that $\log(20.608) \neq -0.072$. When considering the TotalSchools variable an extreme outlier was discovered, and after removing the school district that constituted that outlier the logarithmic transformation resulted in the values in the table above. The only other additional cleaning or preparation required of the data itself was to make a non-Boolean version of the DistrictComplete variable for use in logistic regression (discussed in the next section) along with a numeric version of the same for use in a Bayesian logistic regression (also discussed in the next section).

1.2 Models

With the data explored and prepared, several methods of analysis were pursued. While the results of these analyses will be reported in Section 2, this section will discuss the various techniques and tests used to answer the questions posed in Section 1. With numerous questions to answer that are essentially unrelated, each will addressed individually.

1.2.1 National vaccination rates over time

As mentioned in Section 1.1, there are five vaccinations whose vaccination rates were tracked from 1980 to 2013. The following plots show how the vaccination rate for each varied over the time period.



It is easily noticed from these plots that the vaccination rate for each of these vaccines is very different. For example, DTP1 and Hib3 are steadily increasing over the time frame (except noticably around in the late 1980's, when the CDC stopped recommending vaccinations and cases of many diseases spiked for a time [3]), while MCV1 was quite oscillatory at first and then flattened in the 2000's. In order to better understand how the vaccination rates actually changed, the following series of graphs show changepoint analysis and autocorrelation for each variable. The horizontal red lines show the point where the mean of the time series changed, the vertical red line shows the pointin time where the variation

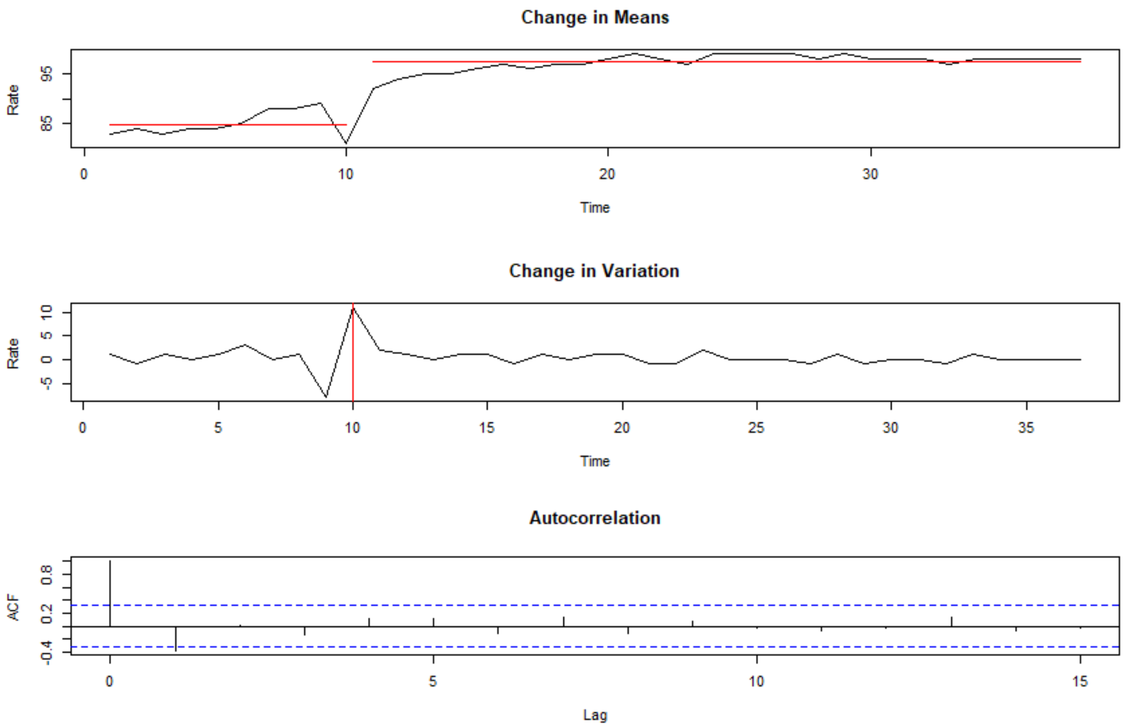


Figure 2: DTP1

shifted, and as always the ACF plot measures the stationarity of the time series (very few significant lagged correlations is the goal).

The goal of these changepoint plots is to see if there is commonality among the mean shift dates or variation shift dates. If so then those times can be studied further to gain insight into the cause. The following table summarizes the mean and variation changepoints. The table also contains the standard deviations for each vaccine, as this is a good way to measure the volatility.

Vaccine	Mean changepoint	Variation changepoint	σ
DTP1	1990	1990	5.868
HepB_BD	2004	NA	22.539
Pol3	1995	1996	15.353
Hib3	1988	1988	7.136
MCV1	1987	1996	4.188

Table 2: Changepoints

Other than the late 1980's as previously discussed, no real commonality is present. Note the "NA" for the variation changepoint for hepatitis. In other words, there is no marked shift in the variation in the vaccination rate for hepatitis. It is no coincidence that hepatitis also has the highest standard deviation, $\sigma = 22.539$, and hence the highest volatility.

As a final word on the nation vaccinate rates, here are the vaccination rates in the final year of the time series (2013). Notice that hepatitis has the lowest rate, while DTP1 has the highest.

