

AULA 7 (24/06) - Ernesto Amaral

1. Estimação de taxas de migração por grupos etários com dados de última etapa e data fixa
2. Modelagem de padrões migratórios
3. Aplicações ao caso brasileiro

AULA 8 (08/07) - Ernesto Amaral

1. Estimação de fluxos migratórios com estoque de crianças
2. Modelagem de migração com modelos log-lineares
3. Modelo híbrido (*model migration schedules & log-lineares*)

**ESTIMAÇÃO DE TAXAS DE MIGRAÇÃO
POR GRUPOS ETÁRIOS COM DADOS DE
ÚLTIMA ETAPA E DATA FIXA
(continuação)**

Ernesto Friedrich de Lima Amaral

**Universidade Federal de Minas Gerais
Faculdade de Ciências Econômicas
Centro de Desenvolvimento e Planejamento Regional
24 de junho de 2008**

EQUAÇÃO DE MACHADO REFORMULADA
(Amaral et al. 2004)

$$TEE_{x,y} = \frac{\sum_{i=0}^4 K_{i,y}}{0,5K_{0,y} + 1,5K_{1,y} + 2,5K_{2,y} + 3,5K_{3,y} + 4,5K_{4,y} + 4,5K_{0,y} + 3,5K_{1,y} + 2,5K_{2,y} + 1,5K_{3,y} + 0,5K_{4,y} + 5K_{m,y}}$$

TAXAS DE EMIGRAÇÃO COM DADOS DE DATA FIXA

$$TEE_{x,y} = \frac{\sum K_{i,y}}{t * \sum \left[\frac{(K_{i,t} + K_{i,t+1}) + (K_{i,t})}{2} \right]}$$

MODIFICAÇÕES NA TÉCNICA PROPOSTA POR MACHADO

- O denominador passou a incluir as pessoas que declararam residir na região i há menos de um ano, que é o componente $K0, .i$.
- Deve-se levar em consideração os efeitos indiretos da migração, ou seja, os nascimentos ocorridos de mães migrantes após o movimento migratório.
- O peso específico para o cálculo do número de pessoas-ano, definido pelo número de anos de residência na região de origem (i), deve levar em consideração o erro de referência dos Censos, que são realizados no meio do ano calendário.

DENOMINADOR DA FÓRMULA

- Os termos $K_{t,.i}$ calculam a quantidade de pessoas-ano que as pessoas que migraram de todas as regiões diferentes de i para a região i estiveram expostas ao risco de migrar em i .
- Os termos $K_{t,i.}$ calculam a quantidade de pessoas-ano que as pessoas que migraram de i para todas as regiões diferentes de i estiveram expostas ao risco de migrar em i .
- O migrante estará no denominador de duas regiões distintas:
 - 1) um sujeito que mudou de uma região diferente de i para uma região de DESTINO i será calculado como $(K_{t,.i})$.
 - 2) ele não fará parte do outro termo $(K_{t,i.})$ nessa mesma região, já que no começo do intervalo ele não estava em i . Assim, ele fará parte também do denominador da região de ORIGEM dele (que é uma região diferente de i).

FILHOS DE MULHERES MIGRANTES

- A idéia básica é a seguinte, considerem uma mulher que migrou de i para j e que teve 1 filho em i e 3 filhos em j . O filho nascido em i será computado como emigrante de i para j . No entanto, os outros 3 filhos serão considerados não-migrantes.
- O exemplo com quantidades inversas também pode acontecer. Suponha uma mulher que migrou de i para j e que teve 3 filhos em i e 1 filho em j . Os três primeiros filhos serão considerados emigrantes de i para j , e o último filho será um não-migrante.
- Há ainda casos de mulheres que têm filhos somente em i ou somente em j .

