11/9/2016

Bhupesh Joshi

Amitesh Saha

Kiran Soibam

Roopam Chabara

Assignment-2

Analysis of Male Labour Market

***Group-05***

***State Excluded- 07 (Delhi)***

We are analysing the male labour market participation. The responses as based on survey are not in labour force, students only, non-wage labour, waged labour or a combination of waged or non-waged labour. We discard females, children’s, students and males not in labour force from our analysis. Further the missing values for of wage and workhour for non-wage worker are set as zero. The summary statistics of all the required variable are as follows:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| VARIABLES | Explanation | N | mean | sd | min | max |
|  |  |  |  |  |  |  |
| exp | Experience | 46,978 | 24.15 | 13.54 | 0 | 55 |
| exp\_sq | Square of experience | 46,978 | 766.5 | 731.3 | 0 | 3,025 |
| Dummy for religion | | | | | | |
| dhindu | Dummy for Hindu | 46,978 | 0.808 | 0.394 | 0 | 1 |
| dmuslim | Dummy for Muslim | 46,978 | 0.127 | 0.333 | 0 | 1 |
| dchrstn | Dummy for Christian | 46,978 | 0.0246 | 0.155 | 0 | 1 |
| dsikh | Dummy for Sikh | 46,978 | 0.0255 | 0.158 | 0 | 1 |
| djain | Dummy for Jain | 46,978 | 0.00262 | 0.0511 | 0 | 1 |
| dorelg | Dummy for other | 46,978 | 0.0121 | 0.109 | 0 | 1 |
| Dummy for Caste | | | | | | |
| duch | Dummy for upper caste | 46,978 | 0.0454 | 0.208 | 0 | 1 |
| dobc | Dummy for OBC | 46,978 | 0.406 | 0.491 | 0 | 1 |
| dsc | Dummy for SC | 46,978 | 0.215 | 0.411 | 0 | 1 |
| dst | Dummy for ST | 46,978 | 0.0900 | 0.286 | 0 | 1 |
| dosgr | Dummy for Other | 46,978 | 0.244 | 0.429 | 0 | 1 |
| Dummy for region | | | | | | |
| region1 | Dummy for Rural | 46,978 | 0.665 | 0.472 | 0 | 1 |
| region2 | Dummy for Non metro | 46,978 | 0.281 | 0.449 | 0 | 1 |
| region3 | Dummy for Metro | 46,978 | 0.0538 | 0.226 | 0 | 1 |
| Dummy for education | | | | | | |
| educd1 | Dummy for Illiterate | 46,978 | 0.183 | 0.387 | 0 | 1 |
| educd2 | Dummy for primary | 46,978 | 0.0851 | 0.279 | 0 | 1 |
| educd3 | Dummy for Middle | 46,978 | 0.274 | 0.446 | 0 | 1 |
| educd4 | Dummy for Secondary | 46,978 | 0.234 | 0.423 | 0 | 1 |
| educd5 | Dummy for HiSec | 46,978 | 0.111 | 0.314 | 0 | 1 |
| educd6 | Dummy for PostHiSec. | 46,978 | 0.114 | 0.318 | 0 | 1 |
| Dummy for States | | | | | | |
| STATEID1 | J&K | 46,978 | 0.0208 | 0.143 | 0 | 1 |
| STATEID2 | Himanchal | 46,978 | 0.0311 | 0.174 | 0 | 1 |
| STATEID3 | Punjab | 46,978 | 0.0460 | 0.209 | 0 | 1 |
| STATEID4 | Chandigarh | 46,978 | 0.00192 | 0.0437 | 0 | 1 |
| STATEID5 | Uttarakhand | 46,978 | 0.0110 | 0.104 | 0 | 1 |
| STATEID6 | Haryana | 46,978 | 0.0483 | 0.214 | 0 | 1 |
| STATEID7 | Rajhasthan | 46,978 | 0.0684 | 0.252 | 0 | 1 |
| STATEID8 | UP | 46,978 | 0.0925 | 0.290 | 0 | 1 |
| STATEID9 | Bihar | 46,978 | 0.0328 | 0.178 | 0 | 1 |
| STATEID10 | Sikkim | 46,978 | 0.00243 | 0.0492 | 0 | 1 |
| STATEID11 | Arunachal Pradesh | 46,978 | 0.00289 | 0.0537 | 0 | 1 |
| STATEID12 | Nagaland | 46,978 | 0.00200 | 0.0447 | 0 | 1 |
| STATEID13 | Manipur | 46,978 | 0.00181 | 0.0425 | 0 | 1 |
| STATEID14 | Mizoram | 46,978 | 0.00183 | 0.0427 | 0 | 1 |
| STATEID15 | Tripura | 46,978 | 0.00492 | 0.0700 | 0 | 1 |
| STATEID16 | Meghalaya | 46,978 | 0.00343 | 0.0584 | 0 | 1 |
| STATEID17 | Assam | 46,978 | 0.0243 | 0.154 | 0 | 1 |
| STATEID18 | West Bengal | 46,978 | 0.0570 | 0.232 | 0 | 1 |
| STATEID19 | Jharkhand | 46,978 | 0.0196 | 0.139 | 0 | 1 |
| STATEID20 | Orissa | 46,978 | 0.0530 | 0.224 | 0 | 1 |
| STATEID21 | Chhattisgarh | 46,978 | 0.0322 | 0.176 | 0 | 1 |
| STATEID22 | MP | 46,978 | 0.0782 | 0.269 | 0 | 1 |
| STATEID23 | Gujarat | 46,978 | 0.0528 | 0.224 | 0 | 1 |
| STATEID24 | Daman & Diu | 46,978 | 0.00149 | 0.0386 | 0 | 1 |
| STATEID25 | Dadar and Nagar Haveli | 46,978 | 0.00177 | 0.0420 | 0 | 1 |
| STATEID26 | Maharashtra | 46,978 | 0.0840 | 0.277 | 0 | 1 |
| STATEID27 | Andhra | 46,978 | 0.0453 | 0.208 | 0 | 1 |
| STATEID28 | Karnataka | 46,978 | 0.102 | 0.302 | 0 | 1 |
| STATEID29 | Goa | 46,978 | 0.00319 | 0.0564 | 0 | 1 |
| STATEID30 | Kerala | 46,978 | 0.0306 | 0.172 | 0 | 1 |
| STATEID31 | Tamil Nadu | 46,978 | 0.0403 | 0.197 | 0 | 1 |
| STATEID32 | Pondicherry | 46,978 | 0.00234 | 0.0483 | 0 | 1 |
| Dummy for Occupation (self) | | | | | | |
| newoccup11 | Agriculture and allied | 46,978 | 0.161 | 0.368 | 0 | 1 |
| newoccup12 | Production | 46,978 | 0.365 | 0.481 | 0 | 1 |
| newoccup13 | Profession | 46,978 | 0.0422 | 0.201 | 0 | 1 |
| newoccup14 | Administrative | 46,978 | 0.0106 | 0.102 | 0 | 1 |
| newoccup15 | Clerical | 46,978 | 0.0510 | 0.220 | 0 | 1 |
| newoccup16 | Sales | 46,978 | 0.0682 | 0.252 | 0 | 1 |
| newoccup17 | Others | 46,978 | 0.302 | 0.459 | 0 | 1 |
| Dummy for Occupation (Father/ father in law) | | | | | | |
| newoccup21 | Agriculture and allied | 46,978 | 0.626 | 0.484 | 0 | 1 |
| newoccup22 | Production | 46,978 | 0.183 | 0.387 | 0 | 1 |
| newoccup23 | Profession | 46,978 | 0.0273 | 0.163 | 0 | 1 |
| newoccup24 | Administrative | 46,978 | 0.0136 | 0.116 | 0 | 1 |
| newoccup25 | Clerical | 46,978 | 0.0307 | 0.173 | 0 | 1 |
| newoccup26 | Sales | 46,978 | 0.0969 | 0.296 | 0 | 1 |
| newoccup27 | Others | 46,978 | 0.0216 | 0.145 | 0 | 1 |
| newoccup28 | NIWF | 46,978 | 0.000788 | 0.0281 | 0 | 1 |
| Dummy for Industry (self) | | | | | | |
| newindus11 | Agriculture and allied | 46,978 | 0.162 | 0.369 | 0 | 1 |
| newindus12 | Production | 46,978 | 0.0554 | 0.229 | 0 | 1 |
| newindus13 | Manufacturing | 46,978 | 0.0613 | 0.240 | 0 | 1 |
| newindus14 | Electronics | 46,978 | 0.202 | 0.402 | 0 | 1 |
| newindus15 | Trade | 46,978 | 0.109 | 0.312 | 0 | 1 |
| newindus16 | Fin & Serv | 46,978 | 0.122 | 0.327 | 0 | 1 |
| newindus17 | Others | 46,978 | 0.287 | 0.453 | 0 | 1 |
| Dummy for Industry (Father/ father in law) | | | | | | |
| newindus21 | Agriculture and allied | 46,978 | 0.201 | 0.401 | 0 | 1 |
| newindus22 | Production | 46,978 | 0.0549 | 0.228 | 0 | 1 |
| newindus23 | Manufacturing | 46,978 | 0.0299 | 0.170 | 0 | 1 |
| newindus24 | Electronics | 46,978 | 0.0843 | 0.278 | 0 | 1 |
| newindus25 | Trade | 46,978 | 0.0984 | 0.298 | 0 | 1 |
| newindus26 | Fin & Serv | 46,978 | 0.0999 | 0.300 | 0 | 1 |
| newindus27 | Others | 46,978 | 0.431 | 0.495 | 0 | 1 |
| log\_wages | Log of wage | 33,600 | 10.49 | 1.128 | 4.605 | 14.70 |
| log\_hr | Log of hours worked | 32,781 | 7.432 | 0.604 | 5.481 | 8.294 |
| log\_hrlywage | Log of hourly wage | 32,781 | 3.135 | 0.714 | -1.099 | 7.048 |
|  |  |  |  |  |  |  |

