

Universidade de Brasília
Programa de Pós-Graduação em Economia
Econometria 2
2021
Prof. Victor Gomes

Lista de Exercícios 1

Data de entrega: 25/08.

Considere o modelo de equação salarial e a abordagem feita por Griliches (1976) para o problema. Blackburn e Neumark (1992) também utilizaram um subconjunto desta base para realizar análise similar. A base de dados contém indivíduos em dois pontos do tempo: o período mais cedo em que se observa salários e as observações de 1980. As variáveis principais são:

RNS : dummy para residência em Estados do sul dos EUA

MRT : dummy para estado civil (= 1 se casado)

SMSA : dummy para residência em áreas metropolitanas

MED : educação da mãe em anos

KWW : nota no teste KWW

IQ : nota no teste de QI

AGE : idade do indivíduo

S : anos de escolaridade completados

EXPR : experiência no mercado de trabalho em anos

TENURE : *tenure* em anos

LW : log dos salários

A amostra a ser utilizada é similar a proposta por Blackburn e Neumark (1992) e Hayashi (p. 251): excluindo *missing values* de educação da mãe e população negra, resultando em 758 observações. Considere o modelo de equação salarial:

$$LW_i = \beta S_i + \gamma IQ_i + \delta' \mathbf{h} + \varepsilon_i \quad (1)$$

tal que LW é o log do salário, S é escolaridade, IQ é a nota do teste de QI, e $\mathbf{h} = (EXPR, TENURE, RNS, SMSA, \text{dummies de ano})'$.

Questões:

1. (0.5 ponto) Escreva (com equações) os vies de variável omitida para a equação (1) (veja seção 3.9).
2. (0.5 ponto) Escreva (com equações) os vies de erros-nas-variáveis para a equação (1) (veja seção 3.9). A estimativa de todos os regressores é viesada na presença de erro-nas-variáveis?

3. (0.5 ponto) Faça uma tabela com média e desvio-padrão das variáveis. Faça outra tabela com a correlação entre S e IQ .
4. (0.5 ponto) Apresente uma análise da escolha dos instrumentos para resolver os problemas de viés de variável omitida e erro-nas-variáveis. Inclua nesta análise a regressão de primeiro estágio que achar pertinente.
5. (1 ponto) A taxa salarial observada é diferente entre indivíduos (datas diferentes para o primeiro ano de observação), isto implica na necessidade de formar variáveis dummies para anos. Assim dummies de anos devem ser incluídas para controlar o efeito do ano. Estime o modelo de salário (1) por 2SLS. Faça uma análise de validade dos instrumentos utilizando a estatística de Sargan. O conjunto de instrumentos deve consistir nos regressores pré-determinados e nas demais variáveis pré-determinadas (MED , KWW , MRT , AGE). Além de apresentar as estimativas do modelo por OLS e 2SLS, explique a diferença de magnitude entre elas.
6. (1 ponto) Griliches também argumenta que escolaridade pode ser endógena. Qual o argumento dele? Estime a equação salarial por 2SLS assumindo IQ e S como endógenas. O que acontece com o coeficiente de escolaridade? Como você pode explicar a diferença entre sua estimativa 2SLS de S e a estimativa anterior (sem tratar S como endógena)? Calcule a estatística de Sargan.
7. (2 pontos) Estime o mesmo modelo anterior por GMM,

$$E(\mathbf{g}) = 0,$$

tratando escolaridade S como pré-determinada. Os desvio-padrão da estimativa 2SLS são menores do que por GMM? Teste o quanto S é pre-determinada pela estatística C. Para calcular C você vai precisar dos GMMs com e sem S como instrumento. Use a mesma matriz $\hat{\mathbf{S}}$ nas duas estimativas. Para calcular $\hat{\mathbf{S}}$ use o resíduo da sua estimativa anterior do 2SLS (a estatística C deve ser 58.168).¹

8. (2 pontos) Na equação salarial com S endógena e estimada por 2SLS e por GMM, uma preocupação são as estatísticas J e Sargan elevadas. Neste caso, considere excluir ambas as variáveis MED e KWW da lista de instrumentos. A condição de ordem para identificação ainda é satisfeita? O que ocorre com a estimativa 2SLS?
9. (2 pontos) Considere o modelo de equação salarial como um sistema:

$$LW_i = \beta_1 S_i + \gamma_1 IQ_i + \delta_1' \mathbf{h} + \varepsilon_{1i}$$

$$KWW_i = \phi_2 + \beta_2 S_i + \gamma_2 IQ_i + \delta_2' \mathbf{h} + \varepsilon_{2i}$$

Estime o sistema por GMM supondo parâmetros diferentes entre as duas equações. A estimação aqui é a de um modelo com duas equações (como

¹A estimativa GMM deve ser realizada utilizando um função de minimização. Eventualmente, pode ser útil estimar 2SLS e o modelo GMM utilizando uma rotina de software (ex. `ivreg2` no Stata). Em seguida escreva o código do estimador GMM já sabendo o resultado dos parâmetros que você computou anteriormente. Para a estimação utilizando a função de minimização veja a seção **Estimação GMM 2-estágios**.

descrito no capítulo 4 de Hayashi). Se você encontrar problemas para estimar o sistema, primeiro estime o modelo equação por equação e depois como um sistema. Também estime supondo coeficiente comum para escolaridade: $\beta_1 = \beta_2$.

Estimação GMM 2-estágios

Uma rápida dica de como se estimar um GMM de 2-estágios no Matlab

%2-Step GMM

```
[theta1 fval_fmin] = fminsearch(@(theta) f_GMM_Obj(theta, DATA,
PRICES, Q),theta_0,[],[],[],[],[],options);
```

```
[theta2 fval2_fmin] = fminsearch(@(theta) f_GMM_Obj(theta, DATA,
PRICES, Q, VarMomentsTheta2),theta_0,[],[],[],[],options);
```

```
thetafinal=[theta2 Others(3)];
```

Função alternativa: `fminunc`

Exemplo de função objetivo:

$$(X'(Y - Xb))'(X(Y - Xb))$$

ou

```
obj = (OMEGA_ALL' * ZZ) * (WUSE \ (OMEGA_ALL' * ZZ)');
```

Notas:

Os outputs (parâmetros) serão os `[theta1 fval_fmin]` e `[theta2 fval2_fmin]`.

`@(theta)` significa que a rotina irá minimizar a expressão dada em `f_GMM_Obj` com respeito a `theta`.

O valor inicial da rotina para computar `@(theta)` é definido no vetor `theta_0`.

A rotina irá usar as especificações definidas em `options`. Exemplo:

```
options=optimset('TolFun',10^-12, 'MaxFunEvals', 1000000,'MaxIter',1000)
```

Referências

Blackburn, M e D. Neumark, Unobserved Ability, Efficiency Wages, and Inter-industry Wage Differentials, *Quarterly Journal of Economics*, 1992.
 Griliches, Zvi. Wages of Very Young Man. *Journal of Political Economy*, 84, 1976.