

Aplicação de um modelo de idade-período-coorte para a atividade econômica no Brasil metropolitano*

EDUARDO L. G. RIOS-NETO**
ANA MARIA H. C. OLIVEIRA***

Um método relevante para a projeção da população economicamente ativa (PEA) é a projeção das taxas específicas de participação na PEA integrando o período com a coorte. Este trabalho formaliza essas aplicações, que normalmente são efetuadas mediante estimativa de planilha, através da implementação de modelos log-lineares topológicos de taxa (rate models). A discussão teórica e metodológica sobre os problemas de identificação nos modelos de idade-período-coorte é uma contribuição do trabalho. Já a aplicação empírica do modelo para os dados da Pesquisa Mensal de Emprego (PME) nos anos 80 e 90 mostrou a eficácia da metodologia, apresentando resultados similares aos obtidos pela metodologia tradicional. A vantagem da metodologia aqui proposta é a possibilidade de se desenhar cenários mais rigorosos para a projeção da PEA.

1 - Introdução

Este trabalho visa utilizar os dados referentes ao conjunto das seis regiões metropolitanas estudadas pela Pesquisa Mensal de Emprego (PME) do IBGE, para desenvolver um modelo estatístico de idade-período-coorte. Tal modelo pretende servir de base tanto para analisar as tendências de participação na população economicamente ativa (PEA) metropolitana nas duas últimas décadas quanto para racionalizar a técnica de projeção utilizada em trabalho publicado nesta revista [ver Wajnman e Rios-Neto (1994)]. Os resultados devem possibilitar a configuração de tendências que definirão cenários alternativos para projeções futuras. Em suma, este é um trabalho de desenvolvimento metodológico que aproveita as tendências de participação na PEA metropolitana (masculina e feminina) obtidas a partir da PME para fazer uma aplicação prática do método.

* Este artigo é derivado de um relatório mais amplo, *A demografia do mercado de trabalho brasileiro*, financiado pelo Ministério do Trabalho. Os autores agradecem os comentários de dois pareceristas anônimos. Os resultados do trabalho continuam sendo de responsabilidade exclusiva dos autores.

** Professor titular do Departamento de Demografia do Cedeplar/UFMG.

*** Aluna do curso de doutorado em Demografia no Cedeplar/UFMG.

2 - Aspectos conceituais

A análise de processos demográficos ou mudança social distingue três níveis: *cross-section*, longitudinal e defasagem temporal (*time lag*). Uma análise *cross-section* entre os grupos etários não é só o resultado de processos de idade, mas também de diferenças de coorte. Isso significa que a análise *cross-section* enfatiza os efeitos de idade e de coorte (idade e coorte variam, enquanto o período é fixo). Uma análise longitudinal mede não só as diferenças ao longo do tempo para o mesmo grupo associadas aos processos de envelhecimento, mas também os resultados das mudanças sociais ou ambientais dos vários períodos em que esse processo de envelhecimento ocorre. Em outras palavras, a análise longitudinal capta os efeitos de idade e de período para uma determinada coorte (idade e período variam, enquanto a coorte é fixa). Uma análise de defasagem temporal compara uma mesma transição associada à idade em diferentes períodos e coortes. Nesse sentido, a análise de defasagem temporal capta os efeitos de período e de coorte para uma determinada idade — período e coorte variam, enquanto a idade é fixa [ver Halli e Rao (1992)].

Normalmente, os efeitos de idade em processos demográficos estão associados a processos biológicos, psicológicos e/ou a mudanças nos papéis sociais dos grupos etários. Já os efeitos de período estão geralmente associados às condições ambientais que variam com o tempo, sendo as flutuações econômicas o exemplo mais conspícuo de efeito de período. A análise de período lida com o estudo de diferentes coortes no mesmo ponto no tempo; as mudanças dependem também da composição da população em um período particular e dos processos de coorte naquele período. Os efeitos de coorte estão geralmente associados às mudanças genéticas, de tamanho, de educação dos pais, de educação da coorte e das interações históricas da coorte. Uma coorte se refere a um grupo de pessoas experimentando um evento particular durante o mesmo período de tempo; e a análise de coorte é a descrição quantitativa de ocorrências datadas aplicada ao estudo das variações temporais no comportamento populacional. Isso não é o mesmo que uma análise longitudinal, que lida com mudanças e comportamentos de indivíduos ao longo do tempo. O estudo da coorte se apóia nas noções de que pessoas de idade a no período t são aquelas que tinham idade $a-1$ no período $t-1$, e de que as transformações do mundo social modificam as pessoas de diferentes idades de diferentes maneiras de tal forma que os efeitos dessas transformações são persistentes.

2.1 - Atividade: idade, período e coorte

A divisão da população em idade ativa em dois *status* — ativa e inativa — possibilita o cálculo da taxa de atividade. Essa taxa depende da idade, do período e da coorte. Depende da idade porque a força de trabalho e os retornos do mercado de trabalho são organizados de maneiras distintas no que diz respeito à idade do

indivíduo; depende do período porque as flutuações do mercado de trabalho mudam o risco de participação na força de trabalho; e depende da coorte porque os padrões de socialização do passado produzem coortes que são únicas em seu risco de participação.

A importância da idade e do período sobre a participação na força de trabalho é mais óbvia. Quando a análise é de período — um determinado ano, por exemplo —, o risco de participação na força de trabalho varia de acordo com a idade das pessoas e com as diferenças entre as coortes. Nesse caso, os diferenciais de participação por idade refletem a trajetória longitudinal hipotética de uma coorte sintética, que é a média de todas as coortes nessa análise de tipo *cross-section*. A análise longitudinal de uma coorte reflete as variações no risco de atividade, de acordo com mudanças de período indexadas na variação de idade dos indivíduos. Sendo assim, a análise das taxas específicas de atividade (TEA) privilegia o efeito idade, que pode ser visto na perspectiva de um período fixo ou de uma coorte fixa. Quando o período é fixo, a idade reflete a diferença entre as coortes; no caso em que a coorte é fixa, a idade reflete as transições longitudinais para a atividade observadas para uma determinada coorte ao longo do ciclo de vida.