Observation for all the variables accept log\_wages, log\_hr and log\_hrlywage which has missing values for non-wage workers is 46978.

1. Estimate a model for log of hourly earnings using a linear regression model based on the sample for those who are earning wages or salaries (if wgnw>=4).

Sol. The estimated coefficients of the model are provided in Table 1 under Model 1.

Let the estimated equation be:

Where X1 is the set of explanatory variables.

***Table 1: Estimated Coefficients of Model 1 and Model 2***

|  |  |  |
| --- | --- | --- |
| VARIABLES | Model 1 | Model 2 |
|  |  |  |
| log\_hr | - | 0.999\*\*\* |
|  |  | (0.00617) |
| exp | 0.0270\*\*\* | 0.0270\*\*\* |
|  | (0.000961) | (0.000970) |
| exp\_sq | -0.000329\*\*\* | -0.000329\*\*\* |
|  | (1.77e-05) | (1.79e-05) |
| STATEID2 | -0.104\*\*\* | -0.104\*\*\* |
|  | (0.0286) | (0.0286) |
| STATEID3 | -0.200\*\*\* | -0.200\*\*\* |
|  | (0.0296) | (0.0296) |
| STATEID4 | 0.00339 | 0.00350 |
|  | (0.0776) | (0.0776) |
| STATEID5 | -0.162\*\*\* | -0.162\*\*\* |
|  | (0.0362) | (0.0362) |
| STATEID6 | 0.00766 | 0.00771 |
|  | (0.0266) | (0.0266) |
| STATEID7 | -0.229\*\*\* | -0.230\*\*\* |
|  | (0.0254) | (0.0254) |
| STATEID8 | -0.502\*\*\* | -0.502\*\*\* |
|  | (0.0245) | (0.0245) |
| STATEID9 | -0.465\*\*\* | -0.465\*\*\* |
|  | (0.0288) | (0.0288) |
| STATEID10 | 0.0707 | 0.0708 |
|  | (0.0653) | (0.0653) |
| STATEID11 | 0.723\*\*\* | 0.724\*\*\* |
|  | (0.0656) | (0.0656) |
| STATEID12 | 0.230\*\*\* | 0.230\*\*\* |
|  | (0.0746) | (0.0747) |
| STATEID13 | 0.174\*\*\* | 0.174\*\*\* |
|  | (0.0676) | (0.0676) |
| STATEID14 | -0.00923 | -0.00944 |
|  | (0.0713) | (0.0713) |
| STATEID15 | -0.0512 | -0.0514 |
|  | (0.0472) | (0.0472) |
| STATEID16 | 0.152\*\* | 0.152\*\* |
|  | (0.0609) | (0.0609) |
| STATEID17 | -0.00214 | -0.00205 |
|  | (0.0284) | (0.0284) |
| STATEID18 | -0.440\*\*\* | -0.440\*\*\* |
|  | (0.0256) | (0.0256) |
| STATEID19 | -0.327\*\*\* | -0.327\*\*\* |
|  | (0.0303) | (0.0303) |
| STATEID20 | -0.391\*\*\* | -0.391\*\*\* |
|  | (0.0262) | (0.0262) |
| STATEID21 | -0.554\*\*\* | -0.554\*\*\* |
|  | (0.0283) | (0.0283) |
| STATEID22 | -0.591\*\*\* | -0.591\*\*\* |
|  | (0.0253) | (0.0253) |
| STATEID23 | -0.376\*\*\* | -0.376\*\*\* |
|  | (0.0264) | (0.0264) |
| STATEID24 | 0.0616 | 0.0621 |
|  | (0.0873) | (0.0874) |
| STATEID25 | -0.274\*\*\* | -0.274\*\*\* |
|  | (0.0758) | (0.0758) |
| STATEID26 | -0.190\*\*\* | -0.190\*\*\* |
|  | (0.0251) | (0.0251) |
| STATEID27 | -0.122\*\*\* | -0.122\*\*\* |
|  | (0.0262) | (0.0262) |
| STATEID28 | -0.184\*\*\* | -0.184\*\*\* |
|  | (0.0244) | (0.0244) |
| STATEID29 | 0.0327 | 0.0329 |
|  | (0.0520) | (0.0520) |
| STATEID30 | 0.347\*\*\* | 0.347\*\*\* |
|  | (0.0278) | (0.0278) |
| STATEID31 | 0.0404 | 0.0404 |
|  | (0.0268) | (0.0268) |
| STATEID32 | 0.0965 | 0.0967 |
|  | (0.0757) | (0.0757) |
| educd2 | 0.0447\*\*\* | 0.0446\*\*\* |
|  | (0.0126) | (0.0126) |
| educd3 | 0.132\*\*\* | 0.132\*\*\* |
|  | (0.0101) | (0.0101) |
| educd4 | 0.211\*\*\* | 0.211\*\*\* |
|  | (0.0114) | (0.0114) |
| educd5 | 0.302\*\*\* | 0.302\*\*\* |
|  | (0.0143) | (0.0143) |
| educd6 | 0.609\*\*\* | 0.609\*\*\* |
|  | (0.0162) | (0.0162) |
| region2 | 0.146\*\*\* | 0.146\*\*\* |
|  | (0.00830) | (0.00839) |
| region3 | 0.358\*\*\* | 0.359\*\*\* |
|  | (0.0151) | (0.0151) |
| dosgr | 0.112\*\*\* | 0.112\*\*\* |
|  | (0.0135) | (0.0135) |
| dsc | 0.0843\*\*\* | 0.0844\*\*\* |
|  | (0.0125) | (0.0125) |
| dobc | 0.0549\*\*\* | 0.0549\*\*\* |
|  | (0.0121) | (0.0121) |
| duch | 0.137\*\*\* | 0.137\*\*\* |
|  | (0.0199) | (0.0199) |
| dorelg | -0.100\*\* | -0.100\*\* |
|  | (0.0390) | (0.0390) |
| djain | 0.0730 | 0.0729 |
|  | (0.0841) | (0.0841) |
| dchrstn | -0.00239 | -0.00247 |
|  | (0.0341) | (0.0342) |
| dmuslim | -0.100\*\*\* | -0.100\*\*\* |
|  | (0.0295) | (0.0295) |
| dhindu | -0.0737\*\*\* | -0.0738\*\*\* |
|  | (0.0278) | (0.0278) |
| newindus22 | 0.00260 | 0.00252 |
|  | (0.0331) | (0.0331) |
| newindus23 | 0.0440 | 0.0439 |
|  | (0.0346) | (0.0346) |
| newindus24 | 0.0323 | 0.0323 |
|  | (0.0323) | (0.0323) |
| newindus25 | 0.0118 | 0.0116 |
|  | (0.0307) | (0.0307) |
| newindus26 | 0.00393 | 0.00381 |
|  | (0.0304) | (0.0304) |
| newindus27 | 0.00374 | 0.00356 |
|  | (0.00869) | (0.00876) |
| newindus12 | 0.0659 | 0.0661 |
|  | (0.0407) | (0.0407) |
| newindus13 | 0.128\*\*\* | 0.128\*\*\* |
|  | (0.0408) | (0.0408) |
| newindus14 | 0.123\*\*\* | 0.123\*\*\* |
|  | (0.0400) | (0.0400) |
| newindus15 | 0.0195 | 0.0196 |
|  | (0.0400) | (0.0400) |
| newindus16 | 0.166\*\*\* | 0.166\*\*\* |
|  | (0.0399) | (0.0400) |
| newindus17 | 0.0672 | 0.0673 |
|  | (0.0640) | (0.0640) |
| newoccup12 | 0.0388 | 0.0390 |
|  | (0.0399) | (0.0399) |
| newoccup13 | 0.382\*\*\* | 0.383\*\*\* |
|  | (0.0423) | (0.0423) |
| newoccup14 | 0.642\*\*\* | 0.642\*\*\* |
|  | (0.0458) | (0.0459) |
| newoccup15 | 0.260\*\*\* | 0.261\*\*\* |
|  | (0.0409) | (0.0410) |
| newoccup16 | -0.0857\*\* | -0.0854\*\* |
|  | (0.0407) | (0.0408) |
| newoccup17 | 0.00360 | 0.00374 |
|  | (0.0440) | (0.0441) |
| newoccup22 | -0.0284 | -0.0284 |
|  | (0.0309) | (0.0309) |
| newoccup23 | 0.0828\*\* | 0.0828\*\* |
|  | (0.0346) | (0.0346) |
| newoccup24 | 0.152\*\*\* | 0.152\*\*\* |
|  | (0.0406) | (0.0406) |
| newoccup25 | 0.0858\*\*\* | 0.0859\*\*\* |
|  | (0.0329) | (0.0329) |
| newoccup26 | -0.0134 | -0.0133 |
|  | (0.0308) | (0.0308) |
| newoccup27 | -0.0438 | -0.0437 |
|  | (0.0305) | (0.0305) |
| Constant | 2.578\*\*\* | 2.585\*\*\* |
|  | (0.0401) | (0.0590) |
|  |  |  |
| Observations | 32,714 | 32,714 |
| R-squared | 0.391 | 0.704 |

