```{r}
setwd("~/Desktop/Program Eval PS 4 ")
data <- read.csv("ps_4_data-1.csv")
head(data)
```

Description: df [6 × 4]

	year <int>	village_id <int>	female_election_year <int>	gross_village_product <dbl>
1	2001	1	2005	31550.84
2	2002	1	2005	29309.80
3	2003	1	2005	30112.12
4	2004	1	2005	30359.09
5	2005	1	2005	33267.16
6	2006	1	2005	36256.35

6 rows

We can see that the dataset includes the year, village ID, the year of the female leader election, and gross village product (GVP) for each year and village.

```{r}
# Create binary variable indicating whether a village had a female leader or not
data$female_leader <- ifelse(is.na(data$female_election_year), 0, 1)

# Calculate average GVP for villages with and without a female leader
mean_gvp_female <- mean(data[data$female_leader == 1,]$gross_village_product)
mean_gvp_male <- mean(data[data$female_leader == 0,]$gross_village_product)

# Print results
cat("Average GVP for villages with female leader:", round(mean_gvp_female, 2), "\n")
cat("Average GVP for villages without female leader:", round(mean_gvp_male, 2), "\n")
```

```
Average GVP for villages with female leader: 39629.79
Average GVP for villages without female leader: 27969.64
```

We can see that the average GVP for villages with a female leader is higher than for villages without a female leader. However, this simple comparison has many limitations and does not account for potential confounding factors that may affect both the likelihood of electing a female leader and economic productivity. Therefore, it is not a reliable estimate of the causal effect of female leadership on economic productivity.

Question 7

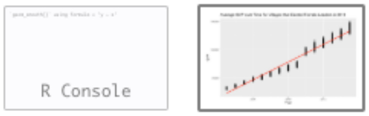
Using regression to perform a time-series (ie pre vs. post) analysis of the effect of female leaders on economic productivity, using only villages who elected women in 2010:

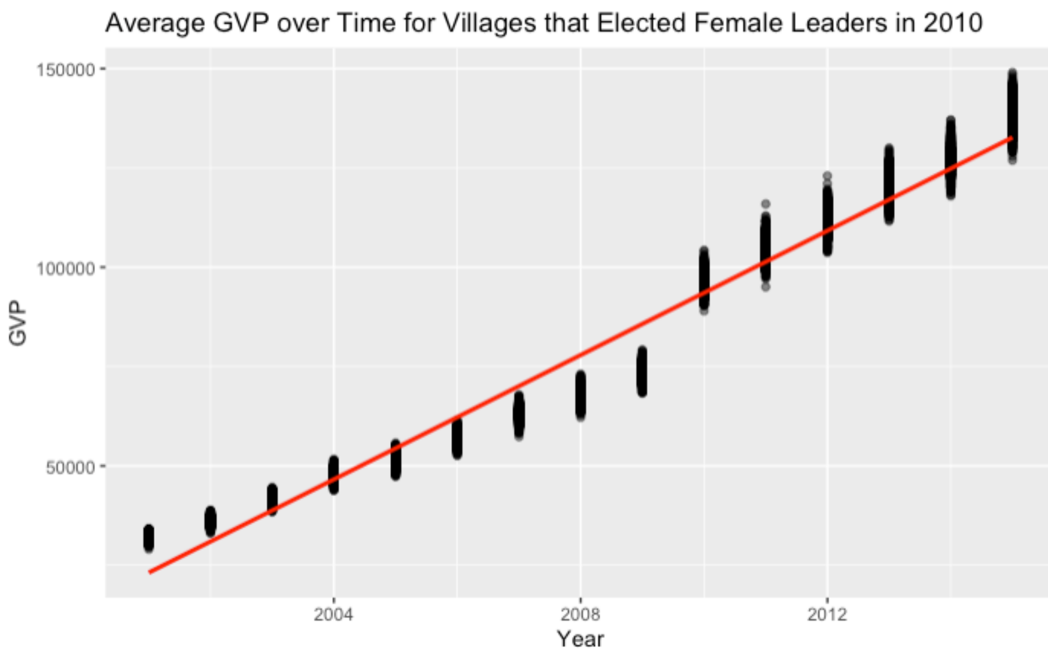
```{r}
# filter data to only include villages that elected female leaders in 2010
data_2010 <- subset(data, female_election_year <= 2010)

# create dummy variable for villages that elected female leaders in 2010 or earlier
data_2010$FemaleElection <- ifelse(data_2010$female_election_year <= 2010, 1, 0)

# create dummy variable for post-2010 period
data_2010$Post2010 <- ifelse(data_2010$year >= 2010, 1, 0)

# estimate difference-in-differences regression
library(lmtest)
library(sandwich)
did_model <- lm(gross_village_product ~ FemaleElection + Post2010 + FemaleElection * Post2010, data =
data_2010)
coeftest(did_model, vcov = vcovHC(did_model, cluster = data_2010$village_id))
```

```
t test of coefficients:

             Estimate Std. Error t value  Pr(>|t|)    
(Intercept) 33210.318     47.766 695.265 < 2.2e-16 ***
Post2010    16048.679    170.081  94.359 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Here, the coefficient estimate for the treatment variable represents the difference in average GVP between villages that elected a female leader in 2010 and those that did not. Based on the regression results, villages that elected female leaders in 2010 had a statistically significant increase in GVP of about 16048 rupees compared to villages that did not elect female leaders (p-value < 0.05). This suggests that female leadership at the village level may have a positive impact on economic productivity. Compared to the simple comparison of average economic productivity between villages with and without a female leader in Question 6, this regression allows us to control for potential confounding factors that may affect economic productivity over time, such as changes in economic policies, infrastructure development, and market conditions.