Um modelo de taxa de atividade na perspectiva de “idade-período” (IP) permite o cálculo das TEAs para cada período, enquanto o modelo de taxa de atividade “idade-coorte” (IC) permite o cálculo das TEAs para cada coorte. Um modelo de “idade-período-coorte” (IPC) seria ideal, pois incorpora as três dimensões relevantes do processo demográfico. O modelo IPC é tautológico porque configura uma identidade, no sentido de que toda vez que duas dimensões variarem, uma terceira será definida. As variações de idade e coorte definem o período; as variações de idade e período definem a coorte; e as variações de período e de coorte definem a idade. As três dimensões não podem variar simultaneamente. Apesar da identidade das três dimensões, há estratégias empíricas que permitem decompor a importância relativa de cada uma delas. Tais estratégias não devem ser vistas como características de estudos de causalidade, mas sim como estimativas dos componentes da identidade, por meio da decomposição.¹

Um modelo estatístico que viabilize a decomposição de IPC permite a replicação dos modelos IP e IC que o antecedem na árvore hierárquica de *recursibilidade*. Sendo assim, caso seja possível estimar um modelo IPC, ao fixar-se o componente P do modelo, o efeito combinado IC gera uma estimativa da TEA de período. Ao fixar-se o componente C do modelo, o efeito combinado IP fornece uma estimativa da TEA de coorte. Por outro lado, os coeficientes estimados para I, P e C fornecem os impactos puros desses vetores sobre as TEAs. Os coeficientes de I mostram a variação relativa nas TEAs decorrentes das mudanças na idade, controlando-se o impacto geral das diferenças entre as coortes e das diferenças entre os períodos; os coeficientes de P mostram o impacto geral dos vários períodos, controlando-se pelos efeitos puros de idade e coorte; e os coefi-

¹ Alguns demógrafos utilizam o termo “determinantes próximos” para esses componentes da decomposição.

cientes de C dão o impacto geral de várias coortes, controlando-se pelo efeito idade puro e o efeito período puro.

Alguns efeitos de coorte são afetados por eventos de período anteriores aos anos controlados pela análise, desde que esses eventos tenham afetado a trajetória de atividade das coortes que estavam sujeitas ao risco de atividade no momento em que a mudança de período prévio ao controle tenha ocorrido. A não ser por esse aspecto, as coortes diferem em sua trajetória de atividade devido a características próprias, como socialização, valores, nível de educação e tamanho da própria coorte. Os efeitos de período puro mostram o impacto do ambiente externo sobre a taxa de atividade. O ambiente externo mais relevante para as flutuações na taxa de atividade é determinado pelas flutuações econômicas, causadas principalmente pelos ajustes (choques de política econômica) e pelo ciclo econômico.

2.2 - Os efeitos de idade, período e coorte sobre a TEA

As variáveis de idade e de período exercem influência sobre a participação na força de trabalho. No caso da idade, a magnitude desse efeito pode ser vista da perspectiva da teoria do capital humano como um indicador de experiência ou como um indicador de posição na estrutura social. Já no caso do período, a variação da participação na força de trabalho é usualmente tomada como resultado de forças puramente econômicas, como por exemplo a mudança na demanda por ocupação ou outros tipos de flutuações no mercado de trabalho. A tentativa de estimar efeitos de período e idade é um modo de resumir os termos empíricos do papel da estrutura etária e da mudança no mercado de trabalho na caracterização da força de trabalho agregada.

A persistência de características da força de trabalho ao longo do tempo para uma determinada coorte é um “efeito de coorte”, indicando uma mudança social indiretamente observável se duas condições se aplicam. As várias coortes podem ser estatisticamente distinguidas uma da outra pelos diferentes padrões de participação, sendo cada um desses padrões característico de uma coorte particular. Essa condição pode ser verificada por meio dos parâmetros que refletem o efeito da coorte sobre a variável dependente U (γ^{CU}). O “efeito de coorte” estatístico não pode ser atribuído aos efeitos do envelhecimento (refletindo progressão da coorte através da idade) e/ou das flutuações de período (refletindo a exposição das coortes às exigências econômicas momentâneas). Esse conceito de um efeito de coorte total é um efeito parcial da categoria de coorte, controlado pelos efeitos das variáveis de idade e período. Quer dizer, os parâmetros γ^{CU} são discutidos no contexto do modelo completo $H_{(I, P, C)}$, no qual os efeitos de idade e de período são controlados estatisticamente. Os efeitos de coorte, propriedades de coortes invariantes ao longo do tempo, são tomados como indicadores da força da socialização, que opera para produzir coortes verdadeiramente distintas; assim, os

parâmetros γ^{CU} , bem como as diferenças entre eles, refletem o ritmo da mudança social.

3 - Aspectos metodológicos

A lógica dos modelos log-lineares topológicos permite identificar a dimensão de coorte numa tabela de contingência em que a margem das linhas representa o efeito principal da idade (I) e a margem das colunas representa o efeito principal do período. Utilizamos a variável idade (I) com categorias $i = 1, \dots, 18$, correspondentes aos grupos etários trienais entre 15 e 68 anos; a variável período (P) com classes $j = 1, \dots, 6$, correspondentes aos anos 1997, 1994, 1991, 1988, 1985 e 1982; e a variável dependente participação na força de trabalho (U). Implícita na tabela abaixo, 18×6 ($I \times P$) está a variável coorte (C), marcando a categoria de coorte cronológica na qual um dado indivíduo se localiza — essa é a base dos modelos topológicos ou de diagonal. A aplicação do método “idade-período-coorte” requer intervalos etários iguais; os 18 intervalos etários e seis intervalos de períodos geraram 23 coortes distintas, como mostra o quadro a seguir:

Período		1982	1985	1988	1991	1994	1997
Idade		P_6	P_5	P_4	P_3	P_2	P_1
15-17	I_{18}	C_{18}	C_{19}	C_{20}	C_{21}	C_{22}	C_{23}
18-20	I_{17}	C_{17}	C_{18}	C_{19}	C_{20}	C_{21}	C_{22}
21-23	I_{16}	C_{16}	C_{17}	C_{18}	C_{19}	C_{20}	C_{21}
24-26	I_{15}	C_{15}	C_{16}	C_{17}	C_{18}	C_{19}	C_{20}
27-29	I_{14}	C_{14}	C_{15}	C_{16}	C_{17}	C_{18}	C_{19}
30-32	I_{13}	C_{13}	C_{14}	C_{15}	C_{16}	C_{17}	C_{18}
33-35	I_{12}	C_{12}	C_{13}	C_{14}	C_{15}	C_{16}	C_{17}
36-38	I_{11}	C_{11}	C_{12}	C_{13}	C_{14}	C_{15}	C_{16}
39-41	I_{10}	C_{10}	C_{11}	C_{12}	C_{13}	C_{14}	C_{15}
42-44	I_9	C_9	C_{10}	C_{11}	C_{12}	C_{13}	C_{14}
45-47	I_8	C_8	C_9	C_{10}	C_{11}	C_{12}	C_{13}
48-50	I_7	C_7	C_8	C_9	C_{10}	C_{11}	C_{12}
51-53	I_6	C_6	C_7	C_8	C_9	C_{10}	C_{11}
54-56	I_5	C_5	C_6	C_7	C_8	C_9	C_{10}
57-59	I_4	C_4	C_5	C_6	C_7	C_8	C_9
60-62	I_3	C_3	C_4	C_5	C_6	C_7	C_8
63-65	I_2	C_2	C_3	C_4	C_5	C_6	C_7
66-68	I_1	C_1	C_2	C_3	C_4	C_5	C_6

A configuração do quadro segue uma organização proposta por Halli e Rao (1992). No quadro, há $k = 1, \dots, 23$ categorias de coorte, correspondentes ao número de diagonais com 18 categorias de idade e seis períodos; o grupo de pessoas que mais recentemente entraram na força de trabalho denota a coorte 23, e assim sucessivamente.

