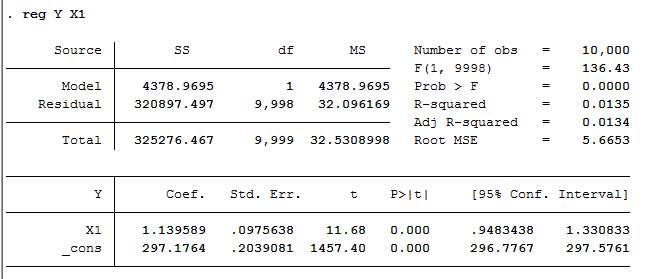
**HW 2 Ordinary Least Square**

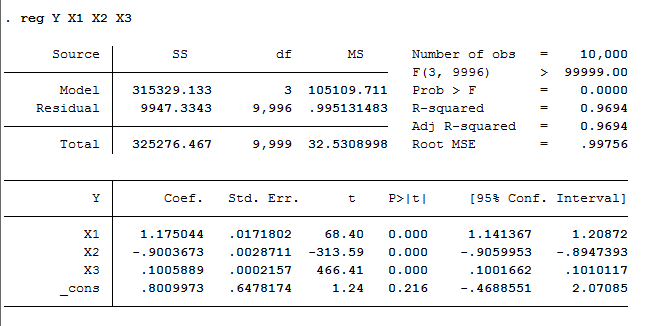
Exercise 2 OLS

* Calculate the correlation between Y and X1



This result is very close to the coefficient (1.2) in the true model.

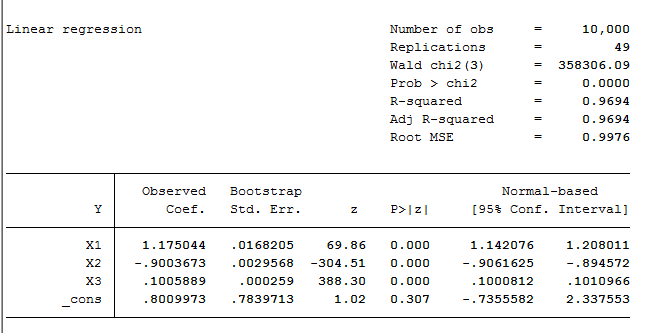
* Calculate the OLS coefficients & standard errors on this regression



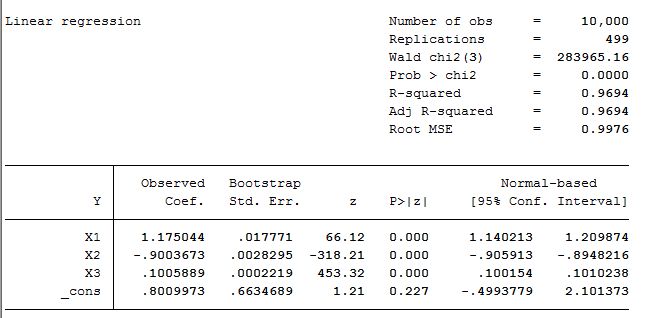
* Calculate the standard errors

(2). Using bootstrap with 49 and 499 replications respectively:

With 49 replications:



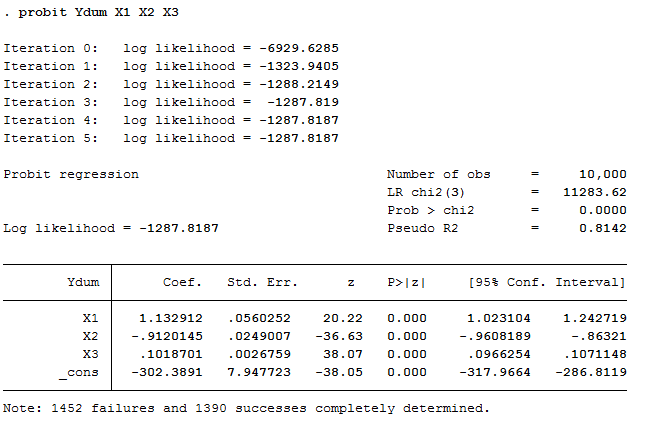
With 499 replications:



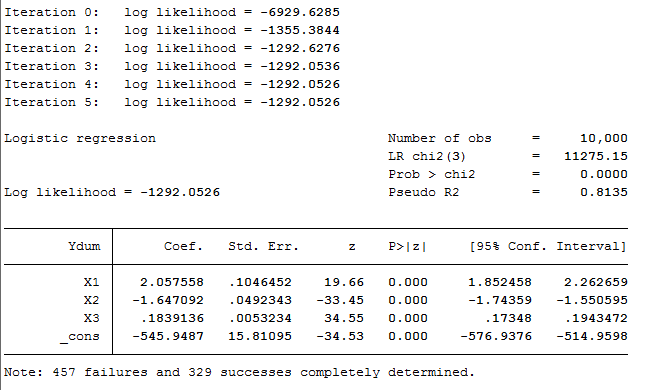
Exercise 4 Discrete Choice

* Interpret and compare the estimated coefficients, how significant are they?

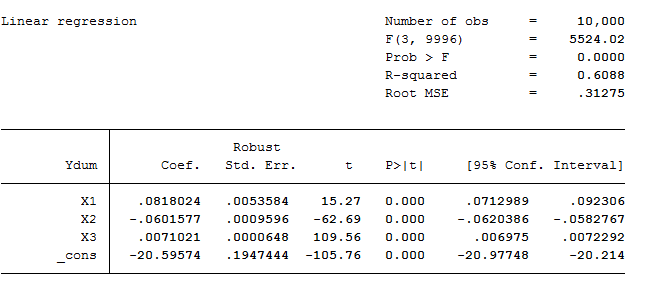
Probit Estimation:



Logit Estimation:



Linear estimation:



According to the test result, coefficient estimations of probit model is closer to the true model than the other two models. The results of significant test are uniform across three models, with all coefficients being strongly significant.