EFEITOS DIRETOS E INDIRETOS DA MIGRAÇÃO

- Quando as crianças mudam acompanhando os pais elas são consideradas migrantes (efeito direto). Quando elas nascem depois que os pais migraram é que elas são consideradas como efeito indireto da migração.
- Dessa forma, ao multiplicar os filhos tidos na origem por 2 é uma forma aproximada de corrigir essa subestimação que pode ser para mais ou para menos. Como essas crianças são multiplicadas por 2 no numerador da taxa de emigração de i para j , elas também devem ser multiplicadas por 2 no denominador dessa taxa de i para j .
- Assim, o multiplicador vai calcular o efeito indireto (crianças nascidas depois da migração dos pais) com base no efeito direto (crianças nascidas antes e que migraram com os pais).

COMPORTAMENTO DEMOGRÁFICO DAS MIGRANTES

- No mais, eu havia comentado que as minhas taxas de emigração do Sudeste para o Nordeste tinham taxas muito altas para as crianças de 0 a 4 anos de idade. Falei que isso talvez seria por causa do diferencial de fecundidade dessas áreas.
- Mas ao realizar a correção do efeito indireto, estamos levando em consideração que uma migrante recente (que mudou no intervalo de 5 anos) ainda possui os comportamentos demográficos da origem (considerando migração, fecundidade e mortalidade). Após um período de tempo maior, ela absorveria os níveis e padrões do destino.
- Além disso, quando multiplicamos os filhos que migraram de i para j por 2, na realidade estamos multiplicando filhos que elas tiveram na origem (que seria o padrão demográfico que a mulher ainda está experienciando, já que a migração foi recente).

POSSÍVEL REFINAMENTO DA TÉCNICA

- O ideal seria saber quais filhos (entre 0 e 4 anos) da mulher migrante na região de destino nasceram depois que ela se mudou. Essa seria uma tarefa bastante complicada.
- Seria preciso primeiro alocar as crianças não-migrantes para as mães (um exercício em que as observações dos bancos de dados são correlacionadas por variáveis de identificação do domicílio), e depois inserir tais crianças no numerador da taxa de emigração da origem.
- Como esse é um exercício muito mais complexo, a simples multiplicação dos 2 anos seria uma aproximação para a taxa de emigração dessas crianças.

CORREÇÃO DO ERRO DE REFERÊNCIA DO CENSO

- Quanto ao período de referência do censo, em que é necessário adicionar os 0,5 anos aos pessoas-ano expostos ao risco de emigrar, aqui vão alguns comentários.
- A data de referência do censo é o meio do ano.
- A pergunta sobre duração de residência atual pergunta à pessoa quanto tempo ela reside na localidade atual.
- Ou seja, a informação de duração da residência será mais relacionada ao ano-calendário, e não ao ano correspondente à data de referência do censo.
- Por isso, estima-se que as pessoas na realidade viveram 0,5 mais anos na região de residência atual do que a quantidade de anos reportada no censo.

MODELAGEM DE TAXAS DE MIGRAÇÃO

Ernesto Friedrich de Lima Amaral

**Universidade Federal de Minas Gerais
Faculdade de Ciências Econômicas
Centro de Desenvolvimento e Planejamento Regional
08 de julho de 2008**