A análise demográfica de tabelas de “idade-período-coorte” começa usualmente com uma análise das distribuições marginais da variável dependente com relação a cada variável I, P e C . As análises informais dessas distribuições marginais apontam uma diferenciação de coorte de participação na força de trabalho, mostrando que essa diferenciação de coorte é diferente para cada sexo. Entretanto, essas comparações informais podem ser equivocadas devido à dependência funcional entre as três variáveis. Cada marginal com relação a cada uma dessas variáveis é geralmente confundida pela influência das outras, não sendo possível distinguir os efeitos de idade, período e coorte pelos métodos informais usuais, porque há uma dependência linear entre essas variáveis. Há uma controvérsia na literatura sobre as possibilidades de se distinguir dois ou três efeitos [ver Mason *et alii* (1973), Glenn (1977), Rodgers (1982), Bloom (1984) e Fienberg e Mason (1985)]. Embora os três efeitos não sejam ortogonais entre si, parte da literatura defende que é possível extrair informação de todos eles.

No caso de um modelo linear de mínimos quadrados seria impossível estimar os parâmetros na forma $U = a + b_1I + b_2P + b_3C + \varepsilon$, devido à dependência linear mencionada. Uma das alternativas seria estimar uma relação de regressão na qual cada uma dessas variáveis fosse quebrada em variáveis *dummies* dentro de cada efeito principal. Em termos da tabela de contingência, a variável conjunta IPC é equivalente à variável conjunta IP ou à IC ou à PC ; a restrição linear nas variáveis independentes da análise de coorte leva ao problema de identificação.

3.1 - O problema de identificação dos efeitos de idade, período e coorte

O problema de identificação no modelo IPC é derivado do fato de que o componente linear de qualquer conjunto de efeitos é a soma ou a diferença dos componentes lineares dos outros dois conjuntos de efeitos, porque coorte é igual a período menos idade. O problema de identificação é criado por essa dependência linear entre a idade, o período e a coorte, em face do esquema de classificação de idade e coorte nos dados de múltiplas *cross-sections*.

Mason *et alii* (1973) sugerem um meio de estimar os componentes lineares pela imposição de uma ou mais restrições lineares sobre a relação entre qualquer uma das variáveis preditoras e a variável dependente. No contexto de um modelo de múltipla classificação, se os efeitos de estar em um de dois grupos etários ou

coortes, ou períodos se restringem a permanecer iguais, o modelo é identificado e os parâmetros podem ser estimados. No contexto geral de todos os possíveis modelos aditivos, qualquer restrição linear é suficiente para assegurar que o modelo seja identificado. Glenn (1977) discute a solução proposta por Mason *et alii* (1973) para o problema de identificação, com objeções à aditividade implicada dos efeitos de idade, período e coorte nesses procedimentos e à exclusão de efeitos interativos.²

Outra possível solução para o problema é a utilização de medidas empíricas disponíveis para qualquer uma das variáveis — idade, período ou coorte —, como por exemplo uma medida de ciclo para período, ou tamanho de coorte para coorte [Fienberg e Mason (1985)]. Pode-se pensar em uma das classificações dos dados sendo mais refinada enquanto se deixa a outra como está; por exemplo, usar um intervalo de idade mais refinado, enquanto o intervalo de período é mantido constante e menos refinado. No entanto, nesse caso, perde-se a capacidade de localizar corretamente os indivíduos ao longo do tempo.

Segundo Heckman e Robb (1985), não seria possível estimar todos os coeficientes da equação sem restrições adicionais devido às interdependências conceituais entre as variáveis. Informações adicionais devem ser utilizadas para quebrar o problema de identificação, de acordo com esses autores. Rodgers (1982) também conclui pela solução do problema se pelo menos um dos indicadores — idade, período ou coorte — puder ser substituído por medidas mais diretas dos conceitos teóricos.

Juhn, Murphy e Pierce (1993), analisando a desigualdade salarial, sugerem a possibilidade de se distinguir medidas dentro de uma coorte (eliminando o efeito de coorte) e comparar essa diferença com coortes adjacentes (eliminando o efeito de idade e deixando somente uma mudança ao longo do tempo). Segundo esses autores, isso seria idêntico a comparar as mudanças médias dentro de coortes ao longo do tempo.

Outro tratamento dado ao problema de identificação é discutido por Halli e Rao (1992). Trata-se da estimativa de modelos lineares por mínimos quadrados, obtendo-se a variável dependente a partir de uma transformação logit da variável dependente com participação na PEA. Os modelos de efeitos principais são obtidos a partir das estimativas de regressões específicas para as variáveis *dummies* de idade, período e coorte — três regressões: (I), (P) e (C). Os modelos de interação de primeira ordem podem ser estimados como (IP), (IC) e (PC). O modelo de interação de segunda ordem (IPC) não produz coeficientes de regressão acurados devido ao problema de identificação. Duas alternativas de identificação são discutidas pelos autores. A primeira implica a restrição de variáveis *dummies* com o

2 Admite-se que os efeitos de período são similares para todas as coortes e que essas envelhecem de maneira semelhante; para vários tipos de variáveis, tais suposições sobre a ausência de efeitos de interação podem não ser razoáveis.

estabelecimento de um coeficiente igual a zero por variável, sendo que no modelo IPC dois coeficientes de coorte têm de ser iguais a zero. A segunda consiste em estimar os coeficientes de idade e período no modelo (IP) e, em seguida, estimar os coeficientes de coorte a partir de uma regressão com o resíduo do modelo (IP).

Como já discutido, o problema de identificação pode ser tratado por meio da “restrição de identificação” nas variáveis *dummies*. Essa restrição leva ao processo de “tradução demográfica” de taxas de período em taxas de coorte e vice-versa. A “restrição de identificação” é imposta nos parâmetros do modelo de coorte. Uma solução clássica seguida na identificação dos parâmetros seria restringir como iguais, na maioria das vezes nulos, os coeficientes estimados para duas coortes adjacentes. No contexto do modelo log-linear,³ essa saída de identificação dos parâmetros permite a estimativa dos efeitos de idade, período e coorte. O teste do ajuste da tabela de contingência permite a avaliação do poder explicativo dos componentes de idade, período e coorte.