Vaccine	Rate
DTP1	98%
HepB_BD	64%
Pol3	94%
Hib3	93%
MCV1	92%

Table 3: Vaccination Rates in 2013

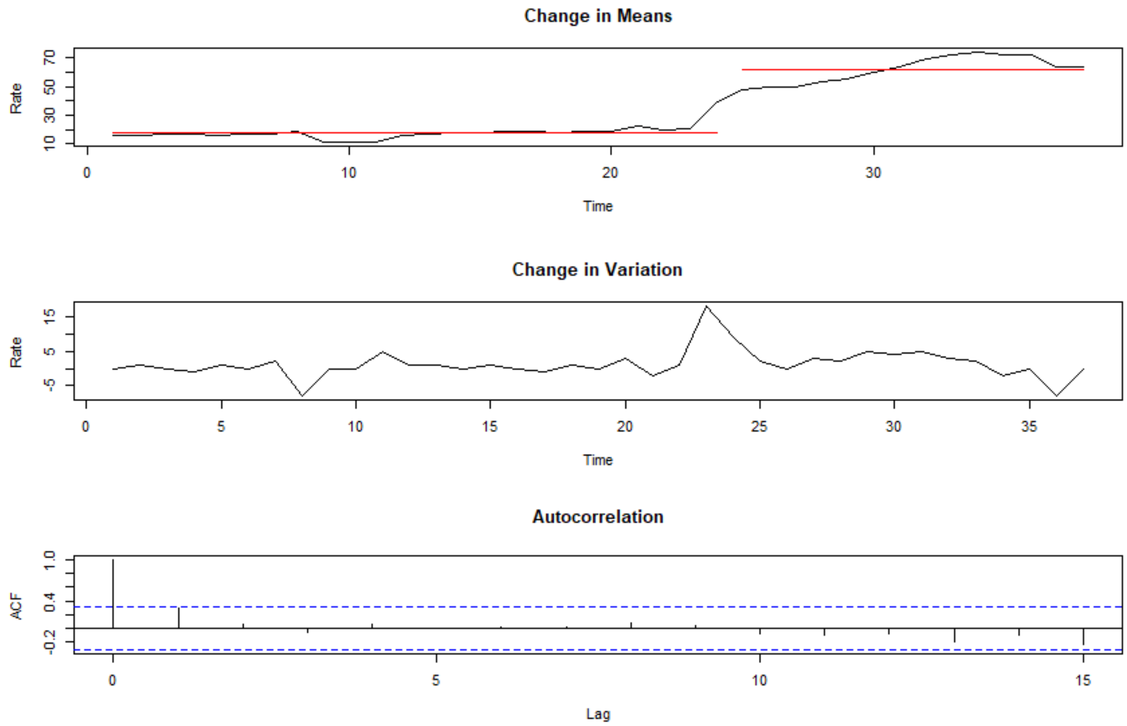


Figure 3: HepB

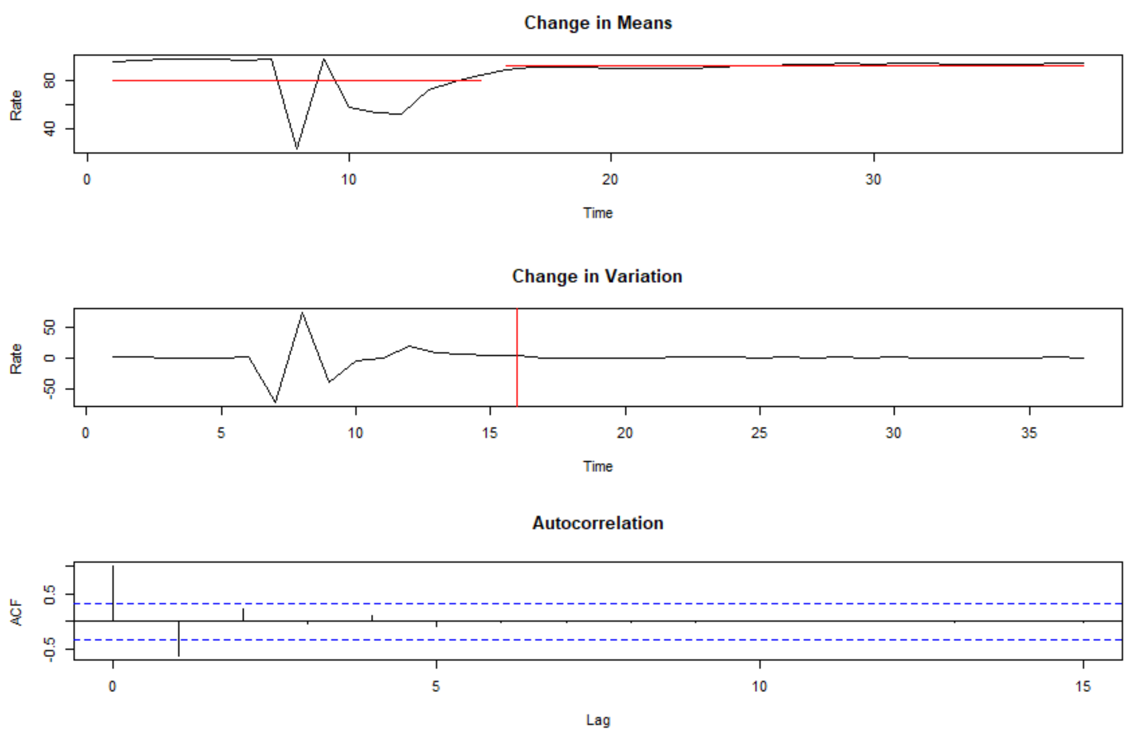


Figure 4: POL3

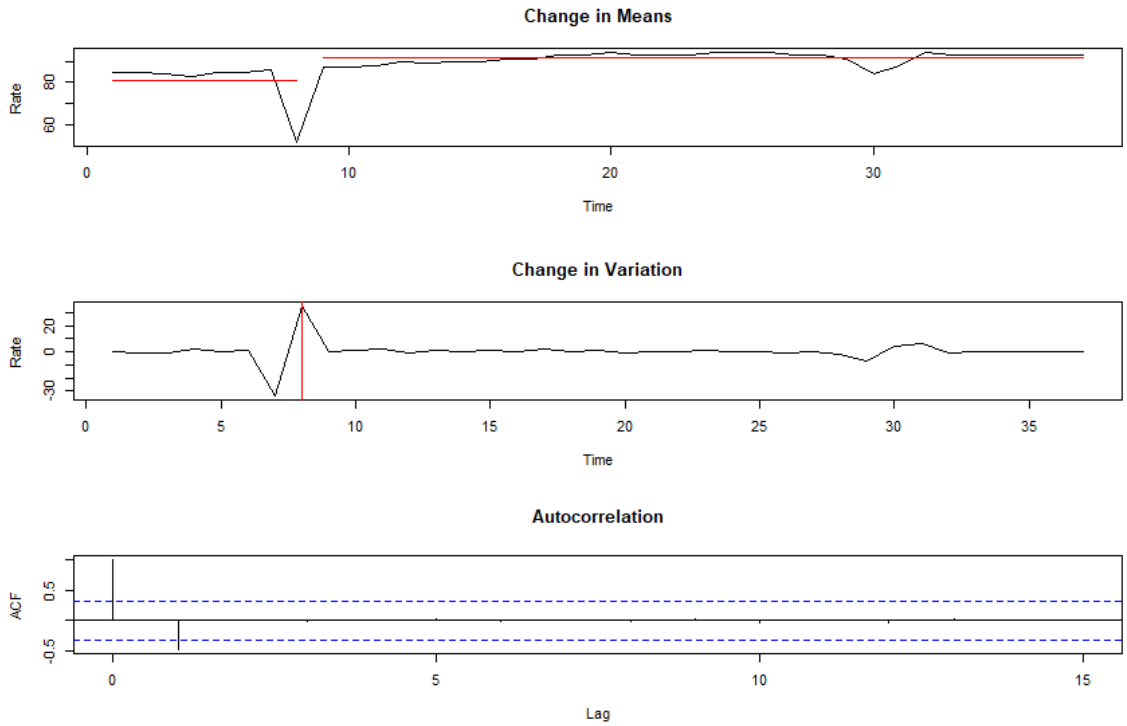


Figure 5: Hib3

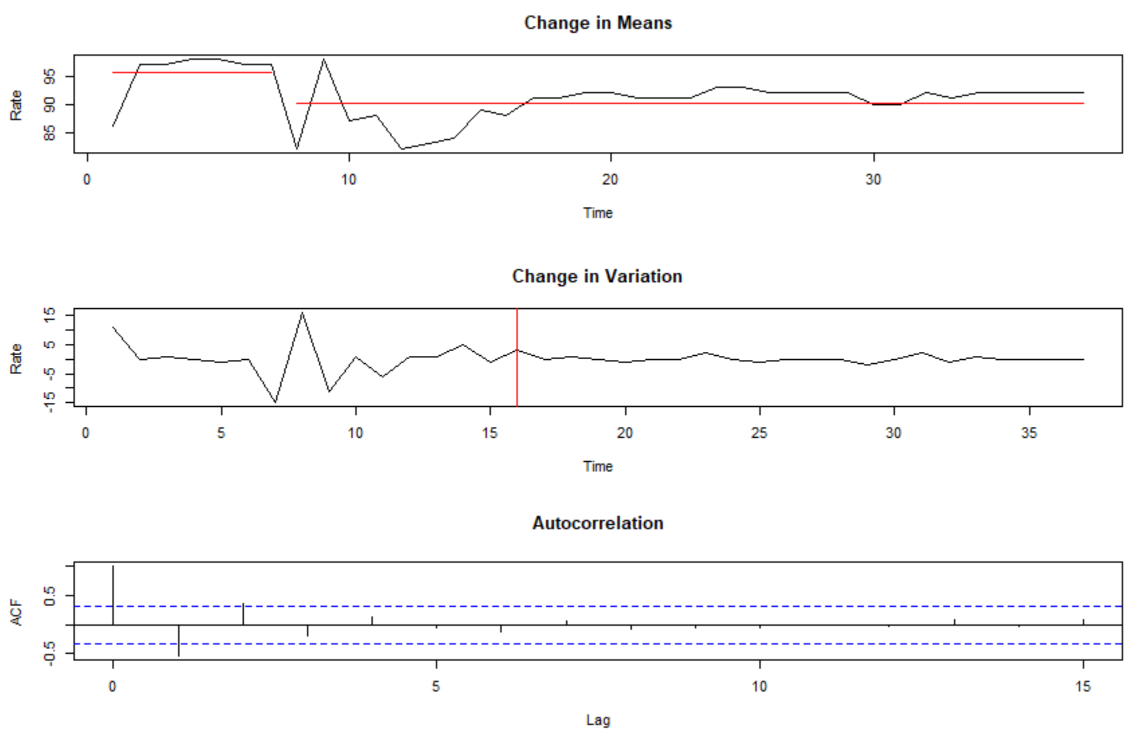


Figure 6: MCV1

1.2.2 School reporting

One of the issues that must be contended with is getting schools to report their vaccination rates. While it is ultimately up to the individiual school to report their vaccination rates, not having information for each school could pose problems with funding allocation. It is not a stretch to posit that private schools report their vaccinations less frequently that public schools.[1] So the next order of business would be to see the rates at which both public and private schools report their vaccination data, and see if there is any real difference in the reporting rates.

The following contigency table shows how public and private schools report their vaccination rates.

	No	Yes	Total
Private	252	1397	1649
Public	148	5584	5732
Total	400	6981	7381

