# ECON 6120I Topics in Empirical Industrial Organization

*Kohei Kawaguchi*

*Last updated: 2019-02-19*

# Contents

# Chapter 1

# Syllabus

## 1.1 Instructor Information

- Instructor:
    - Name: Kohei Kawaguchi.
    - Office: LSK6070, Monday 11:00-12:00.
- All questions related to the class have to be publicly asked on the discussion board of canvas rather than being privately asked in e-mail. The instructor usually does not reply in the evening, weekends, and holidays.

## 1.2 General Information

### 1.2.1 Class Time

- Date: Monday.
- Time: 13:30-17:20.
- Venue: CYTG001.

### 1.2.2 Description

- This is a PhD-level course for empirical industrial organization. This course covers various econometric methods used in industrial organization that is often referred to as the structural estimation approach. These methods have been gradually developed since 1980s in parallel with the modernization of industrial organization based on the game theory and now widely applied in antitrust policy, business strategy, and neighboring fields such as labor economics and international economics.

- This course presumes a good understanding of PhD-level microeconomics and microeconometrics. Participants are expected to understand at least UG-level industrial organization. This course requires participants to write programs mostly in R and sometimes in C++ to implement various econometric methods. In particular, all assignments will involve such a non-trivial programming task. Even though the understanding of these programming languages is not a prerequisite, a sharp learning curve will be required. Some experience in other programming languages will help. Audit without a credit is not admitted for students.

### 1.2.3   Expectation and Goals

- The goal of this course is to learn and practice econometric methods for empirical industrial organization. The lecture covers the econometric methods that have been developed between 80s and 00s to estimate primitive parameters governing imperfect competition among firms, such as production and cost function estimation, demand function estimation, merger simulation, entry and exit analysis, and dynamic decision models. The lecture also covers various new methods to recover model primitives in certain mechanisms such as auction, matching, network, and bargaining. The emphasis is put on the former group of methods, because they are the basis for other new methods. Participants will not only understand the theoretical background of the methods but also become able to implement these methods from scratches by writing their own programs. I will briefly discuss the latter class of new methods through reading recent papers. The participants will become able to understand and use these new methods.

## 1.3   Required Environment

- Participants should bring their laptop to the class. We have enough extension codes for students. The laptop should have sufficient RAM (at least $\geq$ 8GB, $\geq$ 16GB is recommended) and CPU power (at least Core i5 grade, Core i7 grade is recommended). Participants are fully responsible for their hardware issues. Operating System can be arbitrary. The instructor mainly uses OSX High Siera with iMac (Retina 5K, 27-inch, Late 2015) and Macbook Pro (Retina, 15-inch, Early 2017).
- Please install the following software by the first lecture. Technical issues related to the installment should be resolved by yourself, because it depends on your local environment. If you had an error, copy and paste the error message on a search engine, and find a solution. This solves 99.9% of the problems.
    - R: https://www.r-project.org/
    - RStudio: https://www.rstudio.com/
    - LaTeX:
        * MixTex https://miktex.org/
        * TeXLive https://www.tug.org/texlive/
        * MacTeX http://www.tug.org/mactex/

## 1.4   Evaluation

- Assignments (80): In total 8 homework are assigned. Each homework involves the implementation of the methods covered in the class. Each homework has 10 points. The working hour for each homework will be around 10-20 hours.
- Participation (10): Every time a participant asks a question in the class, after the class, during the office hour, or in the canvas. the participant gets one point, up to 10 points. The participant who asked the question writes the name, ID number, his/her question, and my answer in a discussion board on the course website to claim a point.
- Referee report (10): Toward the end of the semester, a paper in industrial organization is randomly assigned to each participant. Each participant writes a critical referee report of the assigned paper in A4 2 pages that consists of the summary, contribution, strong and weak points of the paper.
- Grading is based on the absolute scores: A+ with more than 80 points, A with more than 70 points, A- with more than 60 points, B+ with more than 50 points, B with more than 40 points, B- with more than 30 points and C otherwise.

## 1.5  Academic Integrity

Without academic integrity, there is no serious learning. Thus you are expected to hold the highest standard of academic integrity in the course. You are encouraged to study and do homework in groups. However, no cheating, plagiarism will be tolerated. Anyone caught cheating, plagiarism will fail the course. Please make sure adhere to the HKUST Academic Honor Code at all time (see http://www.ust.hk/vpaao/integrity/).

## 1.6  Schedule

- Introduction to structural estimation, R and RStudio
- Production function estimation I
- Production function estimation II
- Demand function estimation I
- Demand function estimation II
- Merger Analysis
- Entry and Exit
- Single-Agent Dynamics I
- Single-Agent Dynamics II, I change date due to my business trip
- Dynamic Game I
- Dynamic Game II
- Auction
- Other Mechanisms

## 1.7  Course Materials

### 1.7.1  R and RStudio

- Grolemund, G., 2014, Hands-On Programming with R, O'Reilly.
  - Free online version is available: https://rstudio-education.github.io/hopr/.
- Wickham, H., & Grolemund, G., 2017, R for Data Science, O'Reilly.
  - Free online version is available: https://r4ds.had.co.nz/.
- Boswell, D., & Foucher, T., 2011, The Art of Readable Code: Simple and Practical Techniques for Writing Better Code, O'Reilly.

### 1.7.2  Handbook Chapters

- Ackerberg, D., Benkard, C., Berry, S., & Pakes, A. (2007). "Econometric tools for analyzing market outcomes". Handbook of econometrics, 6, 4171-4276.
- Athey, S., & Haile, P. A. (2007). "Nonparametric approaches to auctions". Handbook of Econometrics, 6, 3847-3965.
- Berry, S., & Reiss, P. (2007). "Empirical models of entry and market structure". Handbook of Industrial Organization, 3, 1845-1886.
- Bresnahan, T. F. (1989). "Empirical studies of industries with market power". Handbook of Industrial Organization, 2, 1011-1057.
- Hendricks, K., & Porter, R. H. (2007). "An empirical perspective on auctions". Handbook of Industrial Organization, 3, 2073-2143.
- Matzkin, R. L. (2007). "Nonparametric identification". Handbook of Econometrics, 6, 5307-5368.
- Newey, W. K., & McFadden, D. (1994). "Large sample estimation and hypothesis testing". Handbook of Econometrics, 4, 2111-2245.

- Reiss, P. C., & Wolak, F. A. (2007). "Structural econometric modeling: Rationales and examples from industrial organization". Handbook of Econometrics, 6, 4277-4415.

### 1.7.3  Books

- Train, K. E. (2009). Discrete Choice Methods with Simulation, Cambridge university press.
- Davis, P., & Garces, E. (2010). Quantitative Techniques for Competition and Antitrust Analysis, Princeton University Press.
- Tirole, J. (1988). The Theory of Industrial Organization, The MIT Press.

### 1.7.4  Papers

- The list of important papers are occasionally given during the course.

# Chapter 2

# Introduction

## 2.1 Structural Estimation and Counterfactual Analysis

### 2.1.1 Example

- Igami [2017] "Estimating the Innovator's Dilemma: Structural Analysis of Creative Destruction in the Hard Disk Drive Industry, 1981-1998".

- **Question**:

- Does "Innovator's Dilemma" [Christensen, 1997] or the delay of innovation among incumbents exist?

- Christensen argued that old winners tend to lag behind entrants even when introducing a new technology is not too difficult, with a case study of the HDD industry.

    - Apple's smartphone vs. Nokia's feature phones.
    - Amazon vs. Borders.
    - Kodak's digital camera.

- If it exists, what is the reason for that?

- How do we empirically answer this question?

- **Hypotheses**:

- Identify potentially competing hypotheses to explain the phenomenon.

    1. Cannibalization: Because of cannibalization, the benefits of introducing a new product are smaller for incumbents than for entrants.
    2. Different costs: The incumbents may have higher costs for innovation due to the organizational inertia, but at the same time they may have some cost advantage due to accumulated R&D and better financial access.
    3. Preemption: The incumbents have additional incentive for innovation to preempt potential rivals.
    4. Institutional environment: The impacts of the three components differ across different institutional contexts such as the rules governing patents and market size.

- Casual empiricists pick up their favorite factors to make up a story.

- Serious empiricists should try to separate the contributions of each factor from data.

- To do so, the author develops an economic model that explicitly incorporates the above mentioned factors, while keeping the model parameters flexible enough to let the data tell the sign and size of the effects of each factor on innovation.

## Cumulative Number of Innovators



## Innovators as Fractions of All Firms including Fringe



FIG. 1.—The incumbent-entrant innovation gap. The top panel plots the timing of the initial shipment of 3.5-inch HDDs separately for incumbents (i.e., firms already active in the 5.25-inch generation) and entrants (i.e., firms that appeared for the first time as producers of 3.5-inch HDDs). The bottom panel expresses similar numbers in terms of the fractions of all firms that could have innovated. See the text for details.

Figure 2.1: Figure 1 of Igam (2017)

- **Economic model**:

- The time is discrete with finite horizon $t = 1, \cdots, T$.

- In each year, there is a finite number of firms indexed by $i$.

- Each firm is in one of the technological states:

$$s_{it} \in \{\text{old only, both, new only, potential entrant}\}, \tag{2.1}$$

  where the first two states are for incumbents (stick to the old technology or start using the new technology) and the last two states are for actual and potential entrants (enter with the new technology or stay outside the market).

- In each year:
    - Pre-innovation incumbent ($s_{it} = $ old): exit or innovate by paying a sunk cost $\kappa^{inc}$ (to be $s_{i,t+1} = $ both).
    - Post-innovation incumbent ($s_{it} = $ both): exit or stay to be both.
    - Potential entrant ($s_{it} = $ potential entrant): give up entry or enter with the new technology by paying a sunk cost $\kappa^{net}$ (to be $s_{i,t+1} = $ new).
    - Actual entrant ($s_{it} = $ new): exit or stay to be new.

- Given the industry state $s_t = \{s_{it}\}_i$, the product market competition opens and the profit of firm $i$, $\pi_t(s_{it}, s_{-it})$, is realized for each active firm.

- As the product market competition closes:
    - Pre-innovation incumbents draw private cost shocks and make decisions: $a_t^{pre}$.
    - Observing this, post-innovation incumbents draw private cost shocks and make decisions: $a_t^{post}$.
    - Observing this, actual entrants draw private cost shocks and make decisions: $a_t^{act}$.
    - Observing this, potential entrants draw private cost shocks and make decisions: $a_t^{pot}$.

- This is a dynamic game. The equilibrium is defined by the concept of **Markov-perfect equilibrium** [Maskin and Tirole, 1988].

- The representation of the competing theories in the model:
    - The existence of cannibalization is represented by the assumption that an incumbent maximizes the joint profits of old and new technology products.
    - The size of cannibalization is captured by the shape of profit function.
    - The difference in the cost of innovation is captured by the difference in the sunk costs of innovation.
    - The preemptive incentive for incumbents are embodied in the dynamic optimization problem for each incumbent.

- **Econometric model**:

- The author then turns the economic model into an econometric model.

- This amounts to specify which part of the economic model is observed/known and which part is unobserved/unknown.

- The author collects the data set of the HDD industry during 1977-99.

- Based on the data, the author specify the identities of active firms and their products and the technologies embodied in the products in each year to code their **state variables**.

- Moreover, by tracking the change in the state, the author code their **action variables**.

- Thus, the state and action variables, $s_t$ and $a_t$. These are the **observables**.

- The author does not observe:

- Profit function $\pi_t(\cdot)$.
- Sunk cost of innovation for pre-innovation incumbents $\kappa^{inc}$.
- Sunk cost of entry for potential entrants $\kappa^{net}$.
- Private cost shocks.

- These are the **unobservables**.

- Among the unobservables, the profit function and sunk costs are the **parameter of interets** and the private cost shocks are **nuissance parameters** in the sense only the knowledge about the distribution of the latter is demanded.

- **Identification**:

- Can we infer the unobservables from the observables and the restrictions on the distribution of observable by the economic theory?

- The profit function is identified from estimating the demand function for each firm's product, and estimating the cost function for each firm from using their price setting behavior.

- The sunk costs of innovation are identified from the conditional probability of innovation across various states. If the cost is low, the probability should be high.

- **Estimation**:

- The identification established that in principle we can uncover the parameters of interests from observables under the restrictions of economic theory.

- Finally, we apply a statistical method to the econometric model and infer the parameters of interest.

- **Counterfactual analysis**:

- If we can uncover the parameters of interest, we can conduct **comparative statics**: study the change in the endogenous variables when the exogenous variables including the model parameters are set different. In the current framework, this exercise is often called the **counterfactual analysis**.

- What if there was no cannibalization?:

    - An incumbents separately maximizes the profit from old technology and new technology instead of jointly maximizing the profits. Solve the model under this new assumption everything else being equal.
    - Free of cannibalization concerns, 8.95 incumbents start producing new HDDs in the first 10 years, compared with 6.30 in the baseline.
    - The cumulative numbers of innovators among incumbents and entrants differ only by 2.8 compared with 6.45 in the baseline.
    - Thus cannibalization can explain a significant part of the incumbent-entrant innovation gap.

- What if there was no preemption?:

    - A potential entrant ignores the incumbents' innovations upon making entry decisions.
    - Without the preemptive motives, only 6.02 incumbents would innovate in the first 10 7ears, compared with 6.30 in the baseline.
    - The cumulative incumbent-entrant innovation gap widen to 8.91 compared with 6.45 in the baseline.

- The sunk cost of entry is smaller for incumbents than for entrants in the baseline.

- **Interpretations and policy/managerial implication**:

- Despite the cost advantage and the preemptive motives, the speed of innovation is slower among incumbents due to the strong cannibalization effect.

- Incumbents that attempt to avoid the "innovator's dilemma" should separate the decision makings between old and new sections inside the organization so that it can avoid the concern for cannibalization.

## 2.1.2 Recap

- The structural approach in empirical industrial organization consists of the following components:

1. Research question.
2. Competing hypotheses.
3. Economic model.
4. Econometric model
5. Identification.
6. Data collection.
7. Data cleaning.
8. Estimation.
9. Counterfactual analysis.
10. Coding.
11. Interpretations and policy/managerial implications.

- The goal of this course is to be familiar with the standard methodology to complete this process.
- The methodology covered in this class is mostly developed to analyze the standard framework to dynamic or oligopoly competition.
- The policy implications are centered around competition policies.
- But the basic idea can be extend to different class of situations such as auction, matching, voting, contract, marketing, and so on.
- Note that the depth of the research question and the relevance of the policy/managerial implications are the most important part of the research.
- Focusing on the methodology in this class is to minimize the time to allocate to less important issues and maximize the attention and time to the most valuable part in the future research.
- Given a research question, what kind of data is necessary to answer the question?
- Given data, what kind of research questions can you address? Which question can be credibly answered? Which question can be an over-stretch?
- Given a research question and data, what is the best way to answer the question? What type of problem can you avoid using the method? What is the limitation of your approach? How will you defend the possible referee comments?
- Given a result, what kinds of interpretation can you credibly derive? What kinds of interpretation can be contested by potential opponents? What kinds of contribution can you claim?
- To address these issues is **necessary** to publish a paper and it is **necessary** to be familiar with the methodology to do so.

## 2.1.3 Historical Remark

- The words **reduced-form** and **structural-form** date back to the literature of estimation of simultaneous equations in macroeconomics [Hsiao, 1983].
- Let $y$ be the vector of observed endogenous variables, $x$ be the vector of observed exogenous variables, and $\epsilon$ be the vector of unobserved exogenous variables.
- The equilibrium condition for $y$ on $x$ and $\epsilon$ is often written as:

$$Ay + Bx = \Sigma\epsilon. \tag{2.2}$$

- These equations **implicitly** determine the vector of endogenous variables $y$ .
- If $A$ is invertible, we can solve the equations for $y$ to obtain:

$$y = -A^{-1}Bx + A^{-1}\Sigma\epsilon. \tag{2.3}$$

- These equations **explicitly** determine the vector of endogenous variables $y$.
- Equation (2.2) is the **structural-form** and (2.3) is the **reduced-form**.

- If $y$ and $x$ are observed and $x$ is of full column rank, then $A^{-1}B$ and $A^{-1}\Sigma A^{-1}$ will be estimated by regression for (2.3). But this does not mean that $A, B$ and $\Sigma$ are separately estimated.
- This was the traditional identification problems.
- Thus, reduced-form does not mean either of:
  - Regression analysis;
  - Statistical analysis free from economic assumptions.
- Recent development in this line of literature of identification is found in Matzkin [2007].
- In econometrics, the idea of imposing restrictions from economic theories seems to have been formalized by the work of Manski [1994] and Matzkin [1994].

## 2.2   Setting Up The Environment

- Assume that R, RStudio and LaTex are all installed in the local computer.

