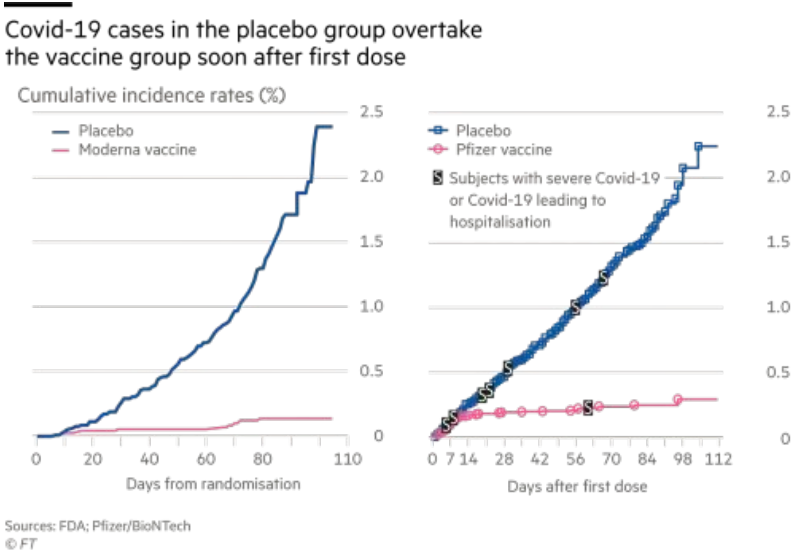
Vaccination Rates & New Case Counts

University of California, Berkeley
4/13/2021

Dera C., Pavan E., Greg T., Kayla W.

Research Question

Are vaccinations causing a decrease in new COVID-19 cases?



We see from controlled trials that these vaccines prevent cases.

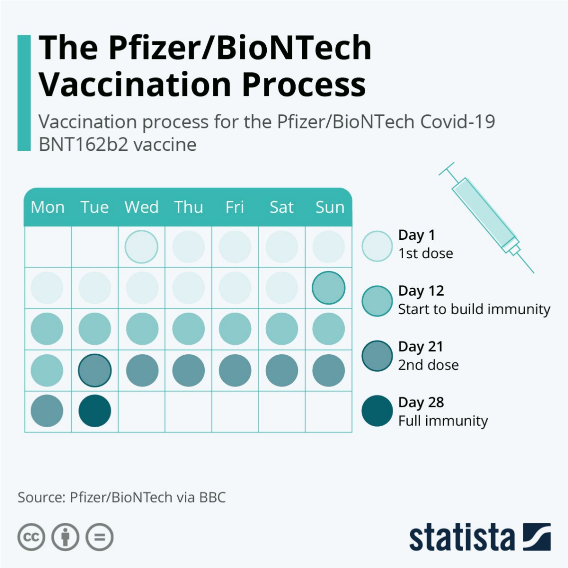
What do we see in the real world?

Other factors driving cases:

- “Risky Behaviors”
- Not social distancing
- Not adhering to wearing masks
- Creating space for viral spread

Data & Preparation

2+ Week Lag for Efficacy



CDC Data on New Covid Cases

Aggregated weekly and presented as new cases per 100k

OWID Data on Vaccinations

Reported daily and presented as people per 100, or % vaccinated (sourced from CDC)