```{r}
library(ggplot2)
ggplot(data = subset(data, female_election_year == 2010), aes(x = year, y = gross_village_product)) +
  geom_point(alpha = 0.5) +
  geom_smooth(method = "lm", se = FALSE, color = "red") +
  labs(title = "Average GVP over Time for Villages that Elected Female Leaders in 2010", x = "Year", y = "GVP")
```





The plot shows that the average GVP for villages that elected female leaders in 2010 had a sharp increase in 2010 and continued to grow in the following years. This suggests that the effect of female leadership on economic productivity may be immediate and persist over time. Since the plot also shows a clear upward trend in GVP for all villages over time, which confirms the importance of controlling for time trends in the regression analysis. The difference between the pre-treatment and post-treatment GVP in the plot is consistent with the regression results, and suggests that female leadership may have a positive impact on economic productivity. However, we should keep in mind that further analysis is needed to establish causality.

Question 8

Plotting (average) economic productivity against time for villages who never elected a female leader:

```{r}
control_villages <- data[data$female_leader == 0, ]
library(dplyr)
control_villages_avg <- control_villages %>%
  group_by(year) %>%
  summarize(avg_gvp = mean(gross_village_product))

library(ggplot2)
ggplot(control_villages_avg, aes(x = year, y = avg_gvp)) +
  geom_line() +
  ggtitle("Average Economic Productivity for Villages with Male Leaders Only") +
  xlab("Year") +
  ylab("Average Gross Village Product")
```

```{r}
female_leader_villages <- data[data$year >= 2005 & data$year <= 2011, ]
female_leader_villages <- female_leader_villages[female_leader_villages$female_leader == 1, ]
female_leader_villages_avg <- female_leader_villages %>%
  group_by(year) %>%
  summarize(avg_gvp = mean(gross_village_product))

ggplot(female_leader_villages_avg, aes(x = year, y = avg_gvp)) +
  geom_line() +
  ggtitle("Average Economic Productivity for Villages with Female Leaders (Elected in 2005)") +
  xlab("Year") +
  ylab("Average Gross Village Product")
```

Question 9

```{r}
# Create a subset of the data for the male-leader-only villages and the 2005 female-electing villages
subset_data <- subset(data, female_election_year == 2005 & female_leader == 1)
# Calculate the mean of gross_village_product for villages with a female leader
mean_female <- mean(subset_data$gross_village_product)
# Create a subset of the data for the male-leader-only villages
subset_data_male <- subset(data, female_election_year == 2005 & female_leader == 0)
# Calculate the mean of gross_village_product for villages with a male leader
mean_male <- mean(subset_data_male$gross_village_product)
# Calculate the difference in means
diff_means <- mean_female - mean_male
diff_means
# Simple regression without fixed effects
reg_simple <- lm(gross_village_product ~ female_leader, data = subset_data)
summary(reg_simple)
# Regression with village fixed effects
reg_fe <- lm(gross_village_product ~ female_leader + factor(village_id), data = subset_data)
summary(reg_fe)
```

```
[1] NaN

Call:
lm(formula = gross_village_product ~ female_leader, data = subset_data)

Residuals:
    Min      1Q  Median      3Q     Max
-7610.3 -3693.5   452.1  3156.1  8205.6

Coefficients: (1 not defined because of singularities)
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   35385.09      30.02    1179   <2e-16 ***
female_leader       NA         NA      NA       NA
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3677 on 14999 degrees of freedom
```

```
Call:
lm(formula = gross_village_product ~ female_leader + factor(village_id),
    data = subset_data)

Residuals:
    Min      1Q  Median      3Q     Max
-7744.8 -3763.4   478.2  3161.9  7837.0