### 2.2.1   RStudio Project

- The assignments should be conducted inside a project folder for this course.
- `File > New Project...> New Directory > New Directory > R Package using RcppEigen`.
- Name the directory `ECON6120I` and place in your favorite location.
- You can open this project from the upper right menu of RStudio or by double clicking the `ECON6120I.Rproj` file in the `ECON6120I` directory.
- This navigates you to the root directory of the project.
- In the root directory, make folders named:
  - `assignment`.
  - `input`.
  - `output`.
  - `figuretable`.
- We will store R functions in R folder, C/C++ functions in `src` folder, and data in `input` folder, data generated from the code in `output`, and figures and tables in `figurtable` folder.
- Open `src/Makevars` and erase the content. Then, write: `PKG_CPPFLAGS = -w -std=c++11 -O3`
- Open `src/Makevars.win` and erase the content. Then, write: `PKG_CPPFLAGS = -w -std=c++11`

### 2.2.2   Basic Programming in R

- `File > New File > R Script` to open `Untitled` file.
- `Ctrl (Cmd) + S` to save it with `test.R` in `assignment` folder.
- In the console, type and push enter:

```
1 + 1
```

```
## [1] 2
```
```
100:130
```

```
##  [1] 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116
## [18] 117 118 119 120 121 122 123 124 125 126 127 128 129 130
```

- This is the interactive way of using R functionalities.
- In `test.R`, write:

```
1 + 1
```

- Then, save the file and push `Run`.

- Alternatively, place the mouse over the `1 + 1` line in `test.R` file.

- Then, `Ctrl (Cmd) + Enter` to run the line.
- In this way, we can write procedures in the file and send to the console to run.
- There are functions to conduct basic calculations:

```
1 + 2
```

```
## [1] 3
```
```
2 * 3
```

```
## [1] 6
```
```
4 - 1
```

```
## [1] 3
```
```
6 / 2
```

```
## [1] 3
```
```
2^3
```

```
## [1] 8
```

- We can define objects and assign values to them.

```
a <- 1
a
```

```
## [1] 1
```
```
a + 2
```

```
## [1] 3
```

- In addition to scalar object, we can define a vector by:

```
2:10
```

```
## [1]  2  3  4  5  6  7  8  9 10
```
```
3:20
```

```
##  [1]  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20
```
```
c(2, 3, 5, 9, 10)
```

```
## [1]  2  3  5  9 10
```
```
seq(1, 10, 2)
```

```
## [1] 1 3 5 7 9
```

- `seq` is a function with initial value, end values, and the increment value.
- By typing `seq` in the `help`, we can read the manual page of the function.
- `seq {base}` means that this function is named `seq` and is contained in the library called `base`.
- Some libraries are automatically called when the R is launched, but some are not.
- Some libraries are even not installed.
- We can install a library from a repository called `CRAN`.

```
install.packages("ggplot2")
```

- To use the package, we have to load by:

```
library(ggplot2)
```

- Use `qplot` function in `ggplot2` library to draw a scatter plot.

```r
x <- c(-1, -0.8, -0.6, -0.4, -0.2, 0, 0.2, 0.4, 0.6, 0.7, 1)
y <- x^3
qplot(x, y)
```
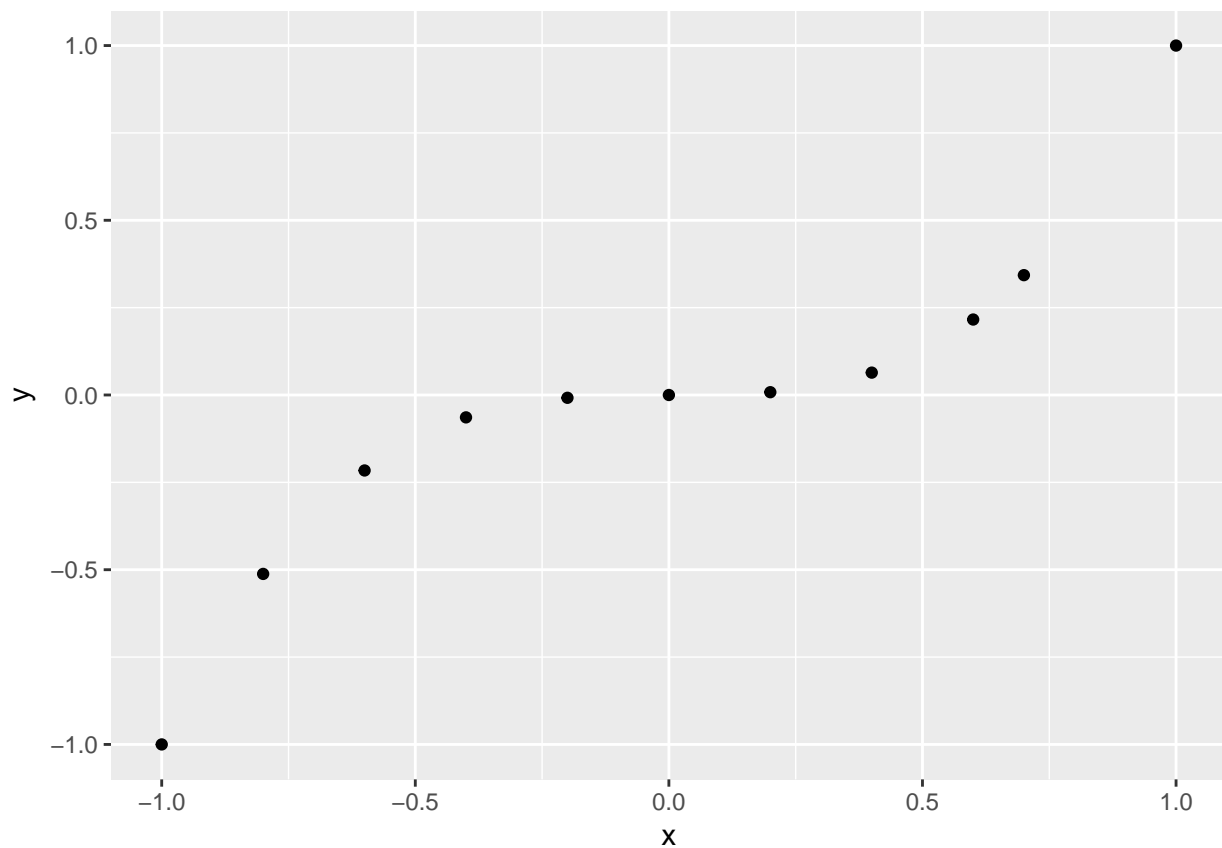


Instead of loading a package by `library`, you can directly call it as:

```r
ggplot2::qplot(x, y)
```
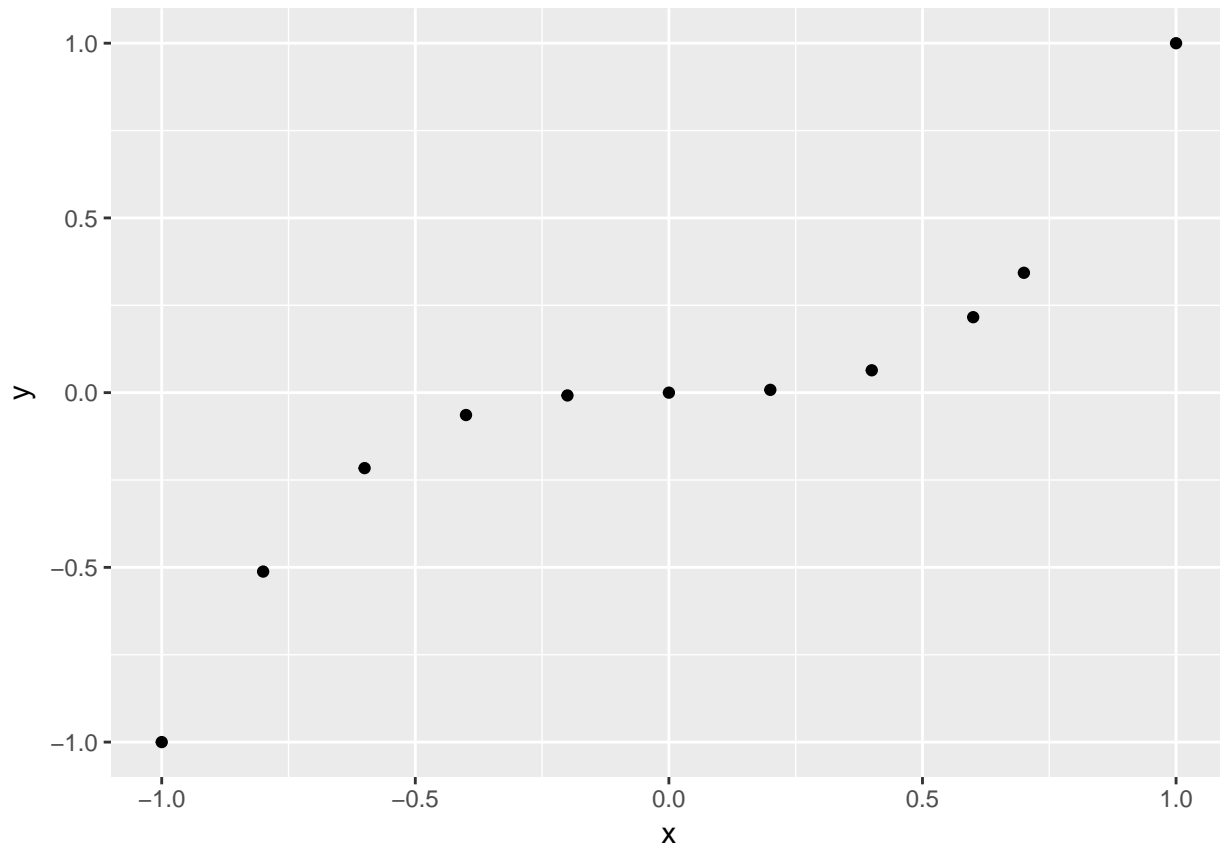
- We can write own functions.

```r
roll <- function(n) {
  die <- 1:6
  dice <- sample(die, size = n, replace = TRUE)
  y <- sum(dice)
  return(y)
}
roll(1)
```

```
## [1] 6
```

```r
roll(2)
```

```
## [1] 9
```

```r
roll(10)
```

```
## [1] 46
```

```r
roll(10)
```

```
## [1] 31
```

- We can set.seed to obtain the same realization of random variables.

```r
set.seed(1)
roll(10)
```

```
## [1] 38
```

```r
set.seed(1)
roll(10)
```

```
## [1] 38
```

- When a variable used in a function is not given as its argument, the function calls the variable in the global environment:

```r
y <- 1
plus_1 <- function(x) {
  return(x + y)
}
plus_1(1)
```

```
## [1] 2
```

```r
plus_1(2)
```

```
## [1] 3
```

- However, you should NOT do this. All variables used in a function should be given as its arguments:

```r
y <- 1
plus_2 <- function(x, y) {
  return(x + y)
}
plus_2(2, 3)
```

```
## [1] 5
```

```r
plus_2(2, 4)
```

```
## [1] 6
```

- The best practice is to use `findGlobals` function in `codetools` to check global variables in a funciton:

```r
library(codetools)
findGlobals(plus_1)
```

```
## [1] "{"      "+"      "return" "y"
```

```r
findGlobals(plus_2)
```

```
## [1] "{"      "+"      "return"
```

- This function returns the list of global variables used in a function. If this returns a global variable other than the system global gariables, you should include it as the argument of the function.

- You can write the functions in the files with executing codes.

- But I recommend you to separate files for writing functions and executing codes.

- `File > New File > R Rcript` and name it as `functions.R` and save to `R` folder.

- Cut the function you wrote and paste it in `functions.R`.

- There are two ways of calling a function in `functions.R` from `test.R`.

- One way is to use `source` function.

```r
source("R/functions.R")
```

- When this line is read, the codes in the file are executed.
- The other way is to bundle functions as a package and load it.

- Choose `Build > Clean and Rebuild`.
- This compiles files in `src` folder and bundle functions in `R` folder and build a package named `ECON6120I`.
- Now, the functions in `R` folder and `src` folder can be used by loading the package by:

```r
library(ECON6120I)
```

- Best practice:
  1. Write functions in the scratch file.
  2. As the functions are tested, move them to `R/functions.R`.
  3. Clean and rebuild and load them as a package.

### 2.2.3  Reproducible Reports using Rmarkdown

- Reporting in empirical studies involves:

1. Writing texts;
2. Writing formulas;
3. Writing and implementing programs;
4. Demonstrating the results with figures and tables.

- Moreover, this has to be done in a **reproducible** manner: Whoever can reproduce the output from the scratch.

- "Whoever" includes yourself in the future. Because the revision process of structural papers is usually lengthy, you often have to remember the content few weeks or few months later. It is inefficient if you cannot recall what you have done.

- We use `Rmarkdown` to achieve this goal.

- This assumes that you have LaTex installed.

- Install package `Rmarkdown`:

```r
install.packages("rmarkdown")
```

- `File > New File > R Markdown... > HTML` with title `Test`.
- Save it in `assignment` folder with name `test.Rmd`.
- From `Knit` tab, choose `Knit to HTML`.
- This outputs the content to html file.
- You can also choose `Knit to PDF` from `Knit` tab to obtain output in pdf file.
- Reports should be knit to pdf to submit.
- But you can use html output while writing a report because html is lighter to compile.
- Refer to the help page for further information.

# Chapter 3

# Production and Cost Function Estimation

## 3.1 Motivations

- Estimating **production and cost functions** of producers is the cornerstone of economic analysis.
- Estimating the functions includes to separate the contribution of observed inputs and the other factors, which is often referred to as the **productivity**.
- "What determines productivity?" [Syverson, 2011]-type research questions naturally follow.
- The methods covered in this chapter are widely used across different fields.
- Some of them are variants from the standard methods.

### 3.1.1 IO

- Olley and Pakes [1996]:
  - How much did the deregulation in the U.S. telecommunication industry, in particular the divestiture of AT&T in 1984, spurred the productivity growth of the incumbent, facilitated entries, and increased the aggregate productivity?
  - To do so, the authors estimate the plant-level production functions and productivity in the telecommunication industry.
- Doraszelski and Jaumandreu [2013]:
  - What is the role of R&D in determining the differences in productivity across firms and the evolution of firm-level productivity over time?
  - To do so, the authors estimate the firm-level production functions and productivity of Spanish manufacturing firms during 1990s in which the transition probability of a productivity is a function of the R&D activities.

### 3.1.2 Development

- Hsieh and Klenow [2009]:
  - How large is the misallocation of inputs across manufacturing firms in China and India compared to the U.S? How will the aggregate productivity of China and India change if the degree of misallocation is reduced to the U.S. level?
  - To do so, the authors measure the revenue productivity of firms, which should be the same across firms within an industry if there were no distortion, and the measurement of the revenue productivity requires to estimate the production function.

- Gennaioli et al. [2013]:
  - What are the determinants of regional growth? Do geographic, institutional, cultural, and human capital factors explain the difference across regions?
  - To do so, the authors construct the data set that covers 74% of the world's surface and 97% of its GDP and estimate the production function in which the above mentioned factors could affect the productivity.

### 3.1.3  Trade

- Haskel et al. [2007]:
  - Are there spillovers from FDI to domestic firms?
  - To do so, the authors estimate the plant-level production function of the U.K. manufacturing firms during 1973 and 1992 and study how the foreign presence in the U.K. affected the productivity.
- De Loecker [2011]:
  - Does the removal of trade barriers induces efficiency gain for producers?
  - To do so, the author estimate the production functions of Belgian textile industry during 1994-2002 in which the degree of trade protection can affect the productivity level.

### 3.1.4  Management

- Bloom and Van Reenen [2007]:
  - How do management practices affect the firm productivity?
  - To do so, the authors first estimate the production function and productivity of manufacturing firms in developed countries, and then study how the independently measured management practices of the firms affect the estimated productivity.
- Braguinsky et al. [2015]:
  - How do changes in ownership affect the productivity and profitability of firms?
  - To do so, the authors estimate the production function for various outputs including the physical output, return on capital and labor, and the utilization rate, price level, using the cotton spinners data in Japan during 1896 and 1920.

### 3.1.5  Education

- Cunha et al. [2010]:
  - How do childhood and schooling interventions "produce" the cognitive and non-cognitive skills of children?
  - To do so, the authors estimate the mapping from childhood and schooling interventions to children's cognitive and non-cognitive skills, the "production function" of childhood environment and education.

## 3.2  Analyzing Producer Behaviors

- There are several levels of parameters that govern the behavior of firms:
- **Production function**
  - Add factor market structure.
  - Add cost minimization.
- → **Cost function**
  - Add product market structure.

– Add profit maximization.

- $\rightarrow$ **Supply function (Pricing function)**

  – Combine cost and supply (pricing) functions.

- $\rightarrow$ **Profit function**

- Which parameter to identify?

- Primitive enough to be invariant to relevant policy changes.

  – e.g. If you conduct a policy experiment that changes the factor market structure, identifying cost functions is not enough.

- As reduced-form as possible among such specifications.

  – A reduced-form parameter usually can be rationalized by a class of underlying structural parameters and institutional assumptions. Thus, the analysis becomes robust to some misspecifications.
  – e.g. A non-parametric function $C(q, w)$ can represent a cost function of a producer who is not necessarily minimizing the cost. If we derive a cost function from a production function and a factor market structure, then the cost function cannot represent such a non-optimization behavior.

