Department of Statistics

Course STATS 330

Model Answer for Assignment 3, 2003

Note: Unlike assignments 1 and 2, your answer for this assignment was expected to be in the form of a report. A sample report follows.

The Effect on Employment of Changes in the Minimum Wage.

Report by Alan Lee, XYZ Consultants

Executive Summary

On April 1, 1992, the state of New Jersey raised its minimum wage from $4.25 per hour to $5.05. This report examines the effect of this change on employment in the fast food industry. Despite conventional wisdom, no decrease in employment was found. Rather a small but significant increase in employment occurred. This may of course have been due to other economic factors.

Introduction

A survey was conducted in 367 fast food restaurants in New Jersey and Pennsylvania each restaurant being visited before and after the change in the New Jersey minimum wage. The change in employment was noted between the two surveys, and also the change in wages required to bring wages up to the new minimum. (Some restaurants paid above the minimum wage on both surveys.) Also recorded were the name of the franchise and the ownership of the store.

A regression model was fitted to the data to explore the relationship between the wages change and the employment change, and the effect, if any, of the different franchises and ownership types.

Data

The data consisted of 367 records, one per store. Each record contained data on the following variables:

Change in employment = number of employees after – number employees before, variable **`emp.diff`**

Gap = the proportional change in wages needed to bring wages up to the new minimum, (this is zero for Pennsylvania stores and those NJ stores that were paying more than $5.05 before the new minimum wage), variable **`gap`**

Chain = the restaurant chain, one of Burger King, KFC, Roy Rodgers or Wendy's, variable **`chain`**

Chain Owned = indicates if the restaurant is owned by the chain (1) or locally owned (0), Variable **`chain.owned`**

Analysis

A coplot of the data was drawn, with one scatterplot of employment change (**`emp.diff)`** versus wage gap (**`gap`**) for each of the 8 possible franchise/onwership combinations. This plot revealed an outlier, which was identified as point 37 (recorded as -233 in the data file, correct value -23). The point was corrected and the data re-plotted. The corrected plot is shown in Figure 1.

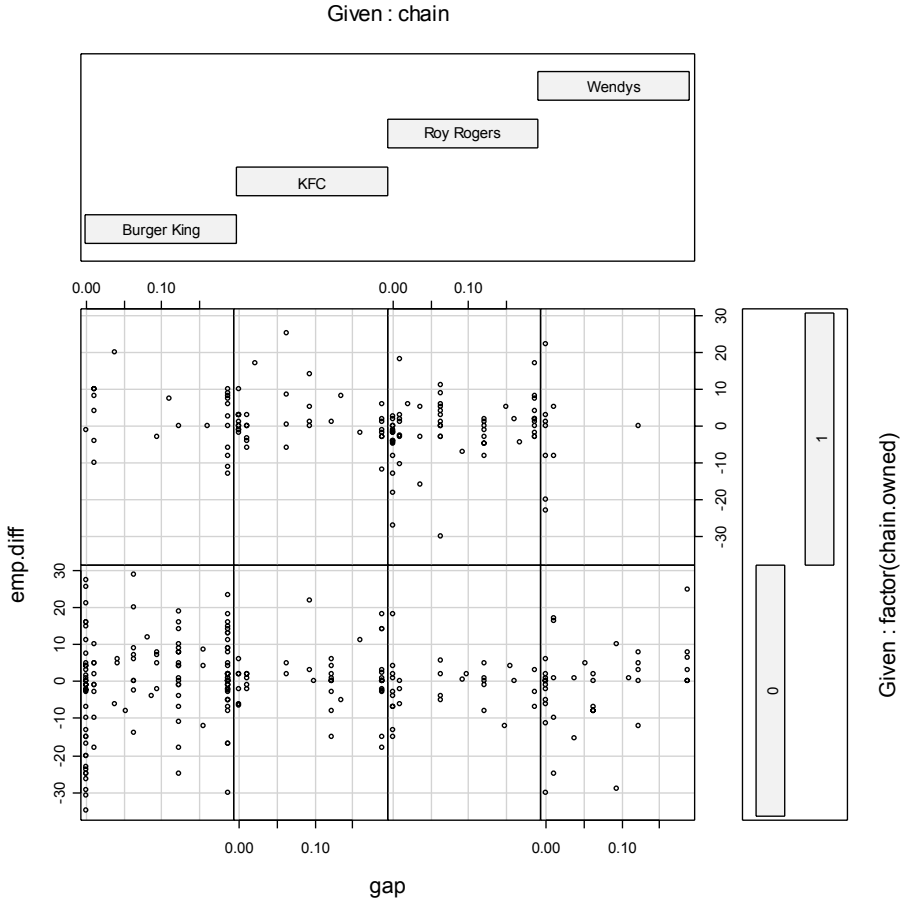

Figure 1: Coplots of the fast food data.