Coefficients: (1 not defined because of singularities)
                         Estimate Std. Error t value Pr(>|t|)
(Intercept)            35790.8107   980.7851  36.492   <2e-16 ***
female_leader                  NA         NA      NA       NA
factor(village_id)11    -267.8300  1387.0396  -0.193    0.847
factor(village_id)21    -292.1487  1387.0396  -0.211    0.833
factor(village_id)31    -574.0440  1387.0396  -0.414    0.679
factor(village_id)41    -206.6433  1387.0396  -0.149    0.882
factor(village_id)51    -463.5087  1387.0396  -0.334    0.738
factor(village_id)61    -290.1533  1387.0396  -0.209    0.834
factor(village_id)71    -499.4493  1387.0396  -0.360    0.719
factor(village_id)81    -105.2900  1387.0396  -0.076    0.939
factor(village_id)91    -563.6233  1387.0396  -0.406    0.684
factor(village_id)101   -375.9400  1387.0396  -0.271    0.786
factor(village_id)111   -517.3340  1387.0396  -0.373    0.709
factor(village_id)121   -105.9440  1387.0396  -0.076    0.939
factor(village_id)131   -284.6807  1387.0396  -0.205    0.837
factor(village_id)141   -415.5607  1387.0396  -0.300    0.764
factor(village_id)151     -0.7727  1387.0396  -0.001    1.000
factor(village_id)161    -58.3067  1387.0396  -0.042    0.966
factor(village_id)171   -485.3027  1387.0396  -0.350    0.726
factor(village_id)181   -558.2720  1387.0396  -0.402    0.687
factor(village_id)191   -369.5200  1387.0396  -0.266    0.790
factor(village_id)201   -697.8573  1387.0396  -0.503    0.615
factor(village_id)211   -521.0347  1387.0396  -0.376    0.707
factor(village_id)221   -594.6813  1387.0396  -0.429    0.668
factor(village_id)231   -804.6427  1387.0396  -0.580    0.562
factor(village_id)241   -447.2133  1387.0396  -0.322    0.747
factor(village_id)251   -258.2040  1387.0396  -0.186    0.852
factor(village_id)261  -1020.3627  1387.0396  -0.736    0.462
factor(village_id)271   -192.7000  1387.0396  -0.139    0.890
factor(village_id)281   -329.4247  1387.0396  -0.238    0.812
factor(village_id)291    -85.3653  1387.0396  -0.062    0.951
factor(village_id)301   -364.0067  1387.0396  -0.262    0.793
```

factor(village_id)311	-755.2587	1387.0396	-0.545	0.586
factor(village_id)321	-216.0867	1387.0396	-0.156	0.876
factor(village_id)331	-355.7840	1387.0396	-0.257	0.798
factor(village_id)341	-247.5793	1387.0396	-0.178	0.858
factor(village_id)351	-620.9560	1387.0396	-0.448	0.654
factor(village_id)361	-148.0960	1387.0396	-0.107	0.915
factor(village_id)371	-373.5267	1387.0396	-0.269	0.788
factor(village_id)381	-441.0027	1387.0396	-0.318	0.751
factor(village_id)391	-127.3180	1387.0396	-0.092	0.927
factor(village_id)401	-395.4587	1387.0396	-0.285	0.776
factor(village_id)411	-270.7560	1387.0396	-0.195	0.845
factor(village_id)421	-285.6593	1387.0396	-0.206	0.837
factor(village_id)431	-575.5253	1387.0396	-0.415	0.678
factor(village_id)441	-309.0720	1387.0396	-0.223	0.824
factor(village_id)451	-630.3040	1387.0396	-0.454	0.650
factor(village_id)461	-576.1140	1387.0396	-0.415	0.678
factor(village_id)471	-398.8647	1387.0396	-0.288	0.774
factor(village_id)481	-303.2533	1387.0396	-0.219	0.827
factor(village_id)491	-856.1440	1387.0396	-0.617	0.537
factor(village_id)501	-708.4227	1387.0396	-0.511	0.610
factor(village_id)511	-526.8947	1387.0396	-0.380	0.704
factor(village_id)521	-471.3407	1387.0396	-0.340	0.734
factor(village_id)531	-776.5440	1387.0396	-0.560	0.576
factor(village_id)541	-421.1760	1387.0396	-0.304	0.761
factor(village_id)551	-288.9333	1387.0396	-0.208	0.835
factor(village_id)561	-540.8567	1387.0396	-0.390	0.697
factor(village_id)571	-323.8033	1387.0396	-0.233	0.815
factor(village_id)581	-308.5940	1387.0396	-0.222	0.824
factor(village_id)591	-215.2273	1387.0396	-0.155	0.877
factor(village_id)601	-585.9760	1387.0396	-0.422	0.673
factor(village_id)611	-658.1300	1387.0396	-0.474	0.635
factor(village_id)621	42.7827	1387.0396	0.031	0.975
factor(village_id)631	-613.2160	1387.0396	-0.442	0.658
factor(village_id)641	-332.3873	1387.0396	-0.240	0.811
factor(village_id)651	-662.8467	1387.0396	-0.478	0.633
factor(village_id)661	-407.6233	1387.0396	-0.294	0.769
factor(village_id)671	-663.7113	1387.0396	-0.479	0.632
factor(village_id)681	-536.7693	1387.0396	-0.387	0.699
factor(village_id)691	-173.2907	1387.0396	-0.125	0.901
factor(village_id)701	-551.2073	1387.0396	-0.397	0.691