## 3.3   Production Function Estimation

### 3.3.1   Cobb-Douglas Specification as a Benchmark

- Most of the following argument carries over to a general model.
- For firm $j = 1, \cdots, J$ and time $t = 1, \cdots, T$, we observe output $Y_{jt}$, labor $L_{jt}$, and capital $K_{jt}$.
- We consider an asymptotic of $J \rightarrow \infty$ for a fixed $T$.
- Assume Cobb-Douglas production function:

$$Y_{jt} = A_{jt} L_{jt}^{\beta_l} K_{jt}^{\beta_k}, \tag{3.1}$$

  where $A_{jt}$ is firm $j$ and time $t$ specific unobserved heterogeneity in the model.
- Taking the logarithm gives:

$$y_{jt} = \beta_0 + \beta_l l_{jt} + \beta_k k_{jt} + \epsilon_{jt}, \tag{3.2}$$

  where lowercase symbols represent natural logs of variables and $\ln(A_{jt}) = \beta_0 + \epsilon_{jt}$.
- This can be regarded as a first-order log-linear approximation of a production function.
- Linear regression model! May OLS work?

### 3.3.2   Potential Bias I: Endogeneity

- $\epsilon_{jt}$ contains everything that cannot be explained by the observed inputs: better capital may be employed, a worker may have obtained better skills, etc.
- When the manager of a firm makes an input choice, she should have some information about the realization of $\epsilon_{jt}$.
- Thus, the input choice can be correlated with $\epsilon_{jt}$; for example under static optimization of $L_{jt}$ given $K_{jt}$:

$$L_{jt} = \left[ \frac{p_{jt}}{w_{jt}} \beta_l \exp^{\beta_0 + \epsilon_{jt}} K_{jt}^{\beta_k} \right]^{\frac{1}{1-\beta_l}}. \tag{3.3}$$

- In this case, OLS estimator for $\beta_l$ is *positively* biased, because when $\epsilon_{jt}$ is high, $l_{jt}$ is high and thus the increase in output caused by $\epsilon_{jt}$ is captured as if caused by the increase in labor input.

- The endogeneity problem was already recognized by Marschak and Andrews [1944].

### 3.3.3   Potential Bias II: Selection

- Firms freely enter and exit market.
- Therefore, a firm that had low $\epsilon_{jt}$ is likely to exit.
- However, if firms have high capital $K_{jt}$, it can stay in the market even if the realization of $\epsilon_{jt}$ is very low.
- Therefore, conditional on being in the market, there is a *negative* correlation between the capital $K_{jt}$ and $\epsilon_{jt}$.
- This problem occurs even if the choice of $K_{jt}$ itself is not a function of $\epsilon_{jt}$.

### 3.3.4   How to Resolve Endogeneity Bias?

- Temporarily abstract away from entry and exit.
- The data is balanced.

1. Panel data.
2. First-order condition for inputs.
3. Instrumental variable.
4. Olley-Pakes approach and its followers/critics.

- Griliches and Mairesse [1998] is a good survey of the history up to Olley-Pakes approach.
- Ackerberg et al. [2015] also offer a good survey and clarify problems and implicit assumptions in Olley-Pakes approach.

### 3.3.5   Panel Data

- Assume that $\epsilon_{jt} = \mu_j + \eta_{jt}$, where $\eta_{jt}$ is uncorrelated with input choices up to period $t$:

$$y_{jt} = \beta_0 + \beta_l l_{jt} + \beta_k k_{jt} + \mu_j + \eta_{jt}. \tag{3.4}$$

- Then, by differentiating period $t$ and $t-1$ equations, we get:

$$y_{jt} - y_{j,t-1} = \beta_l(l_{jt} - l_{j,t-1}) + \beta_k(k_{jt} - k_{j,t-1}) + (\eta_{jt} - \eta_{j,t-1}). \tag{3.5}$$

- Then, because $\eta_{jt} - \eta_{j,t-1}$ is uncorrelated either with $l_{jt} - l_{j,t-1}$ or $k_{jt} - k_{j,t-1}$, we can identify the parameter.
- Problem:
  - Restrictive heterogeneity.
  - When there are measurement errors, fixed-effect estimator can generate higher biases than OLS estimator, because measurement errors more likely to survive first-difference and within-transformation.

### 3.3.6   First-Order Condition for Inputs

- Use the first-order condition for inputs as the moment condition [McElroy, 1987].
- Closely related to the cost function estimation literature.
- Need to specify the factor market structure and the nature of the optimization problem for a firm.
- Recently being center of attention again as one of the solutions to the "collinearity problem" discussed below.

### 3.3.7 Instrumental Variable

- Borrow the idea from the first-order condition approach that the input choices are affected by some exogenous variables.

- If we have instrumental variables that affect inputs but are uncorrelated with errors $\epsilon_{jt}$, then we can identify the parameter by an instrumental variable method.

- One candidate for the instrumental variables: **input prices**.

- Input price affect input decision.

- Input price is not correlated with $\epsilon_{jt}$ if the factor product market is competitive and $\epsilon_{jt}$ is an idiosyncratic shock to a firm.

- Problems:

    - Input prices often lack cross-sectional variation.
    - Cross-sectional variation is often due to unobserved input quality.

- Another candidate for the instrumental variables: **lagged inputs**.

- If $\epsilon_{jt}$ does not have auto-correlation, lagged inputs are not correlated with the current shock.

- If there are adjustment costs for inputs, then lagged inputs are correlated with the current inputs.

- Problem:

    - If $\epsilon_{jt}$ has auto-correlation, all lagged inputs are correlated with the errors: For example, if $\epsilon_{jt}$ is AR(1), $\epsilon_{jt} = \alpha \epsilon_{j,t-1} + \nu_{j,t-1} = \cdots \alpha^l \epsilon_{j,t-l} + \nu_{j,t-1} + \cdots, \alpha^{l-1} \nu_{j,t-l}$ for any $l$.

### 3.3.8 Olley-Pakes Approach

- Exploit restrictions from the economic theory [Olley and Pakes, 1996].
- Write $\epsilon_{jt} = \omega_{jt} + \eta_{jt}$, where $\omega_{jt}$ is an anticipated shock and $\eta_{jt}$ is an ex-post shock.
- Inputs are correlated with $\omega_{jt}$ but not with $\eta_{jt}$
- The model is written as:

$$y_{jt} = \beta_0 + \beta_l l_{jt} + \beta_k k_{jt} + \omega_{jt} + \eta_{jt}. \tag{3.6}$$

- OP use economic theory to derive a valid proxy for the anticipated shock $\omega_{jt}$.

### 3.3.9 Assumption I: Information Set

- The firm's information set at $t$, $I_{jt}$, includes current and past productivity shocks $\{\omega_{j\tau}\}_{\tau=0}^{t}$ but does not include future productivity shocks $\{\omega_{j\tau}\}_{\tau=t+1}^{\infty}$.
- The transitory shocks $\eta_{jt}$ satisfy $\mathbb{E}\{\eta_{jt}|I_{jt}\} = 0$.

### 3.3.10 Assumption II: First Order Markov

- Productivity shocks evolve according to the distribution:

$$p(\omega_{j,t+1}|I_{jt}) = p(\omega_{j,t+1}|\omega_{jt}), \tag{3.7}$$

and the distribution is known to firms and stochastically increasing in $\omega_{jt}$.

- Then:

$$\omega_{jt} = \mathbb{E}\{\omega_{jt}|\omega_{j,t-1}\} + \nu_{jt}, \tag{3.8}$$

  and:

$$\mathbb{E}\{\nu_{jt}|I_{j,t-1}\} = 0, \tag{3.9}$$

  by construction.

### 3.3.11  Assumption III: Timing of Input Choices

- Firms accumulate capital according to:

$$k_{jt} = \kappa(k_{j,t-1}, i_{j,t-1}), \tag{3.10}$$

  where investment $i_{j,t-1}$ is chosen in period $t-1$.
- Labor input $l_{jt}$ is non-dynamic and chosen at $t$.
- This assumption characterizes and distinguishes labor and capital.
- Intuitively, it takes a full period for new capital to be ordered, delivered, and installed.

### 3.3.12  Assumption IV: Scalar Unobservable

- Firms' investment decisions are given by:

$$i_{jt} = f_t(k_{jt}, \omega_{jt}). \tag{3.11}$$

- This assumption places strong implicit restrictions on additional firm-specific unobservables.
    - No **across firm** unobserved heterogeneity in adjustment cost of capital, in demand and labor market conditions, or in other parts of the production function.
    - Okay with **across time** unobserved heterogeneity.

### 3.3.13  Assumption IV: Strict Monotonicity

- The investment policy function $f_t(k_{jt}, \omega_{jt})$ is strictly increasing in $\omega_{jt}$.
- This holds if the realization of higher $\omega_{jt}$ implies higher expectation for future productivity (Assumption III) and if the marginal product of capital is increasing in the expectation for future productivity.
- To verify the latter condition in a given game is often not easy.

### 3.3.14  Two-step Approach: The First Step

- Insert $\omega_{jt} = h(k_{jt}, i_{jt})$ to the original equation to get:

$$y_{jt} = \beta_l l_{jt} + \underbrace{\beta_0 + \beta_k k_{jt} + h(k_{jt}, i_{jt})}_{\text{unknown function of } k_{jt} \text{ and } i_{jt}} + \eta_{jt}$$
$$\equiv \beta_l l_{jt} + \phi(k_{jt}, i_{jt}) + \eta_{jt}. \tag{3.12}$$

- This is a **partially linear model**: see Ichimura and Todd [2007] for reference.

- Because $l_{jt}, k_{jt}$ and $i_{jt}$ are uncorrelated with $\eta_{jt}$, we can identify $\beta_l$ and $\phi(\cdot)$ by exploiting the moment condition:

$$
\begin{aligned}
&\mathbb{E}\{\eta_{jt}|l_{jt}, k_{jt}, i_{jt}\} = 0 \\
&\Leftrightarrow \mathbb{E}\{y_{jt} - \beta_l l_{jt} - \phi(k_{jt}, i_{jt})|l_{jt}, k_{jt}, i_{jt}\} = 0.
\end{aligned}
\tag{3.13}
$$

  **if there is enough variation** in $l_{jt}, k_{jt}$ and $i_{jt}$.
- This "if there is enough variation" part is actually problematic. Discuss later.
- Let $\beta_l^0$ and $\phi^0$ be the identified true parameters.

### 3.3.15 Two-step Approach: The Second Step

- Note that:

$$
\omega_{jt} \equiv \phi(k_{jt}, i_{jt}) - \beta_0 - \beta_k k_{jt}.
\tag{3.14}
$$

- Therefore, we have:

$$
\begin{aligned}
&y_{jt} - \beta_l^0 l_{jt} \\
&= \beta_0 + \beta_k k_{jt} + \omega_{jt} + \eta_{jt} \\
&= \beta_0 + \beta_k k_{jt} + g(\omega_{j,t-1}) + \nu_{jt} + \eta_{jt} \\
&= \beta_0 + \beta_k k_{jt} + g[\phi^0(k_{j,t-1}, i_{j,t-1}) - (\beta_0 + \beta_k k_{j,t-1})] + \nu_{jt} + \eta_{jt}.
\end{aligned}
\tag{3.15}
$$

- $\nu_{jt}$ and $\eta_{jt}$ are independent of the covariates.
- This is a **multiple-index model** with indices $\beta_0 + \beta_1 k_{jt}$ and $\beta_0 + \beta_1 k_{j,t-1}$ where parameters of two indices are restricted to be the same: see Ichimura and Todd [2007] for reference.
- We can identify $\beta_0, \beta_k$ and $g$ by exploiting the moment condition:

$$
\begin{aligned}
&\mathbb{E}\{\nu_{jt} + \eta_{jt}|k_{jt}, k_{j,t-1}, i_{j,t-1}\} = 0 \\
&\Leftrightarrow \mathbb{E}\{\nu_{jt} + \eta_{jt}|k_{jt}, k_{j,t-1}, i_{j,t-1}\} = 0.
\end{aligned}
\tag{3.16}
$$

### 3.3.16 Identification of the Anticipated Shocks

- If $\phi, \beta_0, \beta_k$ are identified, then $\omega_{jt}$ is also identified by:

$$
\omega_{jt} \equiv \phi(k_{jt}, i_{jt}) - \beta_0 - \beta_k k_{jt}.
\tag{3.17}
$$

### 3.3.17 Two-Step Estimation of Olley and Pakes [1996].

- **First step**: Estimate $\beta_L$ and $\phi$ in :

$$
y_{jt} = \beta_l l_{jt} + \phi(k_{jt}, i_{jt}) + \eta_{jt}.
\tag{3.18}
$$

  by approximating $\phi$ with some basis functions, say, polynomials or splines:

$$
\begin{aligned}
y_{jt} &= \beta_l l_{jt} + \sum_{p=1}^{P} \gamma_p \phi_p(k_{jt}, i_{jt}) + \left[ \phi(k_{jt}, i_{jt}) - \sum_{n=1}^{N} \gamma_n \phi_n(k_{jt}, i_{jt}) \right] + \eta_{jt} \\
&= \beta_l l_{jt} + \sum_{p=1}^{P} \gamma_p \phi_p(k_{jt}, i_{jt}) + \tilde{\eta}_{jt}
\end{aligned}
\tag{3.19}
$$

where $P \to \infty$ when the sample size goes to infinity.

- e.g. second-order polynomial approximation:

$$\begin{aligned}
&\phi_1(k_{jt}, i_{jt}) = k_{jt}, \phi_2(k_{jt}, i_{jt}) = i_{jt} \\
&\phi_3(k_{jt}, i_{jt}) = k_{jt}^2, \phi_4(k_{jt}, i_{jt}) = i_{jt}^2 \\
&\phi_5(k_{jt}, i_{jt}) = k_{jt} i_{jt}.
\end{aligned} \tag{3.20}$$

- Once the basis functions are fixed, estimation is the same as the linear model.
- But the inference (the computation of the standard deviation) is difference, because of the approximation error.
- See Chen [2007] for reference.
- Let $\hat{\beta}_l$ and $\hat{\phi}$ be the estimates from the first step.
- **Second step**: Estimate $\beta_0$, $\beta_k$, and $g$ in:

$$\begin{aligned}
y_{jt} - \hat{\beta}_l l_{jt} &= \beta_0 + \beta_k k_{jt} + g[\hat{\phi}(k_{j,t-1}, i_{j,t-1}) - (\beta_0 + \beta_k k_{j,t-1})] + \nu_{jt} + \eta_{jt} \\
&\quad + [\beta_l - \hat{\beta}_l] l_{jt} \\
&\quad + \left\{ g[\phi(k_{j,t-1}, i_{j,t-1}) - (\beta_0 + \beta_k k_{j,t-1})] - g[\hat{\phi}(k_{j,t-1}, i_{j,t-1}) - (\beta_0 + \beta_k k_{j,t-1})] \right\} \\
&= \beta_0 + \beta_k k_{jt} + g[\hat{\phi}(k_{j,t-1}, i_{j,t-1}) - (\beta_0 + \beta_k k_{j,t-1})] + \nu_{jt} + \tilde{\eta}_{jt}
\end{aligned} \tag{3.21}$$

by approximating $g$ by some basis functions, say, polynomials or splines.

### 3.3.18   From An Economic Models to An Econometric Model

- Starting from economic model with some unobserved heterogeneity, we reach some reduced-form model.
- If the resulting model belongs to a class of econometric models whose identification and estimation are established, we can simply apply the existing methods.

### 3.3.19   How to Resolve Selection Bias

- Use propensity score to correct selection bias: Ahn and Powell [1993].
- At the beginning of period $t$, after observing $\omega_{jt}$, firm $j$ decides whether to continue the business ($\chi_{jt} = 1$) or exit ($\chi_{jt} = 0$).
- Assume that the difference between continuation and exit values is strictly increasing in $\omega_{jt}$.
- Then, there is a threshold $\underline{\omega}(k_{jt})$ such that:

$$\chi_{jt} = \begin{cases} 1 & \text{if } \omega_{jt} \geq \underline{\omega}(k_{jt}) \\ 0 & \text{otherwise.} \end{cases} \tag{3.22}$$

- We can only observe firms that satisfy $\chi_{jt} = 1$.

### 3.3.20   Correction in the First Step

- In the first step, we need no correction because:

$$\begin{aligned}
&\mathbb{E}\{y_{jt} | l_{jt}, k_{jt}, i_{jt}, \chi_{jt} = 1\} \\
&= \beta_l l_{jt} + \phi(k_{jt}, i_{jt}) + \mathbb{E}\{\eta_{jt} | \chi_{jt} = 1\} \\
&= \beta_l l_{jt} + \phi(k_{jt}, i_{jt}).
\end{aligned} \tag{3.23}$$

- Ex-post shock $\eta_{jt}$ is independent of continuation/exit decision. Therefore, we can identify $\beta_l$ and $\phi(\cdot)$ as in the previous case.

### 3.3.21 Correction in the Second Step I: The Source of Bias

- One the other hand, we need correction in the second step, because:

$$
\begin{aligned}
&\mathbb{E}\{y_{jt} - \beta_l^0 l_{jt} | k_{jt}, i_{jt}, k_{j,t-1}, l_{j,t-1}, \chi_{jt} = 1\} \\
&= \beta_0 + \beta_k k_{jt} + g[\phi^0(k_{jt}, i_{jt}) - (\beta_0 + \beta_k k_{jt})] \\
&\quad + \mathbb{E}\{\nu_{jt} + \eta_{jt} | k_{jt}, i_{jt}, k_{j,t-1}, l_{j,t-1}, \chi_{jt} = 1\} \\
&= \beta_0 + \beta_k k_{jt} + g[\phi^0(k_{j,t-1}, i_{j,t-1}) - (\beta_0 + \beta_k k_{j,t-1})] \\
&\quad + \mathbb{E}\{\nu_{jt} | k_{jt}, i_{jt}, k_{j,t-1}, l_{j,t-1}, \chi_{jt} = 1\}.
\end{aligned}
\tag{3.24}
$$

and

$$
\mathbb{E}\{\nu_{jt} | k_{jt}, i_{jt}, k_{j,t-1}, l_{j,t-1}, \chi_{jt} = 1\} \neq 0,
\tag{3.25}
$$

since anticipated shock matters continuation/exit decision in period $t$.