Policy Actions by KFF

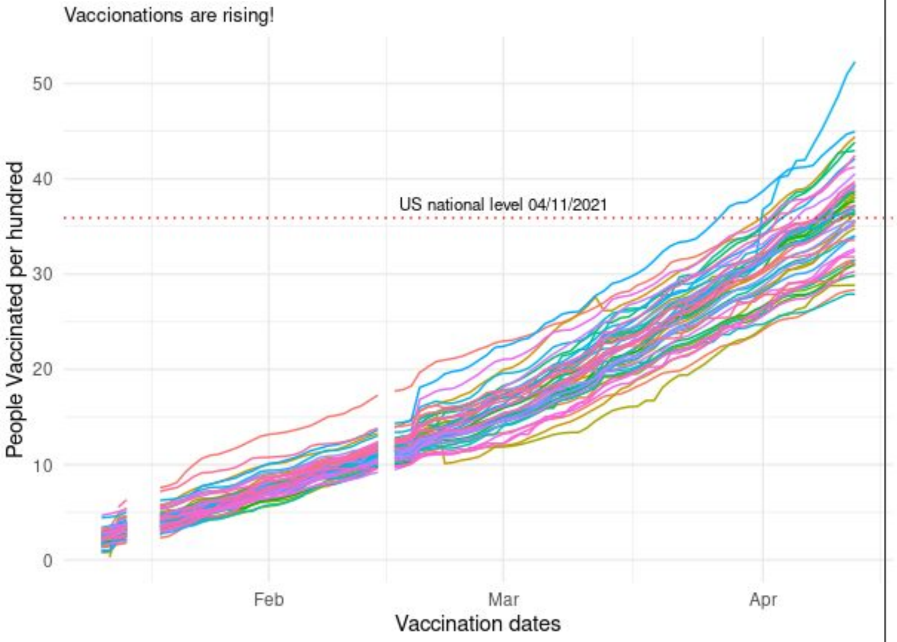
Taken for the last week in March 2021

Community Movement from Google

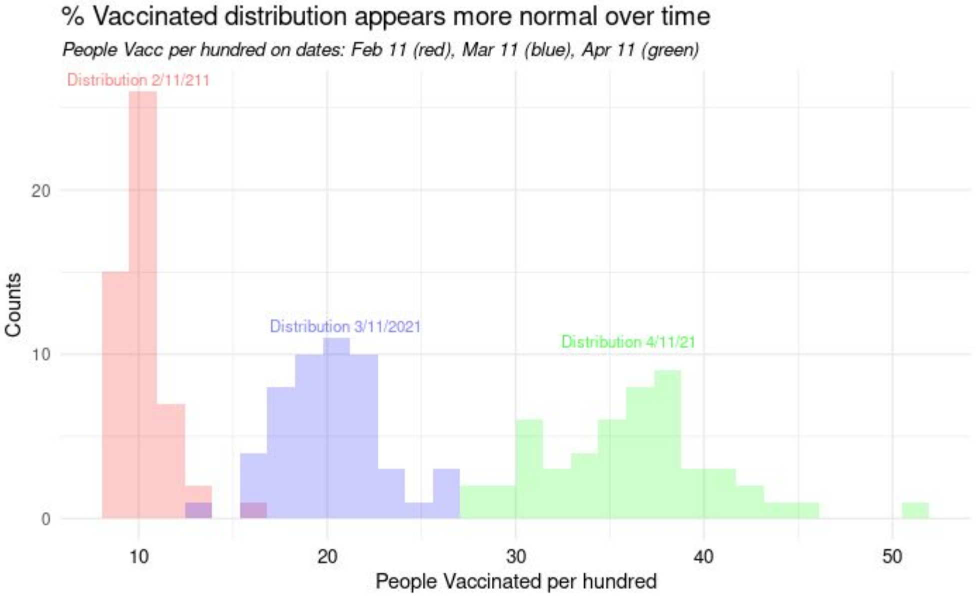
Logged as percent change from baseline for given community. Categorized by High, Med, Low based on quartiles.

EDA of % Population Vaccinated

% Vaccinated is rising!



Distribution becomes more normal over time



Base Model

$$f(\text{Log}(\text{Count New Cases})) = \beta_0 + \beta_1(\%\ \text{Population Vaccinated})$$

T test

```
t test of coefficients:

                             Estimate Std. Error t value  Pr(>|t|)
(Intercept)                  3.690186   0.556829  6.6271 2.509e-08 ***
people_vaccinated_per_hundred 0.054373   0.028701  1.8945   0.06407 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

% Population Vaccinated is **<u>not statistically significant</u>**

Improved Model
- Add population density & “risky behavior” variables

$$f(\text{Log(Count New Cases)}) = \beta_0 + \beta_1(\text{\% Population Vaccinated}) + \beta_2(\log(\text{Density}) + \beta_3(\text{Workplace Movement}) + \beta_4(\text{Retail Movement}) + \beta_5(\text{Transit Movement})$$

T test

```
t test of coefficients:

                            Estimate Std. Error t value  Pr(>|t|)
(Intercept)                 2.665557   0.635621  4.1936  0.000127 ***
people_vaccinated_per_hundred 0.068029 0.020325  3.3471  0.001657 **
log(Density)                0.165464   0.050027  3.3075  0.001857 **
Workplace_flag              0.391674   0.157921  2.4802  0.016939 *
Retail_flag                -0.168079   0.200767 -0.8372  0.406913
Transit_flag               -0.236087   0.142132 -1.6610  0.103658
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

% Population Vaccinated, Density and Workplace movement are **statistically significant**

