Section A

1 We can not judge such relationship here by only checking the big or small of the estimates. We could use t–test here with different critical values. The results shows that the author’s statement is not statistically significant.

1、

$H_0: \beta_1 = \beta_2 \quad H_1: \beta_1 > \beta_2$

$$t = \frac{0.8-0.7}{0.07} = \frac{0.1}{0.07} = \frac{10}{7} = 1.42857$$

$1.42857 < 1.96$

$1.42857 < 1.645$

$\therefore$ For both 95%, 97.5% 1-tailed test

It shows we can not reject the null hypothesis

which means $\beta_1 > \beta_2$ is not statistically significant

2 The one–step ahead forecasts are recommended to be used though the multiple–step–ahead forecasts can also be applied but more complicated and less precise. The forecasts of one variable are generated through a linear regression including lagged values of the target variable and the other variables in the model. Additionally the expected error terms are 0 conditional to lagged values of independent variables. The sum of squared residuals here might be used to measure the uncertainty part. It is to see how many deviations of the forecasted values away from the real value.

2 $y_t = k + a_1 y_{t-1} + \cdots + a_n y_{t-n} + b_1 z_{t-1} + \cdots + b_n z_{t-n} + u_t$

$E(u_t \mid y_{t-1}, z_{t-1}, y_{t-2}, z_{t-2}, \ldots) = 0$

3

(a) we need to know the standard deviations of beta 1, beta 1's value and critical values with corresponding confidence intervals.

3 $[\beta_1 - \hat{L} \cdot se_{\beta_1}, \beta_1 + \hat{U} \cdot se_{\beta_1}]$

(b) The coverage probability means the probability that the confidence interval contains the true mean in this sample. It means the fixed true beta 1 is included in the [L,U] interval with 95% probability.

(c) It is common to have different estimates for different instrumental variables since they are different variables in nature even though they could be used as the IV here. Therefore, if they are both good IV for the regression, they must have different error terms. The least squared form containing two IV, its coefficients are not the same either.

Section B

Question 1

The coefficient on an individual's log of annual income (000) 0.15 measure how much contribution of this factor to y holding all the other variables constant. Meanwhile, it means when there's one unit of log of annual income increases, assuming the other variables are 0, the yi here will increase 0.15 and closer to 1 which means this individual is more capable/likely to have the health insurance.

However, there might be some other important variables missing here like education. The education variable is also correlated with the income factor. It violates the assumption of Cov(ui, xi)=0 which makes the estimation inconsistent. It is also possible that the sample taken are not random enough. Maybe people with good body condition likes to stay together and vice-versa.

2

2

$$Pr(y_i=1)=\frac{1}{1+e^{-(\beta_0+\beta_1 X_{1i}+\beta_2 X_{2i}+\beta_3 X_{2i}^2+u_i)}}$$

$$pr(y_i=0)=1-\frac{1}{1+e^{-(\beta_0+\beta_1 X_{1i}+\beta_2 X_{2i}+\beta_3 X_{2i}^2+u_i)}}$$

$$\log\left(\frac{Pr(y_i=1)}{Pr(y_i=0)}\right)=\log\left[\frac{1}{1+e^{-(\cdots)}}\Big/\frac{e^{-(\cdots)}}{1+e^{-(\cdots)}}\right]$$

$$=\log\left[\frac{1}{e^{-(\beta_0+\beta_1 X_{1i}+\beta_2 X_{2i}+\beta_3 X_{2i}^2+u_i)}}\right]$$

3

The coefficient on log income contributes positively to the answer to say yes which contributes 0.48 units.