Seguindo a metodologia proposta por Clogg (1979), a variável dependente (U) é discutida no contexto de um modelo multiplicativo ou log-linear com distribuição estatística binomial. Sendo f_{ijkl} a frequência observada na categoria de idade i , de período j , de coorte k e de classe l da variável dependente referente à participação na força de trabalho, uma razão que descreva as chances de U assumir o valor 1 relativo à classe l' é $\omega_{ijkl'} = f_{ijkl}/f_{ijkl'}$. Sendo F_{ijkl} a frequência esperada na célula (i, j, k, l) , a razão correspondente à última equação expressa em termos das frequências esperadas F_{ijkl} é $\Omega_{ijkl'} = F_{ijkl}/F_{ijkl'}$.

Considerando uma série de modelos relacionando $\Omega_{ijkl'}$ aos parâmetros que denotam os principais efeitos multiplicativos de idade, período e coorte, cada modelo a ser considerado é um caso especial do modelo completo $\Omega_{ijkl'} = \gamma_{l'}^U \gamma_{il'}^{IU} \gamma_{jl'}^{PU} \gamma_{kl'}^{CU}$ ($i = 1, \dots, 18; j = 1, \dots, 6; k = 1, \dots, 23$). Os parâmetros no lado direito dessa equação (exceto $\gamma_{l'}^U$) representam os principais efeitos multiplicativos da idade i , do período j e da coorte k sobre a variável dependente U . Os parâmetros $\gamma_{il'}^{IU}$ denotam o principal efeito da categoria etária i sobre a razão particular $\Omega_{ijkl'}$ no que se refere à variável U , e os outros parâmetros são interpretados da mesma forma. Uma vez que esse modelo completo permite os principais efeitos das três variáveis independentes, o denotamos por $H_{(l, p, c)}$.

Um modo alternativo e equivalente de se ver o modelo $H_{(l, p, c)}$ é sugerido pela equação que relaciona as frequências esperadas F_{ijkl} aos outros parâmetros multiplicativos relacionados àqueles do lado direito da última equação $F_{ijkl} = \tau_l^U \tau_{ij}^{IP} \tau_{il}^{IU} \tau_{jl}^{PU} \tau_{kl}^{CU}$, sendo $\gamma_{il'}^{IU} = \tau_{il}^{IU} / \tau_{il'}^{IU}$, e expressões similares para

3 Halli e Rao (1992) aplicam essa restrição de identificação na estimativa de modelos lineares por mínimos quadrados.

$\gamma_{jl'}^{PU} \gamma_{kl'}^{CU}$. Os parâmetros γ são usados para estimar as frequências esperadas sob os modelos e os parâmetros τ são compelidos a obedecer à restrição $\prod_l \tau_{il}^{AU} = \prod_l \tau_{jl}^{PU} = \prod_l \tau_{kl}^{CU} = 1$.

Outros modelos a serem testados para os dados são:

$$\Omega_{ijkl'} = \gamma_{l'}^U \gamma_{il'}^{IU} \quad (H_{(I)}) \quad \Omega_{ijkl'} = \gamma_{l'}^U \gamma_{il'}^{IU} \gamma_{jl'}^{PU} \quad (H_{(I,P)})$$

$$\Omega_{ijkl'} = \gamma_{l'}^U \gamma_{jl'}^{PU} \quad (H_{(P)}) \quad \Omega_{ijkl'} = \gamma_{l'}^U \gamma_{il'}^{IU} \gamma_{kl'}^{CU} \quad (H_{(I,C)})$$

$$\Omega_{ijkl'} = \gamma_{l'}^U \gamma_{kl'}^{CU} \quad (H_{(C)}) \quad \Omega_{ijkl'} = \gamma_{l'}^U \gamma_{jl'}^{PU} \gamma_{kl'}^{CU} \quad (H_{(P,C)})$$

O modelo $H_{(I)}$ expressa a *odds* $\Omega_{ijkl'}$ somente em termos dos efeitos principais de idade, fixando os parâmetros γ^{PU} e γ^{CU} em 1 e similarmente para os outros modelos; esses podem ser representados em termos dos parâmetros τ . Os modelos apresentados são identificados e podemos usar as estatísticas χ^2 de cada modelo para obter as contribuições totais dos parâmetros γ^U , γ^{PU} e γ^{CU} . O problema de identificação, mencionado anteriormente, também aparece aqui. Devido à restrição da categoria de coorte em relação à idade e ao período, um problema especial surge com o modelo completo $H_{(I,P,C)}$, porque os parâmetros desse modelo não são identificados, mesmo se as restrições usuais são impostas.

3.2 - O método de estimativa utilizado

O modelo de estimativa utilizado segue a tradição dos modelos log-lineares generalizados e a perspectiva de especificação proposta por Clogg (1979), conforme a discussão apresentada. O modelo utilizado é de contagem (*count model*), com a função de ligação logarítmica e uma distribuição de Poisson.⁴

A classe dos modelos lineares generalizados é uma extensão dos modelos lineares tradicionais que permitem que a média de uma população dependa de um preditor linear por meio de uma função de ligação não-linear e que a distribuição de probabilidade da resposta seja membro de uma família exponencial de distri-

4 Em vez da binomial apresentada por Clogg (1979). Um modelo logito com distribuição binomial também foi estimado com resultados similares. O modelo de Poisson foi escolhido para valorizar uma aplicação do "modelo taxa". A distribuição de Poisson desse modelo de contagem foi originalmente aplicada na análise de eventos raros. Em termos de mercado de trabalho, as taxas de desemprego são particularmente adequadas.

buições.⁵ No caso, os dados não têm distribuição normal, que é contínua e não adequada para modelar contagens ou proporções discretas. Um modelo linear generalizado consiste dos seguintes componentes: um preditor linear definido como nos modelos lineares tradicionais: $\eta_i = x_i'\beta$; e uma função de ligação g que descreve como o valor esperado de y_i se relaciona com o preditor linear η_i : $g(\mu_i) = x_i'\beta$. As variáveis-resposta y_i são independentes para $i = 1, 2, \dots$, e têm uma distribuição de probabilidade pertencente à família exponencial.⁶

Um modelo linear generalizado é construído a partir da decisão sobre as variáveis-resposta e as variáveis explicativas e da escolha de uma função de ligação apropriada e uma distribuição de probabilidade da resposta. A regressão de Poisson em modelos log-lineares tem como variável-resposta uma contagem, com distribuição de Poisson e uma função de ligação logarítmica $\eta = \log(\mu)$. Isso quer dizer que o parâmetro médio de Poisson μ é relacionado ao preditor linear através de $\log(\mu) = x_i'\beta$. A distribuição de Poisson é usada para modelar a distribuição de contagens nas caselas em uma tabela de contingência de múltiplas entradas. O logaritmo de N (população) é usado como um *offset*, ou seja, uma variável da regressão com um coeficiente constante de 1 para cada observação. Nesse caso, a variável de *offset* serve para normalizar as médias ajustadas das caselas a uma base individual. A relação log-linear especificada entre a média e os preditores por meio da função de ligação logarítmica assegura resultados positivos. O ajuste do modelo linear generalizado aos dados usa uma estimação de máxima verossimilhança para o vetor de parâmetros β . Cada termo no modelo é chamado de efeito, que pode ser uma variável categórica ou discreta ou uma variável contínua.