**Standard errors in parentheses**

**\*\*\* p<0.01, \*\* p<0.05, \* p<0.1**

2. Suppose the model is estimated in log wages and log hours is an explanatory variable in the model. How will you compare the two models? Can we have a test of hypothesis to compare the two models? That is, the model in 1 nested in model 2 or vice-versa? If so, what is your conclusion and based on the result which model will you choose as best fitting one?

Sol. Let the estimated equation be

Where X1 is the set of explanatory variables and log\_hr is the log of number of hours worked.

The estimated coefficients for the model with log wages as dependent and log hours as explanatory variable are presented as Model 2 in Table 1.

The R-sq is very high for model two this is because of the presence of log\_hr which explains log\_wage. We cannot compare Model 1 and Model 2 using R-sq as both the Models have different dependent variable.

As observed the estimated coefficients for most of the explanatory variable is same.

If we test the hypothesis

HO: = 1

Ha: ≠ 1

We used t-test to test the hypothesis. If HO is accepted, then we can say that Model 1 is same as Model 2, otherwise different.

We accept fail to reject the null at 5% level of significance. Therefore, we can say that model one is similar to model two.

We will use Model 1 for our analysis as we can easily apply tobit model to that as there is no observations are missing for dependent variable in the model. Whereas in Model 2 value for the variable log of workhours (log\_hr) are missing which is an explanatory variable.

For further analysis and to use the complete data we set the missing values of the dependent variable log of hourly earnings (log\_hrlywage) as zero.

3. Based on the result in question2 for the best fitting model, estimate either a model log of hourly earnings or log earnings as may be the case, including those in the non-wage sector that is for wgnw>=3. Estimate the model using the discussions in section 16.4 of book1. Estimate the tobit model, OLS model and model diagnostics as suggested in that section. Given appropriate interpretations based your data.

Sol. Model 1 is the best fitted model.

We use the Tobit model to estimate the data after censoring the missing values to zero i.e. log\_hrlywages=0. We also estimate the OLS model post censoring the data and we then compare it with the tobit model. The results are presented in Table 2.