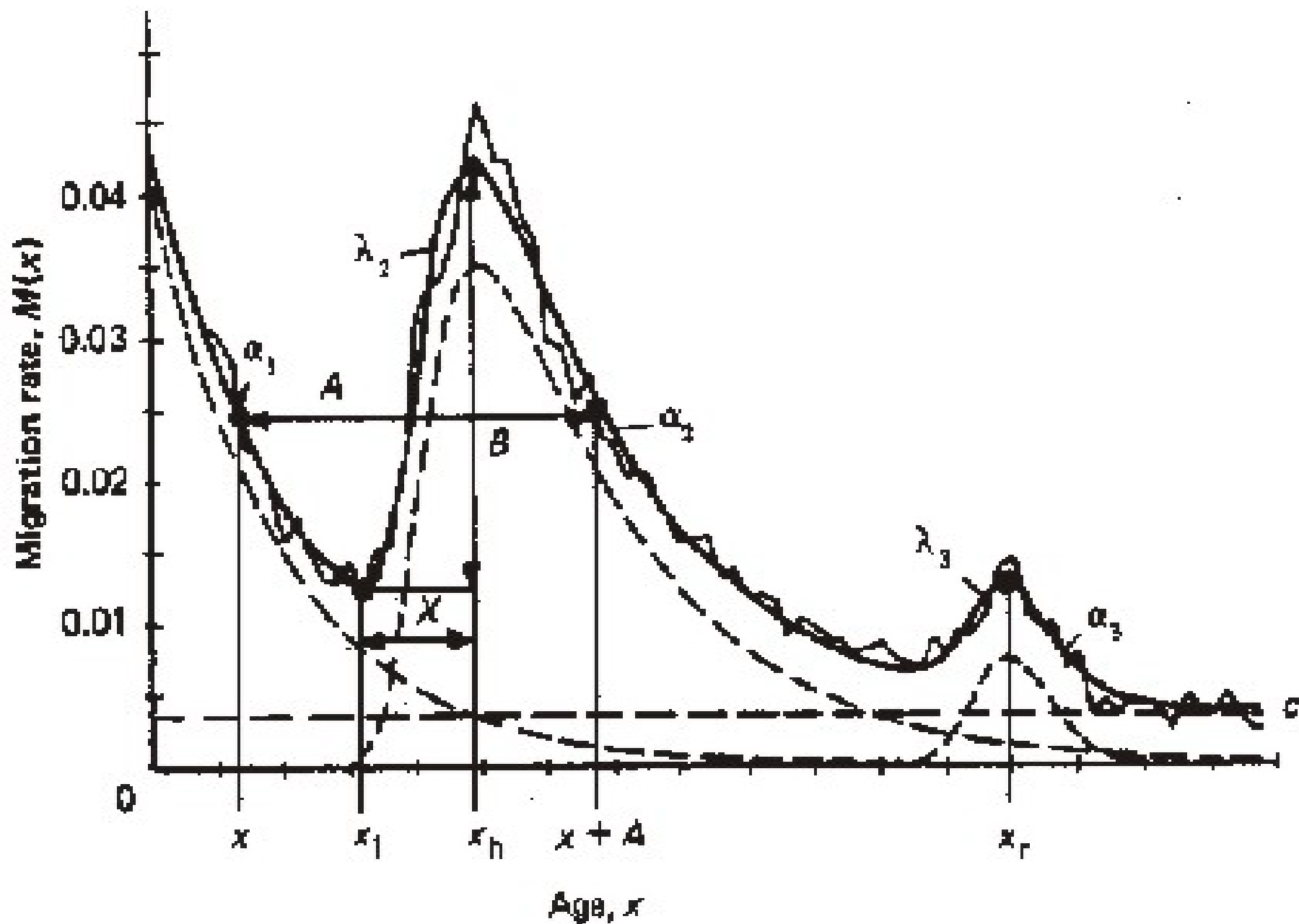
MODELAGEM DE TAXAS DE MIGRAÇÃO

- A modelagem das taxas é uma tarefa complexa, em virtude da definição dos parâmetros iniciais e variabilidade amostral.
- Essa variação ocorre porque os dados de migração são captados por amostragem, e porque a migração é um evento raro que não ocorre com a mesma frequência que a fecundidade e a mortalidade.
- Essas dificuldades ocorrem também por questões substantivas, relacionadas ao deslocamento da migração da esfera do mercado de trabalho e às limitações do modelo para se ajustar aos novos condicionantes.

FUNÇÕES DE MIGRAÇÃO DE ROGERS E CASTRO

- Rogers e Castro (1981) elaboraram modelos de migração, com a justificativa de que os comportamentos por idade variam dentro de limites para as populações.
- No estudo desses autores, a migração está bastante influenciada pelos condicionantes do mercado de trabalho, já que os padrões propostos indicam momentos diferentes da inserção dos indivíduos na economia.
- Esse modelo é muito robusto para modelar a estrutura etária das taxas de migração.
- Mesmo com valores ausentes (“missing”), ou com grande variação das taxas de um grupo etário para outro, essa metodologia é capaz de modelar as taxas para todas as idades.