### 3.3.22 Correction in the Second Step II: Conditional Exit Probability

- Let's see that the conditional expectation:

$$
\begin{aligned}
&\mathbb{E}\{\omega_{jt} | k_{jt}, i_{jt}, k_{j,t-1}, l_{j,t-1}, \chi_{jt} = 1\} \\
&= \mathbb{E}\{\omega_{jt} | k_{jt}, i_{jt}, k_{j,t-1}, l_{j,t-1}, \omega_{jt} \geq \underline{\omega}(k_{jt})\} \\
&= \int_{\underline{\omega}(k_{jt})} \omega_{jt} \frac{p(\omega_{jt} | \omega_{j,t-1})}{\int_{\underline{\omega}(k_{jt})} p(\omega | \omega_{j,t-1}) d\omega} d\omega_{jt} \\
&\equiv \tilde{g}(\omega_{j,t-1}, \underline{\omega}(k_{jt})),
\end{aligned}
\tag{3.26}
$$

is a function of $\omega_{j,t-1}$ and $\underline{\omega}(k_{jt})$.

### 3.3.23 Correction in the Second Step III: Invertibility in Threshold

- The propensity of continuation conditional on observed information up to period $t-1$:

$$
\begin{aligned}
P_{jt} &\equiv \mathbb{P}\{\chi_{jt} = 1 | \mathcal{I}_{j,t-1}\} \\
&= \mathbb{P}\{\omega_{jt} \geq \underline{\omega}(k_{jt}) | \mathcal{I}_{j,t-1}\} \\
&= \mathbb{P}\{g(\omega_{j,t-1}) + \nu_{jt} \geq \underline{\omega}[(1-\delta)k_{j,t-1} + i_{j,t-1}] | \mathcal{I}_{j,t-1}\} \\
&= \mathbb{P}\{\chi_{jt} = 1 | i_{j,t-1}, k_{j,t-1}\}.
\end{aligned}
\tag{3.27}
$$

- $\rightarrow$ It suffices to condition on $i_{j,t-1}, k_{j,t-1}$.
- We also have:

$$
P_{jt} = \mathbb{P}\{\chi_{jt} = 1 | \omega_{j,t-1}, \underline{\omega}(k_{jt})\},
\tag{3.28}
$$

and it is invertible in $\underline{\omega}(k_{jt})$, that is,

$$
\underline{\omega}(k_{jt}) \equiv \psi(P_{jt}, \omega_{j,t-1}).
\tag{3.29}
$$

### 3.3.24   Correction in the Second Step IV: Controlling the Threshold

- Now, he have:

$$
\begin{aligned}
&\mathbb{E}\{y_{jt} - \beta_l^0 l_{jt} | k_{jt}, i_{jt}, k_{j,t-1}, l_{j,t-1}, \chi_{jt} = 1\} \\
&= \beta_0 + \beta_k k_{jt} + \mathbb{E}\{\omega_{jt} | k_{jt}, i_{jt}, k_{j,t-1}, l_{j,t-1}, \chi_{jt} = 1\} \\
&= \beta_0 + \beta_k k_{jt} + \tilde{g}(\omega_{j,t-1}, \underline{\omega}(k_{jt})) \\
&= \beta_0 + \beta_k k_{jt} + \tilde{g}(\omega_{j,t-1}, \psi(P_{jt}, \omega_{j,t-1})) \\
&\equiv \beta_0 + \beta_k k_{jt} + \tilde{\tilde{g}}(\omega_{j,t-1}, P_{jt}) \\
&= \beta_0 + \beta_k k_{jt} + \tilde{\tilde{g}}[\phi^0(k_{j,t-1}, i_{j,t-1}) - (\beta_0 + \beta_k k_{j,t-1}), P_{jt}].
\end{aligned}
\tag{3.30}
$$

- At the end, the only difference is to include $P_{jt}$ as a covariate.
- $P_{jt}$ is a **known** function of $i_{j,t-1}$ and $k_{j,t-1}$.
- Even if we condition on $P_{jt} = p$, there are still many combinations of $i_{j,t-1}$ and $k_{j,t-1}$ that gives $P_{jt} = p$.
- With this remaining variation, we can identify $\beta_0$, $\beta_k$, and $\tilde{\tilde{g}}$ by the same argument as the case without selection, for each $P_{jt} = p$.

### 3.3.25   Three Step Estimation of Olley and Pakes [1996]

- **Zero step**: Estimate the propensity score:

$$
P_{jt} = 1\{\chi_{jt} = 1 | i_{j,t-1}, k_{j,t-1}\},
\tag{3.31}
$$

  by a kernel estimator.
- Insert the resulting estimates $\widehat{P}_{jt}$ into the first and second steps.

### 3.3.26   Zero Investment Problem

- One of the key assumptions in OP method was invertibility between anticipated shock and investment:

$$
\omega_{jt} = i^{-1}(k_{jt}, i_{jt}) \equiv h(k_{jt}, i_{jt}).
\tag{3.32}
$$

- However, in micro data, zero investment is a rule rather than exceptions.
- Then, the invertibility does not hold globally: there are some region of the anticipated shock in which the investment takes value zero.

### 3.3.27   Tackle Zero Investment Problem I: Discard Some Data

- Discard a data $(j, t)$ such that $i_{j,t-1} = 0$.
- Use a data $(j, t)$ such that $i_{j,t-1} > 0$.
- Then, invertibility recovers on this selected sample.
- This **does not** cause bias in the estimator because $\nu_{jt}$ in :

$$
\beta_0 + \beta_l k_{jt} + g[\phi^0(k_{j,t-1}, i_{j,t-1}) - (\beta_0 + \beta_k k_{j,t-1})] + \nu_{jt} + \eta_{jt},
\tag{3.33}
$$

  is independent of the event up to $t - 1$, including $i_{j,t-1}$.
- However, this **does** cause information loss. The loss is high if the proportion of the sample such that $i_{j,t-1} = 0$ is high.

### 3.3.28 Tackle Zero Investment Problem II: Use Another Proxy

- Investment is just a possible proxy for the anticipated shock.
- Intermediate inputs can be used as proxies as well [Levinsohn and Petrin, 2003].
- The problem is that these intermediate inputs are included in the gross production function, whereas investment is excluded.
- Let $m_{jt}$ be the log material input, and assume that the production function takes the form of:

$$y_{jt} = \beta_0 + \beta_l l_{jt} + \beta_k k_{jt} + \beta_m m_{jt} + \omega_{jt} + \eta_{jt}. \tag{3.34}$$

- In addition, assume that the **optimal policy function** for $m_{jt}$ is strictly monotonic in the ex-ante shock, and hence is invertible:

$$m_{jt} = m(k_{jt}, \omega_{jt}) \Leftrightarrow \omega_{jt} = m^{-1}(m_{jt}, k_{jt}) \equiv h(m_{jt}, k_{jt}). \tag{3.35}$$

- **First step**:

$$\begin{aligned}
y_{jt} &= \beta_0 + \beta_l l_{jt} + \beta_k k_{jt} + \beta_m m_{jt} + h(m_{jt}, k_{jt}) + \eta_{jt} \\
&= \beta_l l_{jt} + \phi(m_{jt}, k_{jt}) + \eta_{jt}.
\end{aligned} \tag{3.36}$$

- We can identify $\beta_l$ and $\phi$ by exploiting the moment condition:

$$\begin{aligned}
&\mathbb{E}\{\eta_{jt}|l_{jt}, m_{jt}, k_{jt}, i_{jt}\} = 0 \\
\Leftrightarrow &\mathbb{E}\{y_{jt} - \beta_0 - \beta_l l_{jt} - \phi(m_{jt}, k_{jt})|l_{jt}, m_{jt}, k_{jt}\} = 0,
\end{aligned} \tag{3.37}$$

  if **there is enough variation** in $l_{jt}, m_{jt}, k_{jt}$.
- **Second step**:

$$\begin{aligned}
&y_{jt} - \beta_l^0 l_{jt} \\
&= \beta_k k_{jt} + \beta_m m_{jt} + g[\phi^0(m_{j,t-1}, k_{j,t-1}) - \beta_k k_{j,t-1} - \beta_m m_{j,t-1}] \\
&+ \nu_{jt} + \eta_{jt}.
\end{aligned} \tag{3.38}$$

- We can identify $\beta_k$, $\beta_m$, and $g$ by exploiting the moment condition:

$$\mathbb{E}\{\nu_{jt} + \eta_{jt}|k_{jt}, m_{j,t-1}, k_{j,t-1}\} = 0. \tag{3.39}$$

- Because $m_{jt}$ is correlated with $\nu_{jt}$, the moment should not condition on $m_{jt}$.
- The identification of $\beta_m$ comes from $\beta_m m_{j,t-1}$.

### 3.3.29 One-step Estimation of Olley and Pakes [1996] and Levinsohn and Petrin [2003]

- Levinsohn and Petrin [2003] can be estimated in the similar two-step method.
- We can jointly estimate the parameters in first and second steps to improve the efficiency [Wooldridge, 2009].
- We estimate under the assumptions of Olley and Pakes [1996]:

$$y_{jt} = \beta_0 + \beta_1 l_{jt} + \beta_k k_{jt} + \omega_{jt} + \eta_{jt}. \tag{3.40}$$

- The first step exploits the following moment:

$$\mathbb{E}\{\eta_{jt}|l_{jt}, k_{jt}, i_{jt}\} = 0, \tag{3.41}$$

that is:

$$\mathbb{E}\{y_{jt} - \beta_1 l_{jt} - \beta_0 - \beta_k k_{jt} - \omega(k_{jt}, i_{jt}) | l_{jt}, k_{jt}, i_{jt}\} = 0. \tag{3.42}$$

- We can reinforce the moment condition as:

$$\mathbb{E}\{\eta_{jt} | l_{jt}, k_{jt}, i_{jt}, \cdots, l_{j1}, k_{j1}, i_{j1}\} = 0 \tag{3.43}$$

  if we assume that lagged inputs are correlated with the current inputs and $\eta_{jt}$ is independent.
- The second step exploits the following moment:

$$\mathbb{E}\{\nu_{jt} | k_{jt}, i_{j,t-1}, l_{j,t-1}\} = 0, \tag{3.44}$$

  that is:

$$\mathbb{E}\{y_{jt} - \beta_0 - \beta_1 l_{jt} - \beta_k k_{jt} - g[\omega(k_{j,t-1}, i_{j,t-1})] | k_{jt}, i_{j,t-1}, l_{j,t-1}\} = 0. \tag{3.45}$$

- We can reinforce the moment condition as:

$$\mathbb{E}\{\nu_{jt} | k_{jt}, i_{j,t-1}, l_{j,t-1}, \cdots, k_{j1}, i_{j1}, l_{j1}\} = 0, \tag{3.46}$$

  if we assume that lagged input are correlated with the current inputs and $\nu_{jt} + \eta_{jt}$ are independent.
- We can construct a GMM estimator based on equations (3.42) and (3.45).
- The one-step estimator can be more efficient but can be computationally heavier than the two-step estimator.

### 3.3.30   Scalar Unobservable Problem: Finite-order Markov Process

- Borrow the idea of using the first-order condition to resolve the collinearity problem [Gandhi et al., 2017].
- We have assumed that anticipated shocks follow a first-order Markov process:

$$\omega_{jt} = g(\omega_{j,t-1}) + \nu_{jt}. \tag{3.47}$$

- However, it may be true that it has more than one lags, for example:

$$\omega_{jt} = g(\omega_{j,t-1}, \omega_{j,t-2}) + \nu_{jt}. \tag{3.48}$$

- Then, we need proxies as many as the number of unobservables:

$$\begin{pmatrix} i_{jt} \\ m_{jt} \end{pmatrix} = \Gamma(k_{jt}, \omega_{jt}, \omega_{j,t-1}), \tag{3.49}$$

  such that the policy function for the proxies is a bijection in $(\omega_{jt}, \omega_{j,t-1})$.
- Then, we can have:

$$\omega_{jt} = \Gamma_1^{-1}(k_{jt}, i_{jt}, m_{jt}). \tag{3.50}$$

- The reminder goes as in the standard OP method.

### 3.3.31   Scalar Unobservable Problem: Demand and Productivity Shocks

- There may be a demand shock $\mu_{jt}$ that also follows first-order Markov process.
- Then, the policy function depend both on $\mu_{jt}$ and $\omega_{jt}$.
- We again need proxies as many as the number of unobservable.
- Suppose that we can observe the price of the firm $p_{jt}$.

- Inverting the policy function:

$$\begin{pmatrix} i_{jt} \\ p_{jt} \end{pmatrix} = \Gamma(k_{jt}, \omega_{jt}, \mu_{jt}). \tag{3.51}$$

  yields:

$$\omega_{jt} = \Gamma_1^{-1}(k_{jt}, i_{jt}, p_{jt}). \tag{3.52}$$

- If $\omega_{jt}$ only depends on $\omega_{j,t-1}$ but not on $\mu_{j,t-1}$, then the second step of the modified OP method is to estimate:

$$\begin{aligned} y_{jt} - \hat{\beta}_l l_{jt} &= \beta_0 + \beta_k k_{jt} \\ &\quad + g(\omega_{j,t-1}) + \nu_{jt} + \eta_{jt} \\ &= \beta_0 + \beta_k k_{jt} \\ &\quad + g(\hat{\phi}_{j,t-1} - \beta_0 - \beta_k k_{j,t-1}) + \nu_{jt} + \eta_{jt}. \end{aligned} \tag{3.53}$$

- It goes as in the standard OP method.
- If $\omega_{jt}$ depends both on $\omega_{j,t-1}$ and $\mu_{j,t-1}$, the second step regression equation will be:

$$\begin{aligned} y_{jt} - \hat{\beta}_l l_{jt} &= \beta_0 + \beta_k k_{jt} \\ &\quad + g(\omega_{j,t-1}, \mu_{j,t-1}) + \nu_{jt} + \eta_{jt} \\ &= \beta_0 + \beta_k k_{jt} \\ &\quad + g(\hat{\phi}_{j,t-1} - \beta_0 - \beta_k k_{j,t-1}, \mu_{j,t-1}) + \nu_{jt} + \eta_{jt}. \end{aligned} \tag{3.54}$$

- We still have to control $\mu_{j,t-1}$ in the second step.
- Invert the policy function for $\mu_{j,t-1}$ to get:

$$\mu_{j,t-1} = \Gamma_2^{-1}(k_{j,t-1}, i_{j,t-1}, p_{j,t-1}), \tag{3.55}$$

  and plug it into the second step regression equation to get:

$$\begin{aligned} &y_{jt} - \hat{\beta}_l l_{jt} \\ &= \beta_0 + \beta_k k_{jt} \\ &\quad + g(\hat{\phi}_{j,t-1} - \beta_0 - \beta_k k_{j,t-1}, \Gamma_2^{-1}(k_{j,t-1}, i_{j,t-1}, p_{j,t-1})) + \nu_{jt} + \eta_{jt}. \end{aligned} \tag{3.56}$$

- The parameters $\beta_0$ and $\beta_k$ **cannot** be identified only with this observation, because $\Gamma_2^{-1}$ is **unknown non-parametric** function: it can mean any function of $(k_{j,t-1}, i_{j,t-1}, p_{j,t-1})$.
- To estimate such a model, we jointly estimate the demand function along with the production function.
- At this point, we do not investigate it further because we have not yet learned how to estimate the demand function.
- For now just keep in mind that:
    - There has to be as many proxies as the dimension of the unobservable state variables.
    - It is okay that the unobservable state variable includes a demand shock.
    - It can be problematic when the unobservable demand shock affect the evolution of the anticipated productivity shock.

## 3.3.32   Collinearity Problem

- The collinearity problem is formally pointed out by Ackerberg et al. [2015].
- This paper is finally published in 2015, but has been circulated since 2005.

- We assumed that $k_{jt}$ and $\omega_{jt}$ are state variables.
- Then the policy function for labor input should take the form of:

$$l_{jt} = l(k_{jt}, \omega_{jt}). \tag{3.57}$$

- However, because $\omega_{jt} = h(i_{jt}, k_{jt})$, we have:

$$l_{jt} = l[k_{jt}, h(i_{jt}, k_{jt})] = \tilde{l}(i_{jt}, k_{jt}). \tag{3.58}$$

- Therefore, in the first stage, we encounter a multicollinearity problem:

$$
\begin{aligned}
y_{jt} &= \beta_0 + \beta_l \tilde{l}(i_{jt}, k_{jt}) + \phi(i_{jt}, k_{jt}) + \eta_{jt} \\
&\equiv \tilde{\phi}(i_{jt}, k_{jt}).
\end{aligned}
\tag{3.59}
$$

- Thus, $\beta_l$ cannot be identified in the first step.
- The second step becomes:

$$y_{jt} = \beta_0 + \beta_l l_{jt} + \beta_k k_{jt} + g[\tilde{\phi}(i_{j,t-1}, k_{j,t-1}) - \beta_0 - \beta_l l_{j,t-1} - \beta_k k_{jt}] + \nu_{jt} + \eta_{jt} \tag{3.60}$$

- Because $l_{jt}$ is correlated with $\nu_{jt}$, moment can only condition on $l_{j,t-1}$.
- However, conditioning on $k_{j,t-1}$ and $i_{j,t-1}$, again there is no remaining variation in $l_{j,t-1}$.
- Therefore, $\beta_l$ cannot be identified either in the second step.
- $\beta_l$ **cannot be identified!**

### 3.3.33   Tackle Collinearity Problem: Peculiar Assumptions

- To make Olley-Pakes/Levinsohn-Petrin approach workable, we need peculiar data generating process for $l_{jt}$.
- Consider Levinsohn-Petrin framework.