Inspection of the plots revealed considerable scatter, so any relationships present are too weak to be obvious from the plots, although there is a sligh hint of an increasing relationship. There seem to be no obvious differences due to franchises or ownership types. There are no obvious problems with the regression in terms of departures from standard assumptions once the outlier has been corrected.

We then fitted some models to the data, details of which are given in the Appendix. The final model decided on was a simple linear regression, which gives the relationship

`emp.diff` = -1.89 + 19.94 `gap`

Inspection of the data for gap indicates that the amount wages had to be raised to comply with the new minimum wage law ranged from 0% to 18.8% (with corresponding values of gap 0.0 and 0.188). Thus, those restaurants making a maximum increase of 20% actually *raised* employment by approximately $4 \approx 19.94 * 0.188$ persons.

A confidence interval for the slope parameter is 19.94 ± 13.12 i.e. (6.82, 33.06). The slope is significantly different from zero ($p = 0.003$).

A plot of `emp.diff` versus `gap`, with the fitted line drawn on, is shown in Figure 2.

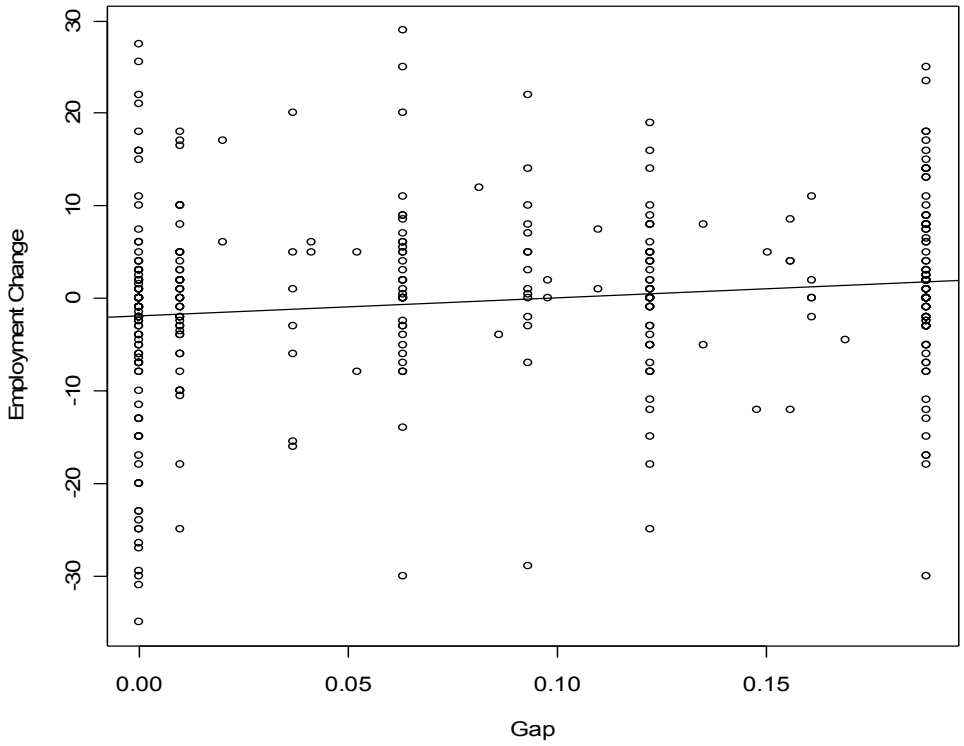

Figure 2. Plot of employment difference versus gap, with least squares line fitted.