The Kitchen Sink Model
- Add more “risky behavior” and policy mandates

$$f(\text{Log(Count New Cases)}) = \beta_0 + \beta_1(\text{\% Population Vaccinated}) + \beta_2(\text{log(Density)}) + \beta_3(\text{Workplace Movement}) + \beta_4(\text{Retail Movement}) +$$
$$\beta_5(\text{Transit Movement}) + \beta_6(\text{Bussiness Closure Policy}) + \beta_7(\text{Bar Closures}) + \beta_8(\text{Mask Policy}) + \beta_9(\text{Parks Movement}) + \beta_1 0(\text{Grocery Movement})$$

T test

```
t test of coefficients:

                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                   2.556280   0.904969  2.8247 0.007346 **
people_vaccinated_per_hundred 0.075011   0.028445  2.6371 0.011851 *
log(Density)                  0.147084   0.058204  2.5270 0.015559 *
Workplace_flag                0.283076   0.179545  1.5766 0.122758
Retail_flag                  -0.213018   0.221253 -0.9628 0.341443
Transit_flag                 -0.343816   0.174081 -1.9750 0.055194 .
Bussiness_Flag                0.347654   0.290626  1.1962 0.238652
Bar_Flag                     -0.232790   0.217567 -1.0700 0.291048
Mask_Flag                    -0.022192   0.284057 -0.0781 0.938118
Parks_flag                    0.088718   0.130827  0.6781 0.501594
Grocery_flag                  0.090847   0.178501  0.5089 0.613585
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

% Population Vaccinated and Density are **statistically significant**

Model Comparisons

Improved Model is the best performing model:

- As seen in Regression Table, all variables have have **lowest Standard error**