Exercise 5 Marginal Effects

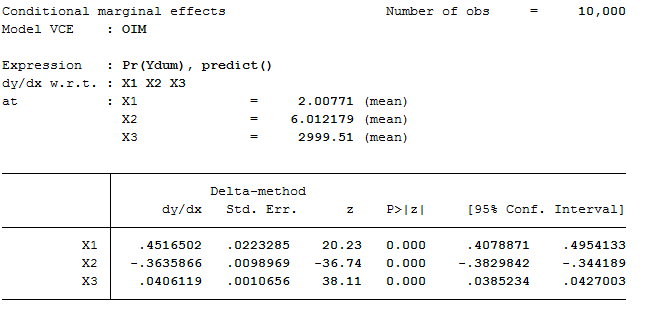
* Compute the marginal effect of X on Y according to the probit and logit models

Please check the excel output “probit marginal effect” and “logit marginal effect”

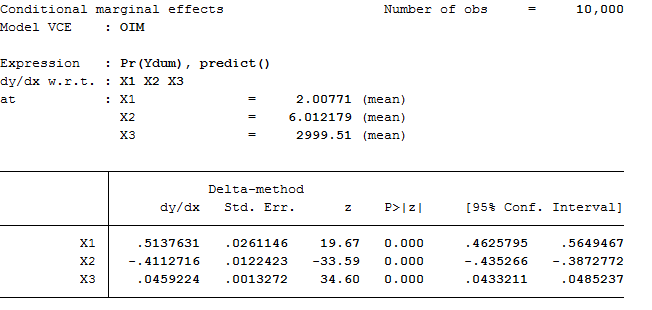
* Compute the standard deviations using

1. Delta method

(a). probit variance-covariance

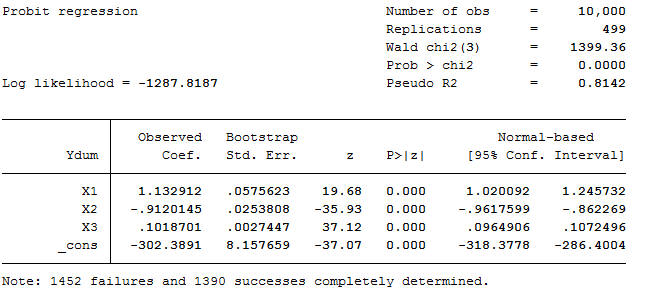


(b). logit variance-covariance

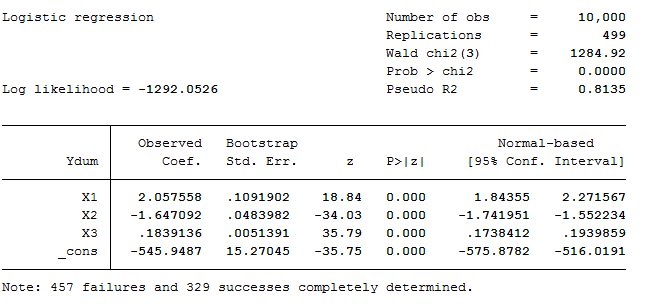


1. Bootstrap

(a). probit



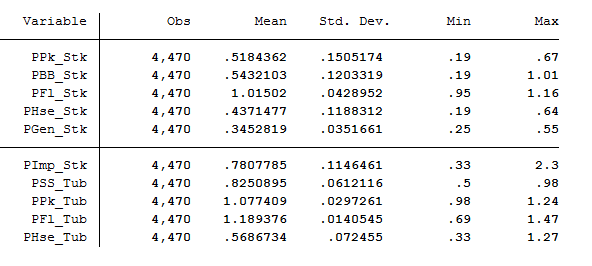
(b). logit



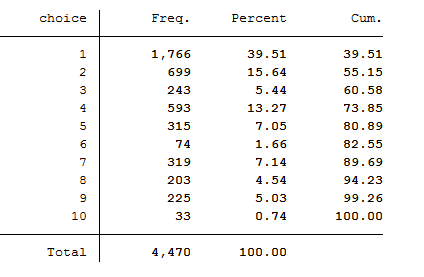
**HW 3 Multinomial Choices**

Exercise 1 Data Description

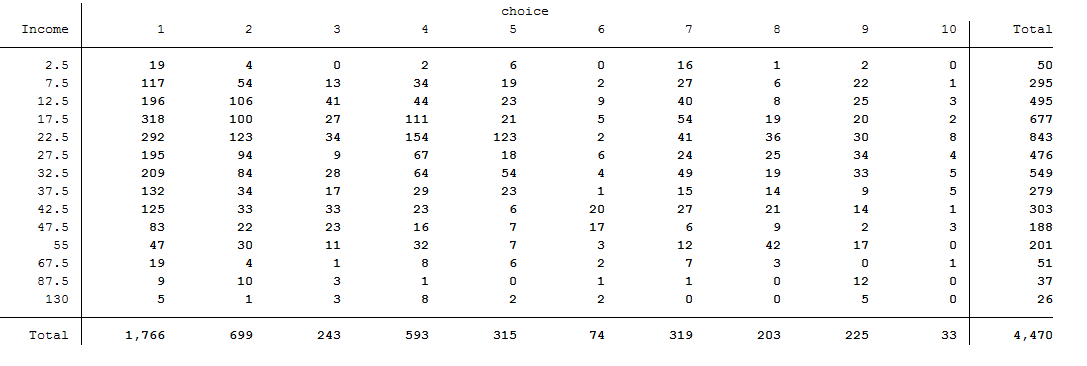
* Average and dispersion in product characteristic



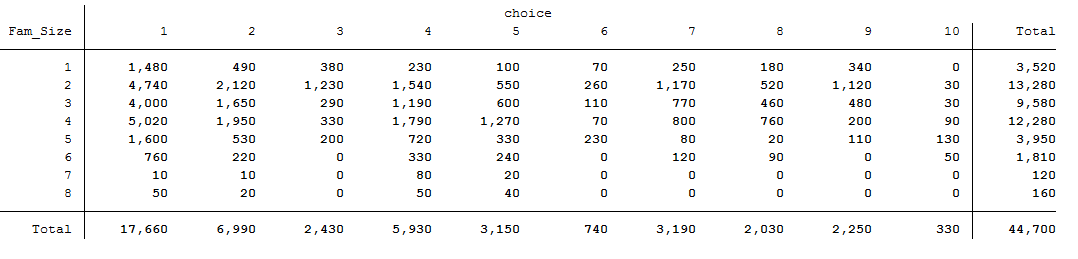
According to the table above, PFl\_Tub has the highest average price among ten products. In general, tub appears to be more expensive than stick, and the Fleischmann’s cheese produce the products with the highest average price in both tub and stick category. Besides, the price of ten products are relative stable, with small variance, standard deviation and price range.