From this it is easy to see that the reporting rate for private schools is $1397/1649 = 0.847$ and the reporting rate for public schools is $5584/5732 = 0.974$. As excepted, the public school reporting rate is higher. But is the difference significant? To answer this, a χ^2 test for independence was performed, and with $df = 1, \chi^2 = 400.49, p < 2.2 \times 10^{-16}$ the null hypothesis that public and private school vaccination rates are independent is REJECTED. Additionally, a Bayesian analysis shows that the odds of the alternative hypothesis of nonindependence is 1.150548×10^{69} : 1, which is further evidence in support of rejecting the null. Additionally, as seen in the following plot, the difference in reporting ratios between public and private schools lies in the highest density interval of $[1.1369, 1.8363]$. The red line indicates the mean (1.4606).

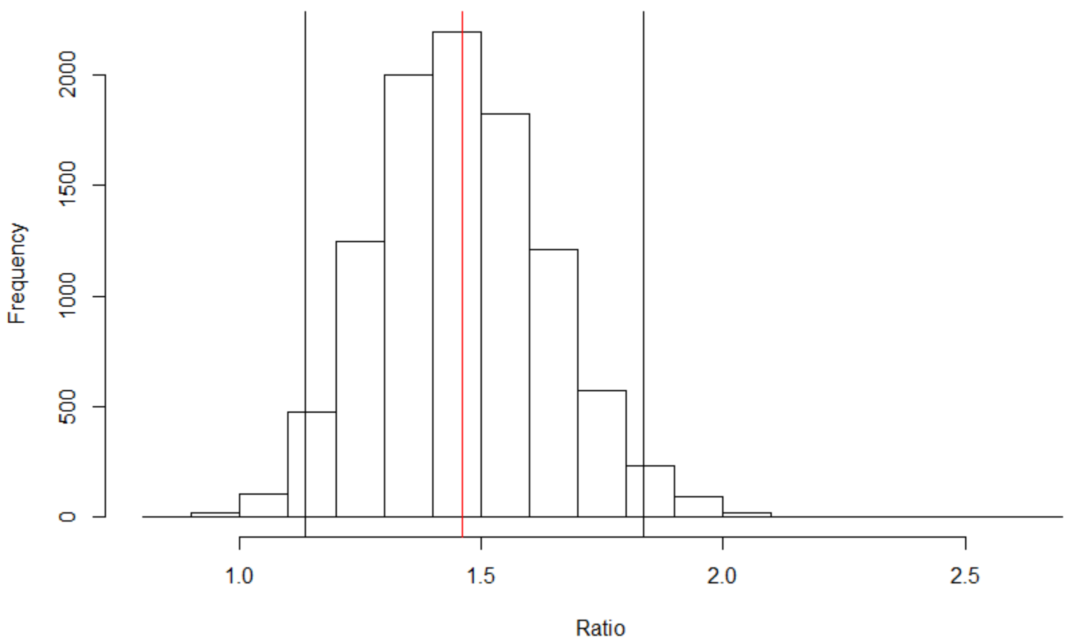


Figure 7: Histogram of posterior distribution of differences in proportions between reporting and not reporting

This lends additional credible evidence that there is an association between whether a school is public or private and whether it reports its vaccination rates.