CURVA DO MODELO MIGRATÓRIO DE ROGERS E CASTRO



Fonte: Rogers e Castro (1981, p.6).

FÓRMULA DA CURVA-PADRÃO DE MIGRAÇÃO DE ROGERS E CASTRO

$$M(x) = a_1 * e^{(-\alpha_1 x)} + a_2 * e^{\{-\alpha_2(x-\mu_2) - e^{[-\lambda_2(x-\mu_2)]}\}} + a_3 * e^{\{-\alpha_3(x-\mu_3) - e^{[-\lambda_3(x-\mu_3)]}\}} + c$$

- Uma curva exponencial negativa nas idades pré-laborais, com um indicador descendente α_1
- Uma parábola posicionada nas idades da população economicamente ativa, com uma média de idade μ_2 , um indicador ascendente λ_2 e descendente α_2
- Uma pequena parábola nas idades pós-laborais com uma média de idade μ_3 , um indicador ascendente λ_3 e descendente α_3
- Uma curva constante c , que é um ajuste da expressão matemática para os valores observados.

COMPONENTES DA CURVA DE ROGERS E CASTRO

– Primeiro componente é a curva pré-laboral (0 a 19 anos).

$$a_1 * e^{(-\alpha_1 x)}$$

– Segundo é uma parábola em idades laborais (20 a 40 anos).

$$a_2 * e^{\{-\alpha_2(x-\mu_2) - e^{[-\lambda_2(x-\mu_2)]}\}}$$

– Terceiro é relativo às idades pós-laborais (65 anos).

Família 1: $a_3 * e^{\{-\alpha_3(x-\mu_3) - e^{[-\lambda_3(x-\mu_3)]}\}}$ Família 3: $a_3 * e^{(\alpha_3 x)}$

– Quarto é uma curva constante que ajusta o padrão (**C**).

MEDIDAS DESCRITIVAS DAS FUNÇÕES DE MIGRAÇÃO

1.1. MEDIDAS BÁSICAS (11 parâmetros):

– **Níveis de migração em diferentes idades:**

a_1 (pré-laboral), a_2 (laboral), a_3 (pós-laboral), c (constante)

– **Localizações das médias de idades:**

μ_2 (laboral), μ_3 (pós-laboral)

– **Inclinações:**

α_1 (descendente pré-laboral),

α_2 (descendente laboral), λ_2 (ascendente laboral),

α_3 (descendente pós-laboral), λ_4 (ascendente pós-laboral)

MEDIDAS DESCRITIVAS DAS FUNÇÕES DE MIGRAÇÃO

1.2. MEDIDAS BÁSICAS (razões dos 11 parâmetros):

$\delta_{1c}=a_1/c$ Nível de crianças em relação à constante.

$\delta_{12}=a_1/a_2$ Razão de dependência jovem. Quanto menor o valor, menor o valor de dependência de crianças, e maior o domínio “laboral” da curva.

$\delta_{32}=a_3/a_2$ Razão de dependência idosa.

$\beta_{12}=\alpha_1/\alpha_2$ Índice de regularidade de mudança parental. Próximo de 1, migração familiar, crianças migrando com pais.

$\sigma_2=\lambda_2/\alpha_2$ Índice de assimetria laboral.

$\sigma_3=\lambda_3/\alpha_3$ Índice de assimetria de aposentadoria.

MEDIDAS DESCRITIVAS DAS FUNÇÕES DE MIGRAÇÃO

2. MEDIDAS DERIVADAS (propriedades padrão migratório):

– Áreas:

Taxa de Migração Total (TMT) ou “Gross Migraproduction Rate”

%(0-14) (Percentual da TMT nas idades 0-14)

%(15-64) (Percentual da TMT nas idades 15-64)

%(65+) (Percentual da TMT de 65 anos e mais)

– Localizações:

n (idade média da migração), x_l (ponto baixo), x_h (ponto alto), x_r (pico de aposentadoria)

– Distâncias:

X (diferença em anos entre x_h e x_l), B (diferença na taxa de migração entre x_h e x_l), A (mudança parental)

APLICAÇÃO DAS FUNÇÕES DE MIGRAÇÃO

- A grande dificuldade de usar a modelagem de Rogers e Castro (1981) é causada pela escolha inadequada de valores iniciais dos parâmetros (ver sugestão na tabela 17, página 42).
- Há o módulo de regressão não-linear do SPSS (Menu “Analyze”/“Regression”/“Non-linear...”), com o método de estimação Levenberg-Marquardt.
- É possível também fazer essa modelagem com o TableCurve 2D (<http://www.systat.com/products/TableCurve2D>) e o Matlab.
- No TableCurve 2D é possível:
 - * Utilizar uma interface gráfica amigável que auxilia a definição dos valores iniciais dos parâmetros.
 - * Usar os valores dos parâmetros da última tentativa, ao contrário do SPSS que mantêm os mesmos valores iniciais.
 - * Testar valores para parâmetros antes de rodar o modelo.

MODELOS LOG-LINEARES E ESTIMAÇÃO DA MIGRAÇÃO

- Raymer e Rogers (2007) propõem a estimação das estruturas por idade e espaciais de migração inter-regional com estimativas indiretas.
- O intuito é de propiciar técnicas para estimação de migração para aqueles que utilizam dados de baixa qualidade.
- Além disso, países podem apresentar a perda de informações detalhadas de migração, como é o caso dos Estados Unidos que não terão mais informações sobre fluxos populacionais nos dados censitários do questionário ampliado, o qual será substituído por uma pesquisa mensal contínua (“American Community Survey”).
- Como a migração possui fortes regularidades espaciais e por idade, um modelo multiplicativo log-linear é proposto.

MODELOS LOG-LINEARES E ESTIMAÇÃO DA MIGRAÇÃO

– O componente multiplicativo pode ser dividido em:

- * Efeito geral (“overall”): componente que representa o nível da migração.
- * Origem: componente que representa as “repulsões” relativas de cada região.
- * Destino: componente que representa as “atrações” relativas de cada região.
- * Interação origem-destino: componente que representa os impactos de distância física ou social entre lugares (e não explicados pelos efeitos acima).
- * Informação de idade do migrante também pode ser utilizada nessa modelagem.

MODELO DE COMPONENTE MULTIPLICATIVO

$$n_{ijx} = (T)(O_i)(D_j)(A_x)(OD_{ij})(OA_{ix})(DA_{jx})(ODA_{ijx})$$

- n_{ijx} : migração observada da região i para a região j na idade x
- T : numero total de migrantes (n_{++})
- O_i : proporção de todos os migrantes deixando a região i (n_{i+}/n_{++})
- D_j : proporção de todos os migrantes mudando para a região j (n_{+j}/n_{++})
- OD_{ij} : razão entre migração observada e migração esperada $n_{ij}/[(T)(O_i)(D_j)]$
- A_x : proporção de todos os migrantes no grupo etário x
- Pode ser expresso como um modelo log-linear saturado:

$$\ln(n_{ijx}) = \lambda + \lambda_i^O + \lambda_j^D + \lambda_x^A + \lambda_{ij}^{OD} + \lambda_{ix}^{OA} + \lambda_{jx}^{DA} + \lambda_{ijx}^{ODA}$$

MODELO DE COMPONENTE MULTIPLICATIVO

- Estes modelos são hierárquicos: se uma interação de dois termos está presente, os efeitos principais também devem ser incluídos; e assim por diante.
- Tabelas com fluxos de migrantes são complicadas por misturar migrantes com não-migrantes (zeros estruturais podem ser incluídos para remover os não-migrantes).
- Geralmente as interações com três termos não são incluídas:
 - * A estrutura da migração é capturada pelos outros termos.
 - * Interpretação é muito difícil e aplicação complicada.