Conclusions

There is no evidence of a *negative* relationship between increased wages and employment: in fact the reverse is true: there is a small but significant positive relationship, with employment increasing with increased wages. There is no evidence that this relationship is different for the different franchises and for different ownership types. Of course, we cannot argue from these data that the increase in the minimum wage in New Jersey has *caused* an increase in employment. However, the restaurants having to make the largest increases in wages due to the minimum wage increase did tend to have the biggest increases in employment, for whatever reason. It is pretty clear that the minimum wage increase didn't cause a *drop* in employment.

Technical Appendix

In this appendix we detail the model fitting process. The first model to try is a model that fits 8 separate lines to the 8 franchise/ownership combinations:

```
model1<-lm(emp.diff~gap*chain*factor(chain.owned),data=fastfood.df)
summary(model1)
Call:
lm(formula = emp.diff ~ gap * chain * factor(chain.owned), data =
fastfood.df)

Residuals:
     Min       1Q   Median       3Q      Max
-32.5692  -4.5457   0.1512   5.3203  30.3667

Coefficients:
                                        Estimate Std. Error t value Pr(>|t|)
(Intercept)                              -2.8667     1.3178  -2.175  0.03026
gap                                      28.8782    10.8773   2.655  0.00829
chainKFC                                  3.3191     3.0872   1.075  0.28304
chainRoy Rogers                           0.8849     2.9274   0.302  0.76262
chainWendys                              -1.4378     2.5678  -0.560  0.57587
factor(chain.owned)1                      7.2094     3.8690   1.863  0.06323
gap:chainKFC                            -25.7447    23.6454  -1.089  0.27698
gap:chainRoy Rogers                     -23.9319    28.8368  -0.830  0.40714
gap:chainWendys                          18.1383    25.9078   0.700  0.48431
gap:factor(chain.owned)1                -48.9945    28.1021  -1.743  0.08211
chainKFC:factor(chain.owned)1            -5.0131     5.3023  -0.945  0.34505
chainRoy Rogers:factor(chain.owned)1     -8.5135     5.0125  -1.698  0.09029
chainWendys:factor(chain.owned)1         -6.1042     5.5882  -1.092  0.27541
gap:chainKFC:factor(chain.owned)1        34.5064    41.4645   0.832  0.40585
gap:chainRoy Rogers:factor(chain.owned)1 72.6738    42.5357   1.709  0.08840
gap:chainWendys:factor(chain.owned)1     30.0532    94.2039   0.319  0.74989

Residual standard error: 9.939 on 360 degrees of freedom
Multiple R-Squared: 0.05291,    Adjusted R-squared: 0.01344
F-statistic: 1.341 on 15 and 360 DF,  p-value: 0.1751
```

From this, it seems that one line might be sufficient, since the interaction terms all seem non-significant. A residual analysis (shown in Figure 3) indicates no real problems. The Cooks D for point 255 is not a problem ($F_{13,360}(0.5)= 0.9509804$, $F_{13,360}(0.1)= 1.542679$). The normality is not great but won't affect the significance of the coefficients.

To test if the single line is adequate (this corresponds to testing if franchise and ownership affect the relationship between employment and gap) we fit the straight line