* Mapping between observed attributes and choices
  + Frequency of buying each choice for each income level



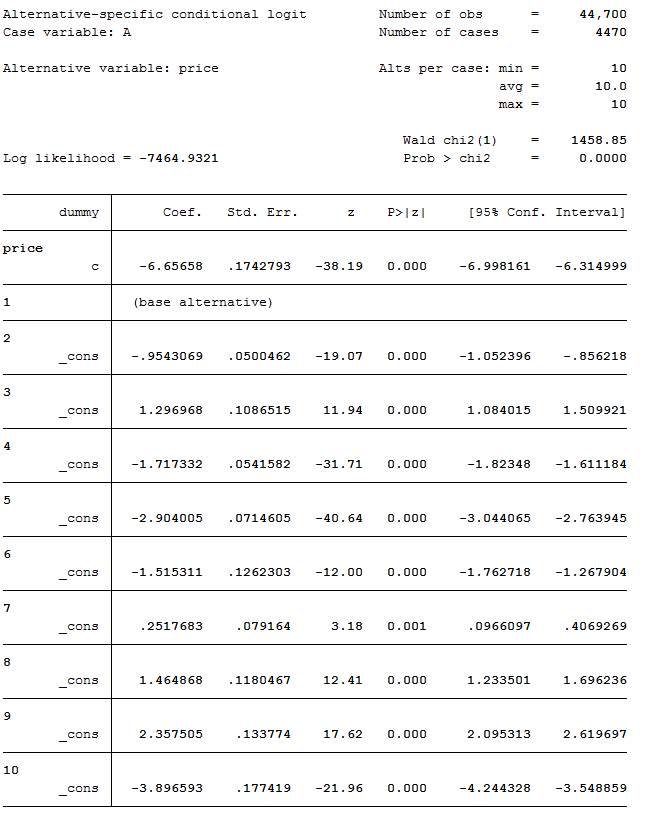
* + Frequency of buying each choice for each family size



Exercise 2 First Model

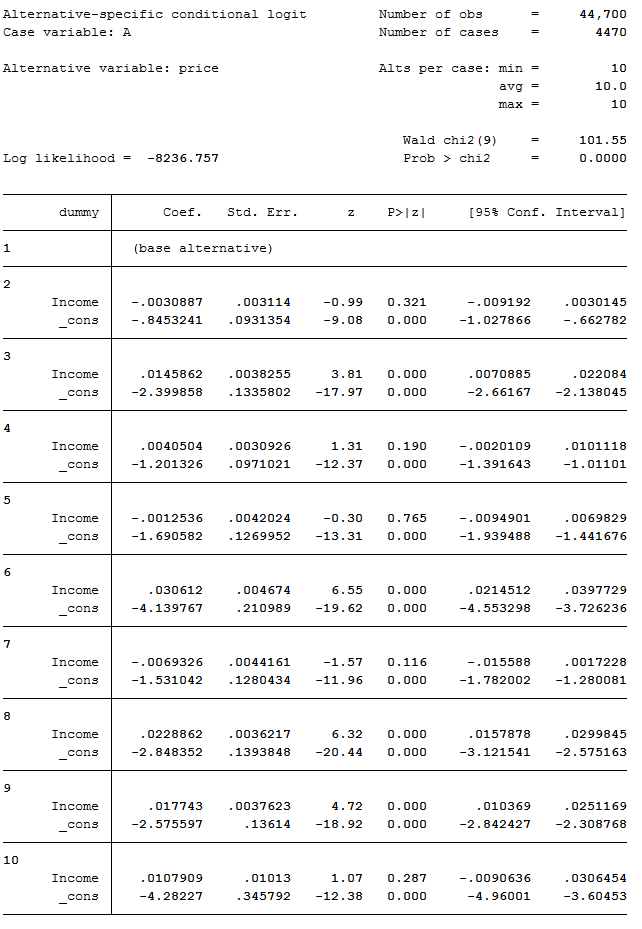
* Apply conditional logit model to capture the choice’s characteristics---the effect of price on demand (since price varies over alternatives)
* Interpret the coefficient on price
  + optimized coefficients on price: ()

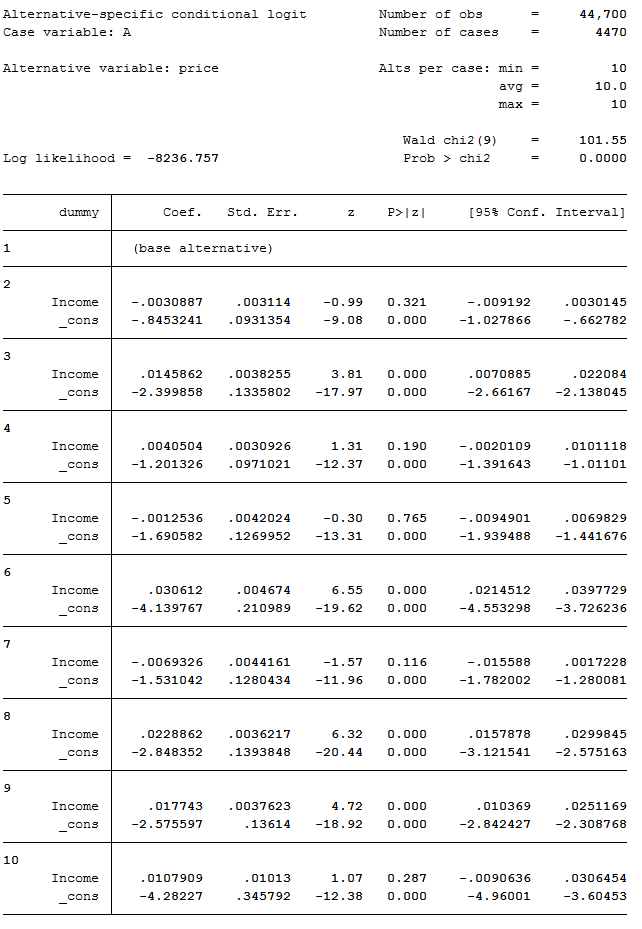
I think the negative sign of the estimated coefficient indicates that holding other product characteristics constant, consumers will be less likely to buy cheese products for an increase in price.