1. There is an optimization error in $l_{jt}$.
    - If it is not i.i.d over time, it becomes a state variable and enters to the policy for $m_{jt}$, violating the scalar unobserved heterogeneity assumption of $m_{jt}$.
    - If there is an optimization error for $m_{jt}$, this again violates the scalar unobserved heterogeneity assumption.
2. $k_{jt}$ is realized, $\omega_{jt}$ is observed, $m_{jt}$ and $i_{jt}$ are determined, a new i.i.d. unexpected shock is observed, $l_{jt}$ is determined, and $\eta_{jt}$ is observed.
    - If it is not i.i.d over time, it becomes a state variable and enters to the policy for $m_{jt}$, violating the scalar unobserved heterogeneity assumption.
3. $k_{jt}$ is realized, an unexpected shock is observed, $l_{jt}$ is determined, $\omega_{jt}$ is observed, $m_{jt}$ and $i_{jt}$ are determined, and $\eta_{jt}$ is observed (Ackerberg [2016] recommends this assumption).
    - In this case, the unexpected shock can be serially correlated, because it suffices to know $k_{jt}$, $i_{jt}$, $l_{jt}$ to decide $m_{jt}$. It does not have to predict the future unexpected shock based on the realization of the current shock because $m_{jt}$ is a static decision.
    - This changes the optimal policy function of $m_{jt}$ (3.35) to:

$$m_{jt} = m(k_{jt}, \omega_{jt}, l_{jt}). \tag{3.61}$$

    - The first step:

$$
\begin{aligned}
y_{jt} &= \beta_0 + \beta_l l_{jt} + \beta_k k_{jt} + h(k_{jt}, m_{jt}, l_{jt}) + \eta_{jt} \\
&= \psi(k_{jt}, m_{jt}, l_{jt}) + \eta_{jt}. \\
&\Rightarrow \mathbb{E}\{y_{jt} - \psi(k_{jt}, m_{jt}, l_{jt}) | k_{jt}, m_{jt}, l_{jt}\} = 0.
\end{aligned}
\tag{3.62}
$$

- The second step:

$$y_{jt} = \beta_0 + \beta_l l_{jt} + \beta_k k_{jt} + g[\psi(k_{j,t-1}, m_{j,t-1}, l_{j,t-1}) - \beta_0 - \beta_l l_{j,t-1} - \beta_k k_{j,t-1}] + \nu_{jt} + \eta_{jt}$$
$$\Rightarrow \mathbb{E}\{y_{jt} - \beta_0 - \beta_l l_{jt} - \beta_k k_{jt} - g[\psi(k_{j,t-1}, m_{j,t-1}, l_{j,t-1}) - \beta_0 - \beta_l l_{j,t-1} - \beta_k k_{j,t-1}] | k_{j,t-1}, i_{j,t-1}, l_{j,t-1}, m_{j,t-1}\}$$

(3.63)

- $m_{jt}$ has to be excluded from the production function, i.e., it has to be a value-added production function. Otherwise, $\beta_m m_{jt}$ and $\beta_m m_{j,t-1}$ appear in the second step. Because $m_{jt}$ is correlated with $\nu_{jt}$, the only hope is to vary $m_{j,t-1}$. But there is no additional variation in $m_{j,t-1}$ conditional on $k_{j,t-1}$, $i_{j,t-1}$, and $l_{j,t-1}$.

### 3.3.34 Tackle Collinearity Problem: Share Regression

- How to avoid the peculiar assumptions on shocks and timing of decisions?

- How to identify gross production function avoiding the third assumption by Ackerberg et al. [2015]?

- Return to the old literature using the first-order condition.

- Let $w_t$ be wage and $p_t$ be the product price.

- Assume that the factor market is competitive.

- Then, the first-order condition for profit maximization with respect to $L_{jt}$ is:

$$P_t F_L(L_{jt}, K_{jt}) e^{\omega_{jt}} \mathbb{E} e^{\eta_{jt}} = w_t$$
$$\Leftrightarrow \frac{P_t F_L(L_{jt}, K_{jt}) e^{\omega_{jt}} \mathbb{E} e^{\eta_{jt}}}{F(L_{jt}, K_{jt})} = \frac{w_t}{F(L_{jt}, K_{jt})}$$
$$\Leftrightarrow \frac{F_L(L_{jt}, K_{jt}) L_{jt}}{F(L_{jt}, K_{jt}) e^{\eta_{jt}}} = \frac{w_t L_{jt}}{P_t \underbrace{F(L_{jt}, K_{jt}) e^{\omega_{jt}} e^{\eta_{jt}}}_{Y_{jt}}},$$

(3.64)

  where the right hand side is expenditure share to the labor, which is observed.

- Furthermore, on the left hand side, we only have $\eta_{jt}$, which is independent of inputs.

- Let $s_{jt}$ be the log of expenditure share to the labor, and take a log of the previous equation gives:

$$s_{jt} = \log[F_L(L_{jt}, K_{jt}) L_{jt} \mathbb{E} e^{\eta_{jt}} / F(L_{jt}, K_{jt})] - \eta_{jt}$$
$$= \log(\beta_l) + \ln \mathbb{E} e^{\eta_{jt}} - \eta_{jt}.$$

(3.65)

- Remember that the coefficient in the Cobb-Douglas function is equal to the expenditure share.

- In general, share regression provides additional variation to identify the elasticity of anticipated production with respect to the labor. Then we can follow the standard OP method to recover other parameters.

## 3.4 Cost Function Estimation

### 3.4.1 Cost Function: Duality

- Given a function $y = F(x)$ such that:

- – Add factor market structure.
- – Add cost minimization.

- $\to$ There exists a unique **cost function** $c = C(y, p)$:

  - **Positivity**: positive for positive input prices and a positive.
  - **Homogeneity**: homogeneous of degree one in the input prices.
  - **Monotonicity**: increasing in the input prices and in the level of output.
  - **Concavity**: concave in the input prices.

- Given a function $c = C(y, p)$ such that:

  - **Positivity**: positive for positive input prices and a positive.
  - **Homogeneity**: homogeneous of degree one in the input prices.
  - **Monotonicity**: increasing in the input prices and in the level of output.
  - **Concavity**: concave in the input prices.

- $\to$ There exists a unique production function $F(x)$ that yields $C(y, p)$ as a solution to the cost minimization problem:

$$C(y, p) = \min_x p'x \text{ s.t. } F(x) \geq y. \tag{3.66}$$

- If the latter condition holds, the function $C$ is said to be **integrable**.

- It is rare that you can find a closed-form cost function of a production function.

- It makes sense to start from cost function.

- The duality ensures that there is a one-to-one mapping between a class of cost function and a class of production function.

- If you accept competitive factor markets and cost minimization, identifying a cost function is equivalent to identifying a production function.

- We used this idea in the last slides to identify the parameters regarding static decision variables.

- See Jorgenson [1986] for the literature in this topic up to the mid 80s.

### 3.4.2   Translog Cost Function

- One of the popular specifications:

$$
\begin{aligned}
\ln c = \alpha_0 + \alpha_p' \ln p + \alpha_y \ln y + \frac{1}{2} \ln p' B_{pp} \ln p \\
+ \ln p' \beta_{py} \ln y + \frac{1}{2} \beta_{yy} (\ln y)^2.
\end{aligned}
\tag{3.67}
$$

- It assumes that the first and second order elasticities are constant.
- A second-order (log) Taylor approximation of a general cost function.

### 3.4.3   Translog Cost Function: Integrability

- Translog cost function is known to be integrable if the following conditions hold:
- **Homogeneity**: the cost shares and the cost flexibility are homogeneity of degree zero: $B_{pp}1 = 0$, $\beta_{py}'1 = 0$.
- **Cost exhaustion**: the sum of cost shares is equal to unity: $\alpha_p'1 = 1$, $B_{pp}'1 = 0$, $\beta_{py}'1 = 0$.

- **Symmetry**: the matrix of share elasticities, biases of scale, and the cost flexibility elasticity is symmetric:

$$\begin{pmatrix} B_{pp} & \beta_{py} \\ \beta'_{py} & \beta_{yy} \end{pmatrix} = \begin{pmatrix} B_{pp} & \beta_{py} \\ \beta'_{py} & \beta_{yy} \end{pmatrix}'. \tag{3.68}$$

- **Monotonicity**: The matrix of share elasticities $B_{pp} + vv' - diag(v)$ is positive semi-definite.

### 3.4.4 Two Approaches

1. Cost data approach.
    - Use accounting cost data.
    - It does not depend on behavioral assumption.
    - One can impose restrictions of assuming cost minimization.
    - The accounting cost data may not represent economic cost.
2. Revealed preference approach.
    - Assume decision problem for firms.
    - Assume profit maximization.
    - Reveal the costs from firm's equilibrium strategy.
    - It depends on structural assumptions.
    - It reveals the cost as perceived by firms.

### 3.4.5 Cost Data Approach

- Estimating a cost function using cost data from accounting data.
- McElroy [1987] is one of the most flexible and robust frameworks.
- The approach is somewhat getting less popular in IO researchers.
- Recently, the approach is not popular among IO researchers.
- I *conjecture* one of the reasons for this is that IO researchers believe cost data taken from accounting information does not capture all the costs firms face.
- However, it is good to know the classical literature because it sometimes gives a new insight.
- cf. Byrne et al. [2015] : Propose a novel method to combine accounting cost data to estimate demand and cost function jointly without using instrumental variable approach.

### 3.4.6 Revealed Preference Approach

- Another approach is to **reveal** the marginal cost from firm's price/quantity setting behavior assuming it is maximizing profit.
    - A parameter affects economic agent's action.
    - Therefore, economic agent's action **reveals** the information about the parameter.
    - See Bresnahan [1981] and Bresnahan [1989] for reference.
- We have shown that the assumption on the factor market and cost function minimization gives restriction on the cost parameters.
- We may further assume the product market structure and profit maximization to identify cost parameters.
- Example: In a competitive market, the equilibrium price is equal to the marginal cost. Therefore, the marginal cost is identified from prices.
- What if the competition is imperfect?

### 3.4.7   Single-product Monopolist

- This approach requires researcher to specify the decision problem of a firm.

- Assume that the firm is a single-product monopolist.

- Let $D(p)$ be the demand function.

- Let $C(q)$ be the cost function.

- Temporarily, assume that we **know** the demand function.

- We learn how to estimate demand functions in coming weeks.

- The only unknown parameter is the cost function.

- The monopolist solves:

$$\max_p D(p)p - C(D(p)). \tag{3.69}$$

- The first-order condition w.r.t. $p$ for profit maximization is:

$$D(p) + pD'(p) - C'(D(p))D'(p) = 0.$$
$$\Leftrightarrow C'(D(p)) = \underbrace{\frac{D(p) + pD'(p)}{D'(p)}}_{p \text{ is observed and } D(p) \text{ is known.}} \tag{3.70}$$

- This identifies the marginal cost *at the equilibrium quantity.*

- To trace out the entire marginal cost function, you need a demand shifter $Z$ that changes the equilibrium: $D(p, Z)$.

$$C'(D(p,z)) = \frac{D(p,z) + pD'(p,z)}{D'(p,z)} \tag{3.71}$$

- This identifies the marginal cost function *at the equilibrium quantity when $Z = z$.*

- If the equilibrium quantities cover the domain of the marginal cost function when the demand shifter $Z$ moves around, then it identifies the entire marginal cost function.

### 3.4.8   Unobserved Heterogeneity in the Cost Function

- Previously we did not consider any unobserved heterogeneity in the cost function.

- Now suppose that the cost function is given by:

$$C(q) = \tilde{C}(q) + q\epsilon + \mu, \tag{3.72}$$

and $\epsilon$ and $\mu$ are not observed.

- Moreover, because it includes anticipated shocks, it is likely to be correlated with input decisions and hence the output.

- The first-order condition w.r.t. $p$ for profit maximization is:

$$D(p, z) + pD'(p, z) - [\tilde{C}'(D(p, z)) + \epsilon]D'(p, z) = 0.$$
$$\Leftrightarrow \tilde{C}'(D(p, z)) = \frac{D(p, z) + pD'(p, z)}{D'(p, z)} - \epsilon. \tag{3.73}$$

- Take the expectation conditional on $Z = z$:

$$\tilde{C}'(D(p, z)) = \frac{D(p, z) + pD'(p, z)}{D'(p, z)} - \mathbb{E}\{\epsilon|Z = z\}. \tag{3.74}$$

- If $Z$ and $\epsilon$ is independent, then the last term becomes zero and we can follow the same argument as before to trace out the marginal cost function.

### 3.4.9 Multi-product Monopolist Case

- Demand for good $j$ is $D_j(p)$ given a price vector $p$.
- Cost for producing a vector of good $q$ is $C(q)$.
- Demand function is **known** but cost function is not known.
- The monopolist solves:

$$\max_p \sum_{j=1}^{J} p_j D_j(p) - C(D_1(p), \cdots, D_J(p)). \tag{3.75}$$

- The first-order condition w.r.t. $p_i$ for profit maximization is:

$$D_i(p) + p_i \sum_{j=1}^{J} \frac{\partial D_j(p)}{\partial p_i} = \sum_{j=1}^{J} \frac{\partial C(D_1(p), \cdots, D_J(p))}{\partial q_j} \frac{\partial D_j(p)}{\partial p_i}.$$
$$= \begin{pmatrix} \frac{\partial D_1(p)}{\partial p_i} & \cdots & \frac{\partial D_J(p)}{\partial p_i} \end{pmatrix} \begin{pmatrix} \frac{\partial C(D_1(p), \cdots, D_J(p))}{\partial q_1} \\ \vdots \\ \frac{\partial C(D_1(p), \cdots, D_J(p))}{\partial q_J} \end{pmatrix} \tag{3.76}$$

- Summing up, the first-order condition w.r.t. $p$ is summarized as:

$$\begin{pmatrix} D_1(p) + p_1 \sum_{j=1}^{J} \frac{\partial D_j(p)}{\partial p_1} \\ \vdots \\ D_J(p) + p_J \sum_{j=1}^{J} \frac{\partial D_j(p)}{\partial p_J} \end{pmatrix} = \begin{pmatrix} \frac{\partial D_1(p)}{\partial p_1} & \cdots & \frac{\partial D_J(p)}{\partial p_1} \\ \vdots & & \\ \frac{\partial D_1(p)}{\partial p_J} & \cdots & \frac{\partial D_J(p)}{\partial p_J} \end{pmatrix} \begin{pmatrix} \frac{\partial C(D_1(p), \cdots, D_J(p))}{\partial q_1} \\ \vdots \\ \frac{\partial C(D_1(p), \cdots, D_J(p))}{\partial q_J} \end{pmatrix}$$
$$\Leftrightarrow \begin{pmatrix} \frac{\partial C(D_1(p), \cdots, D_J(p))}{\partial q_1} \\ \vdots \\ \frac{\partial C(D_1(p), \cdots, D_J(p))}{\partial q_J} \end{pmatrix} = \underbrace{\begin{pmatrix} \frac{\partial D_1(p)}{\partial p_1} & \cdots & \frac{\partial D_J(p)}{\partial p_1} \\ \vdots & & \\ \frac{\partial D_1(p)}{\partial p_J} & \cdots & \frac{\partial D_J(p)}{\partial p_J} \end{pmatrix}^{-1} \begin{pmatrix} D_1(p) + p_1 \sum_{j=1}^{J} \frac{\partial D_j(p)}{\partial p_1} \\ \vdots \\ D_J(p) + p_J \sum_{j=1}^{J} \frac{\partial D_j(p)}{\partial p_J} \end{pmatrix}}_{p \text{ is observed and } D(p)\text{s are known.}}. \tag{3.77}$$

- Hence, the cost function is identified.

- Including unobserved heterogeneity in the cost function causes the same problem as in the previous case.

### 3.4.10  Oligopoly

- There are firm $j = 1, \cdots, J$ and they sell product $j = 1, \cdots, J$, that is, firm = product (for simplicity).

- Consider a price setting game. When the price vector is $p$, demand for product $j$ is given by $D_j(p)$.

- The cost function for firm $j$ is $C_j(q_j)$.

- Given other firms' price $p_{-j}$, firm $j$ solves:

$$\max_{p_j} D_j(p)p_j - C_j(D_j(p)). \tag{3.78}$$

- The first-order condition w.r.t. $p_j$ for profit maximization is:

$$D_j(p) + \frac{\partial D_j(p)}{\partial p_j}p_j = \frac{\partial C_j(D_j(p))}{\partial q_j}\frac{\partial D_j(p)}{\partial p_j}.$$

$$\frac{\partial C_j(D_j(p))}{\partial q_j} = \underbrace{\frac{\partial D_j(p)}{\partial p_j}^{-1}[D_j(p) + \frac{\partial D_j(p)}{\partial p_j}p_j]}_{p \text{ is observed and } D_j(p) \text{ is known}}. \tag{3.79}$$

- In Nash equilibrium, these equations jointly hold for all firms $j = 1, \cdots, J$.]