1.2.3 California vs National

While the focus is on schools in California, it would be nice to see how the state is doing compared to rest of the country. Recall that the 2013 vaccination rates for each of the five vaccines were reported at the end of section 1.2.1. The same is presented in Table 4 for California, along with the national rates and a comparison ratio to see how California compares to the nation, however state-level data for Hib3 was not collected, so it is excluded.

Notice that for polio and measles California is right on track with the rest of the nation. And really, even DTP1 is more or less even. The difference really is hepatitis, but in a good way. This is a good indicator that something is going well with measles vaccination awareness, but perhaps some additional focus could be given to DTP1 to bring it more in line with the nation (though, as mentioned, less than 6% below the national rate is not terrible).

[1]since, for example, many private schools are faith based, and as will be discussed later many of the children whose vaccines are not up to date claim religious exemption

Vaccine	CA	US	CA vs. US
DTP1	0.93	0.98	0.947
Pol3	0.97	0.94	1.036
MCV1	0.97	0.92	1.059
HepB_BD	0.97	0.64	1.521

Table 4: 2013 Vaccination rates in CA vs National

1.2.4 Vaccination Rate by District

Much as it was not unrealistic to posit that private schools report their vaccination rates less than public schools, it is also not unreasonable to conjecture that if a student is missing one vaccine then they are likely missing others. In addition to the high correlation between the four vaccine variables, a t-test was conducted between each pair of vaccines, with the following results:

1. The t-test between WithoutDTP and WithoutPolio was not significant, with $t = 0.8759, df = 1396$, and $p = 0.3813$, so the null hypothesis of independence can not be rejected.
2. The t-test between WithoutDTP and WithoutMMR was not significiant, with $t = 0.139, df = 1395.3$, and $p = 0.8895$, so the null hypothesis of independence can not be rejected.
3. The t-test between WithoutDTP and WithoutHepB was significiant, with $t = 5.2351, df = 1395.1$, and $p = 1.0152 \times 10^{-7} < 0.05$, so the null hypothesis of independence is rejected.
4. The t-test between WithoutPolio and WithoutMMR was not significant, with $t = -0.7274, df = 1396$, and $p = 0.4671$, so the null hypothesis of independence can not be rejected
5. The t-test between WithoutPolio and WithoutHepB was significant, with $t = 4.4554, df = 1395$, and $p = 9.0428 \times 10^{-6} < 0.05$, so the null hypothesis of independence is rejected.
6. The t-test between WithoutMMR and WithoutHepB was significant, with $t = 5.1505, df = 1393$, and $p = 2.9712 \times 10^{-7} < 0.05$, so the null hypothesis of independence is rejected.

Notice that every t-test involving WithoutHepB was significant. This is strong evidence that if a student is missing the hepatitis vaccine there is a very good chance they are also missing the other three.

Before pressing on, it seems appropriate to see what factors contribute to students not having the HepB vaccine. To that end, linear regression was performed to predict WithoutHepB from the five demographic variables PctChildPoverty, PctFreeMeal, PctFamilyPoverty, Enrolled, and TotalSchools. Child poverty showed to be insignificant in the resulting model, however regression was significant, with $F(4, 695) = 59.45, R^2 = 0.2509, p < 2.2 \times 10^{-16}$ on the remaining four variables. It is worth noting that the high correlation of Enrolled and TotalSchools should be addressed. Removing TotalSchools results in $F(3, 695) = 72.95, R^2 = 0.2362, p < 2.2 \times 10^{-16}$, so this could be considered the best model since TotalSchools only accounted for an additional 0.0147 variance. The resulting equation has the form

$$Y = 4.907 - 0.018X_2 - 0.213X_3 - 0.196X_4$$

where the X_i's are as defined in Section 1.1.

1.2.5 Complete Reporting

As was discussed at the beginning of Section 1.2.2, successful vaccination awareness and tracking relies on districts reporting having accurate, up-to-date reporting. But other factors may contribute to reporting completeness. In order to determine which variables are contributory, logistic regression with all five demographic variables was performed to predict completeness. A χ^2 omnibus test on the results of the model were significant for the TotalSchools variable with $\chi^2(1) = 45.649, p = 1.414 \times 10^{-11}$ and Wald's z-test $z = -5.916, p = 3.3 \times 10^{-9}$, as well as Enrolled with $\chi^2(1) = 8.916, p = 0.002827$ and Wald's z-test $z = 5.120, p = 3.06 \times 10^{-7}$, however no other variables were significant. So instead of considering every variable together, each indvidual variable was examined to see how well it might predict completion. The following paragraphs contain the findings for each of these.

Logistic regression was performed to predict complete reporting from PctChildPoverty, with MacFadden's $R^2 = 0.0116$. A χ^2 omnibus test on the results was (just barely) not significant, with $\chi^2(1) = 3.692, p = 0.0547$, while the Wald's z-test was significant, but only just, with $z = -1.973, p = 0.0485$. The 95% confidence interval for PctChildPoverty was $I = [0.953, 1.000]$, meaning that increasing PctChildPoverty by 1% decreases the likelihood of completion by up to 4.7%. It's also worth noting that this confidence interval contains the odds of 1:1, which is reflective of the non-significance of the regression. There is not sufficient evidence to reject the null hypothesis that completion is independent from PctChildPoverty. Figure 8 shows both a boxplot of the distribution of completion by PctChildPoverty and a histogram showing the highest density interval ($HDI = [0.954, 0.999]$) of the odds of completion. The mean-odds (0.977) is indicated with the red vertical line.