**Table 2: Estimated coefficients for Tobit and OLS model.**

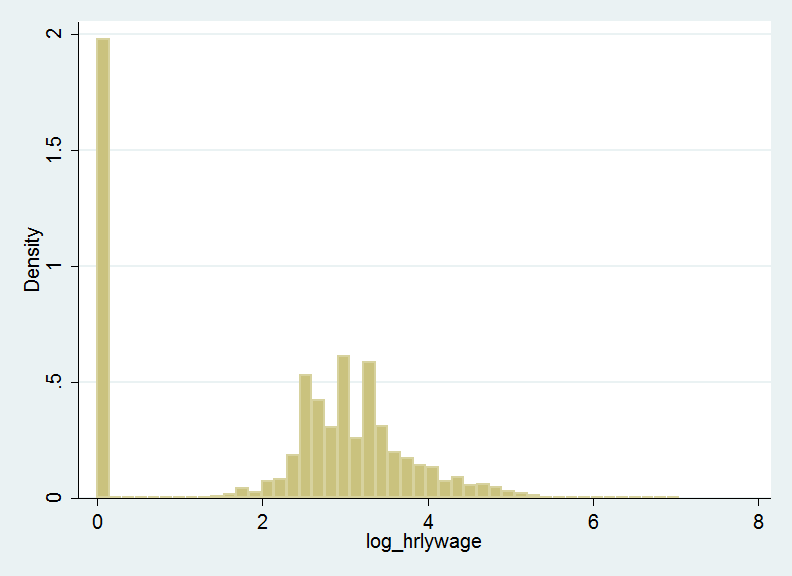
|  |  |  |
| --- | --- | --- |
| VARIABLES | TOBIT Model | OLS Model |
|  |  |  |
| exp | 0.0275\*\*\* | 0.0173\*\*\* |
|  | (0.00126) | (0.000866) |
| exp\_sq | -0.000400\*\*\* | -0.000237\*\*\* |
|  | (2.32e-05) | (1.60e-05) |
| STATEID1= J&K as dummy | | |
| STATEID2 | -0.167\*\*\* | -0.119\*\*\* |
|  | (0.0372) | (0.0268) |
| STATEID3 | -0.263\*\*\* | -0.181\*\*\* |
|  | (0.0387) | (0.0279) |
| STATEID4 | -0.0793 | -0.0717 |
|  | (0.0990) | (0.0702) |
| STATEID5 | -0.255\*\*\* | -0.164\*\*\* |
|  | (0.0476) | (0.0348) |
| STATEID6 | -0.0428 | -0.0300 |
|  | (0.0346) | (0.0250) |
| STATEID7 | -0.279\*\*\* | -0.198\*\*\* |
|  | (0.0331) | (0.0239) |
| STATEID8 | -0.621\*\*\* | -0.416\*\*\* |
|  | (0.0320) | (0.0230) |
| STATEID9 | -0.497\*\*\* | -0.343\*\*\* |
|  | (0.0377) | (0.0266) |
| STATEID10 | -0.0253 | 0.00856 |
|  | (0.0857) | (0.0633) |
| STATEID11 | 0.275\*\*\* | 0.265\*\*\* |
|  | (0.0820) | (0.0609) |
| STATEID12 | 0.611\*\*\* | 0.415\*\*\* |
|  | (0.0942) | (0.0709) |
| STATEID13 | 0.176\* | 0.252\*\*\* |
|  | (0.0899) | (0.0720) |
| STATEID14 | 0.0337 | 0.145\* |
|  | (0.0955) | (0.0758) |
| STATEID15 | 0.0155 | 0.0318 |
|  | (0.0624) | (0.0470) |
| STATEID16 | 0.623\*\*\* | 0.372\*\*\* |
|  | (0.0767) | (0.0568) |
| STATEID17 | -0.000751 | 0.0278 |
|  | (0.0373) | (0.0279) |
| STATEID18 | -0.592\*\*\* | -0.440\*\*\* |
|  | (0.0334) | (0.0246) |
| STATEID19 | -0.346\*\*\* | -0.245\*\*\* |
|  | (0.0400) | (0.0297) |
| STATEID20 | -0.461\*\*\* | -0.325\*\*\* |
|  | (0.0342) | (0.0248) |
| STATEID21 | -0.718\*\*\* | -0.546\*\*\* |
|  | (0.0367) | (0.0271) |
| STATEID22 | -0.743\*\*\* | -0.507\*\*\* |
|  | (0.0329) | (0.0237) |
| STATEID23 | -0.433\*\*\* | -0.297\*\*\* |
|  | (0.0345) | (0.0247) |
| STATEID24 | 0.0740 | 0.0318 |
|  | (0.113) | (0.0787) |
| STATEID25 | -0.279\*\*\* | -0.191\*\*\* |
|  | (0.101) | (0.0729) |
| STATEID26 | -0.230\*\*\* | -0.149\*\*\* |
|  | (0.0327) | (0.0236) |
| STATEID27 | -0.220\*\*\* | -0.122\*\*\* |
|  | (0.0342) | (0.0255) |
| STATEID28 | -0.260\*\*\* | -0.176\*\*\* |
|  | (0.0318) | (0.0232) |
| STATEID29 | 0.0382 | 0.139\*\* |
|  | (0.0698) | (0.0561) |
| STATEID30 | 0.334\*\*\* | 0.351\*\*\* |
|  | (0.0364) | (0.0273) |
| STATEID31 | -0.0252 | 0.0505\* |
|  | (0.0351) | (0.0261) |
| STATEID32 | -0.0539 | -0.0775 |
|  | (0.0950) | (0.0642) |
| Educd1= illiterate as dummy | | |
| educd2 | 0.0304\* | 0.0264\*\* |
|  | (0.0166) | (0.0124) |
| educd3 | 0.0713\*\*\* | 0.0584\*\*\* |
|  | (0.0134) | (0.00983) |
| educd4 | 0.120\*\*\* | 0.0957\*\*\* |
|  | (0.0150) | (0.0109) |
| educd5 | 0.187\*\*\* | 0.142\*\*\* |
|  | (0.0188) | (0.0132) |
| educd6 | 0.433\*\*\* | 0.296\*\*\* |
|  | (0.0210) | (0.0146) |
| Region1= rural as dummy | | |
| region2 | 0.203\*\*\* | 0.137\*\*\* |
|  | (0.0110) | (0.00789) |
| region3 | 0.444\*\*\* | 0.350\*\*\* |
|  | (0.0199) | (0.0149) |
|  |  |  |
| dosgr | -0.0237 | 0.0968\*\*\* |
|  | (0.0230) | (0.0130) |
| dst | -0.133\*\*\* |  |
|  | (0.0259) |  |
| dsc | -0.00164 | 0.125\*\*\* |
|  | (0.0230) | (0.0126) |
| dobc | -0.0631\*\*\* | 0.0806\*\*\* |
|  | (0.0225) | (0.0119) |
| Hindu as base | | |
| dorelg | 0.00591 |  |
|  | (0.0367) |  |
| djain | 0.00260 | -0.0596 |
|  | (0.0962) | (0.0640) |
| dsikh | 0.0446 | 0.0105 |
|  | (0.0362) | (0.0371) |
| dchrstn | 0.0515\* | 0.0260 |
|  | (0.0289) | (0.0348) |
| dmuslim | -0.000339 | -0.0224 |
|  | (0.0142) | (0.0296) |
| newindus22 | -0.158\*\*\* | -0.0949\*\*\* |
|  | (0.0441) | (0.0331) |
| newindus23 | -0.101\*\* | -0.0554 |
|  | (0.0460) | (0.0346) |
| newindus24 | -0.0798\* | -0.0141 |
|  | (0.0431) | (0.0326) |
| newindus25 | -0.175\*\*\* | -0.109\*\*\* |
|  | (0.0408) | (0.0306) |
| newindus26 | -0.157\*\*\* | -0.0866\*\*\* |
|  | (0.0405) | (0.0303) |
| newindus27 | -0.0843\*\*\* | -0.0546\*\*\* |
|  | (0.0115) | (0.00864) |
| newindus12 | 0.114\*\* | 0.142\*\*\* |
|  | (0.0549) | (0.0457) |
| newindus13 | 0.172\*\*\* | 0.184\*\*\* |
|  | (0.0550) | (0.0458) |
| newindus14 | 0.0317 | 0.0343 |
|  | (0.0539) | (0.0449) |
| newindus15 | 0.0815 | 0.114\*\* |
|  | (0.0539) | (0.0449) |
| newindus16 | 0.226\*\*\* | 0.264\*\*\* |
|  | (0.0539) | (0.0449) |
| newindus17 | -4.124\*\*\* | -2.554\*\*\* |
|  | (0.0622) | (0.0494) |
| newoccup12 | 0.0255 | 0.0432 |
|  | (0.0537) | (0.0447) |
| newoccup13 | 0.463\*\*\* | 0.565\*\*\* |
|  | (0.0570) | (0.0472) |
| newoccup14 | 0.727\*\*\* | 0.861\*\*\* |
|  | (0.0618) | (0.0513) |
| newoccup15 | 0.320\*\*\* | 0.406\*\*\* |
|  | (0.0552) | (0.0459) |
| newoccup16 | -0.0624 | -0.0299 |
|  | (0.0549) | (0.0457) |
| newoccup17 | -0.380\*\*\* | -0.249\*\*\* |
|  | (0.0591) | (0.0492) |
| newoccup22 | 0.114\*\*\* | 0.0478 |
|  | (0.0412) | (0.0310) |
| newoccup23 | 0.192\*\*\* | 0.107\*\*\* |
|  | (0.0458) | (0.0340) |
| newoccup24 | 0.260\*\*\* | 0.141\*\*\* |
|  | (0.0531) | (0.0380) |
| newoccup25 | 0.204\*\*\* | 0.139\*\*\* |
|  | (0.0438) | (0.0329) |
| newoccup26 | 0.0878\*\* | 0.0208 |
|  | (0.0408) | (0.0303) |
| newoccup27 | 0.159\*\*\* | 0.0893\*\*\* |
|  | (0.0406) | (0.0311) |
| Constant | 2.568\*\*\* | 2.645\*\*\* |
|  | (0.0418) | (0.0382) |
|  |  |  |
| Observations | 46,978 | 46,978 |
| R-squared |  | 0.836 |

Standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

All the significant slope coefficients estimated for Tobit model are larger than that of the OLS model in absolute terms, whereas the constant is smaller for the Tobit model. This means that the slope coefficients estimated for the Tobit model are overestimated and the constant is underestimated. The slope coefficients are 50% smaller for OLS when compared to Tobit model. The impact of censoring is equal to the proportion of zero observation in the data viz. 44%.

We are not going for upper censoring as the values for log\_hrlywages wages are nearly evenly distributed for both the extreme ends i.e. for log\_hrlywage < 2 & >5.

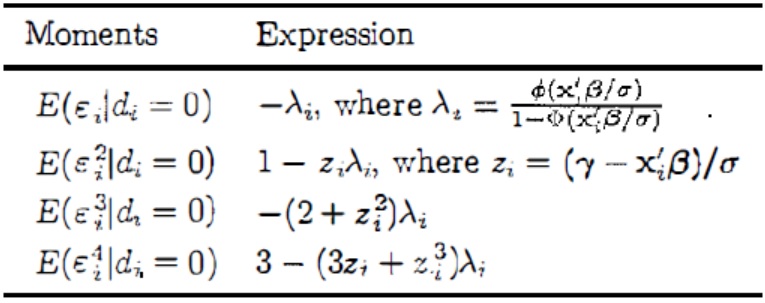


**Model Diagnostics:**

We will check the assumption of normality of error and homoscedasticity. As specified in the book we will use the generalized residuals and scores to test. For uncensored part we estimate the residuals as

We further raise the power of this residual to get the relevant moment.