Section B 3

$H_0: \beta_1 = 0$ $\qquad$ $H_1: \beta_1 \neq 0$

$$\frac{0.48}{0.12} = 4$$

$4 > 1.96$

$4 > 1.65$

It is significantly different from 0 at 2-tailed 0.1 and ~~0.05~~ 0.05 tests.

$$\cancel{Pr(y_i=1)=}\ 0.48 \times 10 + 0.03 \times 40 - 0.008 \times 40^2$$

$$= 1.27$$

$$Pr(y_i=1) = \frac{1}{1+e^{-(1.27)}} = 78.0743\%$$

$$Pr(y_i=0) = 1 - 78.0743\% = 21.9257\%$$

$$\log\left(\frac{Pr(y_i=1)}{Pr(y_i=0)}\right) = \log\left(\frac{78.0743\%}{21.9257\%}\right) = 0.551555$$

4

$$G(x) = \frac{e^x}{1+e^x}$$

$$pr(y_i=1) = \frac{e^x}{1+e^x}$$

$$pr(y_i=0) = \frac{1}{1+e^x}$$

$$\log\left[\frac{pr(y_i=1)}{pr(y_i=0)}\right] = \cancel{e}\log(e^x)$$

$$= x$$

Section C

1 (a) The fixed effects are regressions on time-demanded variables. It requires strict exogeneity. This method is using the differences between Yit and average of y, Xit and the average of x and also the error term and its average. It eliminates the unobserved heterogeneity which makes the estimation consistent. However it can not estimate the time invariant effects.

FE

$$y_{it} - \bar{y}_i = (X_{it} - \bar{X}_i)'\beta + (\varepsilon_{it} - \bar{\varepsilon}_i)$$

$cov(X_{it}, u_{it}) = 0$ contemporaneous exogeneity

Strict exogeneity : $cov(\tilde{\varepsilon}_{it}, \tilde{X}_{it}) = 0$

$$E(\varepsilon_{it} \mid X_{i1}, \ldots X_{iT}) = 0$$

The random effects assumes beta RE is consistent with unobserved heterogeneity kept. It also requires strict exogeneity, no serial correlation constant variance over time. The covariance of unobserved heterogeneity and error term is 0. It needs to be tested whether the initial beta RE is consistent. It can estimate the time invariant effects.

RE

~~Assume $cov(\varepsilon_{it}, x_{it}) = 0$~~

Assume $cov(\varepsilon_{it}, X_{it}) = 0$:

$$\left.\begin{array}{l} cov(\varepsilon_{it}, X_{it}) = 0 \\ E(\varepsilon_{it} \mid X_{i1} \cdots X_{iT}) = 0 \end{array}\right\} \text{strict exogeneity}$$

$var(u_{it}) = \sigma_u^2$ for all t

$cov(u_i, \varepsilon_{it}) = 0$

The First differencing method is using the difference between two periods fir the same group of people by eliminating heterogeneity. Like FE, they. Both can not estimate the time invariant effects. If T=2, beta FD = beta FE, T>2, beta FE is better.

FD:

As long as $cov(X_{it} - X_{i,t-1}, \varepsilon_{it} - \varepsilon_{i,t-1}) = 0$

$\hat{\beta}_{FD}$ is consistent.

(b) Beta 0 is the constant term. Arrest equals to beta 0 when all the other variables are 0 eg. A person not in 1990s and not living in the Florida. The coefficient of y90 is the effect of whether the population is from 1990 group to their ‘arrest’. The beta 1 measures the effect of whether people live in Florida or not to their ‘arrest’. The coefficient on the y90*FL is the effect of people living in the Florida and meanwhile belongs to 1990s group.
The coefficient on the y90*FL measures the effect of law. It is only meaningful to see the effect of the law for someone living in the Florida before the law was passed and still living in the Florida after the law is passed.
The other factors like driving experience/ driving years. For example, an experienced driver may have clear mind about the consequences of drunk driving. The longer the driving years, the less likely he/she would be arrested. Such factor is controlled here due to this LPM focusing on the effect of law passed on drunk driving to ‘arrest’.

(c)
Now there are only two years data for two states. The dependent variable now is fraction less than 1. LPM may create values larger than one which is not suitable to be used.
Probit/Logit model only suffers from negative datas which here is not possible. Under arrest binary variable condition, logit model can estimate the probability of laws passed in 1990s and in 1985s to see the changes.
Maybe using FD model/FE model. Since now there are only two periods data for the same group of people/states. The unobserved heterogeneity can be eliminated. We can use such method to see how well the ‘arrest’ being affected by the laws passed in 1990 compared with 1985s. It will give us a consistent beta FE/FD. It is very possible that the drunk driving is negatively affected by the law.