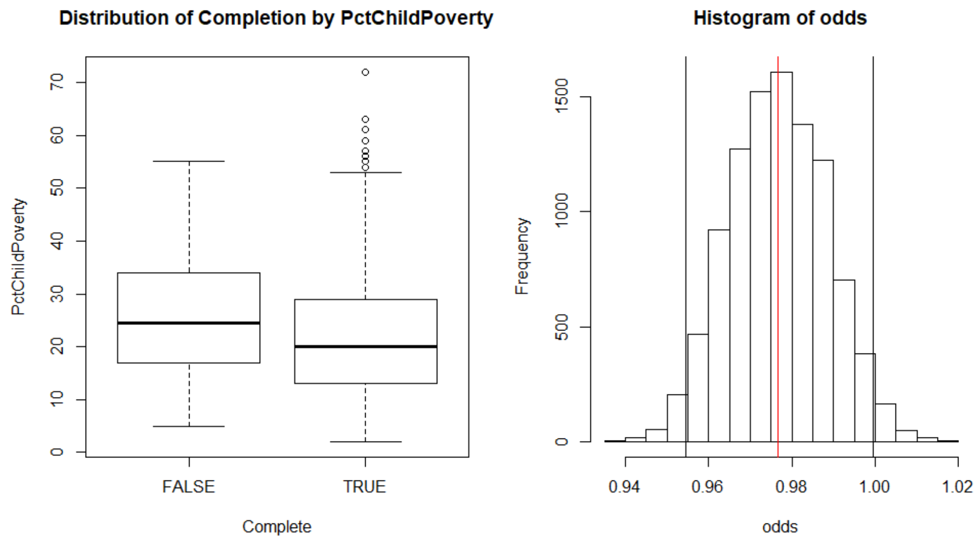


Figure 8: PctChildPoverty regression results

Logistic regression was performed to predict complete reporting from PctFreeMeal, with MacFadden's $R^2 = 0.0247$. A χ^2 omnibus test on the results was significant, with $\chi^2(1) = 7.856, p = 0.00506$, as was the Wald's z-test with $z = -2.713, p = 0.000667$. The 95% confidence interval for PctFreeMeal was $I = [0.967, 0.994]$, meaning that a 1% increase in PctFreeMeal decresases the likelihood of completion by 0.6% to 3.3%. This confidence interval does not contain the odds of 1:1, which lends sufficient evidence to reject the null hypothesis that completion is independent from PctFreeMeal. Figure 9 has both a boxplot of the distribution of completion by PctFreeMeal and a histogram showing the highest density interval ($HDI = [0.968, 0.994]$) of the odds of completion. The mean-odds (0.981) is indicated with the red vertical line.

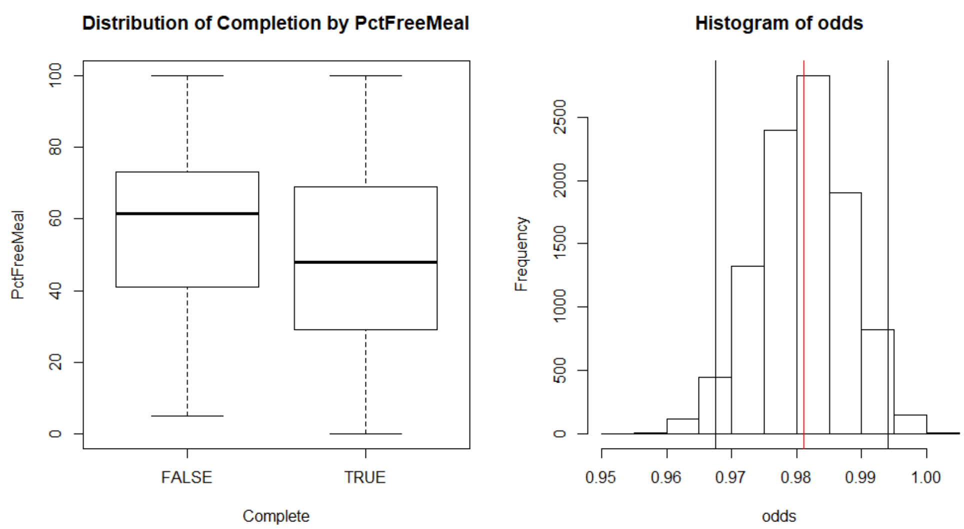


Figure 9: PctFreeMeal regression results

Logistic regression was performed to predict complete reporting from PctFamilyPoverty, with MacFaddens $R^2 = 0.0232$. A χ^2 omnibus test on the results was significant, with $\chi^2(1) = 7.384, p = 0.00658$, as was the Wald's z-test with $z = -2.726, p = 0.00641$. The 95% confidence interval for PctFamilyPoverty was $I = [0.539, 0.904]$, meaning that increasing PctFamilyPoverty by 1% decreases the likelihood of completion by 9.96% to 46.1%. Once again, this confidence interval does not contain the odds of 1:1, and there is thus sufficient evidence to reject the null hypothesis that completion is independent from PctFamilyPoverty. Figure 10 gives both a boxplot of the distribution of completion by PctFamilyPoverty and a histogram showing the highest density interval ($HDI = [0.54, 0.905]$) of the odds of completion. The mean-odds (0.699) is indicated with the red vertical line.

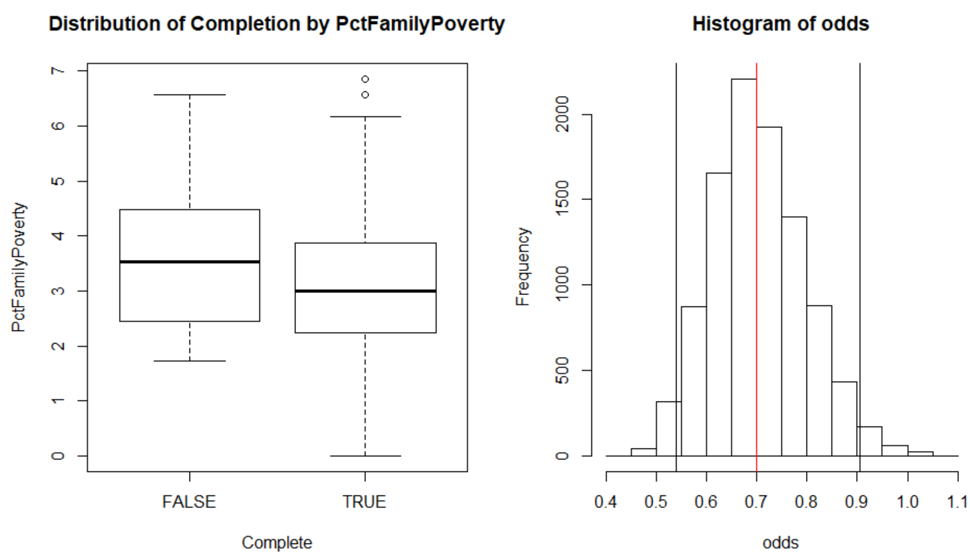


Figure 10: PctFamilyPoverty regression results

Logistic regression was performed to predict complete reporting from enrolled students, with MacFaddens $R^2 = 0.0358$. A χ^2 omnibus test on the results was significant, with $\chi^2(1) = 11.358, p = 0.00075$, as was the Wald's z-test with $z = -3.236, p = 0.00121$. The 95% confidence interval for Enrolled was $I = [0.557, 0.863]$, meaning that increaseing enrolled by merely 1 student could decrease the likelihood of completion by 23.7% to 44.3%. Once again, this confidence interval does not contain the odds of 1:1, and there is thus sufficient evidence to reject the null hypothesis that completion is independent from enrolled students. Figure 11 gives a boxplot of the distribution of completion by Enrolled and a histogram showing the highest density interval ($HDI = [0.554, 0.852]$) of the odds of completion. The mean-odds (0.693) is indicated with the red vertical line.