O modelo estimado foi desenvolvido para estimar taxas, e também é conhecido como “modelo de taxas” (*rate model*). Esse modelo é estimado por meio de uma regressão de Poisson em que o logaritmo das pessoas ativas é a variável dependente, mas o das pessoas em risco entra no lado direito da equação, com uma restrição de *offset*, significando que o coeficiente é assumido constante e com valor unitário para essa variável — essa restrição garante a estimativa das taxas.

Suponhamos que μ seja o valor esperado da variável U (pessoas ativas) de acordo com uma distribuição de Poisson. A variável t é um índice de população em risco de atividade. Agresti (1991) sugere a seguinte especificação para o modelo: $\log(\mu / t) = \alpha + \beta X$, que pode também ser representado pela expressão: $\mu = t \cdot \exp(\alpha + \beta X)$. O modelo de idade-período-coorte com efeitos principais teria a seguinte especificação:

5 Exemplos: modelos lineares clássicos com erros normais, modelos logísticos e probito para dados binários, modelos log-lineares para dados multinomiais.

6 Isso implica que a variância da resposta depende da média μ mediante uma função de variância V : $\text{var } y_i = \phi V(\mu_i)/\omega_i$, onde ϕ é o parâmetro de dispersão constante e ω_i é um peso conhecido para cada observação.

$$\log(\mu/t) = \alpha + \beta^I X^I + \beta^P X^P + \beta^C X^C$$

Esse modelo está sujeito às mesmas restrições de identificação discutidas anteriormente. No caso desse exercício, o coeficiente das duas coortes mais jovens (23 e 22) é restringido ao valor nulo.⁷

3.3 - Base de dados e variáveis

O estudo empírico aqui apresentado é pautado na elaboração própria de tabulações especiais da PME do IBGE, por intermédio da geração das séries anuais da Taxa de Atividade Econômica controlando-se por sexo e idade, cobrindo o período 1982/97. O corte simultâneo da taxa de atividade por sexo e idade é inédito em relação às estatísticas regularmente publicadas pelo IBGE e pela maioria dos boletins de mercado de trabalho. As variáveis de coorte foram construídas a partir dessa série.

A discussão aqui se restringe a observações de seis períodos de tempo na história da força de trabalho do grupo etário de 15 a 68 anos: 1982, 1985, 1988, 1991, 1994 e 1997. A variável idade é medida em grupos trienais para permitir um maior aproveitamento da série histórica e minimizar as flutuações erráticas que caracterizariam a análise de intervalos etários anuais. Para viabilizar o modelo topológico, o intervalo etário deve correr exatamente ao espaçamento de períodos; portanto, a utilização de todos os anos da série da PME implicaria a opção por intervalos etários anuais e a idade simples tende a ser bastante errática com esse tamanho de amostra. As coortes são caracterizadas pelas pessoas que nasceram num mesmo triênio, para ficar coerente com os intervalos etários e de flutuação do período. O padrão etário de progressão das taxas de atividade de uma coorte poderá ser inferido pela observação da experiência parcial dessa coorte ao longo do tempo, ou seja, ao longo dos seis triênios deste estudo.

O uso de amostras estratificadas pode afetar os resultados estimados. Tal fato é mais relevante no caso de modelos log-lineares dos tipos aplicados aqui, pois as informações utilizadas são obtidas a partir de tabelas de contingência. Por causa disso, as tabelas geradas utilizaram os pesos de expansão amostral sugeridos pelo IBGE.

⁷ A literatura argumenta que é preferível assumir essa igualdade entre as duas coortes mais velhas, principalmente quando há mudanças substantivas entre as mais jovens. A restrição foi imposta nas coortes mais jovens por facilidade computacional; estimativas futuras aplicarão a restrição nas coortes mais velhas.

4 - A escolha do melhor ajuste

Um aspecto importante dos modelos lineares generalizados é a seleção das variáveis explicativas. Mudanças nas estatísticas de qualidade do ajuste são usadas para avaliar a contribuição de subconjuntos de variáveis explicativas para um modelo. O desvio (*deviance*), definido como duas vezes a diferença entre a máxima log verossimilhança atingível e a log verossimilhança do modelo, é geralmente usado como uma medida de qualidade do ajuste. A máxima log verossimilhança atingível é alcançada com um modelo que tem um parâmetro para cada observação. Uma estratégia para seleção das variáveis é ajustar uma seqüência de modelos, começando com um modelo simples com somente um termo de intercepto e depois incluir uma variável explicativa adicional em cada modelo sucessivamente. Pode-se mensurar a importância da variável explicativa adicional por meio da diferença nos desvios ou log verossimilhanças ajustados entre os modelos sucessivos.

Os dados apresentados na Tabela 1 representam todas as tentativas de ajuste do modelo de taxa para a TEA das mulheres, obtida pelo critério do desvio (*deviance*) entre o valor predito e o valor observado em cada célula da tabela de contingência. A significância dos modelos é medida pela diferença na *deviance* entre os dois modelos, de acordo com os valores de dG^2 que seguem uma distribuição qui-quadrado. Os problemas de *overdispersion* e de tamanho da amostra fazem com que os testes de ajuste para determinação do melhor modelo não sejam muito confiáveis. Entretanto, a inspeção visual do valor do falso R^2 já é um indicador dos melhores modelos.

Os resultados mostram que a dimensão de idade (modelo I) apresenta o maior poder preditivo da TEA feminina: 88,6%. O modelo (C) apresenta o segundo maior poder preditivo (66,9%), enquanto a dimensão temporal do modelo de período (P) explica um componente muito baixo das taxas de atividade feminina. O modelo de idade-coorte explica uma grande proporção da dispersão da tabela de contingência (91%). O modelo de independência condicional ou modelo de efeitos principais (IPC) explica 91,2% da dispersão observada na tabela de contingência.

Os resultados de ajuste de modelo obtidos na Tabela 2 relativos aos homens vão na mesma direção dos resultados das mulheres. O poder explicativo de (I) é maior, seguido do modelo (C), e finalmente com o baixo poder explicativo do modelo (P). O modelo de independência condicional ou efeitos principais explica 93,6% da dispersão da tabela. Esse bom ajuste dos dois modelos será muito importante para o sucesso das projeções das TEAs masculina e feminina.

