Applied Econometrics (MSc.) Lecture 2: The OLS Model Applications

Yonas Alem

Department of Economics
University of Gothenburg

December 3, 2014

Application I: Hamermesh & Parker

- Reference: Hamermesh, D.S., & Parker, A. (2005). Beauty in the classroom: Instructors' pulchritude and putative pedagogical productivity. *Economics of Education Review*, 24: 369-376.

Application I: Hamermesh & Parker
Intro.

- The literature in social psychology (see, Hatfield & Sprecher, 1986) has examined the impact of human beauty on a variety of non-economic outcomes
- Economists have also attempted to study the impact of beauty on labor market outcomes
- Evidence shows that academic administrators consider teaching quality when setting salaries
- Higher teaching quality (as proved by higher student evaluations) implies, *ceteris paribus*, increase in salary

Application I: Hamermesh & Parker
Intro. Cont.

- **Question:** Does beauty trigger students evaluate some teachers more favorably than others, so that beauty earns them higher salary?
- The authors answer this question using data collected from undergrad students at the University of Texas, Austin and OLS estimation technique
- Student evaluations of instructors takes place during the last 3 weeks of a 16 week semester.
- Evaluation is administered by a representative student while the instructor is absent from the classroom

Application I: Hamermesh & Parker

Data

- The rating forms include: “Overall, this instructor was very unsatisfactory (1), unsatisfactory(2), satisfactory(3), very good(4), or excellent(5)” and “Overall, this course was very unsatisfactory, unsatisfactory ....”
- The authors concentrate on the second question because it seems more germane to inferring the teacher’s educational productivity, and because, in any event, there is a high correlation between the responses to the two questions $(r = 0.95)$
- Pictures of instructors and their teaching evaluations for the years 2000-2002 were used

Application I: Hamermesh & Parker
Variables

- Course evaluation $(1 = VeryUnsatisfactory)$ to $(5 = Excellent)$
- Beauty rating of instructor: A panel of six students assessed physical appearance of instructors
- Female $=$ 1, a dummy variable for females, and 0 otherwise
- Minority $=$1, a dummy variable if instructor is a non-white and 0 otherwise
- Nonnative English $=$ 1, a dummy variable if instructors first language is not English, 0 otherwise
- Introduction $=$ 1 if course is introductory, 0 otherwise
- One credit course $=$1 if course is a single-credit elective, 0 otherwise
- Age, instructor's age

Application I: Hamermesh & Parker

Simple Linear Regression

- Consider the regression

$$CE = \beta_0 + \beta_1 Beauty + u_i \tag{1}$$

- Where u_i captures all other factors except beauty that affect course evaluation
- Assuming [A1][A5], we can estimate the impact of beauty on course evaluation using simple linear regression

Table: Determinants of course evaluation- Simple Linear Regression Results

Variable	Coef.	SE	t-value	P-value
Beauty	0.133	0.032	4.130	0.000
Intercept	3.998	0.025	157.730	0.000
Obs. = 463				
F(1, 461) = 17.08				
Prob > F = 0.000				
R-sq = 0.036				
Adj R-sq = 0.033				

- How do you interpret the coefficient of beauty (i.e., β_1? Do you see any problem with this regression in addressing the research question?

Application I: Hamermesh & Parker
Simple Linear Regression Cont.

- We have enough evidence to believe that the above regression is potentially biased because of omission of variables
- Undertake the Ramesey RESET (Regression Equation Specification Error Test) test. According to this test, the regression model under consideration is mis-specified if non-linear combinations of the explanatory variables have any power in explaining the response variable.
- Consider the regression model:

$$\hat{y} = E\{y|x\} = \beta x \tag{2}$$

- The command tests if $(\beta x^2), (\beta x^3), ..., (\beta x^k)$ has any power in explaining y

Application I: Hamermesh & Parker

Simple Linear Regression Cont.

- To check for this, one could run the following regression

$$y = \alpha x + \theta_1 \hat{y}^2, ..., \theta_{k-1} \hat{y}^k + u \tag{3}$$

- Test $H_0 = \theta_1$ through $\theta_{k-1} = 0$ using the *F-test* (this is equivalent to saying the model has no omitted variables

Table: Ramsey RESET test

F(3,458)	= 5.960
Prob > F	= 0.001

- As you can see the H_0 is rejected!

Application I: Hamermesh & Parker

Multiple Linear Regression Cont.

- Would the problem of omission variable bias disappear if we add one more important variable, say age of instructors in the regression?

$$CR = \beta_0 + \beta_1 Beauty + \beta_2 Age + u \tag{4}$$

- It is reasonable to assume apriori that beauty and age are likely to be negatively correlated. Consequently, β_1 would be downward biased!

Application I: Hamermesh & Parker

Multiple Linear Regression Cont.

Table: Determinants of course evaluation- Multiple Linear Regression Results

Variable	Coef.	SE	t-value	P-value
Beauty	0.133	0.032	4.130	0.000
Age	0.0003	0.003	0.110	0.916
Intercept	3.984	0.133	29.790	0.000
Obs. = 463				
F(2, 460) = 8.530				
Prob > F = 0.000				
Adj R-sq = 0.032				

Table: Ramsey RESET test

F(3,457) = 5.110
Prob > F = 0.002

Application I: Hamermesh & Parker

Multiple Linear Regression Cont.

