

マイクロデータの実証分析

第3回宿題

2002年6月12日配布 6月26日回収

OLSを用いた研究例

Thomas J. Kane and Cecilia Elena Rouse (1995) "Labor-Market Returns to Two- and Four-Year College," *American Economic Review*, Vol. 85, No. 3, pp. 600-614.

米国には Community College と呼ばれる日本の専門学校や短大に当たる2年制の大学がたくさんある。それらの学校に行くことが所得をどれほど上昇させているかを調べ、そしてそのリターンを4年制大学のものと比較した論文である。

論文を読んで以下の質問に答えよ。

1. 高卒で働く人、2年制大学卒業後に働く人、あるいは4年制大学卒業後に働く人の間には学歴の違い以外に「能力」や「家庭環境」の違いも存在すると考えられる。そしてその「能力」や「家庭環境」が直接所得に影響を与えてしまう Ability Bias の問題が存在するために2年制大学のリターン、4年制大学のリターンはバイアスとともに推定されてしまう可能性がある。この際、憂慮されるバイアスは常識的に考えて上方バイアスか下方バイアスか？またバイアスの絶対値は2年制大学のリターンと4年制大学のリターンのどちらでより深刻になると考えられるか？（2点）
2. 豊富な情報量を持ったデータセットを使って、高校時代の成績、家庭環境などをコントロールして Ability Bias の問題に対処しようとしている。これらの「能力」を示すような変数がデータセットの中にはたくさんあり、高卒、2年制大学卒、4年制大学中退といったグループの変数の分布を比較するのは容易ではない。そのため彼らはいくつもの「能力」や「家庭環境」を示すと考えられる変数を束ねて一つの変数を作り、その変数の分布をグループごとに比べている。その変数とはなにか？（ヒント：Figure 2 でプロットされている変数は何か？）（2点）
3. Table 2 には所得の自然対数値を2年制大学でとった単位数と4年制大学でとった単位数に回帰している。コラム(i)では「能力」、「家庭環境」変数を含めない回帰、コラム(ii)ではそれらの変数を含めた回帰分析を行っている。その後推定された係数の大きさを比較しているが、そのねらいは何か？2年制大学のリターンの推定値はコラム(i)とコラム(ii)でほとんど変化していないが、4年制大学のリターンの推定値は大きく変化している。このことは2年制大学と4年制大学のリターンの Ability Bias の大きさに関して何を示唆しているか？1.への回答と関連づけて回答せよ。（2点）
4. 仮に彼らのデータの中に2年制大学、4年制大学で取得した単位数のデータはあるもの

の「能力」、「家庭環境」を示す変数に欠損がある標本があるとする。その場合コラム(i)のサンプル数はコラム(ii)のサンプル数よりも大きくすることが可能である。しかし彼らはあえて同じサンプルを用いてコラム(i)と(ii)の推定を行っている。コラム(i)とコラム(ii)の推定を違うサンプルを用いて行うことの問題点は何か？ (2点)

5. コラム(i)と(ii)の推定結果を見ると R^2 の値は 0.1485 と 0.1665 である。この数字をみてある統計学出身の研究者が「こんなに当てはまりが悪い回帰分析の結果は信頼に値しない。」といている。この研究の目的が 2 年制大学と 4 年制大学での学習が所得に与える因果関係の推定であることを念頭に置いて、この意見に反論せよ。 (1点)
6. コラム(ii)の 2 つの係数をみても係数の大きさそのものはほとんど同じであるが、2 年制大学の係数の標準誤差(0.0095)は 4 年制大学の係数の標準誤差(0.0042)より 2 倍ほど大きい。これはなぜだろうか？ 授業でふれた標準誤差の計算方法にふれ(ノートのないものは Wooldridge (1999) 98 ページの (3.54) 式を参照すること)可能性を 2 つ指摘せよ。その際 Figure 1 で示された取得単位数の分布も参考にすること。 (1点)