- Including unobserved heterogeneity in the cost function causes the same problem as in the previous case.

# Chapter 4

# Assignment 1: Basic Programming in R

The deadline is the **start time of February 18 class**.

Report the following results in html format using R markdown. In other words, replicate this document. You write functions in a separate R file and put in `R` folder in the project folder. Build the project as a package and load it from the R markdown file. The execution code sholuld be written in R markdown file.

You submit:

- R file containing functions.
- R markdown file containing your answers and executing codes.
- HTML report generated from the R markdown.

## 4.1 Simulate data

Consider to simulate data from the following model and estimate the parameters from the simulated data.

$$y_{ij} = 1\{j = \text{argmax}_{k=1,2}\beta x_k + \epsilon_{ik}\},$$

where $\epsilon_{ik}$ follows i.i.d. type-I extreme value distribution, $\beta = 0.2$, $x_1 = 0$ and $x_2 = 1$.

1. To simulate data, first make a data frame as follows:

```
## # A tibble: 2,000 x 3
##        i     k     x
##    <int> <int> <dbl>
## 1     1     1     0
## 2     1     2     1
## 3     2     1     0
## 4     2     2     1
## 5     3     1     0
## 6     3     2     1
## 7     4     1     0
## 8     4     2     1
## 9     5     1     0
## 10    5     2     1
## # ... with 1,990 more rows
```

2. Second, draw type-I extreme value random variables. Set the seed at 1. You can use `evd` package to draw the variables. You should get exactly the same realization if the seed is correctly set.

```
## # A tibble: 2,000 x 4
##         i     k     x         e
##     <int> <int> <dbl>     <dbl>
## 1     1     1     0   0.281
## 2     1     2     1  -0.167
## 3     2     1     0   1.93
## 4     2     2     1   1.97
## 5     3     1     0   0.830
## 6     3     2     1  -1.06
## 7     4     1     0  -0.207
## 8     4     2     1   0.617
## 9     5     1     0   0.0444
## 10    5     2     1   1.92
## # ... with 1,990 more rows
```

3. Third, compute the latent value of each option to obtain the following data frame:

```
## # A tibble: 2,000 x 5
##         i     k     x         e  latent
##     <int> <int> <dbl>     <dbl>   <dbl>
## 1     1     1     0   0.281    0.281
## 2     1     2     1  -0.167    0.0331
## 3     2     1     0   1.93     1.93
## 4     2     2     1   1.97     2.17
## 5     3     1     0   0.830    0.830
## 6     3     2     1  -1.06    -0.863
## 7     4     1     0  -0.207   -0.207
## 8     4     2     1   0.617    0.817
## 9     5     1     0   0.0444   0.0444
## 10    5     2     1   1.92     2.12
## # ... with 1,990 more rows
```

4. Finally, compute $y$ by comparing the latent values of $k = 1, 2$ for each $i$ to obtain the following result:

```
## # A tibble: 2,000 x 6
##         i     k     x         e  latent     y
##     <int> <int> <dbl>     <dbl>   <dbl> <dbl>
## 1     1     1     0   0.281    0.281      1
## 2     1     2     1  -0.167    0.0331     0
## 3     2     1     0   1.93     1.93       0
## 4     2     2     1   1.97     2.17       1
## 5     3     1     0   0.830    0.830      1
## 6     3     2     1  -1.06    -0.863      0
## 7     4     1     0  -0.207   -0.207      0
## 8     4     2     1   0.617    0.817      1
## 9     5     1     0   0.0444   0.0444     0
## 10    5     2     1   1.92     2.12       1
## # ... with 1,990 more rows
```

## 4.2  Estimate the parameter

Now you generated simulated data. Suppose you observe $x_k$ and $y_{ik}$ for each $i$ and $k$ and estimate $\beta$ by a maximum likelihood estimator. The likelihood for $i$ to choose $k$ ($y_{ik} = 1$) can be shown to be:

$$p_{ik}(\beta) = \frac{\exp(\beta x_k)}{\exp(\beta x_1) + \exp(\beta x_2)}.$$

Then, the likelihood of observing $\{y_{ik}\}_{i,k}$ is:

$$L(\beta) = \prod_{i=1}^{1000} p_{i1}(\beta)^{y_{i1}}[1 - p_{i1}(\beta)]^{1-y_{i1}},$$

and the log likelihood is:

$$l(\beta) = \sum_{i=1}^{1000} \{y_{i1} \log p_{i1}(\beta) + (1 - y_{i1}) \log[1 - p_{i1}(\beta)]\}.$$

1. Write a function to compute the livelihood for a given $\beta$ and data and name the function `loglikelihood_A1`.

2. Compute the value of log likelihood for $\beta = 0, 0.1, \cdots, 1$ and plot the result using `ggplot2` packages. You can use `latex2exp` package to use LaTeX math symbol in the label:
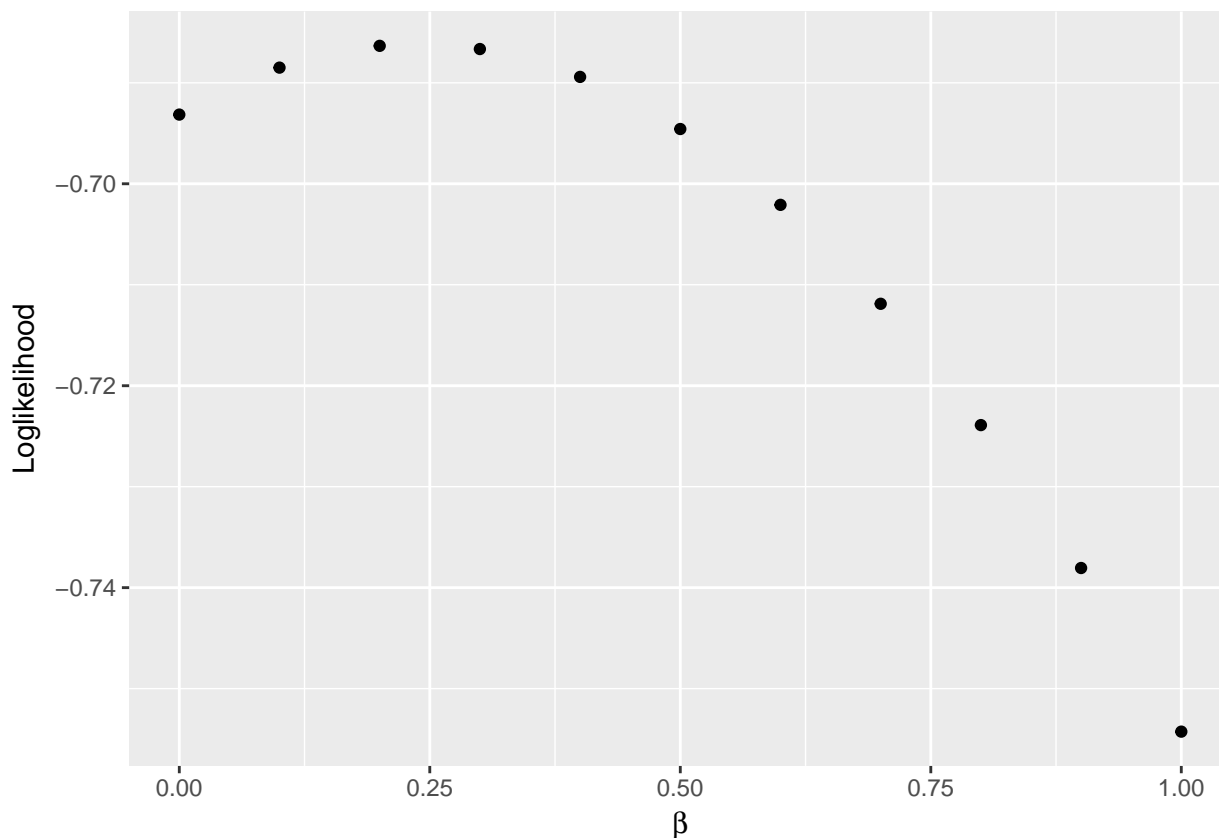


1. Find and report $\beta$ that maximizes the log likelihood for the simulated data. You can use `optim` function to achieve this. You will use `Brent` method and set the lower bound at -1 and upper bound at 1 for the parameter search.

```
## $par
## [1] 0.2371046
##
## $value
## [1] -0.6861689
##
## $counts
## function gradient
##       NA       NA
##
## $convergence
## [1] 0
##
## $message
## NULL
```

# Chapter 5

# Assignment 2: Production Function Estimation

The deadline is the **start time of February 25 class**.

Report the following results in html format using R markdown. In other words, replicate this document. You write functions in a separate R file and put in `R` folder in the project folder. Build the project as a package and load it from the R markdown file. The execution code sholuld be written in R markdown file.

## 5.1 Simulate data

Consider the following production and investment process for $j = 1, \cdots, 1000$ firms across $t = 1, \cdots, 10$ periods.

The log production function is of the form:

$$y_{jt} = \beta_0 + \beta_l l_{jt} + \beta_k k_{jt} + \omega_{jt} + \eta_{jt},$$

where $\omega_{jt}$ is an anticipated shock and $\eta_{jt}$ is an ex post shock.

The anticipated shocks evolve as:

$$\omega_{jt} = \alpha \omega_{j,t-1} + \nu_{jt},$$

where $\nu_{jt}$ is an i.i.d. normal random variable with mean 0 and standard deviation $\sigma_\nu$. The ex post shock is an i.i.d. normal random variable with mean 0 and standard deviation $\sigma_\eta$.

The product price the same across firms and normalized at 1. The price is normalized at 1. The wage $w_t$ is an i.i.d. exponential of normal random variable with mean 0 and standard deviation $\sigma_w$.

Finally, the capital accumulate according to:

$$K_{j,t+1} = (1 - \delta)K_{jt} + I_{jt}.$$

We set the parameters as follows:

| parameter | variable | value |
|---|---|---|
| $\beta_0$ | `beta_0` | 1 |
| $\beta_l$ | `beta_l` | 0.2 |
| $\beta_k$ | `beta_k` | 0.7 |
| $\alpha$ | `alpha` | 0.7 |

| parameter | variable | value |
|---|---|---|
| $\sigma_\eta$ | sigma_eta | 0.2 |
| $\sigma_\nu$ | sigma_nu | 0.5 |
| $\sigma_w$ | sigma_w | 0.1 |
| $\delta$ | delta | 0.05 |

1. Define the parameter variables as above.

2. Write a function that returns the log output given $l_{jt}$, $k_{jt}$, $\omega_{jt}$, and $\eta_{jt}$ under the given parameter values according to the above production function and name it `log_production`.

Suppose that the labor is determined after $\omega_{jt}$ is observed, but before $\eta_{jt}$ is observed, given the log capital level $k_{jt}$.

3. Derive the optimal log labor as a function of $\omega_{jt}$, $\eta_{jt}$, $k_{jt}$, and $w_t$. Write a function to return the optimal log labor given the variables and parameters and name it `log_labor_choice`.

As discussed in the class, if there is no additional variation in labor, the coefficient on the labor $\beta_l$ is not identified. Thus, if we generate labor choice from the previous function, $\beta_l$ will not be identified from the simulated data. To see this, we write a modified version of the previous function in which $\omega_{jt}$ is replaced with $\omega_{jt} + \iota_{jt}$, where $\iota_{jt}$ is an optimization error that follows an i.i.d. normal distribution with mean 0 and standard deviation 0.05. That is, the manager of the firm perceives as if the shock is $\omega_{jt} + \iota_{jt}$, even though the true shock is $\omega_{jt}$.

4. Modify the previous function by including $\iota_{jt}$ as an additional input and name it `log_labor_choice_error`.

Consider an investment process such that:

$$I_{jt} = (\delta + \gamma w_{jt})K_{jt},$$

where $I_{jt}$ and $K_{jt}$ are investment and capital in level. Set $\gamma = 0.1$, i.e., the investment is strictly increasing in $\omega_{jt}$. The investment function should be derived by solving the dynamic problem of a firm. But here, we just specify it in a reduced-form.

5. Define variable $\gamma$ and assign it the value. Write a function that returns the investment given $K_{jt}$, $\omega_{jt}$, and parameter values, according to the previous equation, and name it `investment_choice`.

Simulate the data first using the labor choice without optimization error and second using the labor choice with optimization error. To do so, we specify the initial values for the state variables $k_{jt}$ and $\omega_{jt}$ as follows.

6. Draw $k_{j0}$ from an i.i.d. normal distribution with mean 1 and standard deviation 0.5. Draw $\omega_{j0}$ from its stationary distribution (check the stationary distribution of AR(1) process). Draw a wage. Before simulating the rest of the data, set the seed at 1.

```
## # A tibble: 1,000 x 5
##        j     t     k   omega  wage
##    <int> <dbl> <dbl>   <dbl> <dbl>
## 1     1     1  1 1.09   0.779 0.591
## 2     2     1  1 0.582 -0.610 0.591
## 3     3     1  1 1.80   0.148 0.591
## 4     4     1  1 1.16   0.0486 0.591
## 5     5     1  1 0.590 -1.16  0.591
## 6     6     1  1 1.24   0.568 0.591
## 7     7     1  1 1.37  -1.34  0.591
## 8     8     1  1 1.29  -0.873 0.591
## 9     9     1  1 0.847  0.699 0.591
## 10   10     1  1 1.76  -0.379 0.591
## # ... with 990 more rows
```

7. Compute the labor and investment choice of period 1. For labor choice, compute both types of labor choices.

```
## # A tibble: 1,000 x 9
##        j     t     k   omega   wage     iota      l l_error       I
##    <int> <dbl> <dbl>   <dbl>  <dbl>    <dbl>  <dbl>   <dbl>   <dbl>
##  1     1     1 1 1.09    0.779  0.591 -0.0961   6.20    5.50   0.381
##  2     2     2 1 0.582  -0.610  0.591  0.0810   0.700   0.775 -0.0196
##  3     3     3 1 1.80    0.148  0.591  0.0260   5.23    5.40   0.391
##  4     4     4 1 1.16    0.0486 0.591 -0.00279  2.65    2.65   0.176
##  5     5     5 1 0.590  -1.16   0.591  0.0348   0.352   0.368 -0.120
##  6     6     6 1 1.24    0.568  0.591  0.00268  5.44    5.46   0.370
##  7     7     7 1 1.37   -1.34   0.591 -0.0655   0.560   0.516 -0.330
##  8     8     8 1 1.29   -0.873  0.591 -0.106    0.934   0.818 -0.135
##  9     9     9 1 0.847   0.699  0.591 -0.0104   4.53    4.47   0.280
## 10    10    10 1 1.76   -0.379  0.591 -0.0156   2.61    2.56   0.0702
## # ... with 990 more rows
```

8. Draw ex post shock and compute the output according to the production function for both labor without optimization error and with optimization error.

```
## # A tibble: 1,000 x 12
##        j     t     k   omega   wage     iota      l l_error       I     eta
##    <int> <dbl> <dbl>   <dbl>  <dbl>    <dbl>  <dbl>   <dbl>   <dbl>   <dbl>
##  1     1     1 1 1.09    0.779  0.591 -0.0961   6.20    5.50   0.381  0.0773
##  2     2     2 1 0.582  -0.610  0.591  0.0810   0.700   0.775 -0.0196  0.259
##  3     3     3 1 1.80    0.148  0.591  0.0260   5.23    5.40   0.391  -0.161
##  4     4     4 1 1.16    0.0486 0.591 -0.00279  2.65    2.65   0.176  -0.321
##  5     5     5 1 0.590  -1.16   0.591  0.0348   0.352   0.368 -0.120   0.187
##  6     6     6 1 1.24    0.568  0.591  0.00268  5.44    5.46   0.370   0.361
##  7     7     7 1 1.37   -1.34   0.591 -0.0655   0.560   0.516 -0.330  -0.0113
##  8     8     8 1 1.29   -0.873  0.591 -0.106    0.934   0.818 -0.135   0.377
##  9     9     9 1 0.847   0.699  0.591 -0.0104   4.53    4.47   0.280   0.316
## 10    10    10 1 1.76   -0.379  0.591 -0.0156   2.61    2.56   0.0702  0.100
## # ... with 990 more rows, and 2 more variables: y <dbl>, y_error <dbl>
```

9. Repeat this procedure for $t = 1, \cdots 10$ by updating the capital and anticipated shocks, and name the resulting data frame df_T.

```
## # A tibble: 10,000 x 13
##        j     t     k   omega   wage     iota      l l_error       I     eta
##    <int> <dbl> <dbl>   <dbl>  <dbl>    <dbl>  <dbl>   <dbl>   <dbl>   <dbl>
##  1     1     1 1 1.09    0.779  0.591 -0.0961   6.20    5.50   0.381  0.0773
##  2     2     2 1 0.582  -0.610  0.591  0.0810   0.700   0.775 -0.0196  0.259
##  3     3     3 1 1.80    0.148  0.591  0.0260   5.23    5.40   0.391  -0.161
##  4     4     4 1 1.16    0.0486 0.591 -0.00279  2.65    2.65   0.176  -0.321
##  5     5     5 1 0.590  -1.16   0.591  0.0348   0.352   0.368 -0.120   0.187
##  6     6     6 1 1.24    0.568  0.591  0.00268  5.44    5.46   0.370   0.361
##  7     7     7 1 1.37   -1.34   0.591 -0.0655   0.560   0.516 -0.330  -0.0113
##  8     8     8 1 1.29   -0.873  0.591 -0.106    0.934   0.818 -0.135   0.377
##  9     9     9 1 0.847   0.699  0.591 -0.0104   4.53    4.47   0.280   0.316
## 10    10    10 1 1.76   -0.379  0.591 -0.0156   2.61    2.56   0.0702  0.100
## # ... with 9,990 more rows, and 3 more variables: y <dbl>, y_error <dbl>,
## #   nu <dbl>
```

10. Check the simulated data by making summary table.

|         | N     | Mean         | Sd          | Min        | Max          |
|---------|-------|--------------|-------------|------------|--------------|
| j       | 10000 | 500.5000000  | 288.6894251 | 1.0000000  | 1000.0000000 |
| t       | 10000 | 5.5000000    | 2.8724249   | 1.0000000  | 10.0000000   |
| k       | 10000 | 0.9809540    | 0.5827349   | -1.2846624 | 3.4715184    |
| omega   | 10000 | -0.0042518   | 0.6947047   | -2.4722527 | 2.5482112    |
| wage    | 10000 | 1.2999738    | 1.0034241   | 0.2412463  | 2.9228370    |
| iota    | 10000 | -0.0000634   | 0.0502874   | -0.1841453 | 0.1715419    |
| l       | 10000 | 3.9635027    | 8.0655395   | 0.0095922  | 139.0681340  |
| l_error | 10000 | 3.9729395    | 8.1144243   | 0.0095479  | 151.8869846  |
| I       | 10000 | 0.1801717    | 0.3180590   | -1.3075169 | 5.9394281    |
| eta     | 10000 | 0.0012772    | 0.1999334   | -0.7650371 | 0.7455922    |
| y       | 10000 | 2.4763938    | 2.1954979   | -1.6282224 | 32.1811842   |
| y_error | 10000 | 2.4782811    | 2.2032667   | -1.6281667 | 34.5350569   |
| nu      | 10000 | -0.0013588   | 0.4990855   | -2.1513907 | 1.8253882    |

## 5.2   Estimate the parameters

For now, use the labor choice with optimization error.