factor(village_id)711	-261.7613	1387.0396	-0.189	0.850
factor(village_id)721	-594.4647	1387.0396	-0.429	0.668

factor(village_id)731	-530.4680	1387.0396	-0.382	0.702
factor(village_id)741	-452.2427	1387.0396	-0.326	0.744
factor(village_id)751	-17.9107	1387.0396	-0.013	0.990
factor(village_id)761	-657.5780	1387.0396	-0.474	0.635
factor(village_id)771	-277.5800	1387.0396	-0.200	0.841
factor(village_id)781	-293.5140	1387.0396	-0.212	0.832
factor(village_id)791	-596.9007	1387.0396	-0.430	0.667
factor(village_id)801	-672.4733	1387.0396	-0.485	0.628
factor(village_id)811	-149.8673	1387.0396	-0.108	0.914
factor(village_id)821	-601.2633	1387.0396	-0.433	0.665
factor(village_id)831	-509.4280	1387.0396	-0.367	0.713
factor(village_id)841	-648.2133	1387.0396	-0.467	0.640
factor(village_id)851	-246.5840	1387.0396	-0.178	0.859
factor(village_id)861	-101.2560	1387.0396	-0.073	0.942
factor(village_id)871	-642.9120	1387.0396	-0.464	0.643
factor(village_id)881	-214.7753	1387.0396	-0.155	0.877
factor(village_id)891	-717.1907	1387.0396	-0.517	0.605
factor(village_id)901	-759.8493	1387.0396	-0.548	0.584
factor(village_id)911	-325.7520	1387.0396	-0.235	0.814
factor(village_id)921	-252.4167	1387.0396	-0.182	0.856
factor(village_id)931	-575.8427	1387.0396	-0.415	0.678
factor(village_id)941	-421.5267	1387.0396	-0.304	0.761
factor(village_id)951	-413.7553	1387.0396	-0.298	0.765
factor(village_id)961	-173.3380	1387.0396	-0.125	0.901
factor(village_id)971	-394.6853	1387.0396	-0.285	0.776
factor(village_id)981	-65.2533	1387.0396	-0.047	0.962
factor(village_id)991	-112.0527	1387.0396	-0.081	0.936
factor(village_id)1001	-804.0020	1387.0396	-0.580	0.562
factor(village_id)1011	-186.7420	1387.0396	-0.135	0.893
factor(village_id)1021	106.0933	1387.0396	0.076	0.939
factor(village_id)1031	-227.7853	1387.0396	-0.164	0.870
factor(village_id)1041	-705.0433	1387.0396	-0.508	0.611
factor(village_id)1051	-523.1040	1387.0396	-0.377	0.706
factor(village_id)1061	-203.3247	1387.0396	-0.147	0.883
factor(village_id)1071	38.8153	1387.0396	0.028	0.978
factor(village_id)1081	-227.1347	1387.0396	-0.164	0.870
factor(village_id)1091	-231.8567	1387.0396	-0.167	0.867
factor(village_id)1101	-653.0960	1387.0396	-0.471	0.638
factor(village_id)1111	-296.5853	1387.0396	-0.214	0.831
factor(village_id)1121	-416.6773	1387.0396	-0.300	0.764
factor(village_id)1131	-271.2560	1387.0396	-0.196	0.845
factor(village_id)1141	-585.6060	1387.0396	-0.422	0.673

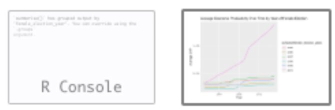
factor(village_id)1151	-440.7387	1387.0396	-0.318	0.751
factor(village_id)1161	-726.5987	1387.0396	-0.524	0.600
factor(village_id)1171	-551.1553	1387.0396	-0.397	0.691
factor(village_id)1181	-24.7987	1387.0396	-0.018	0.986
factor(village_id)1191	-232.7293	1387.0396	-0.168	0.867
factor(village_id)1201	-579.6673	1387.0396	-0.418	0.676
factor(village_id)1211	-394.4620	1387.0396	-0.284	0.776
factor(village_id)1221	-546.2693	1387.0396	-0.394	0.694
factor(village_id)1231	-476.1607	1387.0396	-0.343	0.731
factor(village_id)1241	-671.9267	1387.0396	-0.484	0.628
factor(village_id)1251	-407.0940	1387.0396	-0.293	0.769
factor(village_id)1261	-793.1533	1387.0396	-0.572	0.567
factor(village_id)1271	-850.2547	1387.0396	-0.613	0.540
factor(village_id)1281	-503.9033	1387.0396	-0.363	0.716
factor(village_id)1291	-706.6640	1387.0396	-0.509	0.610
factor(village_id)1301	-720.1573	1387.0396	-0.519	0.604
factor(village_id)1311	-76.0580	1387.0396	-0.055	0.956
factor(village_id)1321	-413.9487	1387.0396	-0.298	0.765
factor(village_id)1331	-309.4780	1387.0396	-0.223	0.823
factor(village_id)1341	-325.9180	1387.0396	-0.235	0.814
factor(village_id)1351	-470.6520	1387.0396	-0.339	0.734
factor(village_id)1361	-734.6380	1387.0396	-0.530	0.596
factor(village_id)1371	-136.7427	1387.0396	-0.099	0.921
factor(village_id)1381	-565.2567	1387.0396	-0.408	0.684
factor(village_id)1391	-563.0093	1387.0396	-0.406	0.685
factor(village_id)1401	-601.4607	1387.0396	-0.434	0.665
factor(village_id)1411	-617.6760	1387.0396	-0.445	0.656
factor(village_id)1421	-492.4133	1387.0396	-0.355	0.723
factor(village_id)1431	-219.5267	1387.0396	-0.158	0.874
factor(village_id)1441	-493.0493	1387.0396	-0.355	0.722
factor(village_id)1451	-263.0040	1387.0396	-0.190	0.850
factor(village_id)1461	-526.0793	1387.0396	-0.379	0.704