For the uncensored part we calculated the residuals as following:



The values for the first, second, third and the fourth moments are marked as gres1, gres2, gres3 and gres4. The summary of these variables are presented below in Table 3.

**Table 3: Summary of generalised residual**



The mean of gres1 is close to zero however the rest of the variables us far from zero.

Therefore, we can say that the mean of the residual is close to zero.

Residuals gres3 and gres4 satisfies zero mean property only if model is correctly specified. Thus, the Tobit model we tried to fit is not correctly specified.

**Normality test:** We applied LR test to test for normality. To do this we calculate the likelihood scores for all the explanatory variables using gres1 which is the score with respect to the intercept of slopes. We then perform an auxiliary regression of 1 (constant) on these scores along with the generalised residuals. We get the NR2 statistics as the output and its p value

The result is presented as follows:

**N R^2 = 7165.6444 with p-value = 0**

The outcome of the test is a very strong rejection of the normality hypothesis.

**Homoscedasticity Test:** We test the following hypothesis:

H0: variance = constant

Ha: variance= constant\* exp(wb)

Here we generate the scores using gres2 which is the score with respect to the standard deviation. We then perform an auxiliary regression of 1 (constant) on these scores along with the generalised residuals. We get the NR2 statistics as the output and its p value

The result is presented as follows:

**N R^2 = 7165.6444 with p-value = 0**

The outcome of the test is a very strong rejection of the null hypothesis.

4. Follow book1 and complete the assignment for the modelling and prediction issues discussed in Sections 16.5-16.7.

Sol.

**Two-Part Model**

The Tobit regression makes a strong assumption that the same probability mechanism generates both the zeros and the positives (i.e. censored and uncensored data).

If we allow for the flexibility to this assumption i.e. different probability models estimates the decision to participate in the wage market and the outcome of the wage market viz. the hourly earnings. We estimate the model in two parts

Part 1. A binary outcome for the participation or non-participation in wage market using the Probit model.

Part 2. The outcome of the wage market i.e. the lof\_hrlywage is estimated using a OLS model of estimation for E(log\_hrlywages| hrly\_wages >0).

**Assumptions:**

1. We assume minimum hourly wage to be 0 and anyone who earns less than or equal to this is not in labour force.
2. The two parts (part 1 and part 2) are independent and are estimated separately.

We create a dummy variable

Then two-part model for y is given by

where is the conditional density of log\_hrlywage.

The results of Part 1 and Part 2 are presented in Table 4.

***Table 4. Estimated Coefficients from two-part model***

|  |  |  |
| --- | --- | --- |
| VARIABLES | Part 1 | Part 2 |
|  |  |  |
| exp | 0.00766\* | 0.0270\*\*\* |
|  | (0.00464) | (0.000943) |
| exp\_sq | -0.000381\*\*\* | -0.000330\*\*\* |
|  | (8.19e-05) | (1.74e-05) |
| STATEID2 | -0.394\*\*\* | -0.0973\*\*\* |
|  | (0.153) | (0.0281) |
| STATEID3 | -0.175 | -0.199\*\*\* |
|  | (0.190) | (0.0291) |
| STATEID4 | -0.444 | 0.00476 |
|  | (0.524) | (0.0761) |
| STATEID5 | -0.558\*\*\* | -0.160\*\*\* |
|  | (0.205) | (0.0355) |
| STATEID6 | -0.188 | 0.00974 |
|  | (0.160) | (0.0261) |
| STATEID7 | -0.344\*\* | -0.227\*\*\* |
|  | (0.143) | (0.0249) |
| STATEID8 | -0.814\*\*\* | -0.495\*\*\* |
|  | (0.137) | (0.0241) |
| STATEID9 | -0.363\*\* | -0.452\*\*\* |
|  | (0.160) | (0.0283) |
| STATEID10 | -0.377 | 0.0706 |
|  | (0.508) | (0.0641) |
| STATEID11 | -1.182\*\*\* | 0.723\*\*\* |
|  | (0.257) | (0.0644) |
| STATEID12 | 1.494\*\*\* | 0.228\*\*\* |
|  | (0.302) | (0.0732) |
| STATEID13 | 0.491 | 0.176\*\*\* |
|  | (0.670) | (0.0663) |
| STATEID14 | -0.508 | 0.0959 |
|  | (0.346) | (0.0707) |
| STATEID15 | 0.447 | -0.0480 |
|  | (0.306) | (0.0463) |
| STATEID16 | 1.479\*\*\* | 0.152\*\* |
|  | (0.243) | (0.0598) |
| STATEID17 | -0.0413 | -0.000791 |
|  | (0.168) | (0.0279) |
| STATEID18 | -0.925\*\*\* | -0.436\*\*\* |
|  | (0.140) | (0.0251) |
| STATEID19 | -0.112 | -0.325\*\*\* |
|  | (0.193) | (0.0297) |
| STATEID20 | -0.509\*\*\* | -0.386\*\*\* |
|  | (0.144) | (0.0257) |
| STATEID21 | -0.859\*\*\* | -0.545\*\*\* |
|  | (0.144) | (0.0277) |
| STATEID22 | -0.865\*\*\* | -0.587\*\*\* |
|  | (0.138) | (0.0248) |
| STATEID23 | -0.514\*\*\* | -0.368\*\*\* |
|  | (0.150) | (0.0259) |
| STATEID24 | 0.250 | 0.0654 |
|  | (0.480) | (0.0857) |
| STATEID25 | -0.0328 | -0.272\*\*\* |
|  | (0.505) | (0.0744) |
| STATEID26 | -0.215 | -0.186\*\*\* |
|  | (0.146) | (0.0246) |
| STATEID27 | -0.595\*\*\* | -0.117\*\*\* |
|  | (0.150) | (0.0257) |
| STATEID28 | -0.525\*\*\* | -0.173\*\*\* |
|  | (0.140) | (0.0239) |
| STATEID29 | 0.504 | 0.0335 |
|  | (0.632) | (0.0510) |
| STATEID30 | 0.347\* | 0.347\*\*\* |
|  | (0.182) | (0.0273) |
| STATEID31 | -0.351\*\* | 0.0471\* |
|  | (0.167) | (0.0263) |
| STATEID32 | -0.580 | 0.0986 |
|  | (0.419) | (0.0743) |
| educd2 | -0.0202 | 0.0408\*\*\* |
|  | (0.0603) | (0.0123) |
| educd3 | -0.231\*\*\* | 0.129\*\*\* |
|  | (0.0481) | (0.00992) |
| educd4 | -0.386\*\*\* | 0.208\*\*\* |
|  | (0.0561) | (0.0112) |
| educd5 | -0.460\*\*\* | 0.301\*\*\* |
|  | (0.0712) | (0.0140) |
| educd6 | -0.596\*\*\* | 0.603\*\*\* |
|  | (0.0871) | (0.0159) |
| region2 | 0.505\*\*\* | 0.146\*\*\* |
|  | (0.0554) | (0.00814) |
| region3 | 0.793\*\*\* | 0.355\*\*\* |
|  | (0.117) | (0.0148) |
| dosgr | 0.0604 | 0.113\*\*\* |
|  | (0.0933) | (0.0132) |
| dst | 0.0730 |  |
|  | (0.0984) |  |
| dsc | 0.301\*\*\* | 0.0874\*\*\* |
|  | (0.0929) | (0.0122) |
| dobc | 0.112 | 0.0589\*\*\* |
|  | (0.0901) | (0.0118) |
| dorelg | 0.127 | -0.100\*\*\* |
|  | (0.166) | (0.0382) |
| djain | -0.256 | 0.0667 |
|  | (0.408) | (0.0825) |
| dsikh | -0.278 |  |
|  | (0.176) |  |
| dchrstn | -0.110 | 0.00188 |
|  | (0.127) | (0.0335) |
| dmuslim | 0.121\*\* | -0.0989\*\*\* |
|  | (0.0579) | (0.0290) |
| newindus22 | -0.800\*\*\* | 0.00929 |
|  | (0.193) | (0.0325) |
| newindus23 | -0.720\*\*\* | 0.0531 |
|  | (0.203) | (0.0339) |
| newindus24 | -0.525\*\*\* | 0.0432 |
|  | (0.191) | (0.0317) |
| newindus25 | -0.914\*\*\* | 0.0163 |
|  | (0.174) | (0.0301) |
| newindus26 | -0.735\*\*\* | 0.00814 |
|  | (0.173) | (0.0298) |
| newindus27 | -0.430\*\*\* | 0.00467 |
|  | (0.0445) | (0.00852) |
| newindus12 | 0.379 | 0.0702\* |
|  | (0.250) | (0.0399) |
| newindus13 | 0.243 | 0.131\*\*\* |
|  | (0.243) | (0.0400) |
| newindus14 | -0.474\*\* | 0.123\*\*\* |
|  | (0.234) | (0.0392) |
| newindus15 | 0.504\*\* | 0.0248 |
|  | (0.246) | (0.0392) |
| newindus16 | 0.333 | 0.167\*\*\* |
|  | (0.243) | (0.0392) |
| newindus17 | -3.774\*\*\* | 0.0938 |
|  | (0.243) | (0.0630) |
| newoccup12 | 0.119 | 0.0378 |
|  | (0.234) | (0.0391) |
| newoccup13 | 0.471\* | 0.396\*\*\* |
|  | (0.266) | (0.0415) |
| newoccup14 | 0.739\*\* | 0.643\*\*\* |
|  | (0.366) | (0.0450) |
| newoccup15 | 0.798\*\*\* | 0.266\*\*\* |
|  | (0.276) | (0.0402) |
| newoccup16 | 0.370 | -0.0798\*\* |
|  | (0.254) | (0.0399) |
| newoccup17 | -0.981\*\*\* | -0.000296 |
|  | (0.240) | (0.0432) |
| newoccup22 | 0.716\*\*\* | -0.0373 |
|  | (0.180) | (0.0303) |
| newoccup23 | 0.468\*\* | 0.0775\*\* |
|  | (0.201) | (0.0339) |
| newoccup24 | 0.673\*\*\* | 0.142\*\*\* |
|  | (0.243) | (0.0399) |
| newoccup25 | 0.492\*\* | 0.0840\*\*\* |
|  | (0.200) | (0.0323) |
| newoccup26 | 0.424\*\* | -0.0188 |
|  | (0.173) | (0.0302) |
| newoccup27 | 1.154\*\*\* | -0.0538\* |
|  | (0.183) | (0.0299) |
| Constant | 2.878\*\*\* | 2.573\*\*\* |
|  | (0.174) | (0.0394) |
|  |  |  |
| Observations | 46,978 | 32,750 |
| R-squared |  | 0.401 |

Standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

The count R-sq for the estimated probit model is 0.979 which is very high. The relaven variables are statistically siginificiant.

We further compare the log likelihood for Two-Part and Tobit model. The log likelihood for Two-Part is is the sum of log-likelihood for Part-1 and Part-2 models. The log-likelihood for Two-Part is -30373.851 whereas for Tobit it was -39905.916. Since the log-likelihood for Two-Part is greater than that of TOBIT model therefore it fits the data better.

**Homoscedasticity Test:**

H0: variance= constant

Ha: variance ≠ constant

We perform B-P test for it.

chi2(1) = 2791.17

Prob > chi2 = 0.0000

Therefor we fail to reject null. Heteroscedasticity is present in the model.

**Normality Test:**

For normality of the residual we check the skewness and kurtosis of the residual.

H0: Error is normally distributed

Ha: Error is not normally distributed

Pr(Skewness)=0

Pr(Kurtosis) = 0

We strongly reject the null hypothesis with sktest.

**Selection Model:**

We will now try to relax the assumption 2 of the Two-Part Model which states that Part1 and Part 2 should be independent.

**Heckman Model without exclusion:** In this model the decision equation and the outcome equation will have same set of explanatory variable.

**Heckman MLE:**

If the models are not independent, then we can have selection biasness for Part-2. To overcome this, we use the Heckman estimation which assumes that the residuals of decision equation (to be in wage market or not) and the outcome equation (the log of hourly earnings) can be co-related to each other but will jointly follow Normal distribution with zero mean and constant variance.

The results are presented in Table 5.