- F-Tests show that the **Improved model is better performing than Base Model**
  - No evidence that “Kitchen sink” is better

Table 1: The relationship between new covid cases and % of population vaccinated

	Dependent variable:		
	log(cumuluative_new_case_7_per100000)		
	(1)	(2)	(3)
people_vaccinated_per_hundred	0.054	0.068*	0.075*
	(0.028)	(0.027)	(0.029)
log(Density)		0.165**	0.147*
		(0.060)	(0.061)
Workplace_flag		0.392**	0.283
		(0.135)	(0.149)
Retail_flag		-0.168	-0.213
		(0.149)	(0.174)
Transit_flag		-0.236	-0.344
		(0.143)	(0.170)
Bussiness_Flag			0.348
			(0.216)
Bar_Flag			-0.233
			(0.157)
Mask_Flag			-0.022
			(0.225)
Parks_flag			0.089
			(0.109)
Grocery_flag			0.091
			(0.157)
Constant	3.690***	2.666***	2.556**
	(0.905)	(0.754)	(0.828)
Observations	51	51	51
R2	0.069	0.404	0.475
Adjusted R2	0.050	0.338	0.344
Residual Std. Error	0.547 (df = 49)	0.457 (df = 45)	0.454 (df = 40)
Note:	*p<0.05; **p<0.01; ***p<0.001		

Conclusions

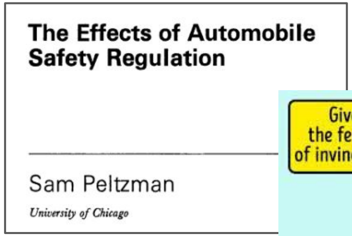
- Improved Model (the best model) indicates that:

> **1 more person out a hundred with a vaccine** is associated with a **6.8% percent increase in Cumulative Cases per 100k in Last 7 Days** (given all else equal)

- **Our original hypothesis was wrong** (with our dataset)

- This is surprising, **but not novel!**

  Could be explained by the **Peltzman Effect**


The Effects of Automobile Safety Regulation

Sam Peltzman

University of Chicago

Source: “The Effects of Automobile Safety Regulation,” Journal of Political Economy, Vol. 83, No. 4 (Aug., 1975),

Next Steps - Address major model limitations

- Test Peltzman Effect