CÁLCULO DO COMPONENTE ORIGEM-DESTINO

Table 3. The Spatial Structure of Interregional Migration in Mexico, 1995–2000

Origin	Destination				
	Border	North Central	Central	South	Total
A. Observed Flows					
Border	0	122,915	69,709	20,883	213,507
North Central	308,712 n_{21}	0	134,961	28,589	472,262 n_{2+}
Central	278,185	219,251	0	199,803	697,239
South	89,973	89,041	201,156	0	380,170
Total	676,870 n_{+1}	431,207	405,826	249,275	1,763,178 n_{++}
B. Multiplicative Components					
Border	0.000	2.354	1.419	0.692	0.121
North Central	1.703	0.000	1.242	0.428	0.268
Central	1.039 OD_{21}	1.286	0.000	2.027	0.395
South	0.616	0.958	2.299	0.000	0.216
Total	0.384	0.245	0.230	0.141	1,763,178

Note: Numbers refer to Mexican-born persons.

Fonte: Raymer e Rogers (2007, p.207).

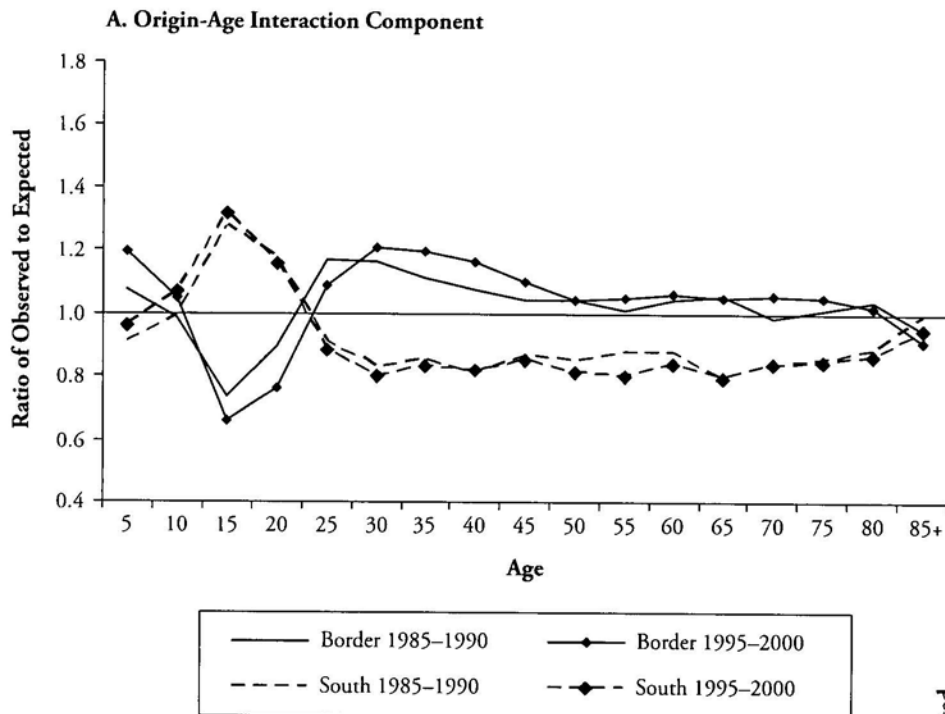
$$n_{21} = (T)(O_2)(D_1)(OD_{21}) = n_{++} \left(\frac{n_{2+}}{n_{++}} \right) \left(\frac{n_{+1}}{n_{++}} \right) \left(\frac{n_{21}}{(n_{++}) \left(\frac{n_{2+}}{n_{++}} \right) \left(\frac{n_{+1}}{n_{++}} \right)} \right) = (1.763.178) \left(\frac{472,262}{1.763.178} \right) \left(\frac{676.870}{1.763.178} \right) \left(\frac{308.712}{181.452} \right)$$

$$n_{21} = (1.763.178)(0,268)(0,384)(1,703) = 308.712$$

OD_{21} informa razão entre migração observada sobre migração esperada. Há 70% mais migrantes do que o esperado.

