15.6 **f.** Column (5) contains the random effects estimates. Which coefficients, apart from the intercepts, show the most difference from the fixed effects estimates? Use the Hausman test statistic (15.36) to test whether there are significant differences between the random effects estimates and the fixed effects estimates in column (3) (Why that one?). Based on the test results, is random effects estimation in this model appropriate?

$$t_{Exper} = \frac{0.0595 - 0.0786}{\sqrt{0.033^2 - 0.022^2}} = -1.67$$

$$t_{Exper^2} = \frac{-0.0012 + 0.0023}{\sqrt{0.0011^2 - 0.0007^2}} = 1.3$$

$$t_{south} = \frac{-0.3261 + 0.2326}{\sqrt{0.1258^2 - 0.0317^2}} = -0.77$$

$$t_{union} = \frac{0.0822 - 0.027}{\sqrt{0.0312^2 - 0.0245^2}} = -1.06$$

$\Rightarrow$ Exper t-ratio 最大，但其結果在統計上不顯著，FE, RE 無明顯差異

15.17

b. Estimate the model $LIQUOR_{it} = \beta_1 + \beta_2 INCOME_{it} + u_i + e_{it}$ using random effects. Construct a 95% interval estimate of the coefficient on *INCOME*. How does it compare to the interval in part (a)?

c. Test for the presence of random effects using the LM statistic in equation (15.35). Use the 5% level of significance.

d. For each individual, compute the time averages for the variable *INCOME*. Call this variable *INCOMEM*. Estimate the model $LIQUOR_{it} = \beta_1 + \beta_2 INCOME_{it} + \gamma INCOMEM_i + c_i + e_{it}$ using the random effects estimator. Test the significance of the coefficient γ at the 5% level. Based on this test, what can we conclude about the correlation between the random effect u_i and *INCOME*? Is it OK to use the random effects estimator for the model in (b)?

b.

```
> confint(re_model)
                  2.5 %     97.5 %
(Intercept) -0.05211904 1.99018381
income       0.01283111 0.04031983
```

The estimate for the random effects (RE) coefficient is slightly smaller than the difference estimator's coefficient. However, the standard error of the random effects estimator is approximately 25% of that of the difference estimator. This results in the difference between the two estimates being statistically significant.

c.

```
	Lagrange Multiplier Test - (Breusch-Pagan)

data:  liquor ~ income
chisq = 20.68, df = 1, p-value = 5.429e-06
alternative hypothesis: significant effects
```

We reject the null hypothesis that the variance of the error term is zero and accept the alternative that it is greater than zero, indicating the presence of statistically significant unobserved heterogeneity.

d.

```
Call:
plm(formula = liquor ~ income + income_mean, data = pdata, model = "random")

Balanced Panel: n = 40, T = 3, N = 120

Effects:
                 var std.dev share
idiosyncratic 0.9640  0.9819 0.571
individual    0.7251  0.8515 0.429
theta: 0.4459

Residuals:
     Min.   1st Qu.    Median   3rd Qu.      Max.
-2.300955 -0.703840  0.054992  0.560255  2.257325

Coefficients:
             Estimate Std. Error z-value Pr(>|z|)
(Intercept) 0.9163337  0.5524439  1.6587  0.09718 .
income      0.0207421  0.0209083  0.9921  0.32117
income_mean 0.0065792  0.0222048  0.2963  0.76700
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    126.61
Residual Sum of Squares: 112.79
R-Squared:      0.10917
Adj. R-Squared: 0.093945
Chisq: 14.3386 on 2 DF, p-value: 0.00076987
```

The t-value for income_mean is 0.3, indicating it is not statistically significant. Therefore, there is no evidence to suggest a correlation between income and the unobserved heterogeneity.

15.20

d. Reestimate the model in part (a) with school random effects. Compare the results with those from parts (a) and (b). Are there any variables in the equation that might be correlated with the school effects? Use the LM test for the presence of random effects.

e. Using the t-test statistic in equation (15.36) and a 5% significance level, test whether there are any significant differences between the fixed effects and random effects estimates of the coefficients on *SMALL*, *AIDE*, *TCHEXPER*, *WHITE_ASIAN*, and *FREELUNCH*. What are the implications of the test outcomes? What happens if we apply the test to the fixed and random effects estimates of the coefficient on *BOY*?

f. Create school-averages of the variables and carry out the Mundlak test for correlation between them and the unobserved heterogeneity.

d.

```
Effects:
                  var std.dev share
idiosyncratic 751.43   27.41 0.829
individual    155.31   12.46 0.171
theta:
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.6470  0.7225  0.7523  0.7541  0.7831  0.8153

Residuals:
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
-97.4830 -17.2361  -3.2822   0.0372  12.8027 192.3460

Coefficients:
              Estimate Std. Error  z-value  Pr(>|z|)
(Intercept) 436.126774   2.064782 211.2217 < 2.2e-16
small         6.458722   0.912548   7.0777 1.466e-12
aide          0.992146   0.881159   1.1260    0.2602
tchexper      0.302679   0.070292   4.3060 1.662e-05
boy          -5.512081   0.727639  -7.5753 3.583e-14
white_asian   7.350477   1.431376   5.1353 2.818e-07
freelunch   -14.584332   0.874676 -16.6740 < 2.2e-16
```

```
(d) LM 檢定結果（檢查是否需要使用隨機效果）:
> print(lm_test)

        Lagrange Multiplier Test - (Breusch-Pagan)

data:  readscore ~ small + aide + tchexper + boy + white_asian + freelunch
chisq = 6677.4, df = 1, p-value < 2.2e-16
alternative hypothesis: significant effects
```

Comparing OLS and fixed effects, the RE estimator is somewhat similar. Under the LM test, we reject the null hypothesis of no endogeneity, so the random effects model should be used.

e.

```
(e) Hausman 檢定結果（固定 vs 隨機效果）:
> print(hausman_test)

        Hausman Test

data:  readscore ~ small + aide + tchexper + boy + white_asian + freelunch
chisq = 13.809, df = 6, p-value = 0.03184
alternative hypothesis: one model is inconsistent
```

reject H_0, 使用FE較適合

f.

```
Linear hypothesis test:
small_avg = 0
aide_avg = 0
tchexper_avg = 0
boy_avg = 0
white_asian_avg = 0
freelunch_avg = 0

Model 1: restricted model
Model 2: readscore ~ small + aide + tchexper + boy + white_asian + freelunch +
    small_avg + aide_avg + tchexper_avg + boy_avg + white_asian_avg +
    freelunch_avg

Note: Coefficient covariance matrix supplied.

  Res.Df Df      F  Pr(>F)
1   5695
2   5689  6 2.2541 0.03557 *
---
Signif. codes:
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

∵ p-value < 0.05, reject H_0

因此不適用隨机效果model