Exercise 3 Second Model

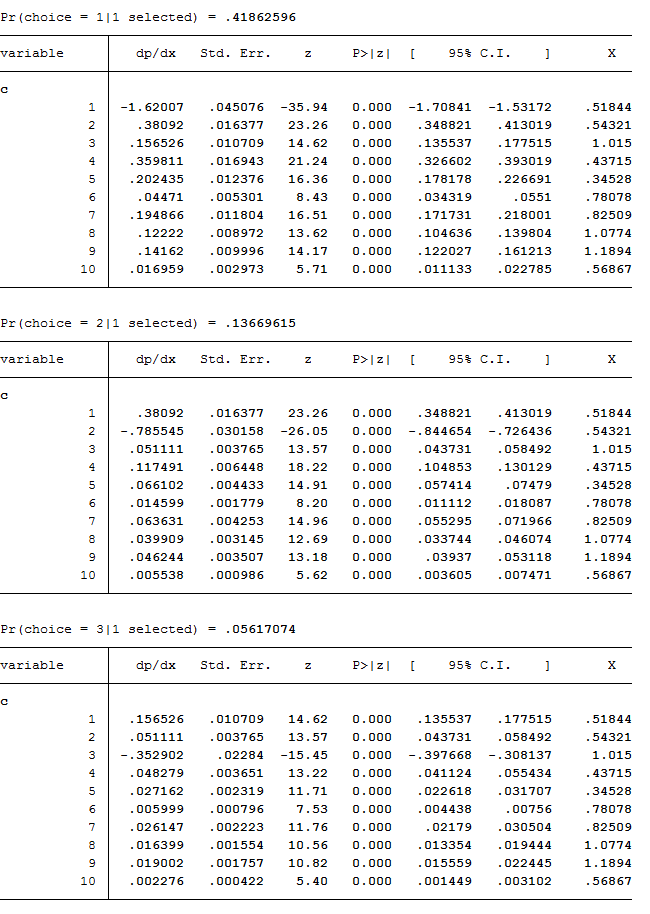
* Apply multinomial logit model to capture the chooser’s characteristics---the effect of family income on demand (since family income do not vary over alternatives)
* Interpret the coefficient on family income
  + optimized coefficients: ( )
  + In particular, the positive sign of 𝑏3 indicates that people are slightly more willing to buy product 3 when their income increase. And for the negative sign of 𝑏5, I think it means that consumers will be less likely to buy this product for an increase in their family income.