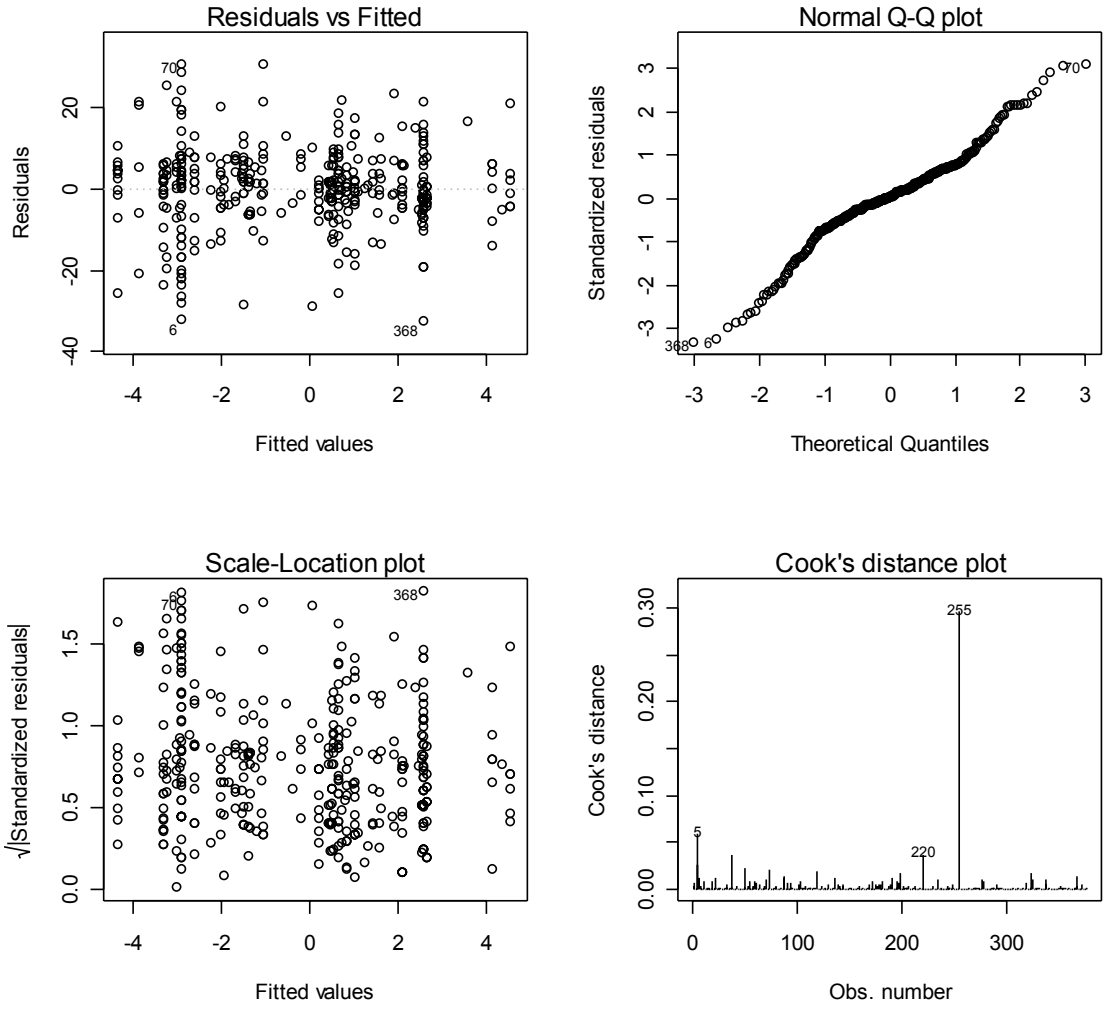

Figure 3: residual analysis for model 1.

model and use the anova function to see if the difference in RSS for the 2 models is significant:

```
> model1<-lm(emp.diff~gap*chain*factor(chain.owned),data=fastfood.df)
> model2<-lm(emp.diff~gap,data=fastfood.df)
> anova(model2,model1)
Analysis of Variance Table

Model 1: emp.diff ~ gap
Model 2: emp.diff ~ gap * chain * factor(chain.owned)
  Res.Df   RSS  Df Sum of Sq      F Pr(>F)
1    374 36674
2    360 35562  14      1112 0.8038 0.6652
```

The large p-value (0.6652) indicates that the simpler model is adequate, and that the relationship between employment difference and gap is the same for all

franchise/ownership combinations. Accordingly, we use model 2 in our analysis. The summary statistics for model 2 are

```
summary(model2)
```

```
> summary(model2)

Call:
lm(formula = emp.diff ~ gap, data = fastfood.df)

Residuals:
     Min       1Q   Median       3Q      Max
-33.1061  -4.1061   0.4569   4.8939  29.6346

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -1.8939     0.7599  -2.492   0.0131 *
gap          19.9384     6.6747   2.987   0.0030 **
---
Signif. codes:  0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1

Residual standard error: 9.902 on 374 degrees of freedom
Multiple R-Squared: 0.0233,     Adjusted R-squared: 0.02069
F-statistic: 8.923 on 1 and 374 DF,  p-value: 0.003001
```

Although the points are scattered widely about the fitted line (as shown by the low R2 (0.0233) and large estimate of error variance ($9.902^2 = 98.4$) the relationship is highly significant and the model fits well, in the sense that the regression assumptions appear to be satisfied.