

Projeto # 20
Migração e seus determinantes no Brasil
Denisard Alves, FEA-USP
Ricardo Avelino, FEA-USP

O projeto estimará as determinantes da migração no Brasil utilizando um modelo de forma reduzida, com erros correlacionados, descrevendo um agente típico que procura maximizar o logaritmo dos ganhos líquidos dos custos de mudança na escolha entre localidades distintas. As próximas seções discutem o processo de decisão e a estimativa dos parâmetros.

Modelo

Suponha que haja K resultados potenciais Y_1^*, \dots, Y_K^* , que podem ser interpretados como o logaritmo dos salários potenciais em cada localidade. O logaritmo do salário potencial na localidade k , Y_k^* , para $k = 1, \dots, K$ é uma função linear aditiva de um conjunto de características observáveis e de um choque idiossincrático u_k , independente de X_k , isto é,

$$Y_k^* = X_k \beta_k + u_k$$

$$X_k \perp\!\!\!\perp u_k$$

Os valores contemporâneos dos choques idiossincráticos são conhecidos pelos agentes no momento em que eles tomam suas decisões, mas não observados pelo economista. De acordo com a literatura da economia do trabalho, adotamos uma especificação de Mincer ampliada para o componente observável, supondo contudo que o logaritmo do salário não necessariamente será linear na educação. Além de educação, experiência e experiência ao quadrado, permitimos que X_k inclua outras variáveis que presumivelmente explicam a dispersão nos ganhos, como sexo, raça e sindicalização.

Se a pessoa experimenta uma baixa oferta salarial na localidade corrente e decide migrar, ela se depara com custos diretos e indiretos. Estes são constituídos pelos ganhos potenciais durante o período da viagem, pelo tempo gasto no destino procurando um novo trabalho e aprendendo a nova função e por custos subjetivos como a falta de familiaridade com o destino, enquanto aqueles incluem um aumento nas despesas com alimentação, acomodação e transporte.

Os custos de mudança da localidade i para a localidade j também são modelados como uma função aditiva linear separável, de acordo com

$$C_{ij} = Z\gamma + \eta$$

$$Z \perp\!\!\!\perp \eta$$

onde η representa um choque idiossincrático conhecido pelo agente, mas não observado pelo economista, e Z é um conjunto de variáveis explicativas. Nós podemos permitir, por exemplo, um custo reduzido de mudança para uma localidade previamente visitada, refletindo uma menor incerteza acerca de variáveis como as condições locais do mercado de trabalho, permitir a possibilidade de que seja mais barato migrar para a região de origem, devido à percepções diferentes de amenidades pelos indivíduos, captada pelas diferenças de temperatura e índices pluviométricos entre as localidades i e j , e à existência de laços familiares e outras ligações, e à dependência do estado, através da inclusão do período na localidade atual entre os componentes de Z .

O salário observado pelo analista é uma função dos salários potenciais e dos custos de mudança, que não são diretamente observados. Seja

$$D_j^* = Y_j^* - \sum_{k \neq j} C_{kj} \quad \Bigg|$$

Nós observamos

$$Y = Y_j^* \Leftrightarrow D_j^* \geq D_k^* \quad \forall k \neq j \quad \Bigg|$$

Em caso de igualdade, nós assumimos que o indivíduo escolhe a localidade com o menor subscrito.

Estimação

Os erros das localidades e dos custos, por hipótese, são correlacionados. Especificamente, o vetor de choques $u = (u_1, \dots, u_k, \eta)$ segue uma distribuição normal multivariada com média zero e matriz de variância-covariância Σ

$$u \sim N(0, \Sigma)$$

Para evitar o problema de dimensionalidade associado com a integração multidimensional, nós exploramos o fato de que o modelo pode ser expresso como um conjunto de restrições lineares de desigualdade envolvendo as variáveis latentes subjacentes e apelamos para o método dos escores simulados. Seja $(Y_1^*, \dots, Y_k^*, C_{1,2}, \dots, C_{K-1,K})$. Lembremos que, supondo que a alternativa j é escolhida, Y é determinado pelo mapa

$$D(Y_j^*) = \{Y_j^* | D_j^* \geq D_k^* \quad \forall j \neq k\}.$$

O estimador de máxima verossimilhança é baseado nos escores da seguinte equação de verossimilhança:

$$l(\theta, Y) = \phi(Y - \mu, \Sigma) \int_{D(Y_j^*)} \phi(Y_{-j} - \mu_{-j}, \Sigma_{-j-j}) dY_{-j}^*$$

em que o subscrito $-j$ ($-j - j$) denota o vetor inteiro (matriz inteira), sem a j -ésima linha (e coluna), μ é a média de Y e ϕ é a função densidade de probabilidade da distribuição normal padrão.

Nós propomos a utilização do simulador de Geweke-Hajivassiliou-Keane (GHK) para gerar erros a partir da distribuição normal truncada. O simulador apresenta algumas vantagens com relação a métodos alternativos. Mesmo para probabilidades pequenas, ele é rápido e possui baixa variância. Além disso, ele é bastante preciso quando a dimensão da matriz de covariância aumenta ou quando os erros das alternativas possuem uma alta correlação. O estimador resultante é consistente desde que o número de simulações cresça mais rapidamente do que \sqrt{N} , onde N representa o tamanho da amostra.

Dados

Serão utilizados dados Censo Demográfico de 1980, 1991 e 2000 e da PNAD para vários anos. Os dados censitários permitem uma visão da evolução de longo prazo da migração no Brasil enquanto que os dados das PNADs permitirão uma análise da dinâmica migratória de curto prazo.

Referências

- Geweke, J. F., M. P. Keane, E.D. Runkle (1997), Statistical inference in the multinomial multiperiod probit model, *Journal of Econometrics* 80, pp. 125-165.
- Greenwood, M. J. (1985). Human migration: theory, models and empirical studies. *Journal of Regional Science*, 25(4), 521-544
- Hajivassiliou, V.A. and D.L. McFadden. 1998. "The Method of Simulated Scores for the Estimation of LDV Models," *Econometrica*, 66, 863-96.
- Pitt, M. M. and D. Millimet, Estimation of Coherent Demand Systems with Many Binding Non-Negativity Constraints, Discussion Paper, Southern Methodist University, Oct. 2003.