- Reduce Scale from States to Counties (increase sample size)

- Find proxies for Omitted Variables:
    - Face Mask Adherence
    - Prevalence of Covid Variants per state
    - Age
    - Returning to school
    - Social gatherings

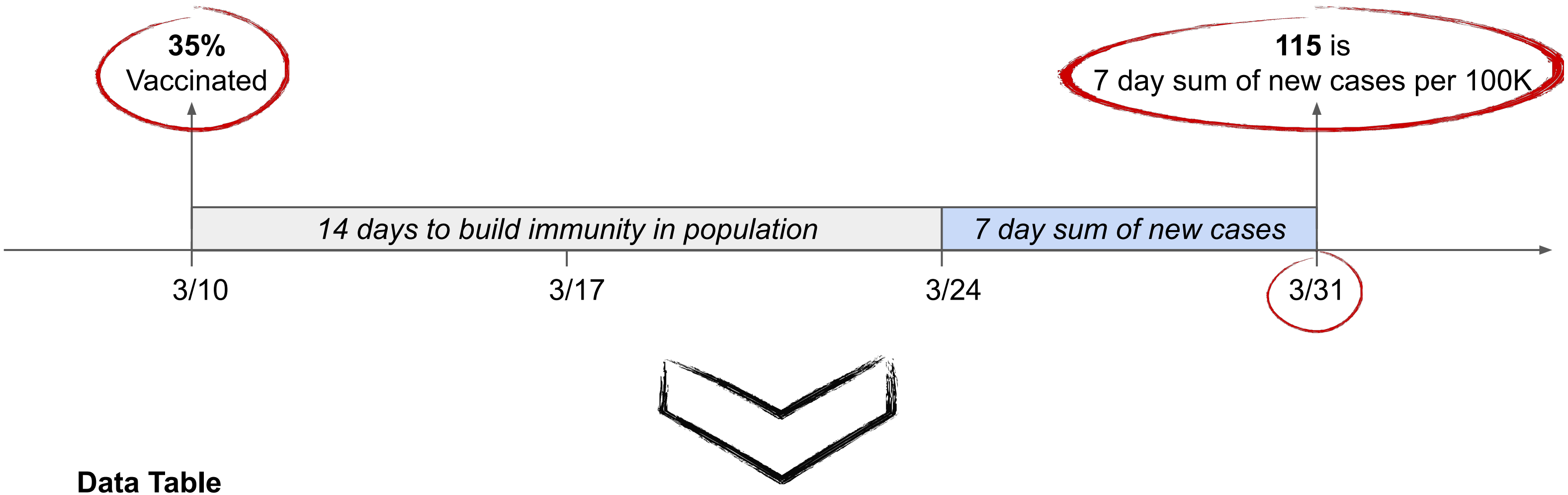
Questions?

University of California, Berkeley
4/13/2021

Dera C., Pavan E., Greg T., Kayla W.

Appendix - Need time to build Immunity and Count Cases

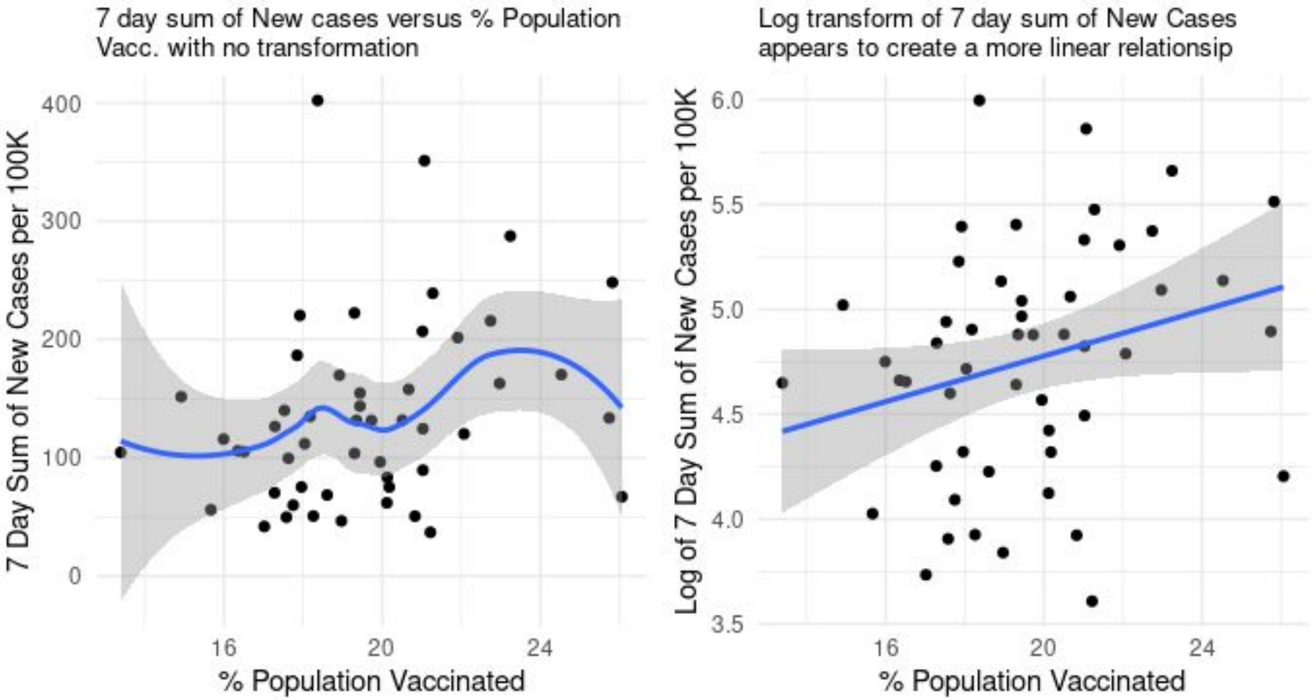
- 14 day buffer from vaccination date to sum of new cases



Data Table

State	Effective Date	% Vaccinated	7 day sum of new cases per 100K
CA	March 31, 2021	35	115

Appendix - Data Transformations: Log-Linear Base Model



Appendix- CLM of Improved Model

1. IID Sampling
2. Linear Conditional Expectation
3. No Perfect Collinearity
4. Homoskedastic Errors
5. Normally Distributed Errors

Appendix- CLM of Improved Model

1. **IID Sampling**
2. Linear Conditional Expectation
3. No Perfect Collinearity
4. Homoskedastic Errors
5. Normally Distributed Errors

Covid case counts

Vaccination Rates

Density

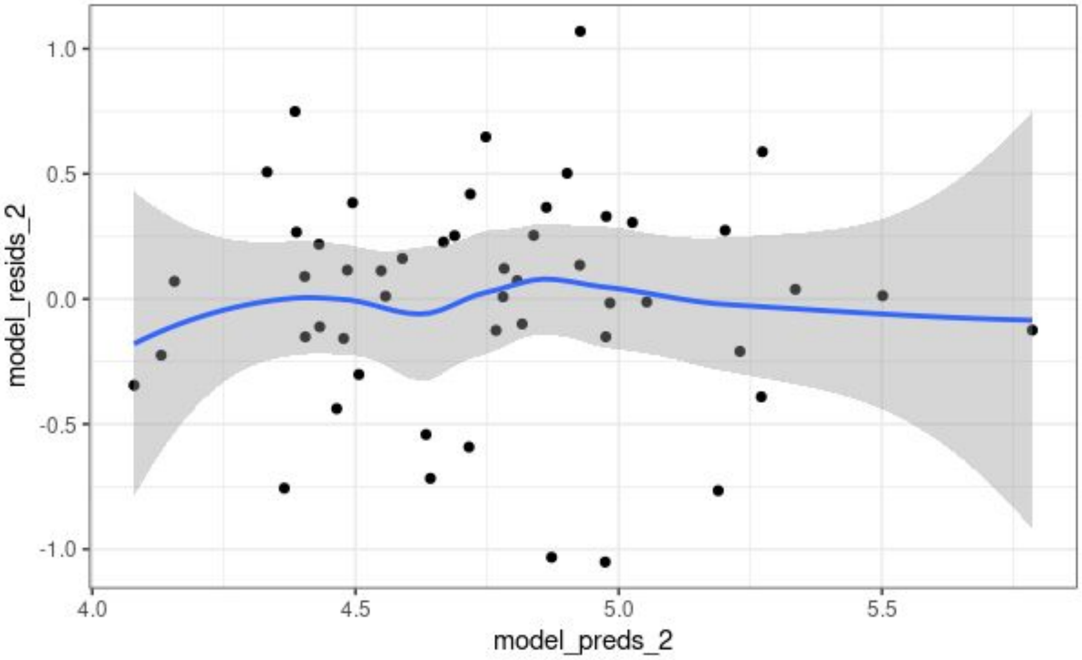
Policy Data

Movement Data

Issues with Independence
No Issues with IID assumptions

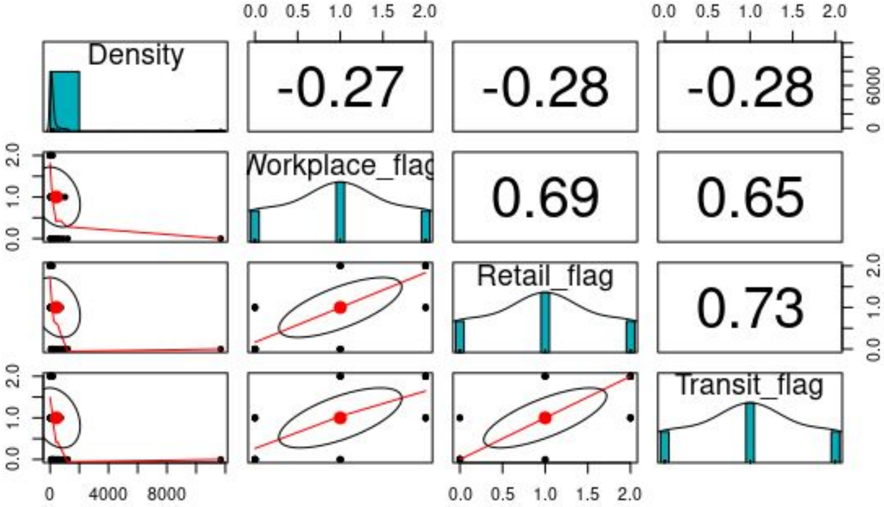
Appendix- CLM of Improved Model

1. IID Sampling
2. **Linear Conditional Expectation**
3. No Perfect Collinearity
4. Homoskedastic Errors
5. Normally Distributed Errors



Appendix- CLM of Improved Model

1. IID Sampling
2. Linear Conditional Expectation
3. **No Perfect Collinearity**
4. Homoskedastic Errors
5. Normally Distributed Errors



VIF Results:

```
people_vaccinated_per_hundred        log(Density)
Workplace_flag
                     1.242642        2.009528
2.262572
                  Retail_flag        Transit_flag
                     2.777849        2.561925
```

Appendix- CLM of Improved Model

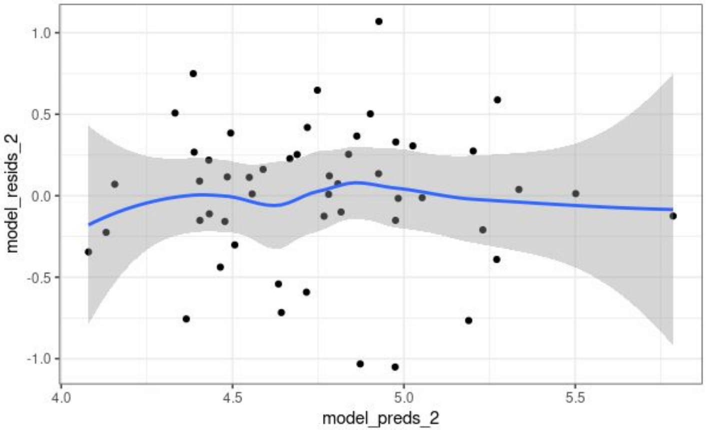
1. IID Sampling
2. Linear Conditional Expectation
3. No Perfect Collinearity
4. **Homoskedastic Errors**
5. Normally Distributed Errors

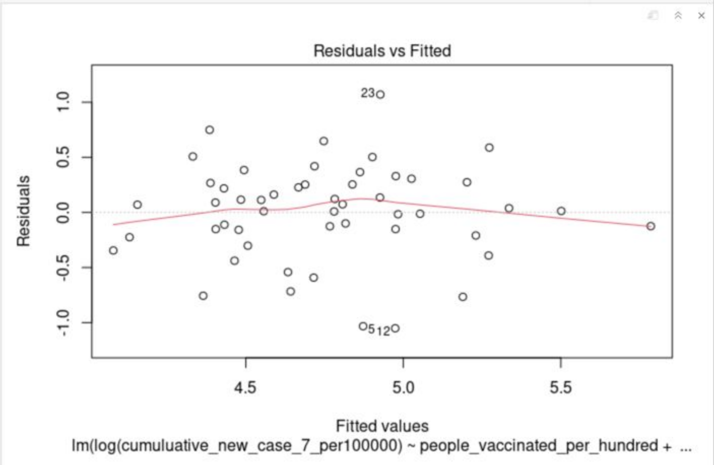
To further test for homoskedastic errors we also ran the Breush Pagan test:

```{r}
bptest(model_two)
```

```
	studentized Breusch-Pagan test

data:  model_two
BP = 5.4459, df = 5, p-value = 0.3639
```





Appendix- CLM of Improved Model

1. IID Sampling
2. Linear Conditional Expectation
3. No Perfect Collinearity
4. Homoskedastic Errors
5. **Normally Distributed Errors**