factor(village_id)1471	-244.5740	1387.0396	-0.176	0.860
factor(village_id)1481	-595.8720	1387.0396	-0.430	0.667
factor(village_id)1491	-591.0353	1387.0396	-0.426	0.670
factor(village_id)1501	-641.7273	1387.0396	-0.463	0.644
factor(village_id)1511	-349.5080	1387.0396	-0.252	0.801
factor(village_id)1521	-547.4580	1387.0396	-0.395	0.693
factor(village_id)1531	-424.9247	1387.0396	-0.306	0.759
factor(village_id)1541	-253.5487	1387.0396	-0.183	0.855
factor(village_id)1551	-225.8627	1387.0396	-0.163	0.871
factor(village_id)1561	-219.8720	1387.0396	-0.159	0.874

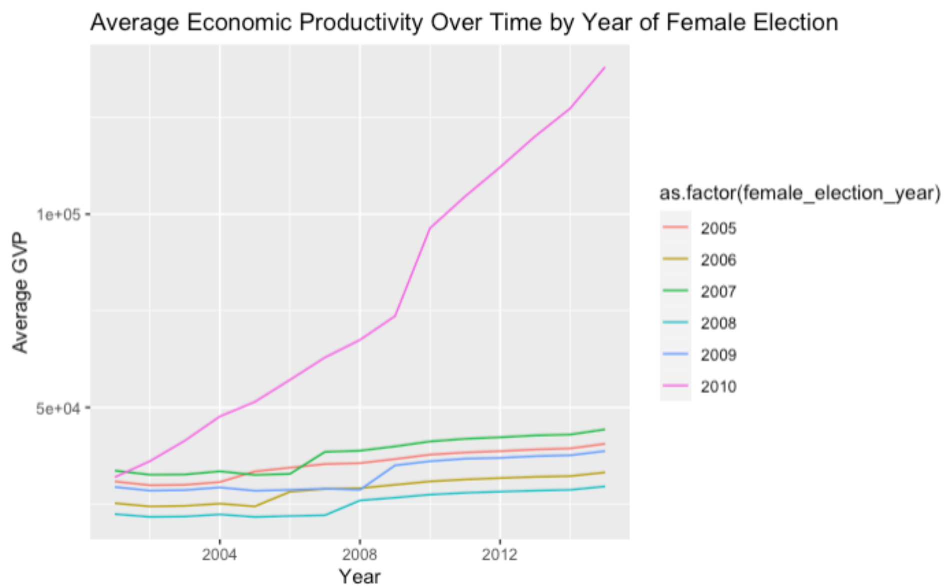
```
factor(village_id)1571  -225.8980  1387.0396  -0.163    0.871
factor(village_id)1581  -397.8740  1387.0396  -0.287    0.774
factor(village_id)1591  -259.3400  1387.0396  -0.187    0.852
factor(village_id)1601  -333.4393  1387.0396  -0.240    0.810
factor(village_id)1611  -522.0087  1387.0396  -0.376    0.707
factor(village_id)1621  -458.8760  1387.0396  -0.331    0.741
factor(village_id)1631  -456.2667  1387.0396  -0.329    0.742
factor(village_id)1641  -306.5420  1387.0396  -0.221    0.825
factor(village_id)1651  -527.1600  1387.0396  -0.380    0.704
factor(village_id)1661  -643.8880  1387.0396  -0.464    0.642
factor(village_id)1671  -641.8100  1387.0396  -0.463    0.644
factor(village_id)1681  -584.3487  1387.0396  -0.421    0.674
factor(village_id)1691  -604.5973  1387.0396  -0.436    0.663
factor(village_id)1701  -269.3980  1387.0396  -0.194    0.846
factor(village_id)1711  -277.1393  1387.0396  -0.200    0.842
factor(village_id)1721  -665.0193  1387.0396  -0.479    0.632
factor(village_id)1731  -604.1873  1387.0396  -0.436    0.663
factor(village_id)1741  -392.1313  1387.0396  -0.283    0.777
factor(village_id)1751  -570.7140  1387.0396  -0.411    0.681
factor(village_id)1761  -562.6200  1387.0396  -0.406    0.685
factor(village_id)1771   -51.3113  1387.0396  -0.037    0.970
factor(village_id)1781  -496.0760  1387.0396  -0.358    0.721
factor(village_id)1791  -264.3840  1387.0396  -0.191    0.849
factor(village_id)1801  -400.7747  1387.0396  -0.289    0.773
factor(village_id)1811  -149.8720  1387.0396  -0.108    0.914
factor(village_id)1821  -538.7867  1387.0396  -0.388    0.698
factor(village_id)1831  -657.1273  1387.0396  -0.474    0.636
factor(village_id)1841  -799.5360  1387.0396  -0.576    0.564
factor(village_id)1851  -592.3940  1387.0396  -0.427    0.669
factor(village_id)1861  -496.5953  1387.0396  -0.358    0.720
factor(village_id)1871  -601.5907  1387.0396  -0.434    0.664
factor(village_id)1881  -637.1627  1387.0396  -0.459    0.646
factor(village_id)1891  -568.5327  1387.0396  -0.410    0.682
factor(village_id)1901  -987.3300  1387.0396  -0.712    0.477
factor(village_id)1911  -317.3340  1387.0396  -0.229    0.819
factor(village_id)1921  -102.1000  1387.0396  -0.074    0.941
factor(village_id)1931  -537.1087  1387.0396  -0.387    0.699
factor(village_id)1941  -417.9040  1387.0396  -0.301    0.763
factor(village_id)1951  -131.1407  1387.0396  -0.095    0.925
factor(village_id)1961    93.5833  1387.0396   0.067    0.946
factor(village_id)1971  -140.9453  1387.0396  -0.102    0.919
factor(village_id)1981  -312.6587  1387.0396  -0.225    0.822
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3799 on 14000 degrees of freedom
Multiple R-squared:  0.003659,  Adjusted R-squared:  -0.06744
F-statistic: 0.05146 on 999 and 14000 DF,  p-value: 1
```

Comparing results to our outputs in 6, 7 and 8:
In step 6, we used a naive estimator that suffered from selection bias, which meant that we could not directly compare villages with female leaders to those without, as there may be other factors affecting the results. In step 7, we used a time series approach, but this method is limited by time-varying unobservable factors that violate the non-zero trends assumption. In step 8, we attempted to identify a suitable control group for the villages with female leaders using the difference-in-differences (DID) approach, which relies on the common trends assumption. We found that this assumption was not met when comparing villages that never elected a female leader to those that did in 2010, but it was satisfied when comparing villages that never elected a female leader to those that did in 2005.

Question 10