TABELA 1
Ajuste dos modelos — mulheres

Modelos	G^2	$g.l.$	dG^2	$dg.l.$	R^2 (%)
0 – Nulo	152.495,2	647			
1 – Idade	17.427,6	625	135.067,6	22	88,6
2 – Período	145.700,2	637	6.795,1	10	4,5
3 – Coorte	50.544,9	620	101.950,3	27	66,9
4 – Idade x período	17.243,8	620	135.251,5	27	88,7
5 – Idade x coorte	13.756,1	603	138.739,1	44	91,0
6 – Período x coorte	49.777,5	615	102.717,7	32	67,4
7 – Idade x período x coorte (independência condicional)	13.472,5	599	139.022,8	48	91,2
1 – Idade	17.427,6	625			
4 – Idade x período	17.243,8	620	183,9	5	1,1
5 – Idade x coorte	13.756,1	603	3.671,5	22	21,1
7 – Idade x período x coorte	13.472,5	599	3.955,2	26	22,7
2 – Período	145.700,2	637			
4 – Idade x período	17.243,8	620	128.456,4	17	88,2
6 – Período x coorte	49.777,5	615	95.922,7	22	65,8
7 – Idade x período x coorte	13.472,5	599	132.227,7	38	90,8
3 – Coorte	50.544,9	620			
5 – Idade x coorte	13.756,1	603	36.788,8	17	72,8
6 – Período x coorte	49.777,5	615	767,4	5	1,5
7 – Idade x período x coorte	13.472,5	599	37.072,4	21	73,3
4 – Idade x período	17.243,8	620			
7 – Idade x período x coorte	13.472,5	599	3.771,3	21	21,9
5 – Idade x coorte	13.756,1	603			
7 – Idade x período x coorte	13.472,5	599	283,6	4	2,1
6 – Período x coorte	49.777,5	615			
7 – Idade x período x coorte	13.472,5	599	36.305,0	16	72,9

FONTE: Elaboração própria a partir de tabulações especiais da PME.

OBS.: G^2 = desvio (*deviance*); $g.l.$ = graus de liberdade; dG^2 = diferença entre os desvios (modelo menos modelo básico); $dg.l.$ = diferença entre os graus de liberdade; $R^2 = dG^2/(G^2$ do modelo básico), representando a melhoria da qualidade do ajuste do modelo em relação ao modelo básico.

TABELA 2
Ajuste dos modelos — homens

Modelos	G^2	$g.l.$	dG^2	$dg.l.$	$R^2(\%)$
0 – Nulo	139.991,8	647			
1 – Idade	12.546,6	625	127.445,2	22	91,0
2 – Período	136.216,6	637	3.775,3	10	2,7
3 – Coorte	53.643,8	620	86.348,0	27	61,7
4 – Idade x período	11.452,3	620	128.539,6	27	91,8
5 – Idade x coorte	8.850,3	603	131.141,5	44	93,7
6 – Período x coorte	53.042,3	615	86.949,5	32	62,1
7 – Idade x período x coorte (independência condicional)	8.727,4	599	131.264,4	48	93,8
1 – Idade	12.546,6	625			
4 – Idade x período	11.452,3	620	1.094,4	5	8,7
5 – Idade x coorte	8.850,3	603	3.696,3	22	29,5
7 – Idade x período x coorte	8.727,4	599	3.819,3	26	30,4
2 – Período	136.216,6	637			
4 – Idade x período	11.452,3	620	124.764,3	17	91,6
6 – Período x coorte	53.042,3	615	83.174,3	22	61,1
7 – Idade x período x coorte	8.727,4	599	127.489,2	38	93,6
3 – Coorte	53.643,8	620			
5 – Idade x coorte	8.850,3	603	44.793,5	17	83,5
6 – Período x coorte	53.042,3	615	601,5	5	1,1
7 – Idade x período x coorte	8.727,4	599	44.916,4	21	83,7
4 – Idade x período	11.452,3	620			
7 – Idade x período x coorte	8.727,4	599	2.724,9	21	23,8
5 – Idade x coorte	8.850,3	603			
7 – Idade x período x coorte	8.727,4	599	122,9	4	1,4
6 – Período x coorte	53.042,3	615			
7 – Idade x período x coorte	8.727,4	599	44.314,9	16	83,5

FONTE: Elaboração própria a partir de tabulações especiais da PME.

OBS.: G^2 = desvio (*deviance*); $g.l.$ = graus de liberdade; dG^2 = diferença entre os desvios (modelo menos modelo básico); $dg.l.$ = diferença entre os graus de liberdade; $R^2 = dG^2/(G^2$ do modelo básico), representando a melhoria da qualidade do ajuste do modelo em relação ao modelo básico.

5 - Análise gráfica dos coeficientes estimados a partir dos modelos

Uma análise comparativa por sexo do impacto da estrutura etária no perfil das taxas de período é apresentada no Gráfico 1. O impacto medido pela razão sobre a taxa do grupo etário mais jovem (15 a 17 anos) é apresentado a partir da estimativa do modelo (IP). Esse gráfico mostra um perfil etário bastante aproximado do perfil de atividade metropolitana observado em média nas décadas de 80 e 90.

A análise do Gráfico 2 mostra o perfil etário das TEAs masculina e feminina estimado a partir do modelo de independência condicional (IPC). O perfil etário das mulheres é bastante mais inclinado do que o estimado para os homens. Esse perfil mais inclinado das mulheres mostra uma crescente entrada feminina na atividade econômica ao longo da sua trajetória longitudinal de vida. O perfil também é crescente para a população masculina, mas o seu ritmo de crescimento é bem mais baixo.

Os dados apresentados no Gráfico 3 mostram o impacto das várias coortes, identificadas pelos anos em que faziam parte do intervalo etário de 15 a 17 anos. O gráfico mostra que as coortes com 15 a 17 anos a partir de meados da década de 60 começam a apresentar níveis crescentes de taxa de atividade, ritmo que