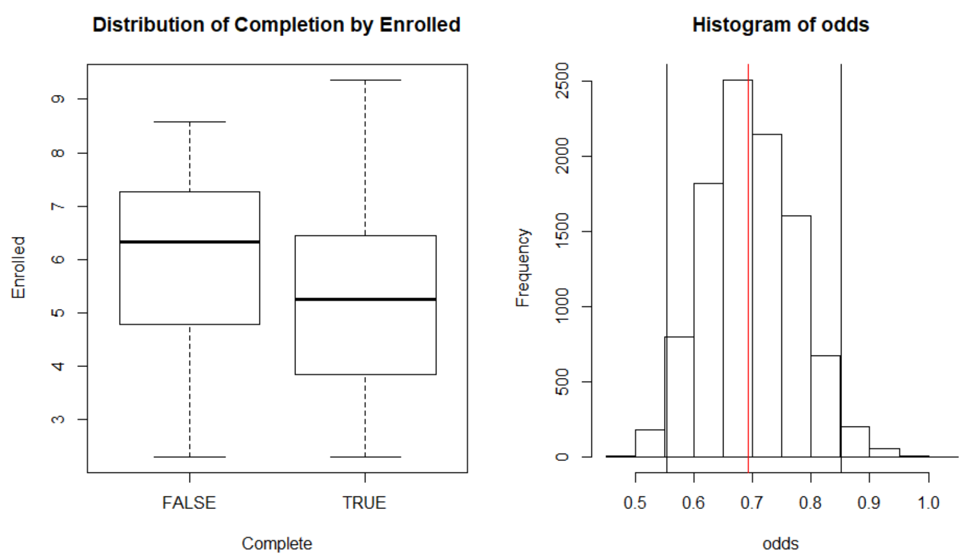


Figure 11: Enrolled regression results

Finally, logistic regression was performed to predict complete reporting from total schools, with MacFaddens $R^2 =$

0.0894. A χ^2 omnibus test on the results was significant, with $\chi^2(1) = 28.416, p < 0.05$, as was the Wald's z-test with $z = -5.12, p = 3.05 \times 10^{-7} < -0.05$. The 95% confidence interval for TotalSchools was $I = [0.357, 0.630]$, meaning that increaseing enrolled by merely 1 student could decrease the likelihood of completion by 47% to 64.3%. Once again, this confidence interval does not contain the odds of 1:1, and there is thus sufficient evidence to reject the null hypothesis that completion is independent from enrolled students. Figure 12 gives a boxplot of the distribution of completion by Enrolled and a histogram showing the highest density interval ($HDI = [0.359, 0.624]$) of the odds of completion. The mean-odds (0.475) is indicated with the red vertical line.

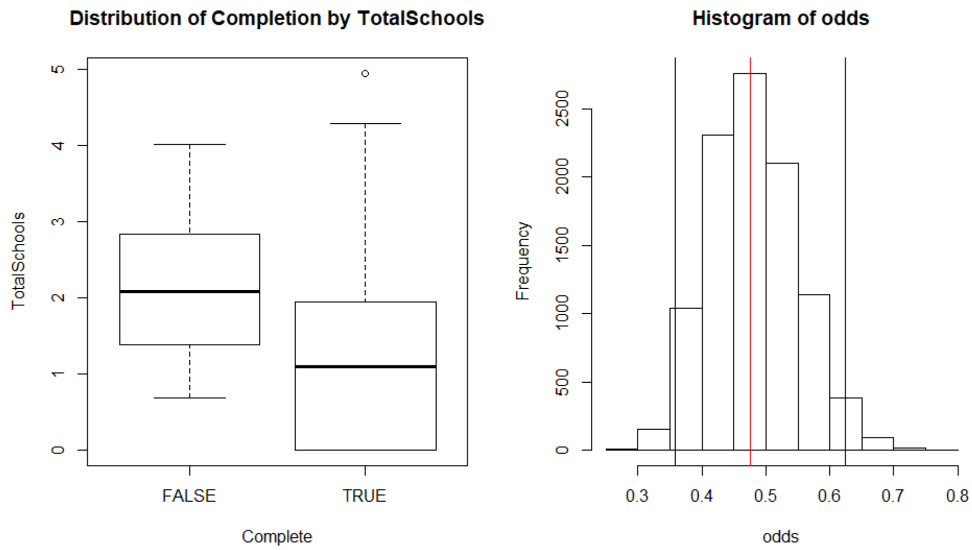


Figure 12: TotalSchools regression results

The big question is what was learned from those regressions? First, it is worth noting that each model did not predict "no completion", and each model's resulting accuracy was .9399. This aside, PctChildPoverty was shown to be less than capable of predicting completion, though none of the other four variables were able to capture even 10% of the variance (as evidenced by the MacFadden R^2 values).

1.2.6 Up-To-Date

Just knowing what contributes to schools reporting their vaccinations is not enough. Knowing what contributes to students not being up-to-date on their vaccinations is just as, if not more important. To that end, linear regression was performed to predict PctUpToDate from the five demographic variables. Regression was significant with all five variables, $F(5, 693) = 27.92, R^2 = 0.1617, p < 0.05$, however not all variables were significant. Below are the regression results for the variables and intercept.

Coefficient	Estimate	Error	t-value	P(> $\|t\|$)
(Intercept)	6.27558	0.370429	16.941	$< 2 \times 10^{-16}$
PctChildPoverty	-0.005024	0.009585	-0.524	0.600297
PctFreeMeal	-0.012623	0.003735	-3.380	0.000766
PctFamilyPoverty	-0.179936	0.098152	-1.833	0.067196
Enrolled	-0.449300	0.092394	-4.863	1.43×10^{-6}
TotalSchools	0.373002	0.127671	2.922	0.003596

Table 5: Linear regression results

It is clear that PctChildPoverty is once again not contributory to the model, nor, in this case, is PctFamilyPoverty (however it just misses the cutoff). As such, at least PctChildPoverty can be dropped, and the resulting model is significant, $F(4, 694) = 34.86, R^2 = 0.1625, p < 0.05$. Coefficient results are in Table 6.

Coefficient	Estimate	Error	t-value	$\Pr(>\lvert t\rvert)$
(Intercept)	6.27598	0.37023	16.951	$< 2 \times 10^{-16}$
PctFreeMeal	-0.01324	0.00354	-3.741	0.000198
PctFamilyPoverty	-0.21357	0.07424	-2.877	0.004139
Enrolled	-0.44497	0.09198	-4.838	1.62×10^{-6}
TotalSchools	0.37435	0.12758	2.934	0.003454

Table 6: Modified linear regression results

The resulting equation thus has the form

$$Y = 6.227598 - 0.01324X_2 - 0.21357X_3 - 0.44497X_4 + 0.37435X_5.$$

Assuming a null hypothesis of $R^2 = 0$ for the overal model, the odds that the null hypothesis is false are 2.051919×10^{23}:1, which is strong, to say the least. Following the same Bayesian line of thinking that provides these odds, a different equation can be constructed using coefficients generated using a Bayesian regression, it takes the form

$$Y_{Bayesian} = 3.0596 - 0.01277X_2 - 0.20829X_3 + 0.36478X_4 - 0.43443X_5.$$

Figure 13 shows distribution of Bayesian B-weights for each of the four variables used in the model, with HDI marked for each. Notice that the mean for each is roughly the associated coefficient in the equation above.

