

Course Title: Quantitative Techniques for Economics

Course code: ECON6002

Topic: The Blinder–Oaxaca decomposition for linear regression models

Ph.D. Economics (1st Semester)

Dr. Kailash Chandra Pradhan

**Mahatma Gandhi Central University,
Department of Economics**

The Blinder–Oaxaca decomposition for linear regression models

- ▶ This method is used to study labor-market outcomes by groups (sex, race, and so on) and this decomposes mean differences in log wages based on linear regression models in a counterfactual manner.
- ▶ The procedure is known in the literature as the Blinder–Oaxaca decomposition (Blinder 1973; Oaxaca 1973).
- ▶ In general, the technique can be employed to study group differences in any (continuous and unbounded) outcome variable.
- ▶ For example, O’Donnell et al. (2008) use it to analyze health inequalities by poverty status.

The Empirical Models

- ▶ In this presentation, we try to make an attempt to find out how the relationship between log per capita expenditure (PCE) and household characteristics (determinants) differs between rural and urban areas.
- ▶ Therefore, we used B-O decomposition to decompose changes in the mean of log real per capita expenditure (RPCE) due to household characteristics between rural and urban areas.
- ▶ The B-O decomposition is done by using the following linear regression model.

$$Y_j = X_j' \beta_j + \varepsilon_j \quad E(\varepsilon_j) = 0 \quad j \in (R, U) \quad (1)$$

Where Y is the log of RPCE, X is vector of household characteristics and a constant, β contains the slope parameters and the intercept, and ε is the error. R and U explain the rural and urban households respectively.

□

The Empirical Models

- ▶ The mean difference of log of RPCE of urban and rural household's can be expressed as:

$$R = E(Y_U) - E(Y_R) = E(X_U)' \beta_U - E(X_R)' \beta_R \quad (2)$$

because

$$E(Y_j) = E(X_j' \beta_j + \varepsilon_j) = E(X_j' \beta_j) + E(\varepsilon_j) = E(X_j)' \beta_j \quad (3)$$

Where $E(\beta_j) = \beta_j$ and $E(\varepsilon_j) = 0$ by assumption

- ▶ To identify the contribution of urban and rural differences in predictors to the overall real per capita consumption difference, equation (2) can be rearranged as:

$$R = \{E(X_U) - E(X_R)\}' \beta_R + E(X_R)' (\beta_U - \beta_R) + \{E(X_U) - E(X_R)\}' (\beta_U - \beta_R) \quad (4)$$

$$R = E + C + I \quad (5)$$

The Empirical Models

- ▶ The first component, $\{E(X_U) - E(X_R)\}'\beta_R$ explains the endowment effects that is due to differences in the predictors.
- ▶ The second components, $E(X_R)'(\beta_U - \beta_R)$ explains the differences in the coefficients (including differences in the intercept).
- ▶ The third component, $\{E(X_U) - E(X_R)\}'(\beta_U - \beta_R)$ explains interaction term accounting for the fact that differences in endowments and coefficients exist simultaneously between the urban and rural households.
- ▶ The decomposition shown in equation is formulated from the viewpoint of rural household.
- ▶ The E component measures the expected change of mean in per capita consumption of rural household if rural households had urban household's predictor levels.
- ▶ The C component measures the expected change of mean in per capita consumption of rural household if rural household had urban household's coefficients.

Data Source and Methodology

- ▶ The determinants of welfare of rural and urban households were identified by using cross section and pooled data obtained from two quinquennial rounds of National Sample Survey (NSS) namely 61st (2004-05) and 68th (2011-12).
- ▶ The study uses per capita household consumption expenditure to measure the welfare of the households
- ▶ Dependent variable –
 - ▶ Per capita household consumption expenditure
- ▶ Explanatory variables –
 - ▶ the numbers of persons living in the household (household size)
 - ▶ the age of the principal earner of the household
 - ▶ the education level of the principal earner (illiterate=1, literate & up to primary=2, secondary=3, under graduation=4, graduation & above=5)
 - ▶ the gender of the principal earner (female=1)
 - ▶ occupations (agriculture=1, industry=2, service=3) of the principal earner
 - ▶ per capita wage of the household
 - ▶ a dummy variable (urban =1) was introduced in the model to estimate the rural/urban consumption gap in case of cross section data
 - ▶ a dummy variable for time period (2011-12 =1) was introduced in the model to estimate the time impact on consumption expenditure in case of pooled data

-
- ▶ We have converted the nominal consumption expenditure and wage earnings of households into real term deflating by consumer price index (CPI, base year 1986-87).
 - ▶ Further, CPI agricultural labourer and CPI industrial workers are used for rural and urban areas respectively to convert nominal consumption expenditure and wages into real term.
 - ▶ We make an attempt to find out how the relationship between log (PCE) and household characteristics (determinants) differs between rural and urban areas.
 - ▶ We used B-O decomposition to decompose changes in the mean of log RPCE due to household characteristics between rural and urban areas.