**Table 5: Heckman MLE**

|  |  |  |
| --- | --- | --- |
| VARIABLES | Outcome Equation | Decision Equation |
|  |  |  |
| exp | 0.0264\*\*\* | -0.00262 |
|  | (0.000954) | (0.00444) |
| exp\_sq | -0.000309\*\*\* | -0.000166\*\* |
|  | (1.76e-05) | (7.88e-05) |
| STATEID2 | -0.0891\*\*\* | -0.480\*\*\* |
|  | (0.0285) | (0.144) |
| STATEID3 | -0.193\*\*\* | -0.321\* |
|  | (0.0295) | (0.178) |
| STATEID4 | 0.0105 | -0.686 |
|  | (0.0773) | (0.555) |
| STATEID5 | -0.148\*\*\* | -0.647\*\*\* |
|  | (0.0360) | (0.193) |
| STATEID6 | 0.0116 | -0.246 |
|  | (0.0265) | (0.151) |
| STATEID7 | -0.218\*\*\* | -0.372\*\*\* |
|  | (0.0252) | (0.132) |
| STATEID8 | -0.471\*\*\* | -0.832\*\*\* |
|  | (0.0244) | (0.127) |
| STATEID9 | -0.445\*\*\* | -0.477\*\*\* |
|  | (0.0286) | (0.149) |
| STATEID10 | 0.0795 | -0.571 |
|  | (0.0651) | (0.498) |
| STATEID11 | 0.760\*\*\* | -1.364\*\*\* |
|  | (0.0648) | (0.253) |
| STATEID12 | 0.172\*\* | 0.998\*\*\* |
|  | (0.0737) | (0.303) |
| STATEID13 | 0.171\*\* | 0.278 |
|  | (0.0675) | (0.736) |
| STATEID14 | 0.119\* | -0.579\* |
|  | (0.0717) | (0.346) |
| STATEID15 | -0.0634 | 0.324 |
|  | (0.0470) | (0.310) |
| STATEID16 | 0.106\* | 1.446\*\*\* |
|  | (0.0599) | (0.225) |
| STATEID17 | -0.00583 | -0.165 |
|  | (0.0283) | (0.160) |
| STATEID18 | -0.408\*\*\* | -0.918\*\*\* |
|  | (0.0254) | (0.129) |
| STATEID19 | -0.324\*\*\* | -0.209 |
|  | (0.0302) | (0.182) |
| STATEID20 | -0.375\*\*\* | -0.667\*\*\* |
|  | (0.0260) | (0.133) |
| STATEID21 | -0.510\*\*\* | -0.952\*\*\* |
|  | (0.0280) | (0.133) |
| STATEID22 | -0.555\*\*\* | -0.919\*\*\* |
|  | (0.0251) | (0.127) |
| STATEID23 | -0.356\*\*\* | -0.522\*\*\* |
|  | (0.0263) | (0.140) |
| STATEID24 | 0.0546 | 0.00755 |
|  | (0.0869) | (0.545) |
| STATEID25 | -0.274\*\*\* | -0.238 |
|  | (0.0755) | (0.512) |
| STATEID26 | -0.183\*\*\* | -0.369\*\*\* |
|  | (0.0249) | (0.137) |
| STATEID27 | -0.103\*\*\* | -0.589\*\*\* |
|  | (0.0260) | (0.139) |
| STATEID28 | -0.160\*\*\* | -0.538\*\*\* |
|  | (0.0243) | (0.129) |
| STATEID29 | 0.0305 | 0.381 |
|  | (0.0519) | (0.626) |
| STATEID30 | 0.347\*\*\* | 0.356\*\* |
|  | (0.0277) | (0.167) |
| STATEID31 | 0.0556\*\* | -0.446\*\*\* |
|  | (0.0267) | (0.158) |
| STATEID32 | 0.110 | -0.663\* |
|  | (0.0753) | (0.400) |
| educd2 | 0.0415\*\*\* | -0.0198 |
|  | (0.0125) | (0.0580) |
| educd3 | 0.138\*\*\* | -0.161\*\*\* |
|  | (0.0100) | (0.0462) |
| educd4 | 0.222\*\*\* | -0.335\*\*\* |
|  | (0.0113) | (0.0540) |
| educd5 | 0.318\*\*\* | -0.384\*\*\* |
|  | (0.0142) | (0.0688) |
| educd6 | 0.621\*\*\* | -0.630\*\*\* |
|  | (0.0161) | (0.0860) |
| region2 | 0.135\*\*\* | 0.693\*\*\* |
|  | (0.00827) | (0.0542) |
| region3 | 0.337\*\*\* | 0.815\*\*\* |
|  | (0.0150) | (0.118) |
| dosgr | -0.0218 | 0.0429 |
|  | (0.0176) | (0.0887) |
| dst | -0.133\*\*\* | 0.0426 |
|  | (0.0197) | (0.0938) |
| dsc | -0.0543\*\*\* | 0.242\*\*\* |
|  | (0.0175) | (0.0883) |
| dobc | -0.0762\*\*\* | 0.0929 |
|  | (0.0172) | (0.0854) |
| dorelg | -0.0310 | 0.112 |
|  | (0.0274) | (0.159) |
| djain | 0.155\*\* | -0.0473 |
|  | (0.0788) | (0.396) |
| dsikh | 0.0799\*\*\* | -0.136 |
|  | (0.0277) | (0.164) |
| dchrstn | 0.0759\*\*\* | -0.123 |
|  | (0.0220) | (0.122) |
| dmuslim | -0.0253\*\* | 0.131\*\* |
|  | (0.0107) | (0.0551) |
| newindus22 | 0.0381 | -0.640\*\*\* |
|  | (0.0330) | (0.188) |
| newindus23 | 0.0797\*\* | -0.560\*\*\* |
|  | (0.0344) | (0.196) |
| newindus24 | 0.0626\* | -0.381\*\* |
|  | (0.0322) | (0.185) |
| newindus25 | 0.0493 | -0.721\*\*\* |
|  | (0.0306) | (0.172) |
| newindus26 | 0.0376 | -0.587\*\*\* |
|  | (0.0303) | (0.172) |
| newindus27 | 0.0239\*\*\* | -0.349\*\*\* |
|  | (0.00864) | (0.0433) |
| newindus12 | 0.0610 | 0.530\*\* |
|  | (0.0406) | (0.236) |
| newindus13 | 0.124\*\*\* | 0.175 |
|  | (0.0407) | (0.227) |
| newindus14 | 0.143\*\*\* | -0.597\*\*\* |
|  | (0.0399) | (0.218) |
| newindus15 | 0.0137 | 0.541\*\* |
|  | (0.0399) | (0.231) |
| newindus16 | 0.158\*\*\* | 0.444\*\* |
|  | (0.0399) | (0.225) |
| newindus17 | 0.847\*\*\* | -3.845\*\*\* |
|  | (0.0613) | (0.229) |
| newoccup12 | 0.0360 | 0.326 |
|  | (0.0398) | (0.218) |
| newoccup13 | 0.387\*\*\* | 1.048\*\*\* |
|  | (0.0422) | (0.246) |
| newoccup14 | 0.632\*\*\* | 1.312\*\*\* |
|  | (0.0458) | (0.345) |
| newoccup15 | 0.255\*\*\* | 1.197\*\*\* |
|  | (0.0409) | (0.252) |
| newoccup16 | -0.0851\*\* | 0.685\*\*\* |
|  | (0.0406) | (0.233) |
| newoccup17 | 0.0610 | -0.761\*\*\* |
|  | (0.0439) | (0.225) |
| newoccup22 | -0.0618\*\* | 0.595\*\*\* |
|  | (0.0308) | (0.175) |
| newoccup23 | 0.0594\* | 0.422\*\* |
|  | (0.0344) | (0.202) |
| newoccup24 | 0.121\*\*\* | 0.548\*\* |
|  | (0.0404) | (0.232) |
| newoccup25 | 0.0634\* | 0.338\* |
|  | (0.0328) | (0.198) |
| newoccup26 | -0.0370 | 0.316\* |
|  | (0.0306) | (0.171) |
| newoccup27 | -0.0888\*\*\* | 0.984\*\*\* |
|  | (0.0303) | (0.177) |
| Constant | 2.629\*\*\* | 2.728\*\*\* |
|  | (0.0319) | (0.163) |
|  |  |  |
| Observations | 46,978 | 46,978 |

Standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**LR test of indep. eqns. (rho = 0): chi2(1) = 396.43 Prob > chi2 = 0.0000**

p-value for LR test comes out to be 0. Thus, the correlation between the errors is significantly different from zero.

**Heckman Two-Step:**

Now we estimate the Heckman two-step model which assumes univariate error than bivariate error. It is based on the conditional expectation

Where

The results for Heckman two-step estimation are present in Table 6.