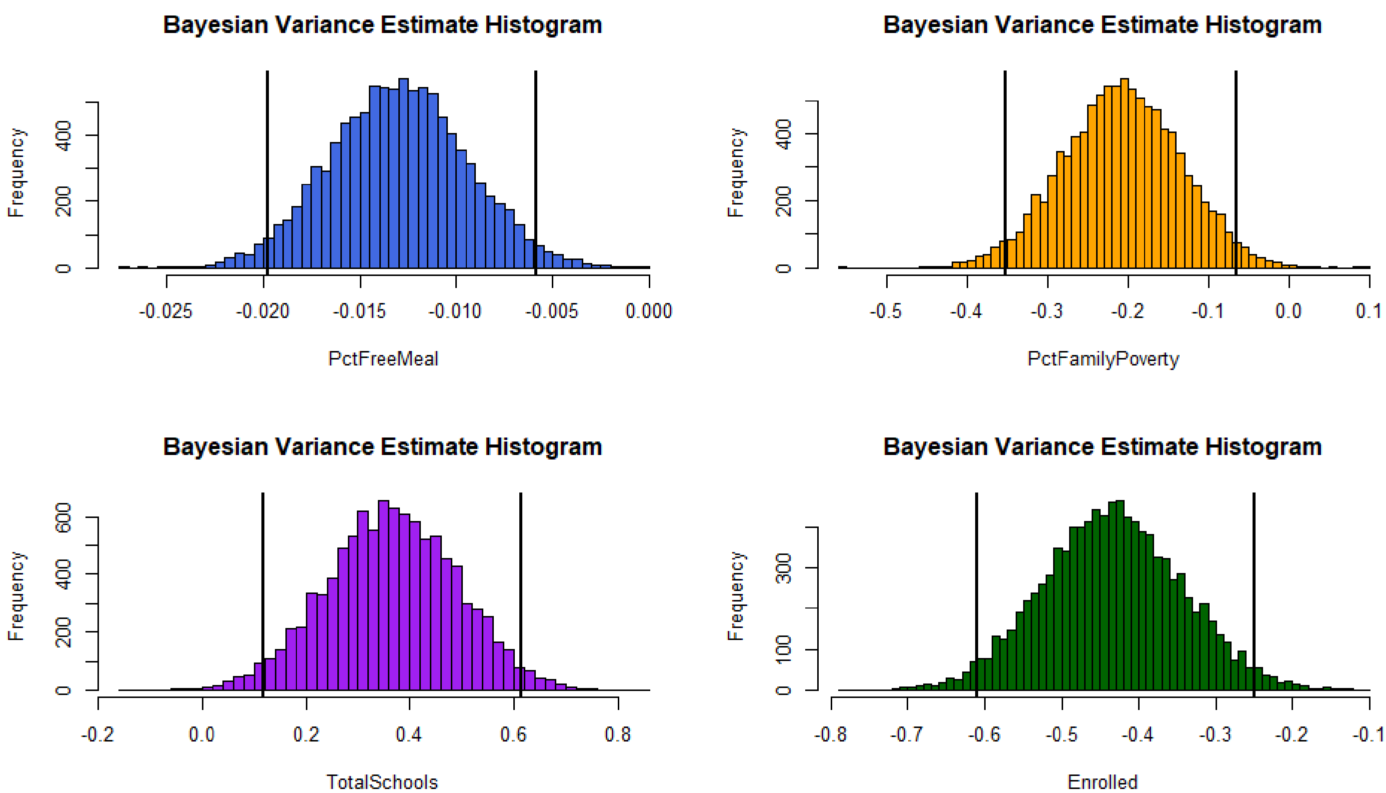


Figure 13: Bayesian coefficient variances

In the end, once again PctChildPoverty does not seem to have much impact on vaccines being up to date, while the other four variables collectively account for $R^2 = 16.25\%$ of the variation. Not great, but it's something.

1.2.7 Religious Exemptions

A similar approach was taken to test predictability of up-to-date vaccinations can be taken to check the percentage of students claiming religious exemption. In this case, only two variables are ultimately beneficial to keep, PctFreeMeal and Enrolled, $F(2, 696) = 102.8, R^2 = 0.2258, p < 0.05$. Table 7 contains the regression results.

Coefficient	Estimate	Error	t-value	Pr($>\|t\|$)
(Intercept)	4.028738	0.203609	19.787	$< 2 \times 10^{-16}$
PctFreeMeal	-0.027566	0.002107	-13.086	$< 2 \times 10^{-16}$
Enrolled	-0.169206	0.033262	-5.087	4.68×10^{-7}

Table 7: Linear Regression results

As before, the odds of the null hypothesis ($R^2 = 0$) are 3.16221×10^{36}:1, so there is clearly sufficient evidence to reject the null hypothesis. Here are the standard and Bayesian equivalent equations:

$$Y = 4.028738 - 0.027566X_2 - 0.169206X_4$$

$$Y_{Bayesian} = 1.80745 - 0.02865X_2 - 0.16724X_4,$$

and here are the distributions for Bayesian B-weights for each variable.