Consumption Function of Rural and Urban Households (2004-05)

VARIABLES	Rural	Urban	Total
	Ln (Real per capita consumption)	Ln (Real per capita consumption)	Ln (Real per capita consumption)
Literate & up to primary	0.107*** (0.00576)	0.125*** (0.00935)	0.109*** (0.00506)
Secondary	0.232*** (0.00621)	0.273*** (0.00910)	0.247*** (0.00523)
Under graduation	0.351*** (0.00890)	0.424*** (0.0111)	0.390*** (0.00693)
Graduate & above	0.457*** (0.00933)	0.531*** (0.0110)	0.518*** (0.00699)
Ln (age)	0.0656*** (0.00647)	0.0160* (0.00922)	0.0478*** (0.00542)
Ln (household size)	-0.182*** (0.00481)	-0.116*** (0.00580)	-0.153*** (0.00374)
Female	0.0732*** (0.00617)	0.137*** (0.00827)	0.105*** (0.00506)
Industry	0.0948*** (0.00553)	0.0897*** (0.0140)	0.0625*** (0.00512)
Service	0.161*** (0.00583)	0.143*** (0.0139)	0.111*** (0.00515)
Ln (Real per capita wage)	0.208*** (0.00315)	0.364*** (0.00382)	0.287*** (0.00246)
Urban			0.0683*** (0.00375)
Constant	4.511*** (0.0246)	4.298*** (0.0349)	4.460*** (0.0202)
Observations	33,485	25,015	58,500
R-squared	0.461	0.605	0.559

Results

- ▶ We find that higher the level education of the principal earner of the household, higher is the per capita consumption of the household.
- ▶ The impact of education on per capital household consumption expenditure is found relatively higher in the case of urban than rural areas, may be due to availability of better quality of education and job opportunities.
- ▶ The impact of the service sector on consumption expenditure is higher than other two sectors.
- ▶ The impact of occupation of the principal earner of the household on household's consumption expenditure in urban shows is slightly lower than rural areas. The reasons could be due to firstly, wage rates of industry and services sectors in rural areas have increased at a higher rate than their urban counterparts and secondly, households in urban areas usually spend less and save more than their rural counterparts.
- ▶ The impact of wage earnings (in real term) of the household on real per capital consumption expenditure of the household is found positive and significant.
- ▶ The impact of per capita real wage rate on per capita real consumption expenditure is however found relatively more in case of the urban households
- ▶ We found that while the sign of coefficients of age of the principal earner and female headed households are positive, the sign of coefficient of household size is negative.
- ▶ Our results support the findings of Michael (1973); Sen, (1999); Grossman (2005) and Hogan and Berning (2012) that education has positive impact on consumption.
- ▶ Urban dummy shows positive sign and statistically significant suggesting that per capita household consumption expenditure of urban households is significantly higher than the per capita consumption expenditure of rural households. The reasons could be due to that urban household's per capita wage earning in absolute value is relatively higher than its rural counterparts.

Blinder-Oaxaca Decomposition of Consumption Functions of Rural and Urban Households (2004-05)

VARIABLES	(1) Differential	(2) Endowments	(3) Coefficients	(4) Interaction
Total		0.225*** (0.00352)	0.0308*** (0.00577)	0.0573*** (0.00504)
Prediction of urban	5.426*** (0.00433)			
Prediction of rural	5.113*** (0.00281)			
Difference	0.313*** (0.00516)			
Observations	58,500	58,500	58,500	58,500

- ▶ The results show the mean of per capita consumption is 5.426 for urban and 5.113 for rural, yielding a consumption gap of 0.313 which is statistically significant at 1% level. It means household consumption in urban was relatively more than rural areas during 2004-05
- ▶ It is found that the predicted mean difference of rural and urban consumption 0.313 is largely (71.9%) attributed to endowment effect, followed by 18.3% by interaction effect coefficients and the remaining of 9.8% by coefficient effect.
- ▶ The mean difference of endowment effect between urban and rural would be minimized if both of them have same characteristics.