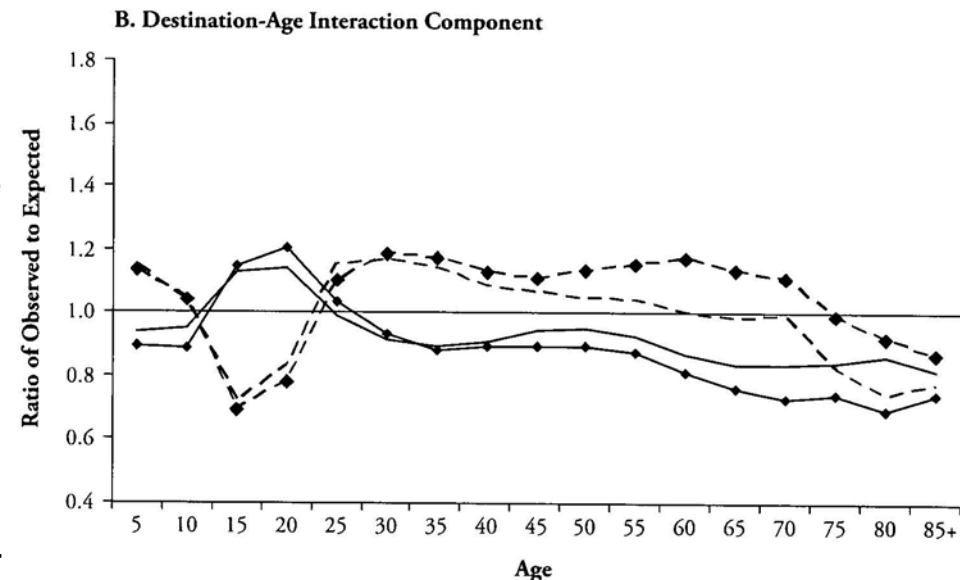
PADRÕES DE MIGRAÇÃO ORIGEM-IDADE E DESTINO-IDADE

Figure 6. The Origin-Age and Destination-Age Interaction Components of Interregional Migration in Mexico, 1985–1990 and 1995–2000: Border and South Regions



- Maior emigração de crianças da região de fronteira (“border”) do que esperado.
- Maior emigração de jovens (10-20) do sul do que esperado.
- Maior emigração da fronteira de pessoas de 25+ anos do que esperado.
- Menor emigração do sul de pessoas de 25+ anos do que esperado.

- Maior imigração de crianças para o sul do que esperado.
- Maior imigração de jovens (15-20) para a fronteira do que esperado.
- Menor imigração para a fronteira de pessoas de 30+ anos do que esperado.
- Maior imigração para o sul de pessoas de 25+ anos do que esperado.



Fonte: Raymer e Rogers (2007, p.209).

ESTIMAÇÃO DE FLUXOS MIGRATÓRIOS COM ESTOQUE DE CRIANÇAS

- Quando dados adequados de migração não estão disponíveis, é preciso estimar os dados com técnicas indiretas.
- Rogers e Jordan (2004) propõem um método em que a população de crianças migrantes no grupo de 0 a 4 anos é usada para estimar as taxas de migração para todos os outros grupos etários.
- As crianças que vivem na data do censo em uma região j , e que nasceram em uma região i , necessariamente migraram no quinquênio anterior ao censo.
- Além disso, há o pressuposto de que essas crianças migraram somente uma vez nesse intervalo, já que nasceram em média há 2 anos e meio.

COMPONENTES DA MODELAGEM COM ESTOQUE DE CRIANÇAS

- Os dados de proporção de crianças (0–4) nascidas em uma região diferente daquela de residência atual [$S_{ij}(-5)$] são usados como a primeira medida de padrão e nível de migração.
- O percentual de pessoas de todas as idades nascidas em uma região i e que residem em uma região j [${}_iK_j(+)%$] também pode ser usado para melhorar as estimativas do modelo.
- Equações de regressão são então usadas para estimar o nível e padrão da migração para os outros grupos etários.
- A equação é então representada da seguinte forma:

$$S_{ij}(x) = a + b(x)S_{ij}(-5) + c(x){}_iK_j(+)% + erro$$

REGRESSÃO LOGÍSTICA

- Para assegurar que as proporções de sobrevivência condicionais estimadas sejam sempre não-negativas, e que fiquem entre zero e um, é utilizada regressão logística.
- Assim, ao invés de estimar uma regressão linear, o logito é estimado e posteriormente convertido em proporções:

$$\ln\left(\frac{S_{ij}(x)}{1 - S_{ij}(x)}\right) = a(x) + b(x)S_{ij}(-5) + c(x)_i K_j(+)\% + erro$$