**Table 6: Heckman Two-Step Model**

|  |  |  |
| --- | --- | --- |
| VARIABLES | Outcome Equation | Decision Equation |
|  |  |  |
| exp | 0.0263\*\*\* | 0.00766\* |
|  | (0.000967) | (0.00464) |
| exp\_sq | -0.000306\*\*\* | -0.000381\*\*\* |
|  | (1.79e-05) | (8.19e-05) |
| STATEID2 | -0.0868\*\*\* | -0.394\*\*\* |
|  | (0.0288) | (0.153) |
| STATEID3 | -0.190\*\*\* | -0.175 |
|  | (0.0299) | (0.190) |
| STATEID4 | 0.0164 | -0.444 |
|  | (0.0782) | (0.524) |
| STATEID5 | -0.144\*\*\* | -0.558\*\*\* |
|  | (0.0365) | (0.205) |
| STATEID6 | 0.0118 | -0.188 |
|  | (0.0268) | (0.160) |
| STATEID7 | -0.216\*\*\* | -0.344\*\* |
|  | (0.0255) | (0.143) |
| STATEID8 | -0.464\*\*\* | -0.814\*\*\* |
|  | (0.0249) | (0.137) |
| STATEID9 | -0.443\*\*\* | -0.363\*\* |
|  | (0.0290) | (0.160) |
| STATEID10 | 0.0841 | -0.377 |
|  | (0.0659) | (0.508) |
| STATEID11 | 0.780\*\*\* | -1.182\*\*\* |
|  | (0.0654) | (0.257) |
| STATEID12 | 0.189\*\* | 1.494\*\*\* |
|  | (0.0745) | (0.302) |
| STATEID13 | 0.171\*\* | 0.491 |
|  | (0.0684) | (0.670) |
| STATEID14 | 0.120\* | -0.508 |
|  | (0.0725) | (0.346) |
| STATEID15 | -0.0669 | 0.447 |
|  | (0.0476) | (0.306) |
| STATEID16 | 0.0964 | 1.479\*\*\* |
|  | (0.0611) | (0.243) |
| STATEID17 | -0.00528 | -0.0413 |
|  | (0.0286) | (0.168) |
| STATEID18 | -0.401\*\*\* | -0.925\*\*\* |
|  | (0.0259) | (0.140) |
| STATEID19 | -0.323\*\*\* | -0.112 |
|  | (0.0305) | (0.193) |
| STATEID20 | -0.370\*\*\* | -0.509\*\*\* |
|  | (0.0264) | (0.144) |
| STATEID21 | -0.497\*\*\* | -0.859\*\*\* |
|  | (0.0287) | (0.144) |
| STATEID22 | -0.545\*\*\* | -0.865\*\*\* |
|  | (0.0257) | (0.138) |
| STATEID23 | -0.353\*\*\* | -0.514\*\*\* |
|  | (0.0266) | (0.150) |
| STATEID24 | 0.0577 | 0.250 |
|  | (0.0877) | (0.480) |
| STATEID25 | -0.273\*\*\* | -0.0328 |
|  | (0.0764) | (0.505) |
| STATEID26 | -0.181\*\*\* | -0.215 |
|  | (0.0252) | (0.146) |
| STATEID27 | -0.0991\*\*\* | -0.595\*\*\* |
|  | (0.0264) | (0.150) |
| STATEID28 | -0.157\*\*\* | -0.525\*\*\* |
|  | (0.0246) | (0.140) |
| STATEID29 | 0.0304 | 0.504 |
|  | (0.0526) | (0.632) |
| STATEID30 | 0.347\*\*\* | 0.347\* |
|  | (0.0280) | (0.182) |
| STATEID31 | 0.0587\*\* | -0.351\*\* |
|  | (0.0270) | (0.167) |
| STATEID32 | 0.113 | -0.580 |
|  | (0.0761) | (0.419) |
| educd2 | 0.0419\*\*\* | -0.0202 |
|  | (0.0126) | (0.0603) |
| educd3 | 0.139\*\*\* | -0.231\*\*\* |
|  | (0.0102) | (0.0481) |
| educd4 | 0.225\*\*\* | -0.386\*\*\* |
|  | (0.0116) | (0.0561) |
| educd5 | 0.321\*\*\* | -0.460\*\*\* |
|  | (0.0145) | (0.0712) |
| educd6 | 0.626\*\*\* | -0.596\*\*\* |
|  | (0.0165) | (0.0871) |
| region2 | 0.129\*\*\* | 0.505\*\*\* |
|  | (0.00853) | (0.0554) |
| region3 | 0.333\*\*\* | 0.793\*\*\* |
|  | (0.0154) | (0.117) |
| dosgr | -0.0215 | 0.0604 |
|  | (0.0178) | (0.0933) |
| dst | -0.132\*\*\* | 0.0730 |
|  | (0.0199) | (0.0984) |
| dsc | -0.0556\*\*\* | 0.301\*\*\* |
|  | (0.0177) | (0.0929) |
| dobc | -0.0758\*\*\* | 0.112 |
|  | (0.0174) | (0.0901) |
| dorelg | -0.0326 | 0.127 |
|  | (0.0277) | (0.166) |
| djain | 0.154\* | -0.256 |
|  | (0.0797) | (0.408) |
| dsikh | 0.0785\*\*\* | -0.278 |
|  | (0.0280) | (0.176) |
| dchrstn | 0.0766\*\*\* | -0.110 |
|  | (0.0222) | (0.127) |
| dmuslim | -0.0265\*\* | 0.121\*\* |
|  | (0.0108) | (0.0579) |
| newindus22 | 0.0429 | -0.800\*\*\* |
|  | (0.0335) | (0.193) |
| newindus23 | 0.0832\*\* | -0.720\*\*\* |
|  | (0.0349) | (0.203) |
| newindus24 | 0.0646\*\* | -0.525\*\*\* |
|  | (0.0326) | (0.191) |
| newindus25 | 0.0544\* | -0.914\*\*\* |
|  | (0.0312) | (0.174) |
| newindus26 | 0.0432 | -0.735\*\*\* |
|  | (0.0308) | (0.173) |
| newindus27 | 0.0291\*\*\* | -0.430\*\*\* |
|  | (0.00907) | (0.0445) |
| newindus12 | 0.0585 | 0.379 |
|  | (0.0412) | (0.250) |
| newindus13 | 0.123\*\*\* | 0.243 |
|  | (0.0412) | (0.243) |
| newindus14 | 0.149\*\*\* | -0.474\*\* |
|  | (0.0405) | (0.234) |
| newindus15 | 0.0114 | 0.504\*\* |
|  | (0.0404) | (0.246) |
| newindus16 | 0.155\*\*\* | 0.333 |
|  | (0.0404) | (0.243) |
| newindus17 | 0.992\*\*\* | -3.774\*\*\* |
|  | (0.109) | (0.243) |
| newoccup12 | 0.0343 | 0.119 |
|  | (0.0403) | (0.234) |
| newoccup13 | 0.382\*\*\* | 0.471\* |
|  | (0.0427) | (0.266) |
| newoccup14 | 0.626\*\*\* | 0.739\*\* |
|  | (0.0464) | (0.366) |
| newoccup15 | 0.250\*\*\* | 0.798\*\*\* |
|  | (0.0414) | (0.276) |
| newoccup16 | -0.0880\*\* | 0.370 |
|  | (0.0411) | (0.254) |
| newoccup17 | 0.0701 | -0.981\*\*\* |
|  | (0.0450) | (0.240) |
| newoccup22 | -0.0655\*\* | 0.716\*\*\* |
|  | (0.0313) | (0.180) |
| newoccup23 | 0.0557 | 0.468\*\* |
|  | (0.0349) | (0.201) |
| newoccup24 | 0.119\*\*\* | 0.673\*\*\* |
|  | (0.0410) | (0.243) |
| newoccup25 | 0.0612\* | 0.492\*\* |
|  | (0.0333) | (0.200) |
| newoccup26 | -0.0392 | 0.424\*\* |
|  | (0.0310) | (0.173) |
| newoccup27 | -0.0941\*\*\* | 1.154\*\*\* |
|  | (0.0310) | (0.183) |
| lambda |  | -0.496\*\*\* |
|  |  | (0.0489) |
| Constant | 2.629\*\*\* | 2.878\*\*\* |
|  | (0.0323) | (0.174) |
|  |  |  |
| Observations | 46,978 | 46,978 |

Standard errors in parentheses

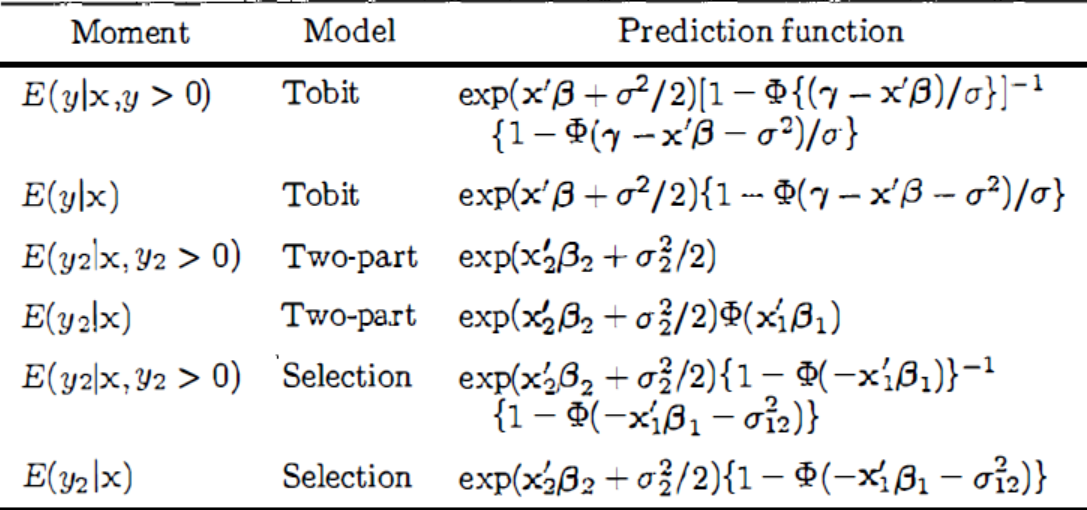
\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

We can use the coefficients of lambda from regression output to know whether there is independence between the two error terms. It follows Z-statistics which is -10.13. It is statistically different from zero as the p value is zero corresponding to lambda. Thus, correlation between the two error terms is statistically different from zero.

**Heckman Model with exclusion:** We try to estimate the Heckman model with introducing an additional variable for the decision equation (participation in wage market) which will be excluded from the outcome equation (hourly earnings). It is done to reduce multi-collinearity problem particularly in small sample. The additional variable chosen must be such that it impacts only the decision equation not the outcome equation.

Since our sample size is pretty large therefore we can afford to skip this estimation.

**Model Prediction:** Till now we have used three model for the analysis of the male labour market. Here we will try to compare the predictions from the three models. The prediction function of the models are as follows.



The predicted value from different models are shown in Table 7 below:



However, if we compare the log likelihood for Tobit, two-step and Heckman they -39905.916, -30373.851 -30175.64 respectively. This result of predicted values is in sharp contrast to the log likelihood for the three models.

**Further analysis**: Further analysis can be done by choosingan exclusion variable in the Heckman model of estimation. For example, we can choose number of individuals in the household which will only impact the decision equation, not the outcome equation.