Consumption Function of Rural and Urban Households (2011-12)

VARIABLES	Rural	Urban	Total
	Ln (Real per capita consumption)	Ln (Real per capita consumption)	Ln (Real per capita consumption)
Literate & up to primary	0.111*** (0.00733)	0.114*** (0.0117)	0.109*** (0.00638)
Secondary	0.225*** (0.00733)	0.261*** (0.0108)	0.237*** (0.00618)
Under graduation	0.333*** (0.0100)	0.375*** (0.0126)	0.355*** (0.00783)
Graduate & above	0.407*** (0.0103)	0.481*** (0.0121)	0.465*** (0.00772)
Ln (age)	0.115*** (0.00816)	0.0423*** (0.0104)	0.0863*** (0.00651)
Ln (household size)	-0.246*** (0.00596)	-0.181*** (0.00642)	-0.217*** (0.00440)
Female	0.104*** (0.00729)	0.0801*** (0.00919)	0.103*** (0.00581)
Industry	0.0709*** (0.00656)	0.0793*** (0.0171)	0.0502*** (0.00626)
Service	0.169*** (0.00750)	0.140*** (0.0171)	0.114*** (0.00662)
Ln (Real per capita wage)	0.196*** (0.00364)	0.347*** (0.00405)	0.275*** (0.00273)
Urban			0.129*** (0.00412)
Constant	4.549*** (0.0321)	4.455*** (0.0416)	4.575*** (0.0252)
Observations	28,423	23,093	51,516
R-squared	0.406	0.561	0.525

Results

- ▶ Similar to the findings for the period 2004-05, the results for 2011-12 indicate higher is the level of education, better is the prospects of household consumption.
- ▶ We found that employment in services in relative to agriculture sector tends to have more positive impact on household consumption than the industry sector.
- ▶ We found that the impact of wage rate of urban households on consumption is two times more than rural ones. Similarly, the impact of education on household consumption is slightly higher in case of urban than rural.
- ▶ Urban dummy, which is included in the aggregate consumption function to measure the differences between rural and urban household consumption shows positive sign and statistically significant.

Blinder-Oaxaca Decomposition of Consumption Functions of Rural and Urban Households (2011-12)

VARIABLES	(1) Differential	(2) Endowments	(3) Coefficients	(4) Interaction
Total		0.208*** (0.00364)	0.0967*** (0.00523)	0.0572*** (0.00406)
Prediction of urban	5.700*** (0.00451)			
Prediction of rural	5.338*** (0.00315)			
Difference	0.362*** (0.00550)			
Observations	51,516	51,516	51,516	51,516

- ▶ The B-O decomposition test shows the rural-urban consumption gap is 0.362 and it is statistically significant at 1% significance level
- ▶ the result provides a vital point here is that the rural-urban consumption gap has slightly widened over the period from 0.313 during 2004-05 to 0.362 during 2011-12.
- ▶ it is attributed largely to endowment effects (57.5%) followed by coefficients effects (26.7%) and interaction effect (15.8%).

Conclusion

- ▶ The data shows that the occupation pattern of rural households is changing fast. Although agriculture remains the primary source of economic activities for rural households, people have chosen non-agriculture activities like construction and low skilled manufacturing jobs as secondary activities.
- ▶ Wage rates in rural areas have also witnessed faster increase than urban areas.
- ▶ The impact of education, occupation and wages on per capita consumption expenditure is higher in case of urban than rural.
- ▶ The results of B-O decomposition test based on 2004-05 data indicate the mean differences of consumption expenditure of urban and rural households is 0.313 and it is largely explained by differences in endowment effect (71.9%).
- ▶ A similar test was also carried on 2011-12 data and found that the mean difference of consumption expenditure of urban and rural households is 0.362. In this case also the urban-rural gap is explained largely by differences in endowment effects (57.5%).
- ▶ But the differences of endowment effect show that it has considerably declined between 2004-05 and 2011-12 suggesting that the sign of convergence of consumption expenditure between the rural and urban households has emerged in India.

References

- ▶ Ben Jann (2008), The Blinder–Oaxaca decomposition for linear regression models, *The Stata Journal*, 8, Number 4, pp. 453–479.
- ▶ PC Parida, KC Pradhan (2018), “Structural Transformation and its impact on Rural-Urban Households in India” in NILERD (eds.) *Reflecting on India’s Development*, Springer Publication, PP 67-89.