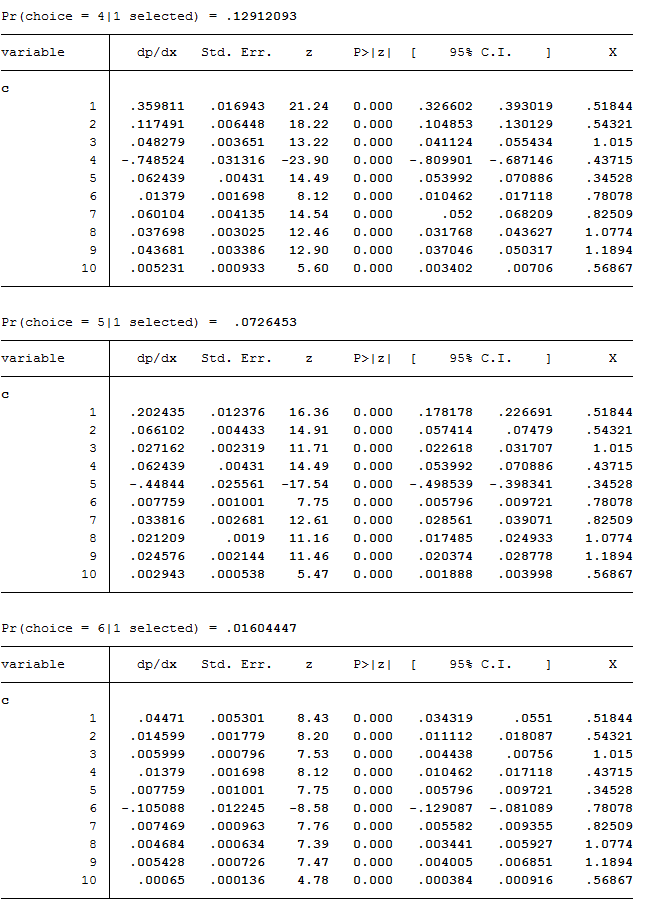


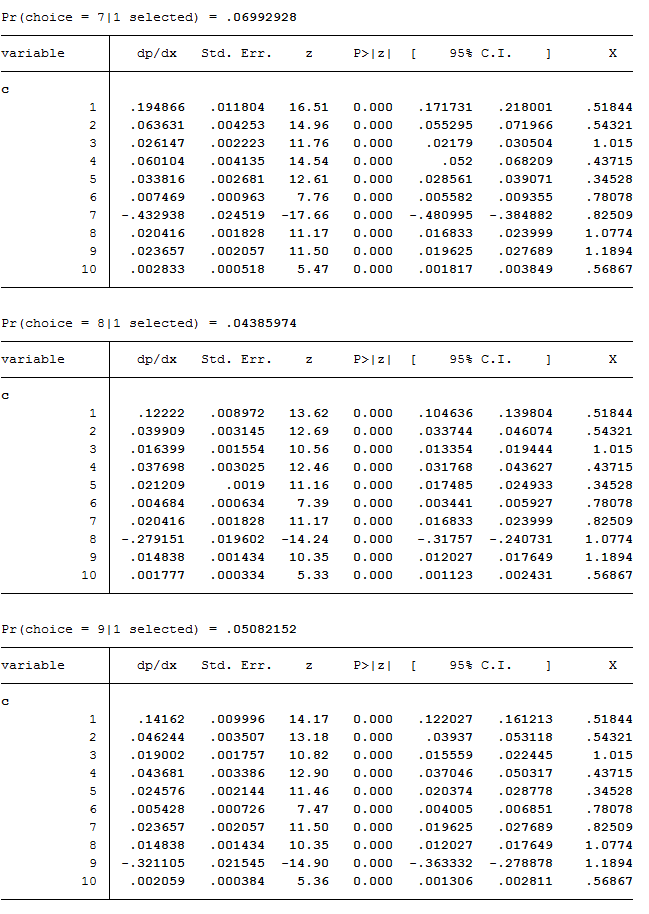


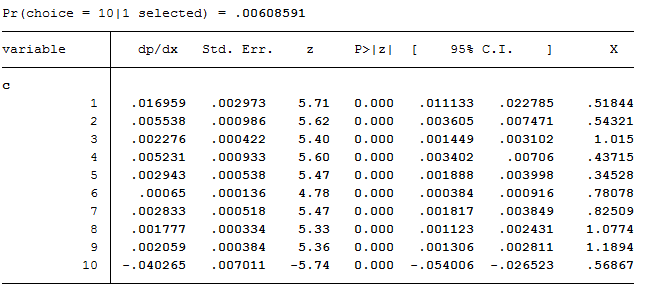
Exercise 4 Marginal Effects

* Compute and interpret the marginal effects for the first (conditional logit) model
  + As we can perceive from the table, marginal effects on the diagonal is negative, it’s consistent with the real life intuition that when the price of the chosen product 1 increase, it's less likely for the person to still choose that product 1; however, when the price of other products increase, it will only increase the probability of choosing product 1 a little bit.

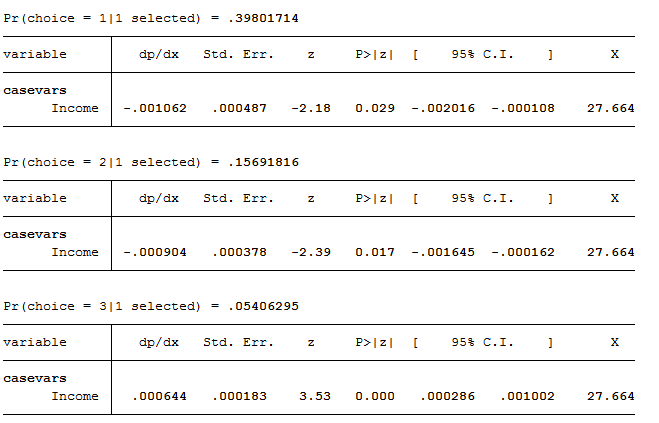


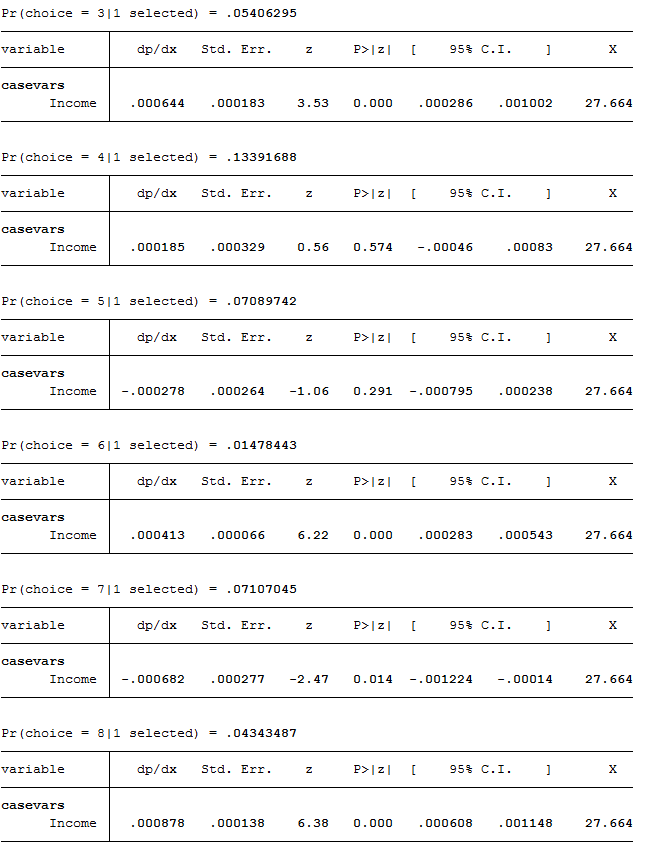


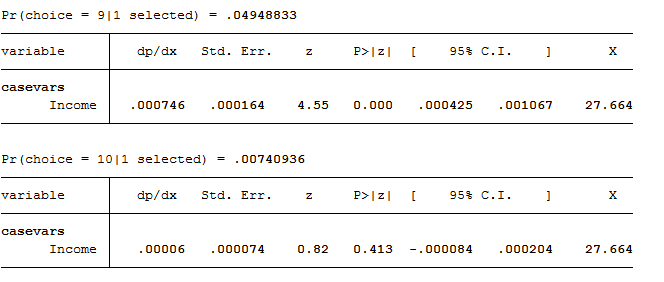




* Compute and interpret the marginal effects for the second (multinomial logit) model
* According to the results, an increase in family income will have slight effects on the product demand, and I noticed that for products with relatively lower average prices, such as product 4, 5, 6, 7, and 10, they tend to encounter a negative marginal effect from the increase in family income. I think it may because that when people are having higher income, they are more eager to live a higher quality life, so that they may want to buy more expensive products to express their willingness.