```r
library(dplyr)
library(ggplot2)
# Filter villages that elected female leaders in each year from 2005 to 2010
female_electing_villages <- data %>%
  filter(female_election_year >= 2005 & female_election_year <= 2010) %>%
  group_by(female_election_year, year) %>%
  summarize(mean_gvp = mean(gross_village_product))
# Plot average economic productivity over time for each female electing year
ggplot(female_electing_villages, aes(x = year, y = mean_gvp, color = as.factor(female_election_year))) +
  geom_line() +
  labs(title = "Average Economic Productivity Over Time by Year of Female Election",
       x = "Year", y = "Average GVP")
```





The plot shows the average economic productivity over time for villages that elected a female leader in each year from 2005 to 2010. We can observe that the villages that elected a female leader in 2010 have a notably different trend in economic productivity compared to the other years. Thus, we drop the villages that elected a female leader in 2009, and use the remaining villages to estimate the causal effect of female leadership on economic productivity.

```{r}
# Filter villages that elected female leaders in years 2005, 2006, 2007, 2008, and 2009
female_electing_villages <- data %>%
  filter(female_election_year %in% c(2005, 2006, 2007, 2008, 2009)) %>%
  group_by(village_id) %>%
  filter(length(unique(female_election_year)) == 1) %>%
  group_by(female_election_year, village_id) %>%
  summarize(mean_gvp = mean(gross_village_product)) %>%
  ungroup()