- Consider the general OLS model of the determinants of course evaluation

$$CE = \beta_0 + \beta_1 Beauty + \beta_2 Female + \beta_3 Minority + \beta_4 Non - Native$$

Table: Determinants of course evaluation- Multiple Linear Regression Results

Variable	Coef.	SE	t-value	P-value
Beauty	0.159	0.031	4.980	0.000
Female	-0.183	0.051	-3.590	0.000
Minority	-0.169	0.076	-2.220	0.027
Non-native	-0.243	0.107	-2.280	0.023
Age	-0.002	0.003	-0.730	0.464
Intro	0.007	0.055	0.150	0.885
One-credit	0.633	0.111	5.68	0.000
Intercept	4.169	0.141	29.400	0.000
Adj R-sq = 0.143				

Application I: Hamermesh & Parker
Multiple Linear Regression Cont.

Table: Ramsey RESET test

F(3,452) = 1.210
Prob > F = 0.305

- You can see that the model now passed the Ramsey RESET test!

Application I: Hamermesh & Parker

Multiple Linear Regression Cont.

- Three other problems that one should worry about when running OLS on a cross-sectional data are Endogeneity, Heteroskedasticity, and Multicollinearity
- **Endogeneity:** A serious problem in affecting consistency of the parameter estimates! We will address this in Lecture 4
- **Heteroskedasticity:** A less serious problem (affects efficiency of parameters). Occurs when the assumption of constant variance is violated. It is easily detectable and remedies are available (E.g. WLS as Hamermesh & Parker).
- **Multicollinearity:** When two or more variables are correlated. Detectable (easily using paire-wise correlations between the variables)

Application II: Deressa & Hassan, 2009

Intro.

- Source: Deressa, T.T., & Hassan R.M. 2009. Economic Impact of Climate Change on Crop Production in Ethiopia: Evidence from Cross-Section Measures. Journal of African Economies, 4: 529-554.

Application II: Deressa & Hassan, 2009
Intro.

- Climate change (CC) affects agriculture more than any other sector
- As CC is global, concerns about its impact on agriculture in developing countries have been increasing
- CC would particularly be damaging to countries of Africa being dependent on rain-fed agriculture and already food insecure
- Understanding the impact and the role of adaptations is therefore important
- **Objective:** Assess the economic impact of CC on Ethiopian farmers, using the **Ricardian** Approach and inform policy-makers on proper adaptation options

Application II: Deressa & Hassan, 2009
Eth. Agr.

- Agr is the most important sector in the Ethiopian economy.
- 85% of the population, 50% of GDP, 88% of export earnings
- Is also major source of food for the population (about 80 million) and hence the prime contributing sector to food security
- Expected to bring growth in the overall economy by generating surplus capital to speed-up overall socio-economic development
- Unfortunately agr. is characterized by small-holder farming with low-input use, low-technology and low output, almost entirely rain-dependent

Application II: Deressa & Hassan, 2009
Eth. Agr. Cont.

- The government has given top priority to this sector and took major steps to bring growth in productivity
- However, various problems are holding this back
  - Declining farm size, high population growth, land degradation (due to inappropriate use of land including cultivation of steep slopes, over cultivation and overgrazing)
  - Tenure insecurity, weak agricultural research and extension services, lack of marketing, inadequate transport network, low level of modern input use
  - However the major causes are drought, which often causes famine, and floods.
  - This ultimately made the country food-aid-dependent

Application II: Deressa & Hassan, 2009

Measuring the economic impacts of CC- alternative approaches

- Two main types of economic impact assessment models (Economy-wide or general equilibrium & Partial Equilibrium models)
- **Economy-wide models:** analytical models **(e.g., Computable General Equilibrium Models or CGE)** which look at the economy as a complete system of interdependent components (industries, factors of production, institutions, and the rest of the world).
  - **Key challenges:** difficulties in model selection, parameter specification and functional forms, data consistency or calibration problems, absence of statistical tests for the model specification
- **Partial Equilibrium Models:** based on analysis of part of the overall economy such as a single market (single commodity) or subsets of markets or sectors
  - Are of three types: Crop suitability, production function & Ricardian approaches

Application II: Deressa & Hassan, 2009

Measuring the economic impacts of CC- alternative approaches Cont.

- **Crop suitability approach (Agro-ecological zoning approach):** used to assess the suitability of various land and biophysical attributes for crop production
- Crop characteristics, existing technology and soil and climate factors are used as determinants of suitability for crop production
- The model enables identification and distribution of potential crop-producing lands.
- As the model includes climate as one determinant of suitability of agr. land, for crop production, it can be used to predict the impact of changing climatic variables on potential agr. output and cropping patterns

Application II: Deressa & Hassan, 2009

Measuring the economic impacts of CC- alternative approaches Cont.