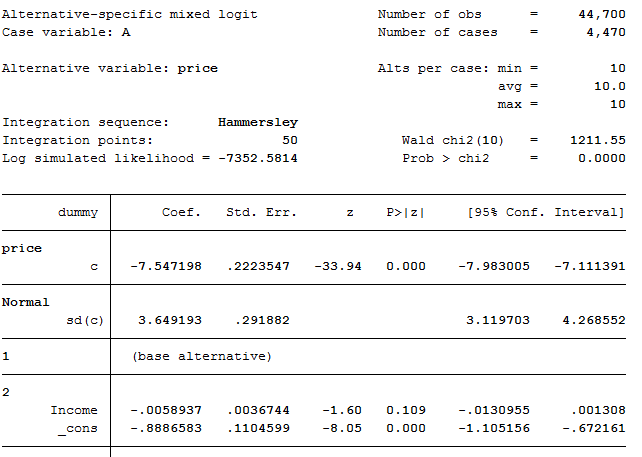


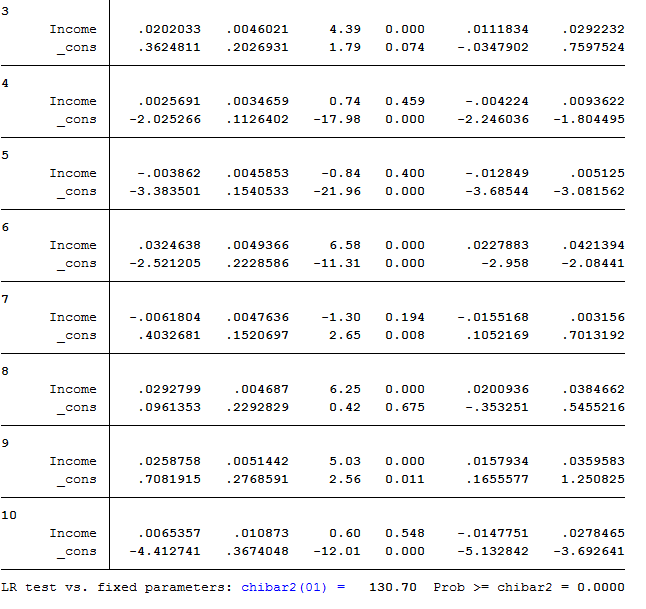




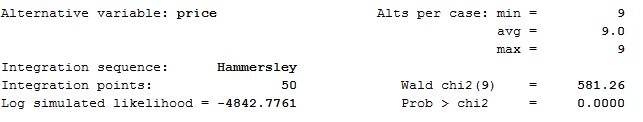
Exercise 5 IIA (mixed logit)

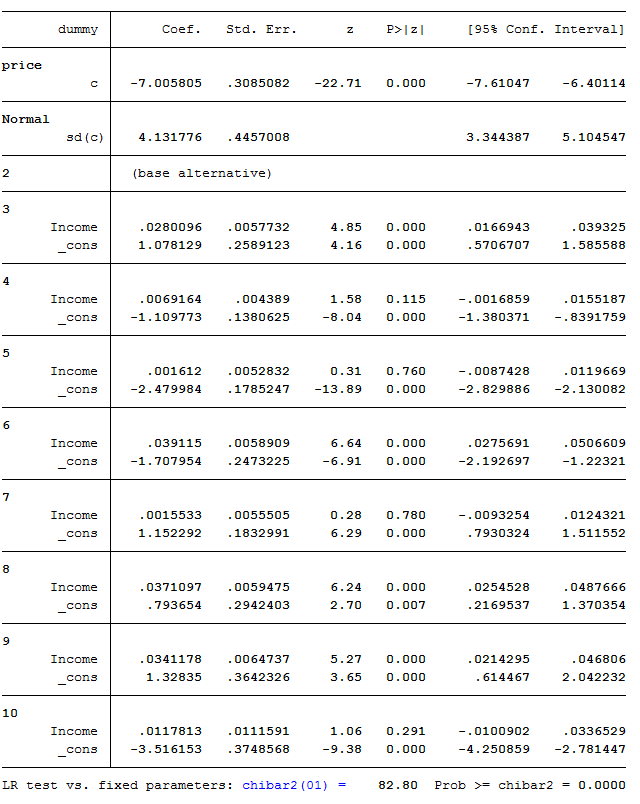
* Apply mixed logit model to capture the effect of price and family income on demand
* : (; )(#30)