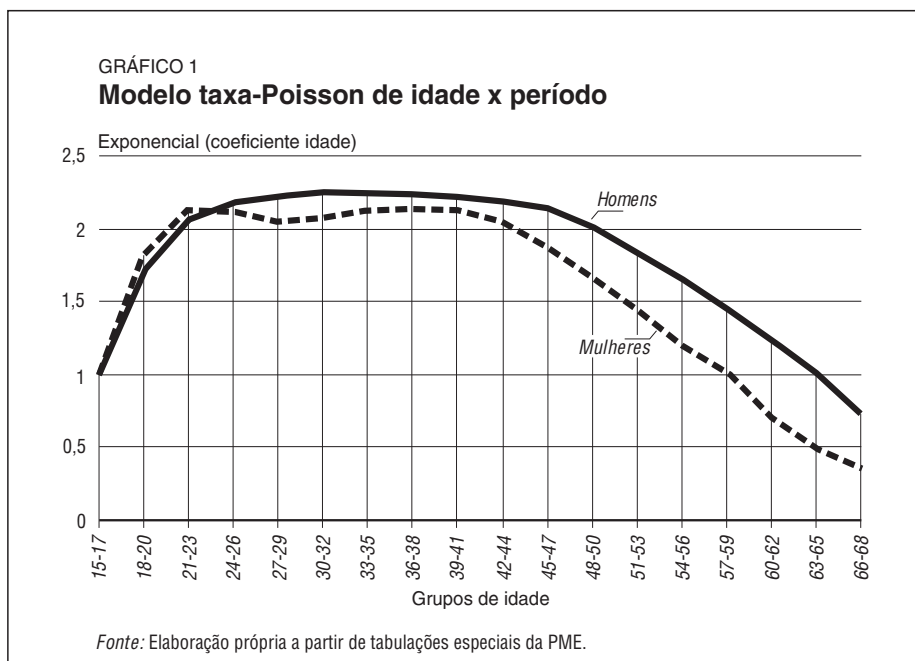
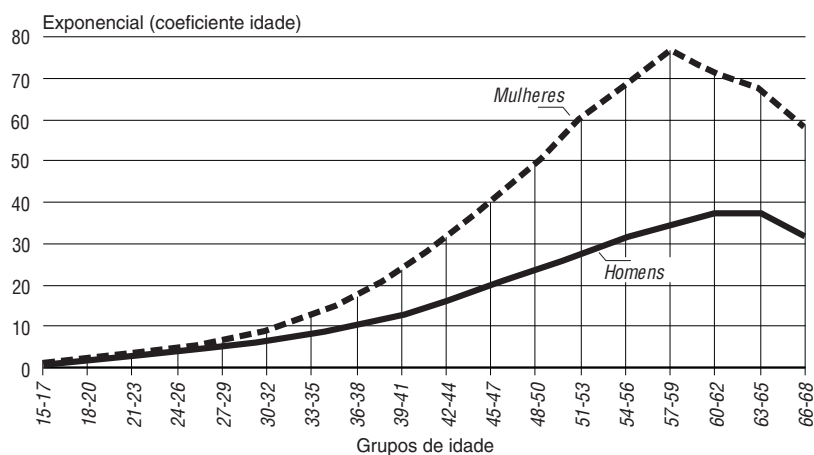
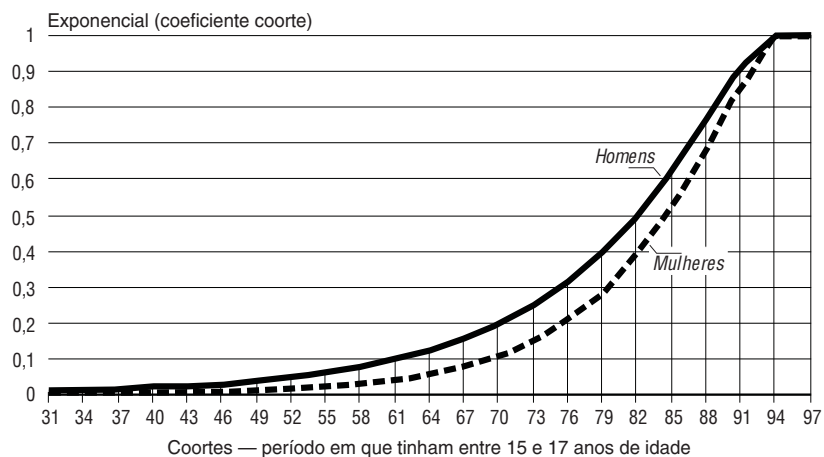


GRÁFICO 2
**Modelo taxa-Poisson de independência condicional
 (idade x período x coorte)**



Fonte: Elaboração própria a partir de tabulações especiais da PME.

GRÁFICO 3
**Modelo de independência condicional
 (idade x período x coorte)**

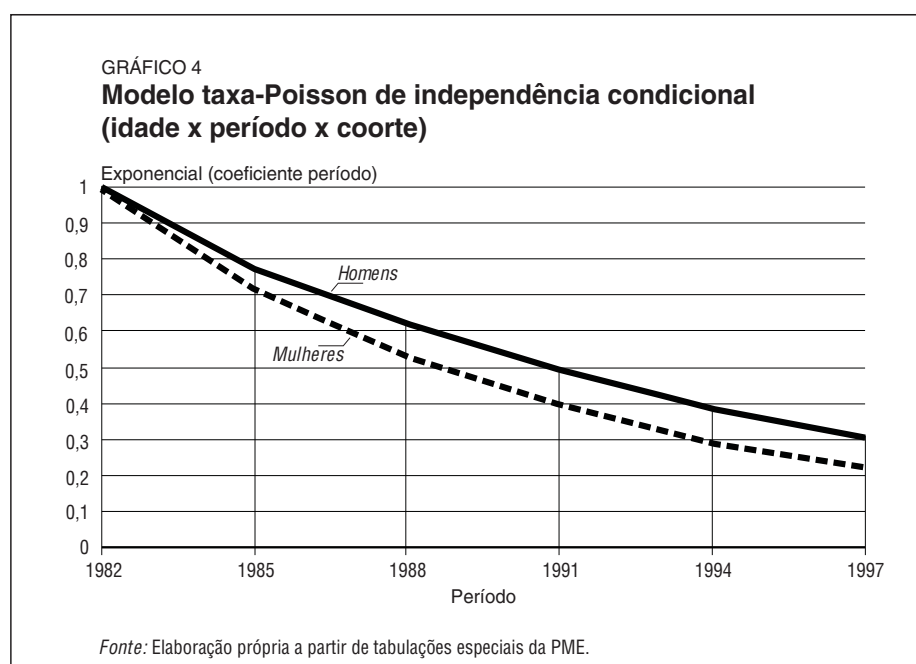


Fonte: Elaboração própria a partir de tabulações especiais da PME.

umenta na década de 70, acelerando mais na seguinte. A partir de meados da década de 70, as coortes mais jovens de mulheres passam a apresentar um ritmo de crescimento no nível de atividade maior do que o das coortes masculinas. A taxa de crescimento dos dois sexos é bastante aproximada em meados da década de 90.

Os dados do Gráfico 4 mostram o efeito de período puro nos anos analisados. Esses resultados revelam que o impacto puro das condições econômicas sobre o nível de atividade tem sido negativo tanto para os homens quanto para as mulheres desde o início dos anos 80. O ritmo do impacto negativo foi maior na primeira metade da década de 80 do que nos demais períodos de queda. Destaca-se o fato de que o impacto negativo do efeito de período no nível de atividade tenha sido maior para as mulheres do que para os homens.