- **Production function approach:** based on empirical or experimental production function that measures the relationship between agr. production and climate change
- A production function which includes environmental variables (temperature and rainfall) is estimated
- Change in yield following change in the climate variables is easily estimated
- Predicts the impact of climatic variables more dependably but does not take adaptation into account

Application II: Deressa & Hassan, 2009

Measuring the economic impacts of CC- alternative approaches Cont.

- **Ricardian Approach**: analyzes a cross-section of farmers under different climatic conditions and examines the relationship between the value of land or net revenue and agro-climatic factors
- Incorporates private adaptations and fully incorporates their costs and benefits (the most important advantage)
- The relevant dependent variable is net revenue of land value (capitalized net revenues)
- Is also cost effective (as it uses secondary data collected from cross-sectional sites)

Application II: Deressa & Hassan, 2009

Ricardian approach

- Pioneered by Mendelsohn et al. 1994 based on David Ricardo's (1817) observation that land values would reflect land productivity at a site under perfect competition.
- Climate affects crop productivity and hence, land values! Easy to understand the marginal impact of a change in climatic variables

$$R = \sum P_i Q_i(X, F, G, Z) - \sum P_X X \tag{6}$$

- Where: $R=$ net revenue per ha; P_i, market price of crop i; Q_i, output of crop i; X, vector of purchased inputs; F, vector of climate variables; G, a set of economic variables; Z, a set of soil variables; and P_X is a vector of input prices.

Application II: Deressa & Hassan, 2009

Ricardian approach Cont.

- Econometric Model (a quadratic specification)

$$R_i = \beta_0 + \beta_1 F + \beta_2 F^2 + \beta_3 G + \beta_4 Z + u_i \qquad (7)$$

- Eq. (7) is not a non-linear model. The model is linear in the βs with quadratic components of the explanatory variables. So one can use OLS.

Application II: Deressa & Hassan, 2009
Data

- A sample of 1000 farmers representing different agro-ecological zones of the country
- Data was collected in 2003/04
- The production features of 95% of the Ethiopian farmers is captured
- Temperature and precipitation data: satelite data provided by the US Department of Defense & the African Rainfall and Temperature Evaluation System (ARTES)
- Soil Data: from Food and Agricultural Organization (FAO). Includes slope and texture
- Hydrological data (flow and run-off) were obtained from the University of Colorado

Application II: Deressa & Hassan, 2009
Results cont.

- Net revenues were regressed on climatic and other control variables (See table 5)
- You can see that most of the climatic variables are significant determinants of agricultural revenue
- So are economic variables
- What would be the effect of an infinitesimal change in temperature and rainfall on Ethiopian farming?

Application II: Deressa & Hassan, 2009

Results

Table 5: *Regression Coefficients of Climatic and Control Variable over Net Revenue per Hectare*

Variable	Coefficient
Winter temperature	384.48
Winter temperature squared	−35.00
Spring temperature	−1740.69*
Spring temperature squared	49.40**
Summer temperature	−4495.21**
Summer temperature squared	84.85*
Fall temperature	6743.39***
Fall temperature squared	−133.40**
Winter precipitation	−1148.63***
Winter precipitation squared	16.11***
Spring precipitation	656.62***
Spring precipitation squared	−2.98***
Summer precipitation	112.30***
Summer precipitation squared	−0.48***
Fall precipitation	−525.18***
Fall precipitation squared	3.06***
Livestock ownership	139.30
Level of education of household head	4.32
Distance of input markets	−1.15
Size of household	−109.42***
Nitosols	659.04
Lithosols	7619.68*
Constant	−384.70
N	550.00
R^2	0.30
F	10.38

*significant at 10%; **significant at 5%; ***significant at 1%.

Application II: Deressa & Hassan, 2009
Results Cont.

- The Marginal Impact of Climate variable (f_i) on net revenue evaluated at the mean:

$$E\frac{dR}{df_i} = \beta_{1,i} + 2 * \beta_{2,i} * E[f_i] \tag{8}$$

- A marginal (one degree) increase in temp during the winter and summer seasons reduces the net revenue per ha by US$997.85 and US$1277.60 respectively)!
- This is a big impact for a poor country like Ethiopia!
- During the fall, a higher temperature is beneficial for harvesting, crops have finished their growth process by fall and higher temp. quickly dries them up and facilitates harvesting.

Application II: Deressa & Hassan, 2009

Results Cont.

Impact of forecasted climate scenarios

- What would be the impact of increasing temperature and precipitation in the future?
- The authors used the coefficients from their Ricardian regression and predicted the impact of future changes in climatic variables
- They show significant negative impacts in the future (see table 9 of their paper)

Application II: Deressa & Hassan, 2009

Results Cont.

Conclusions and Policy Implications

- Increasing temperature during winter and summer reduces net revenue per ha.
- Increasing precipitation during winter significantly reduces net revenue
- Climate change is expected to reduce agricultural yield significantly in the future too!
- Adaptation measures that help to counteract the harmful impact of CC are urgently needed!
- **Possible Extensions:** what is the impact of CC on other sub-sectors such as livestock? What is the magnitude of adaptation in reducing the impact of climate change? Which adaptation measures are more effective?