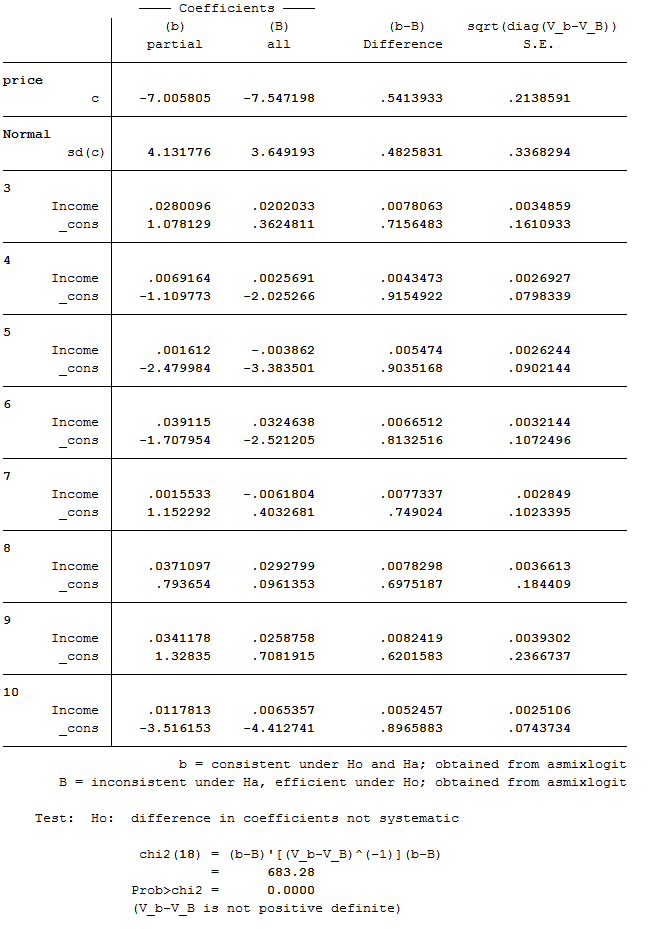


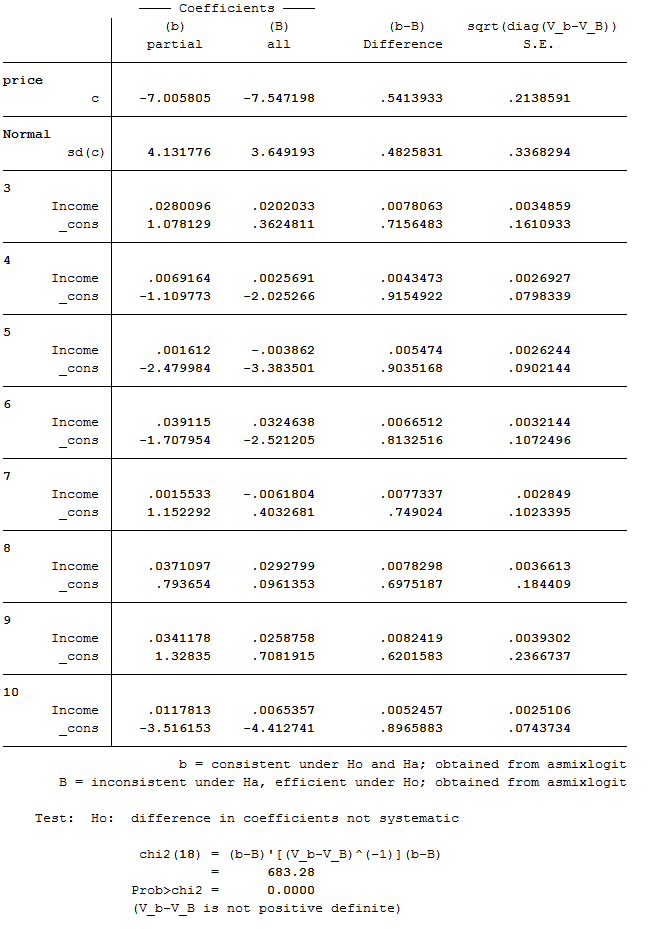
* Consider an alternative specification by removing data from one choice
* : (; )(#27)





* Conclude on IIA



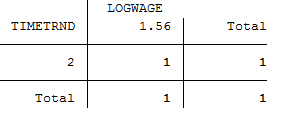


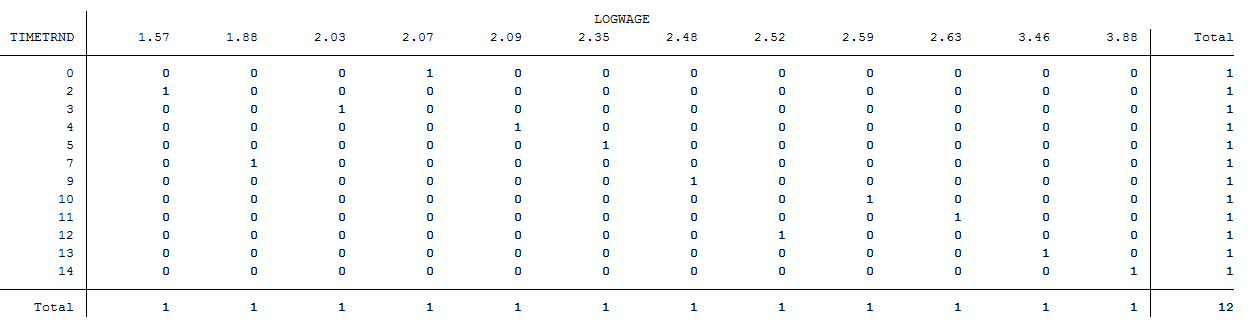
We failed to reject the null hypothesis by removing choice 1 (which takes a large proportion in the data), and able to conclude that the property of independence of irrelevant alternatives does not hold in the model.