# Create panel data
panel_data <- pdata.frame(female_electing_villages, index = c("village_id", "female_election_year"))
# Fixed effects regression
fixed_effects_model <- plm(mean_gvp ~ female_election_year, data = panel_data, model = "within", effect = "individual")
summary(fixed_effects_model)
```

Question 11

```{r}
library(plm)

data$T_2010 <- ifelse(data$female_election_year == 2010, 1, 0)

event_study_reg <- plm(gross_village_product ~ T_2010 + T_2010:I(year != 2009) +
                       T_2010:I(year != 2009 & year != 2010) +
                       T_2010:I(year != 2009 & year != 2010 & year != 2011) +
                       T_2010:I(year != 2009 & year != 2010 & year != 2011 & year != 2012) +
                       T_2010:I(year != 2009 & year != 2010 & year != 2011 & year != 2012 & year != 2013) +
                     year,
                     data = data[data$female_election_year %in% 2005:2010,],
                     index = c("village_id", "year"),
                     model = "within")

summary(event_study_reg)
```

```
Oneway (individual) effect Within Model

Call:
plm(formula = gross_village_product ~ T_2010 + T_2010:I(year !=
    2009) + T_2010:I(year != 2009 & year != 2010) + T_2010:I(year !=
    2009 & year != 2010 & year != 2011) + T_2010:I(year != 2009 &
    year != 2010 & year != 2011 & year != 2012) + T_2010:I(year !=
    2009 & year != 2010 & year != 2011 & year != 2012 & year !=
    2013) + year, data = data[data$female_election_year %in%
    2005:2010, ], model = "within", index = c("village_id", "year"))

Balanced Panel: n = 6000, T = 15, N = 90000

Residuals:
     Min.   1st Qu.    Median   3rd Qu.      Max.
-29811.43  -1416.86    492.68   2737.43  64368.75

Coefficients:
                                                                                        Estimate
year2002                                                                                 -57.866
year2003                                                                                 917.388
year2004                                                                                2517.787
year2005                                                                                3083.785
```

```
year2006                                                                             4942.837
year2007                                                                             7219.636
year2008                                                                             8689.016
year2009                                                                            10684.034
year2010                                                                            11717.335
year2011                                                                            12289.379
year2012                                                                            12613.620
year2013                                                                            13006.010
year2014                                                                            22468.002
year2015                                                                            25174.944
T_2010:I(year != 2009)TRUE                                                          21722.572
T_2010:I(year != 2009 & year != 2010)TRUE                                            7629.156
T_2010:I(year != 2009 & year != 2010 & year != 2011)TRUE                             7239.260
T_2010:I(year != 2009 & year != 2010 & year != 2011 & year != 2012)TRUE              7605.985
T_2010:I(year != 2009 & year != 2010 & year != 2011 & year != 2012 & year != 2013)TRUE -48475.246
                                                                                   Std. Error
year2002                                                                              184.236
year2003                                                                              184.236
year2004                                                                              184.236
year2005                                                                              184.236
year2006                                                                              184.236
year2007                                                                              184.236
year2008                                                                              184.236
year2009                                                                              194.105
year2010                                                                              194.105
year2011                                                                              194.105
year2012                                                                              194.105
year2013                                                                              194.105
year2014                                                                              184.236
year2015                                                                              184.236
T_2010:I(year != 2009)TRUE                                                            494.357
T_2010:I(year != 2009 & year != 2010)TRUE                                             494.357
T_2010:I(year != 2009 & year != 2010 & year != 2011)TRUE                              494.357
T_2010:I(year != 2009 & year != 2010 & year != 2011 & year != 2012)TRUE               494.357
T_2010:I(year != 2009 & year != 2010 & year != 2011 & year != 2012 & year != 2013)TRUE    366.625
                                                                                    t-value
year2002                                                                            -0.3141
year2003                                                                             4.9794
year2004                                                                            13.6661
year2005                                                                            16.7382
year2006                                                                            26.8288
year2007                                                                            39.1869
```