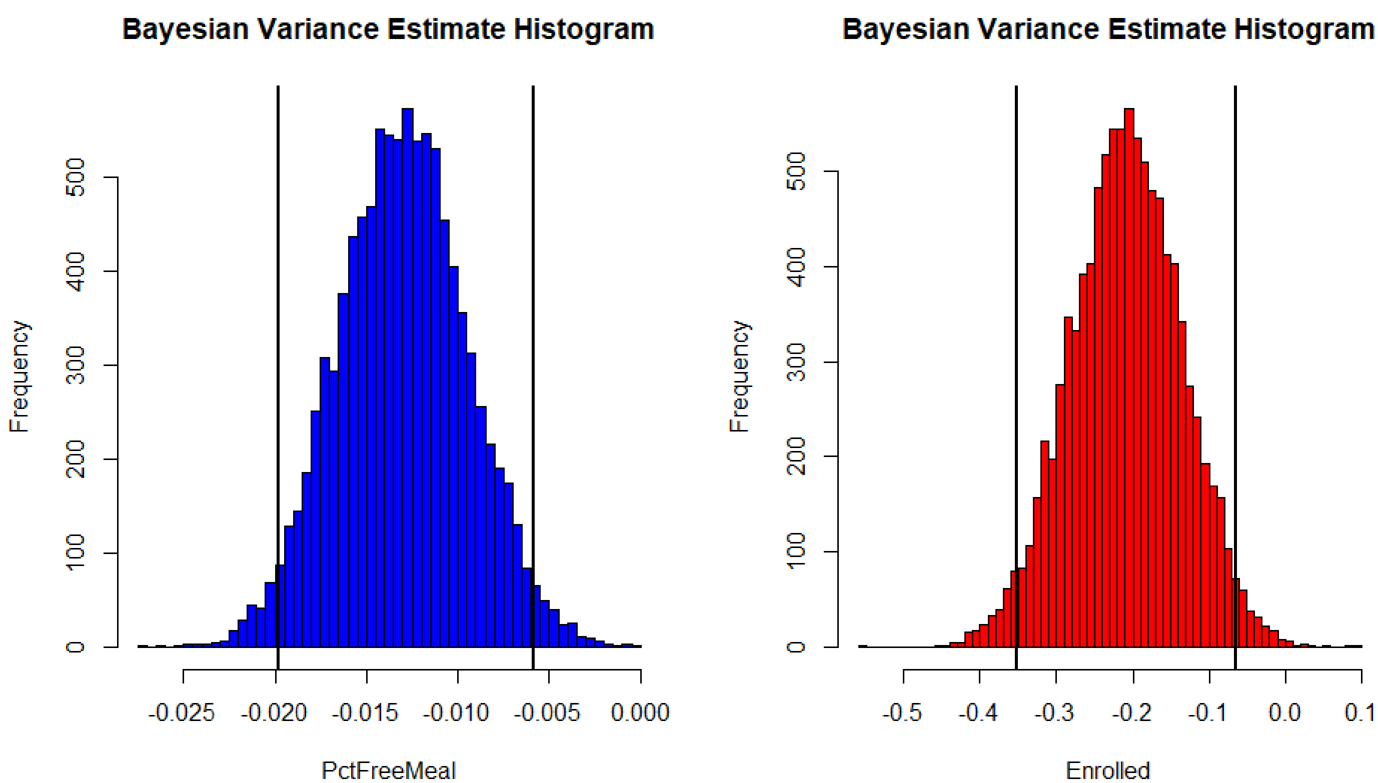


Figure 14: Bayesian coefficient variances

Intuitively, it should not be surprising that only the two variables were significant in this model. Child and family poverty levels should not (again, intuitively) impact whether a religious belief exemption would be present. Nor, for that matter, would the number of schools. However, the numebr of enrolled students certainly would. The perhaps surprising variable was PctFreeMeal. However, it has come up in each test and model conducted thus far.

These results further suggest that PctChildPoverty is of no inherit use. This is not overly surprising, given how incredibly correlated it is with PctFamilyPoverty (0.8385).

2 Results

It seems advantageous at this point to summarize the findings from the previous section.

1. Vaccination rates have, overall, increased over the last 25 years, however some vaccines are more prevalent than others.
2. Even though it has the highest vaccination rate, Hepatitis B seems to be a sticking point insofar as it is indicative of other vaccines being missed.
3. Some focus needs to be placed on private schools, as they are less likely to even report their vaccination rates than public schools.

4. With the exception of DTP, California is ahead of the US in terms of vaccination rates.
5. The percentage of students with free meals, family poverty, and enrolled students in a district can predict the percentage of students without the HepB vaccine, which is a strong indicator that a student does not have the other vaccines.
6. Except for child poverty, each demographic variable (individually) can predict whether a district was complete in its vaccination reporting, though the strength of the prediction is not large.
7. The demographic variables together (again, except for child poverty) can predict the percentage of the student body that is up-to-date in their vaccinations.
8. The number of enrolled students and (surprisingly) the percentage of students with free meals can predict the percentage of students claiming religious exemption to vaccinations.

Obviously, minimizing the percentage of students without Hepatitis B would reduce the overall number of students that are not up-to-date. Conveniently, minimizing that should have an impact on the number of districts that are complete in their reporting, as it not unreasonable to conjecture a district would not be complete in their reporting if they did not have a “sufficiently” good ratio of up-to-date students. This could be done by simultaneously minimizing the models defined in sections 1.2.4 and 1.2.5. Alas, the present author was not able to complete this linear programming in time, so it is left for another study.

Instead, some general recommendations can be made:

1. Every district deserves funding, but those with already high numbers of students with free meals and high numbers of students seem to need the most attention (those variables were present in every model developed).
2. While private schools may or may not qualify for state fudning (depending on the school), some kind of effort should be made to increase awareness in those locations, both to get students up-to-date and increase reporting.
3. Since missing Hepatitis B is a direct indicator of missing other vaccines, it should be prioritized in literature distribution (flyers, posters, mailers, etc.) and parent information/education sessions. This will also increase the number of students with the other vaccines.
4. Religious freedom is important (it is, after all, in the Constitution) and should be respected, however those claiming religious exemption should not be overlooked.

3 Conclusion

It is, ultimately, up to parents to choose to get their children vaccinated, and up to individual schools to report their vaccination rates. Whatever the reason parents choose not to vaccinate their child, care must be taken in how they are approached or there is a good chance the message will be lost. Getting schools to report, on the other hand, is a bit more in the control of the state. Witholding funding is a (albeit draconian) way that might spur districts to get their schools to report. However this would ultimately only hurt the children, so some other method of motivation should be utilized. It is the opinion of the author that the State of California should be aggressive in its vaccination campaign, however not at the expense of the students it is seeking to protect.

There will, of course, always be holdouts (anti-vaxxers) and people that refuse to listen to the opinions of others (anyone “across the aisle”, regardless of where you’re sitting). Indeed this is true of just about anything (flat-Earthers), and it takes a special kind of person to be able to handle all the malarky that comes from them (mediators/negotiators). But perhaps the state can develop a good awareness campaign that can bridge the gap and reach those individuals and schools that don’t want to get with the program. It is the hope of the present author that this work can serve as a foundation for how to build and develop that campaign.

The politics we leave to someone else.

References

[1] Bazin H (2011). *Vaccination: A History.* John Libbey Eurotext. p. 395. ISBN 978-2742007752.

[2] Nelson MC, Rogers J (December 1992). ”The right to die? Anti-vaccination activity and the 1874 smallpox epidemic in Stockholm”. Social History of Medicine. 5 (3): 369–88.

[3] Smith, P. J., Wood, D., & Darden, P. M. (2011). Highlights of historical events leading to national surveillance of vaccination coverage in the United States. Public health reports (Washington, D.C. : 1974), 126 Suppl 2(Suppl 2), 3–12. https://doi.org/10.1177/00333549111260S202