QUALIDADE DOS RESULTADOS ESTIMADOS COM ESTOQUE DE CRIANÇAS

- O categoria de idade que é melhor predita pelo modelo é para aqueles entre 20–24 anos.
- A categoria com a pior predição é para aqueles entre 80–84 anos.
- Esse resultado enfatiza que a migração de crianças (0–4) melhor prediz o fluxo dos pais do que dos avós.
- Esses resultados mostram que as regularidades dos padrões e níveis de migração permitem a estimação indireta das taxas de todo um fluxo populacional usando somente a informação de lugar de nascimento do primeiro grupo etário.

MODELAGEM DE TAXAS DE MIGRAÇÃO COM DIFERENTES METODOLOGIAS

- Há a possibilidade de modelar os dados com os “model migration schedules” antes de estimar os modelos de regressão com base na proporção de crianças.
- O contrário também pode ser realizado, em que as funções de migração de Rogers e Castro são utilizadas para melhorar as estimativas da ESTRUTURA ETÁRIA da migração obtidas com os modelos de estoque das crianças.
- Por fim, modelos loglineares podem ser utilizados após a utilização da metodologia de estoque de crianças para estimar as ESTRUTURAS ESPACIAIS da migração.

MODELAGEM HÍBRIDA

- O modelo log-linear já considera que as tabelas de migração por idade-origem-destino são mais robustas do que informação em menores níveis de desagregação.
- Mesmo com dados agregados, há o receio de que dados pouco confiáveis, ou provenientes de amostras muito pequenas, possam comprometer a estimação das taxas de migração.
- Para obter resultados ainda mais confiáveis, uma matriz com informação auxiliar (“offset”) pode ser usada para aprimorar as estimativas.
- Essas informações auxiliares podem ser fluxos passados agregados de imigração ou emigração, assim como o estoque de crianças entre 0 e 4 anos (Raymer e Rogers 2007).

EXERCÍCIO DE ESTIMAÇÃO E MODELAGEM DE TAXAS DE EMIGRAÇÃO

2. Com os dados do Censo de 2000, estimar e analisar para um fluxo regional brasileiro por sexo:

a) Taxas Específicas de Emigração (TEE_x) e Taxas de Emigração Total (TET) utilizando dados de data fixa.

b) Gráficos com TEE_x e TEE_x proporcionais por grupo etário.

c) Aplique o modelo da família 2 (7 parâmetros) proposto por Rogers e Castro (1981) a um conjunto de taxas de migração por idade estimadas no item “a”. Apresentar os valores dos parâmetros finais e das taxas modeladas em formato de tabela. Ilustrar ainda em um gráfico os valores das taxas originais por grupo etário juntamente com as taxas modeladas.

OBSERVAÇÕES E DICAS QUANTO AO EXERCÍCIO

- Não estimar o fluxo Nordeste–Sudeste e Sudeste–Nordeste apresentado em sala.
- Estimar o fluxo nas duas direções por sexo, com a estimação de quatro grupos de taxas por idade.
- Ignorar migração de crianças de 0 a 4 anos, já que não é possível realizar essa estimativa com dados de data fixa.
- O material distribuído na aula 7 (24/06/2008) apresenta exemplos de estimativas exigidos nesse exercício.
- Verificar em Rogers e Castro (1981) os valores iniciais dos parâmetros propostos por estes autores (Tabela 17, p.42).
- No site www.ernestoamaral.com/cedeplar.html há uma série de notas de aula, programas e textos para auxílio.
- Entregar sintaxes de todos os programas computacionais utilizados.