1. First, simply regress $y_{jt}$ on $l_{jt}$ and $k_{jt}$ using the least square method. This is likely to give an upwardly biased estimates on $\beta_l$ and $\beta_k$. Why is it?

```
##
## Call:
## lm(formula = y_error ~ l_error + k, data = df_T)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.8221 -0.4208  0.0063  0.4281  2.2634
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.793901   0.012732   62.35   <2e-16 ***
## l_error     0.236968   0.000845  280.44   <2e-16 ***
## k           0.757343   0.011766   64.37   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.65 on 9997 degrees of freedom
## Multiple R-squared:  0.913,  Adjusted R-squared:  0.913
## F-statistic: 5.244e+04 on 2 and 9997 DF,  p-value: < 2.2e-16
```

2. Second, take within-transformation on $y_{jt}$, $l_{jt}$, and $k_{jt}$ and let $\Delta y_{jt}$, $\Delta l_{jt}$, and $\Delta k_{jt}$ denote them. Then, regress $\Delta y_{jt}$ on $\Delta l_{jt}$, and $\Delta k_{jt}$ by the least squares method.

```
##
## Call:
## lm(formula = dy_error ~ -1 + dl_error + dk, data = df_T_within)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.45979 -0.34453 -0.00146  0.34631  1.98491
##
## Coefficients:
```

```
##            Estimate Std. Error t value Pr(>|t|)
## dl_error 0.2256116  0.0007631 295.667   <2e-16 ***
## dk       0.0168781  0.0348015   0.485    0.628
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5264 on 9998 degrees of freedom
## Multiple R-squared:  0.8974, Adjusted R-squared:  0.8974
## F-statistic: 4.373e+04 on 2 and 9998 DF,  p-value: < 2.2e-16
```

Next, we attempt to estimate the parameters using Olley-Pakes method. Estimate the first-step model of Olley-Pakes method:

$$y_{jt} = \beta_0 + \beta_1 l_{jt} + \phi(k_{jt}, I_{jt}) + \eta_{jt},$$

by approximating $\phi_t$ by a kernel function.

Remark that $\phi$ in general depends on observed and unobserved state variables. For this reason, in theory, $\phi$ should be estimated for each period. In this exercise, we assume $\phi$ is common across periods because we know that there is no unobserved state variables in the true data generating process. Moreover, we do not include $w_t$ because we know that it is i.i.d. in the true data generating process. Do not forget to consider them in the actual data analysis.

You can use `npplreg` function of `np` package to estimate a partially linear model with a multivariate kernel. You first use `npplregbw` to obtain the optimal band width and then use `npplreg` to estimate the model under the optimal bandwidth. The computation of the optimal bandwidth is time consuming.

   3. Return the summary of the first stage estimation and plot the fitted values against the data points.

```
##
## Partially Linear Model
## Regression data: 10000 training points, in 5 variable(s)
## With 3 linear parametric regressor(s), 2 nonparametric regressor(s)
##
##                   y(z)
## Bandwidth(s): 0.08208261 0.07116851
##
##                   x(z)
## Bandwidth(s): 0.093082611 9.482612e-02
##               0.008783221 2.232608e+06
##               0.805988383 4.831148e-02
##
##              l_error          k        I
## Coefficient(s): 0.1957426 -7.135889 8.365059
##
## Kernel Regression Estimator: Local-Constant
## Bandwidth Type: Fixed
##
## Residual standard error: 0.2414186
## R-squared: 0.9882013
##
## Continuous Kernel Type: Second-Order Gaussian
## No. Continuous Explanatory Vars.: 2
```
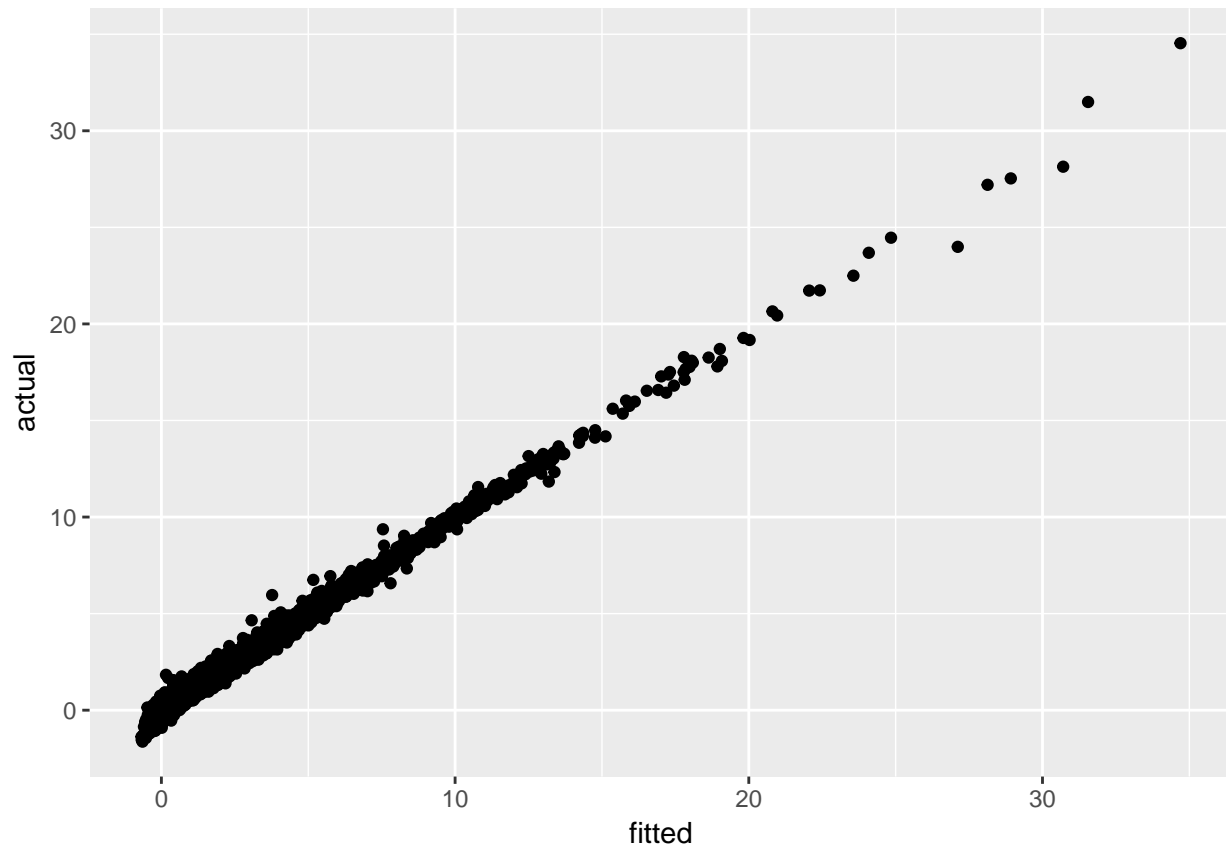
4. Check that $\beta_l$ is not identified with the data without optimization error. Estimate the first stage model of Olley-Pakes with the labor choice without optimization error and report the result.

```
## 
## Partially Linear Model
## Regression data: 10000 training points, in 5 variable(s)
## With 3 linear parametric regressor(s), 2 nonparametric regressor(s)
## 
##                      y(z)
## Bandwidth(s): 0.2245037 0.02527988
## 
##                       x(z)
## Bandwidth(s): 2.854165e-01 0.127411417
##               1.107858e-02 0.472389899
##               2.725632e+05 0.006739356
## 
##                          l         k         I
## Coefficient(s): 0.1787756  1.082508 -56.73451
## 
## Kernel Regression Estimator: Local-Constant
## Bandwidth Type: Fixed
## 
## Residual standard error: 0.3470795
## R-squared: 0.9764468
## 
## Continuous Kernel Type: Second-Order Gaussian
## No. Continuous Explanatory Vars.: 2
```

Then, we estimate the second stage model of Olley-Pakes method:

$$y_{jt} - \hat{\beta}_l l_{jt} = \beta_0 + \beta_k k_{jt} + \alpha[\hat{\phi}(k_{j,t-1}, I_{j,t-1}) - \beta_0 - \beta_k k_{jt}] + \nu_{jt} + \eta_{jt}.$$

In this model, we do not have to non-parametetrically estimate the conditional expectation of $\omega_{jt}$ on $\omega_{j,t-1}$, because we know that the anticipated shock follows an AR(1) process. Remark that we in general have to non-parametrically estimate it.

The model is non-linear in parameters, because of the term $\alpha\beta_0$ and $\alpha\beta_k$. We estimate $\alpha$, $\beta_0$, and $\beta_k$ by a GMM estimator. The moment is:

$$g_{JT}(\alpha, \beta_0, \beta_k) \equiv \frac{1}{JT} \sum_{j=1}^{J} \sum_{t=1}^{T} \{y_{jt} - \hat{\beta}_l l_{jt} - \beta_0 - \beta_k k_{jt} - \alpha[\hat{\phi}(k_{j,t-1}, I_{j,t-1}) - \beta_0 - \beta_k k_{jt}]\} \begin{bmatrix} k_{jt} \\ k_{j,t-1} \\ I_{j,t-1} \end{bmatrix}.$$

5. Using the estimates in the first step, compute:

$$y_{jt} - \hat{\beta}_l l_{jt},$$

and:

$$\hat{\phi}(k_{j,t-1}, I_{j,t-1}),$$

for each $j$ and $t$ and save it as a data frame names df_T_1st.

```
## # A tibble: 10,000 x 4
##          j      t y_error_tilde phi_t_1
##      <int> <dbl>         <dbl>   <dbl>
## 1      1      1          2.64  NA
## 2      1      2          2.22   2.62
## 3      1      3          3.13   2.73
## 4      1      4          2.50   3.21
## 5      1      5          3.08   2.58
## 6      1      6          2.16   3.17
## 7      1      7          1.30   2.54
## 8      1      8          1.17   1.40
## 9      1      9          1.88   0.851
## 10     1     10          1.87   1.56
## # ... with 9,990 more rows
```

6. Compute a function that returns the value of $g_{JT}(\alpha, \beta_0, \beta_k)$ given parameter values, data, and df_T_1st, and name it moment_OP_2nd. Show the values of the first five rows of the moment evaluated at the true parameters.

```
## [1] -0.007100535 -0.006621758 -0.009029895
```

Based on the moment, we can define the objective function of a generalized method of moments estimator with a weighting matrix $W$ as:

$$Q_{JT}(\alpha, \beta_0, \beta_k) \equiv g_{JT}(\alpha, \beta_0, \beta_k)' W g_{JT}(\alpha, \beta_0, \beta_k).$$

7. Write a function that returns the value of $Q_{JT}(\alpha, \beta_0, \beta_k)$ given the vector of parameter values, data, and df_T_1st, and name it objective_OP_2nd. Setting $W$ at the identity matrix, show the value of the objective function evaluated at the true parameters.

```
##              [,1]
## [1,] 0.0001758043
```