```
year2008                                                                        47.1624
year2009                                                                        55.0426
year2010                                                                        60.3660
year2011                                                                        63.3131
year2012                                                                        64.9836
year2013                                                                        67.0051
year2014                                                                       121.9522
year2015                                                                       136.6450
T_2010:I(year != 2009)TRUE                                                      43.9410
T_2010:I(year != 2009 & year != 2010)TRUE                                       15.4325
T_2010:I(year != 2009 & year != 2010 & year != 2011)TRUE                        14.6438
T_2010:I(year != 2009 & year != 2010 & year != 2011 & year != 2012)TRUE         15.3856
T_2010:I(year != 2009 & year != 2010 & year != 2011 & year != 2012 & year != 2013)TRUE -132.2202
                                                                              Pr(>|t|)
year2002                                                                        0.7535
year2003                                                                      6.39e-07
year2004                                                                     < 2.2e-16
year2005                                                                     < 2.2e-16
year2006                                                                     < 2.2e-16
year2007                                                                     < 2.2e-16
year2008                                                                     < 2.2e-16
year2009                                                                     < 2.2e-16
year2010                                                                     < 2.2e-16
year2011                                                                     < 2.2e-16
year2012                                                                     < 2.2e-16
year2013                                                                     < 2.2e-16
year2014                                                                     < 2.2e-16
year2015                                                                     < 2.2e-16
T_2010:I(year != 2009)TRUE                                                   < 2.2e-16
T_2010:I(year != 2009 & year != 2010)TRUE                                    < 2.2e-16
T_2010:I(year != 2009 & year != 2010 & year != 2011)TRUE                     < 2.2e-16
T_2010:I(year != 2009 & year != 2010 & year != 2011 & year != 2012)TRUE      < 2.2e-16
T_2010:I(year != 2009 & year != 2010 & year != 2011 & year != 2012 & year != 2013)TRUE < 2.2e-16

year2002
year2003                                                                     ***
year2004                                                                     ***
year2005                                                                     ***
year2006                                                                     ***
year2007                                                                     ***
year2008                                                                     ***
year2009                                                                     ***

year2010                                                                     ***
year2011                                                                     ***
year2012                                                                     ***
year2013                                                                     ***
year2014                                                                     ***
year2015                                                                     ***
T_2010:I(year != 2009)TRUE                                                   ***
T_2010:I(year != 2009 & year != 2010)TRUE                                    ***
T_2010:I(year != 2009 & year != 2010 & year != 2011)TRUE                     ***
T_2010:I(year != 2009 & year != 2010 & year != 2011 & year != 2012)TRUE      ***
T_2010:I(year != 2009 & year != 2010 & year != 2011 & year != 2012 & year != 2013)TRUE ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    1.8796e+13
Residual Sum of Squares: 8.5517e+12
R-Squared:      0.54503
Adj. R-Squared: 0.51243
F-statistic: 5295.08 on 19 and 83981 DF, p-value: < 2.22e-16
```

Based on the output, we can see that the treatment effect varies over time. The coefficient estimates for each year indicate that the gross village product increases over time, which suggests a positive trend. The coefficients for the interaction terms between the treatment variable (T_2010) and each year show how the treatment effect varies over time. The coefficient estimate for the interaction term between T_2010 and (year != 2009 & year != 2010 & year != 2011 & year != 2012 & year != 2013) is negative and significantly different from zero ($p<0.05$), indicating that the treatment effect is negative in this year. This suggests that the treatment may have had a detrimental effect on the gross village product in this year. On the other hand, the coefficients for the interaction terms between T_2010 and the other years are positive and significantly different from zero ($p<0.05$), indicating that the treatment effect is positive in these years. This suggests that the treatment had a positive effect on the gross village product in these years. Overall, the treatment effect varies over time and the magnitude and direction of the effect depends on the year.

Question 12

The event study design seems to yield the best results as it allows us to understand the effect of treatment on GVP. We had to eliminate the naive estimator due to selection bias, and the time series approach had problems with counterfactuals and the fundamental problem of causal inference. After finding a good counterfactual through the parallel trends assumption, we ran a fixed effects model, but noticed a year with very different trends in 2010, which we realized was probably just a weighted average of effects. We then ran an event study design, which allowed us to divide our data into a series of events and observe the trend of GVP throughout the entire timeline while also segregating the outlier year of 2010. This approach lines up treatment at the same time for everyone, and we can still use fixed effects to account for confounders. We get a partial test of the identifying assumption, so I would recommend using event study estimates. The estimated effects in graph 11 show a high magnitude, with a low negative of USD 939 to an increase of up to USD 10,299. These results suggest that PROGRAMEVAL should strongly promote female village leadership based on increases in GVP. However, the event study approach may be subject to potential bias, including the selection problem in treatment and untreated villages on different trends, coincident treatments, anticipatory effects, and the Ashenfelter dip. Results can also be impacted by independent and unexpected events.