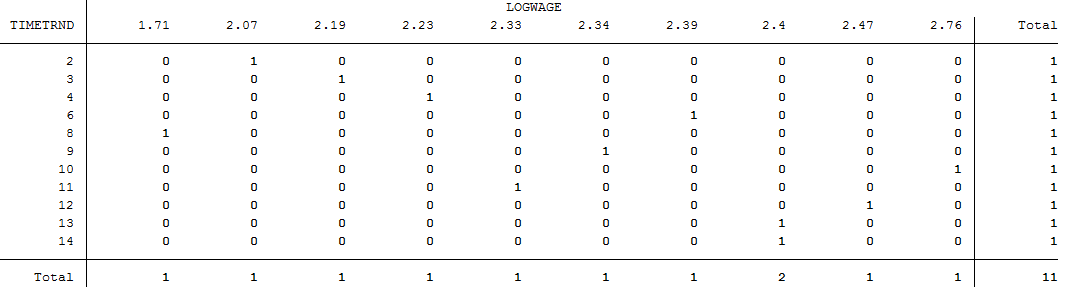
**HW 4 Linear Panel Data**

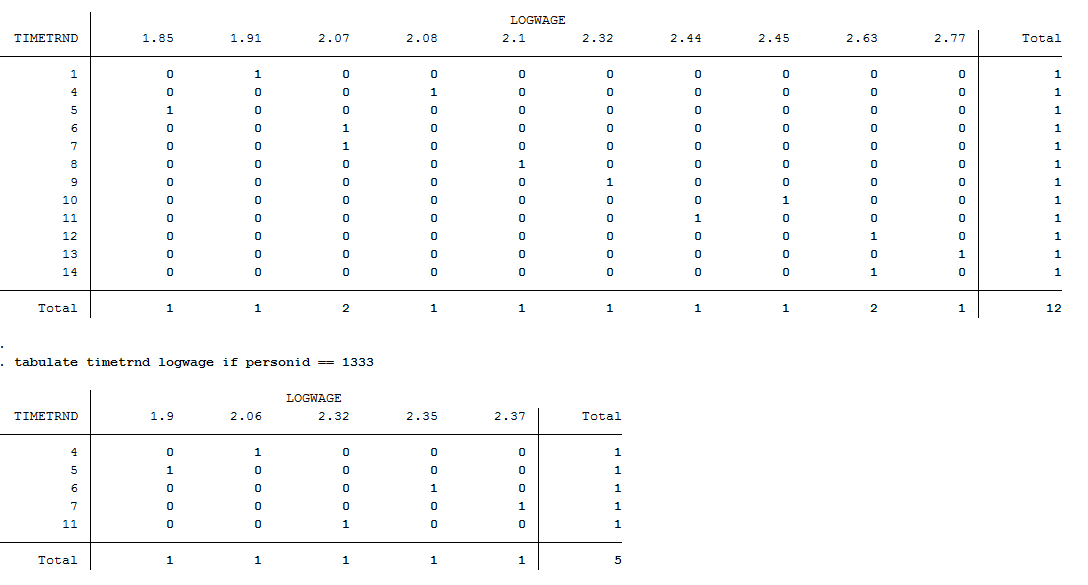
Exercise 1 Data

* Represent the panel dimension of wages for 5 randomly selected individuals



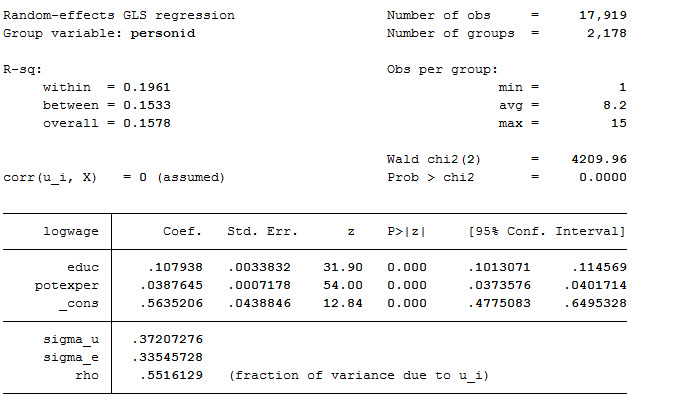






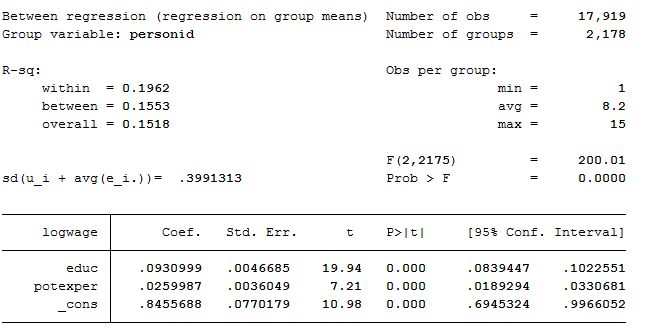
Exercise 2 Random Effects

* Estimate the **random effect model** using GLS

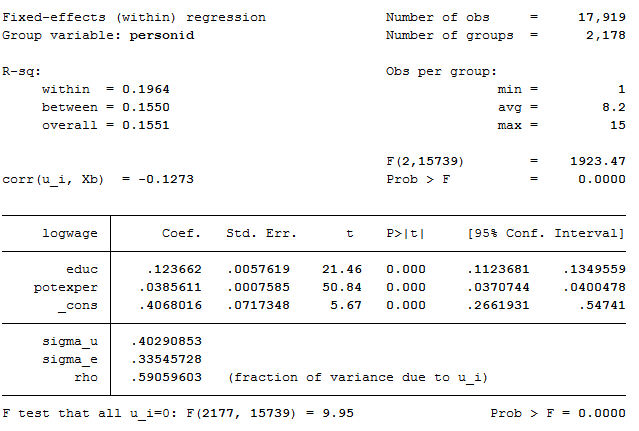


Exercise 3 Fixed Effects Model

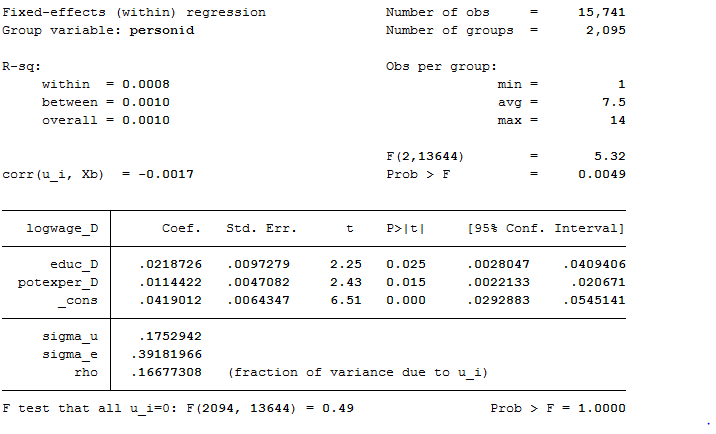
* Estimate **between estimator**



* Estimate **within estimator**



* Estimate **first time difference estimator**



Compare the estimates of and under the different models: among estimates, the estimate coefficient of potential experience is relatively consistent, but the estimate coefficient of education varies a lot, especially for the first time difference estimator. I think it's probably because the time-invariant property of the education variable.