8. Draw the graph of the objective function when one of $\alpha$, $\beta_0$, and $\beta_k$ are changed from 0 to 1 by 0.1 while the others are set at the true value. Is the objective function minimized at around the true value?
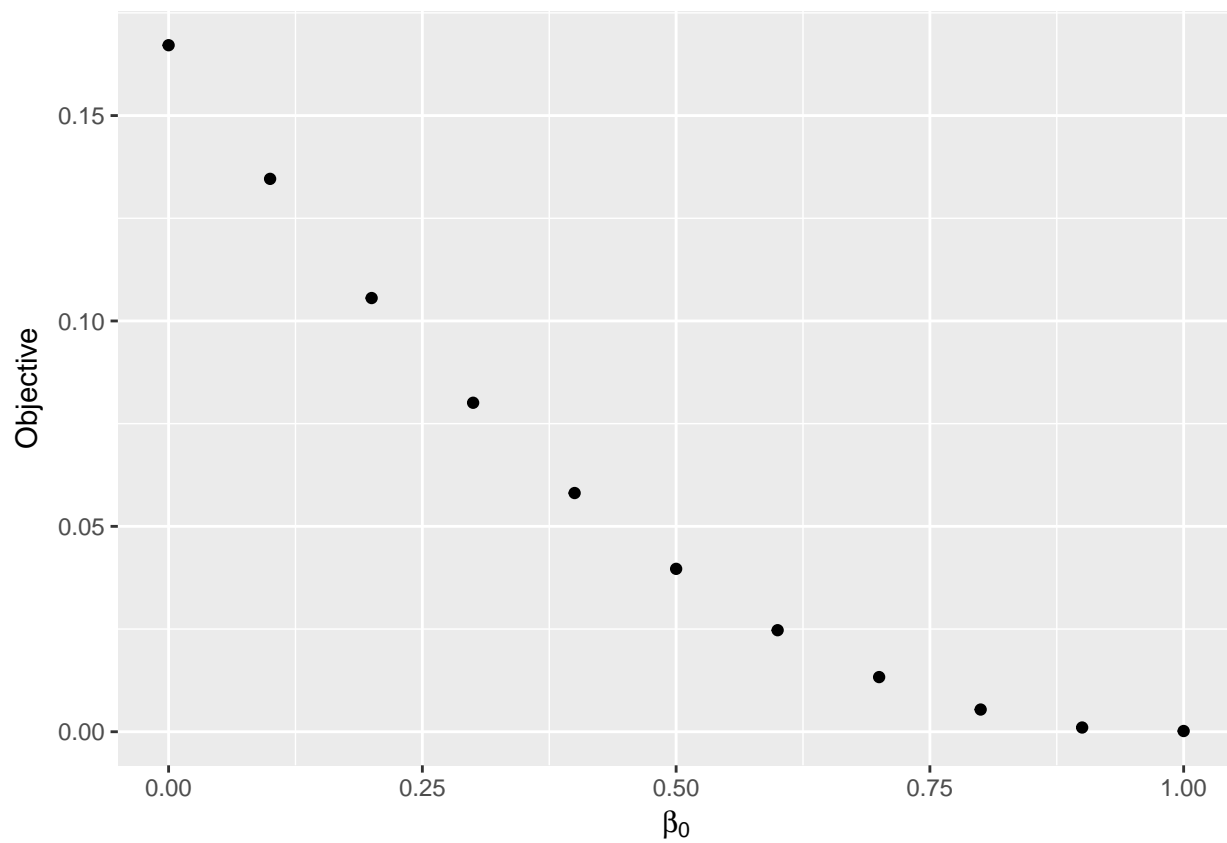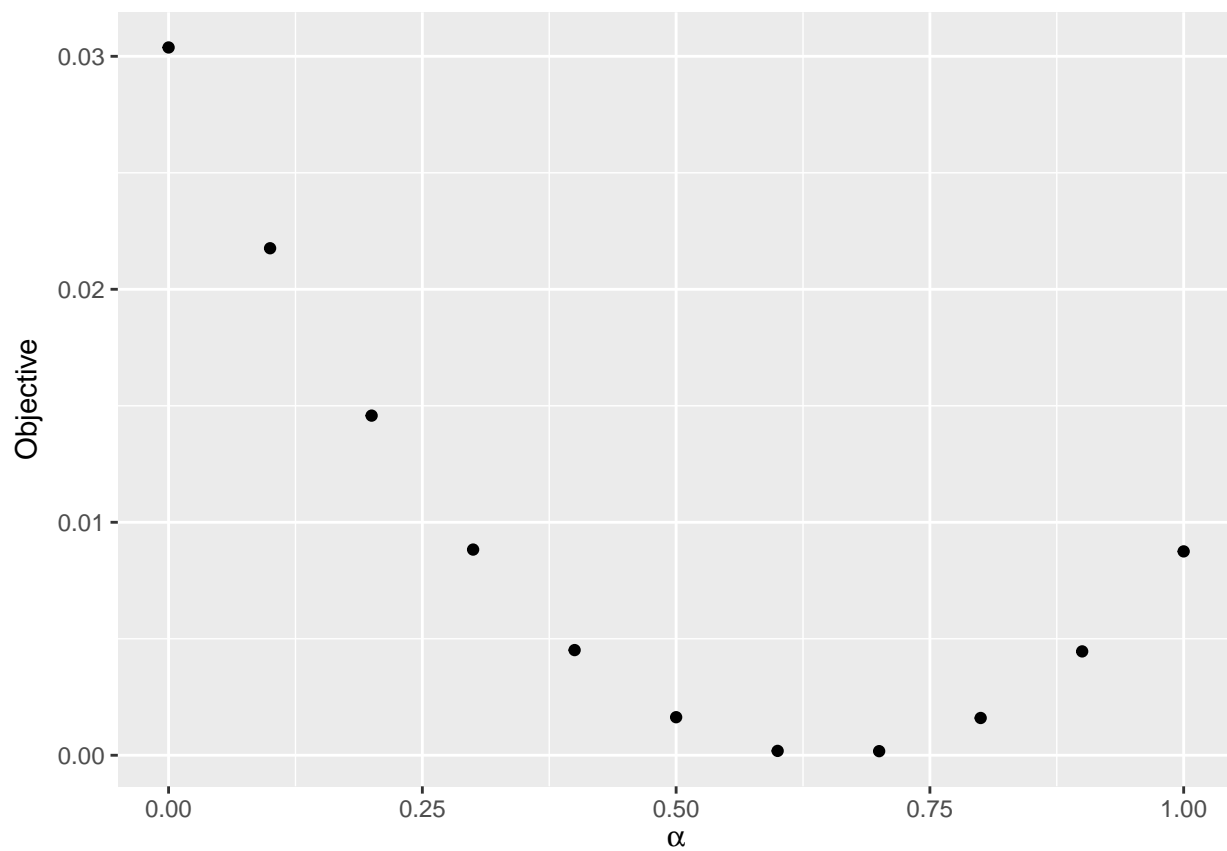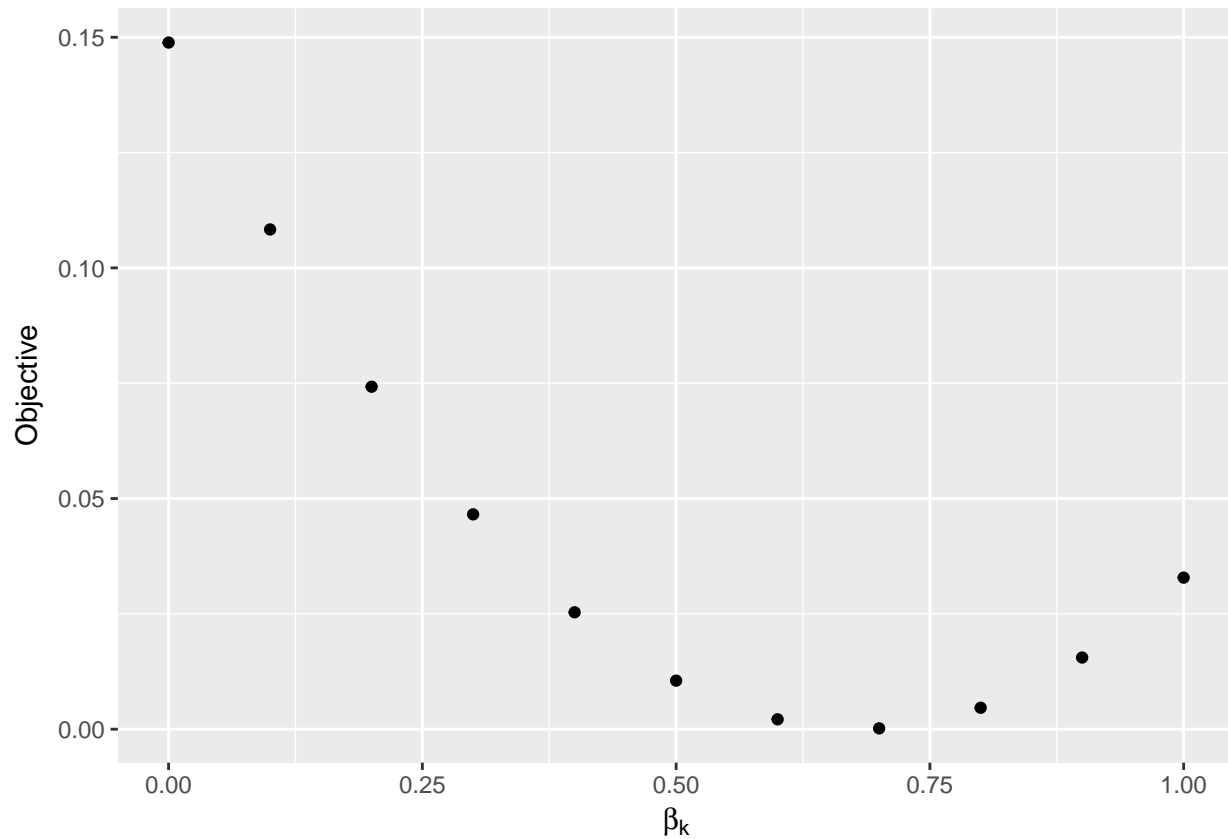
9. Find the parameters that minimize the objective function using `optim`. You may use `L-BFGS-B` method to solve it.

```
## $par
## [1] 0.6539508 1.0021382 0.6968459
##
## $value
## [1] 1.720679e-06
##
## $counts
## function gradient
##       11       11
##
## $convergence
## [1] 0
##
## $message
## [1] "CONVERGENCE: REL_REDUCTION_OF_F <= FACTR*EPSMCH"
```

# Bibliography

Daniel A Ackerberg. Timing Assumptions and E¢ ciency: Empirical Evidence in a Production Function Context. Technical report, 2016. URL http://economics.mit.edu/files/12328.

Daniel A. Ackerberg, Kevin Caves, and Garth Frazer. Identification Properties of Recent Production Function Estimators. *Econometrica*, 83(6):2411–2451, nov 2015. ISSN 0012-9682. doi: 10.3982/ECTA13408. URL https://www.econometricsociety.org/doi/10.3982/ECTA13408.

Hyungtaik Ahn and James L. Powell. Semiparametric estimation of censored selection models with a nonparametric selection mechanism. *Journal of Econometrics*, 58(1-2):3–29, jul 1993. ISSN 0304-4076. doi: 10.1016/0304-4076(93)90111-H. URL https://www.sciencedirect.com/science/article/pii/030440769390111H.

N. Bloom and J. Van Reenen. Measuring and Explaining Management Practices Across Firms and Countries. *The Quarterly Journal of Economics*, 122(4):1351–1408, nov 2007. ISSN 0033-5533. doi: 10.1162/qjec.2007.122.4.1351. URL https://academic.oup.com/qje/article-lookup/doi/10.1162/qjec.2007.122.4.1351.

Serguey Braguinsky, Atsushi Ohyama, Tetsuji Okazaki, and Chad Syverson. Acquisitions, Productivity, and Profitability: Evidence from the Japanese Cotton Spinning Industry. *American Economic Review*, 105(7):2086–2119, jul 2015. ISSN 0002-8282. doi: 10.1257/aer.20140150. URL http://pubs.aeaweb.org/doi/10.1257/aer.20140150.

Timothy F. Bresnahan. Departures from marginal-cost pricing in the American automobile industry: Estimates for 1977–1978. *Journal of Econometrics*, 17(2):201–227, nov 1981. ISSN 0304-4076. doi: 10.1016/0304-4076(81)90027-0. URL https://www.sciencedirect.com/science/article/pii/0304407681900270.

Timothy F. Bresnahan. Chapter 17 Empirical studies of industries with market power. In *Handbook of Industrial Organization*, volume 2, pages 1011–1057. Elsevier, jan 1989. ISBN 9780444704351. doi: 10.1016/S1573-448X(89)02005-4. URL https://www.sciencedirect.com/science/article/pii/S1573448X89020054.

David P Byrne, Susumu Imai, Neelam Jain, Vasilis Sarafidis, and Masayuki Hirukawa. Identification and Estimation of Differentiated Products Models using Market Size and Cost Data. Technical report, 2015. URL https://www.city.ac.uk/{_}{_}data/assets/pdf{_}file/0019/286012/blpcost15{_}final2{_}Byrne{_}Imai{_}Jain{_}Sarafidis{_}Hirukawa.pdf.

Xiaohong Chen. Chapter 76 Large Sample Sieve Estimation of Semi-Nonparametric Models. *Handbook of Econometrics*, 6:5549–5632, jan 2007. ISSN 1573-4412. doi: 10.1016/S1573-4412(07)06076-X. URL https://www.sciencedirect.com/science/article/pii/S157344120706076X.

Clayton M. Christensen. *The innovator's dilemma : when new technologies cause great firms to fail*. HarperBusiness, New York, 1997. ISBN 142219602X. URL https://books.google.co.jp/books?hl=ja{&}lr=lang{_}ja{%}7Clang{_}en{&}id=3JnBAgAAQBAJ{&}oi=fnd{&}pg=PR4{&}dq=innovator{%}27s+dilemma{&}ots=G4y61Qdl{_}2{&}sig=Yy7KYncJzItvyAfACOzITQcg49s{#}v=onepage{&}q=innovator'sdilemma{&}f=false.

Flavio Cunha, James J. Heckman, and Susanne M. Schennach. Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Econometrica*, 78(3):883–931, may 2010. ISSN 0012-9682. doi: 10.3982/ECTA6551. URL http://doi.wiley.com/10.3982/ECTA6551.

Jan De Loecker. Product Differentiation, Multiproduct Firms, and Estimating the Impact of Trade Liberalization on Productivity. *Econometrica*, 79(5):1407–1451, sep 2011. ISSN 0012-9682. doi: 10.3982/ECTA7617. URL http://doi.wiley.com/10.3982/ECTA7617.

U. Doraszelski and J. Jaumandreu. R&D and Productivity: Estimating Endogenous Productivity. *The Review of Economic Studies*, 80(4):1338–1383, oct 2013. ISSN 0034-6527. doi: 10.1093/restud/rdt011. URL https://academic.oup.com/restud/article-lookup/doi/10.1093/restud/rdt011.

Amit Gandhi, Salvador Navarro, David Rivers, Dan Ackerberg, Richard Blundell, Juan Esteban Carranza, Allan Collard-Wexler, Ulrich Do-Raszelski, Steven Durlauf, Jeremy Fox, Silvia Goncalves, Phil Haile, Jinyong Hahn, Joel Horowitz, Jean-Francois Houde, Lance Lochner, Aureo De Paula, Amil Petrin, Mark Roberts, Nicolas Roys, Todd Stinebrickner, Chad Syverson, Chris Taber, and Quang Vuong. On the Identification of Gross Output Production Functions. Technical report, 2017. URL http://publish.uwo.ca/{~}drivers2/research/GNR{_}5{_}17{_}17.pdf.

Nicola Gennaioli, Rafael La Porta, Florencio Lopez-de Silanes, and Andrei Shleifer. Human Capital and Regional Development. *The Quarterly Journal of Economics*, 128(1):105–164, feb 2013. doi: 10.1093/qje/qjs050. URL https://academic.oup.com/qje/article-lookup/doi/10.1093/qje/qjs050.

Zvi Griliches and Jacques Mairesse. Production Functions: The Search for Identification. In Steinar Strom, editor, *Econometrics and economic theory in the twentieth century: The Ragnar Frisch Centennial Symposium*. Cambridge University Press, Cambridge, MA, 1998. doi: 10.3386/w5067. URL http://www.nber.org/papers/w5067.pdf.

Jonathan E Haskel, Sonia C Pereira, and Matthew J Slaughter. Does Inward Foreign Direct Investment Boost the Productivity of Domestic Firms? *Review of Economics and Statistics*, 89(3):482–496, aug 2007. ISSN 0034-6535. doi: 10.1162/rest.89.3.482. URL http://www.mitpressjournals.org/doi/10.1162/rest.89.3.482.

Cheng Hsiao. Chapter 4 Identification. *Handbook of Econometrics*, 1:223–283, jan 1983. ISSN 1573-4412. doi: 10.1016/S1573-4412(83)01008-9. URL https://www.sciencedirect.com/science/article/pii/S1573441283010089.

Chang-Tai Hsieh and Peter J. Klenow. Misallocation and Manufacturing TFP in China and India. *Quarterly Journal of Economics*, 124(4):1403–1448, nov 2009. doi: 10.1162/qjec.2009.124.4.1403. URL https://academic.oup.com/qje/article-lookup/doi/10.1162/qjec.2009.124.4.1403.

Hidehiko Ichimura and Petra E. Todd. Chapter 74 Implementing Nonparametric and Semiparametric Estimators. *Handbook of Econometrics*, 6:5369–5468, jan 2007. ISSN 1573-4412. doi: 10.1016/S1573-4412(07)06074-6. URL https://www.sciencedirect.com/science/article/pii/S1573441207060746.

Mitsuru Igami. Estimating the Innovator's Dilemma: Structural Analysis of Creative Destruction in the Hard Disk Drive Industry, 1981–1998. *Journal of Political Economy*, 125(3):798–847, jun 2017. ISSN 0022-3808. doi: 10.1086/691524. URL https://www.journals.uchicago.edu/doi/10.1086/691524.

Dale W. Jorgenson. Chapter 31 Econometric methods for modeling producer behavior. In *Handbook of Econometrics*, volume 3, pages 1841–1915. Elsevier, jan 1986. ISBN 9780444861870. doi: 10.1016/S1573-4412(86)03011-8. URL https://www.sciencedirect.com/science/article/pii/S1573441286030118.

James Levinsohn and Amil Petrin. Estimating Production Functions Using Inputs to Control for Unobservables. *Review of Economic Studies*, 70(2):317–341, apr 2003. ISSN 0034-6527. doi: 10.1111/1467-937X.00246. URL https://academic.oup.com/restud/article-lookup/doi/10.1111/1467-937X.00246.

Charles F. Manski. Chapter 43 Analog estimation of econometric models. *Handbook of Econometrics*, 4:2559–2582, jan 1994. ISSN 1573-4412. doi: 10.1016/S1573-4412(05)80012-1. URL https://www.sciencedirect.com/science/article/pii/S1573441205800121.

Jacob Marschak and William H Andrews. Random Simultaneous Equations and the Theory of Production Author. Technical Report 4, 1944. URL https://www.jstor.org/stable/pdf/1905432.pdf?refreqid=excelsior{%}3A226527735e61586bdd17a4d63a8e7ade.

Eric Maskin and Jean Tirole. A Theory of Dynamic Oligopoly, I: Overview and Quantity Competition with Large Fixed Costs. Technical Report 3, 1988. URL https://www.jstor.org/stable/1911700?origin=crossref.

Rosa L. Matzkin. Chapter 42 Restrictions of economic theory in nonparametric methods. *Handbook of Econometrics*, 4:2523–2558, jan 1994. ISSN 1573-4412. doi: 10.1016/S1573-4412(05)80011-X. URL https://www.sciencedirect.com/science/article/pii/S157344120580011X.

Rosa L. Matzkin. Chapter 73 Nonparametric identification. *Handbook of Econometrics*, 6:5307–5368, jan 2007. ISSN 1573-4412. doi: 10.1016/S1573-4412(07)06073-4. URL https://www.sciencedirect.com/science/article/pii/S1573441207060734.

Marjorie B. McElroy. Additive General Error Models for Production, Cost, and Derived Demand or Share Systems. *Journal of Political Economy*, 95(4):737–757, aug 1987. ISSN 0022-3808. doi: 10.1086/261483. URL https://www.journals.uchicago.edu/doi/10.1086/261483.

G Steven Olley and Ariel Pakes. The Dynamics of Productivity in the Telecommunications Equipment Industry. Technical Report 6, 1996.

Chad Syverson. What Determines Productivity? *Journal of Economic Literature*, 49(2):326–365, jun 2011. ISSN 0022-0515. doi: 10.1257/jel.49.2.326. URL http://pubs.aeaweb.org/doi/10.1257/jel.49.2.326.

Jeffrey M Wooldridge. On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters*, 104:112–114, 2009. doi: 10.1016/j.econlet.2009.04.026. URL https://ac.els-cdn.com/S0165176509001487/1-s2.0-S0165176509001487-main.pdf?{_}tid=de86c166-4bc8-4227-96e2-26ca915d7f2f{&}acdnat=1548981733{_}720cc65c80b173f5ecdbd9db838e1d33.