Concluindo, observa-se que a aplicação do modelo de taxa com distribuição de Poisson aos dados da PME, concernentes às regiões metropolitanas brasileiras, representou um ajuste bastante adequado dos dados observados na tabela de contingência. Além disso, a inspeção visual do exponencial dos coeficientes estimados mostrou as tendências dos componentes de idade, período e coorte. Os componentes de idade mostram uma tendência crescente nas taxas de atividade ao longo do ciclo de vida, sendo esse crescimento relativo maior no caso das mulheres. Os componentes de coorte mostram um crescimento histórico nas taxas

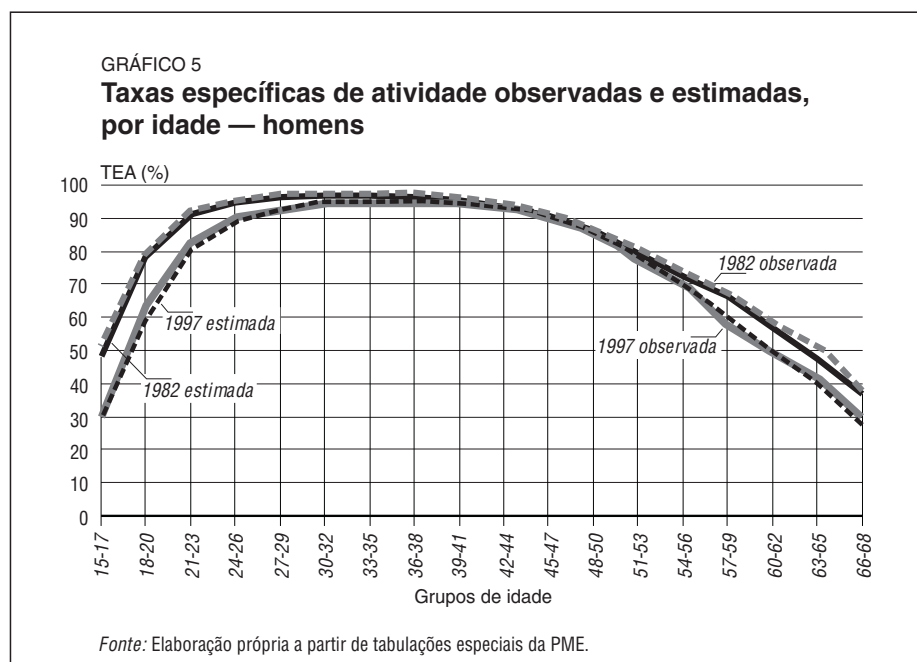


de atividade das coortes mais jovens, sendo que as coortes femininas apresentam uma taxa de crescimento maior do que a dos homens no período mais recente. Finalmente, o efeito de período mostra que o ambiente econômico tem favorecido uma queda nas taxas de atividade, que é maior entre as mulheres.

6 - A projeção da TEA a partir do modelo

Uma das vantagens da estimativa de um modelo empírico aos dados das TEAs é a possibilidade de extrapolar previsões futuras de seu comportamento a partir dos parâmetros estimados na equação.⁸ A primeira inspeção visual relevante é a comparação das TEAs estimadas e observadas nos anos extremos da série de dados (1982 e 1997).

Os dados representados no Gráfico 5 confirmam o excelente ajuste do modelo de taxas aos dados observados. Esse primeiro exercício de estimação de perfis da TEA masculina, obtidos a partir de intervalos etários trianuais, confirma o potencial de extrapolação apresentado pelo modelo.

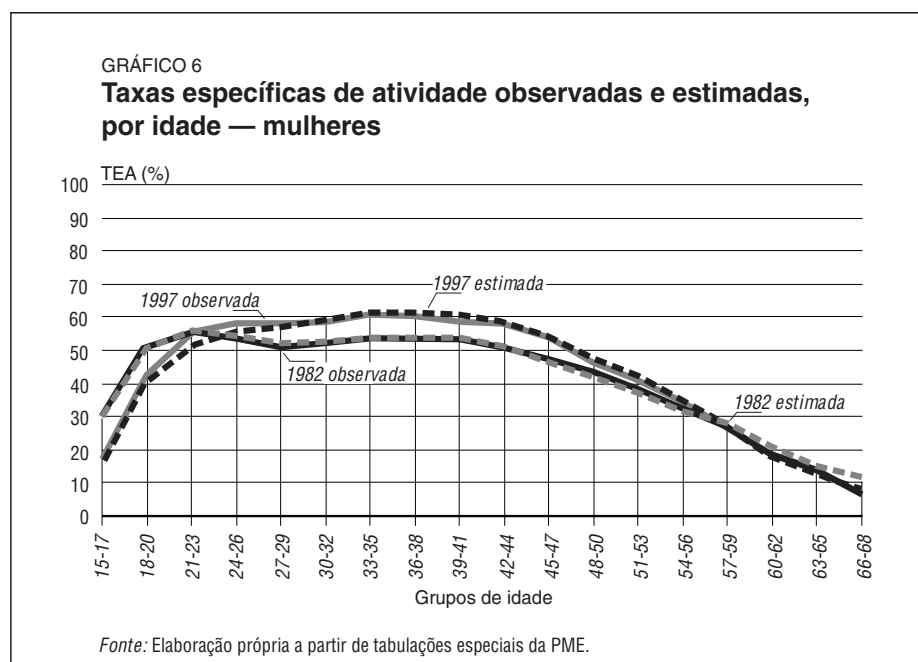


8 Ver dados completos no Apêndice.

O Gráfico 6 ilustra o mesmo tipo de resultado para as mulheres. As taxas observadas e estimadas estão muito próximas uma da outra, e a mudança no perfil observado das TEAs femininas metropolitanas é captada pelo modelo estimado. Tal resultado é forte indicativo do uso potencial desse modelo de estimativas em técnicas projetivas.

A análise anterior mostrou a capacidade preditiva do modelo tanto para homens quanto para mulheres, com relação aos dados observados. O problema agora é tentar estabelecer tendências futuras a partir dos parâmetros estimados no modelo (IPC). Admitimos que o perfil etário longitudinal da TEA é constante e igual aos coeficientes estimados tanto para os homens quanto para as mulheres (Gráfico 2). Portanto, os cenários de projeção serão traçados de acordo com as tendências observadas nos efeitos principais de coorte e de período (Gráficos 3 e 4).

Os cenários serão os mesmos para as projeções das TEAs masculina e feminina. O efeito de coorte exerce uma influência positiva sobre o nível de atividade tanto masculina quanto feminina. A tendência histórica do efeito coorte sobre a taxa de atividade masculina foi crescente até chegar a aproximadamente 27%, arrefecendo muito esse crescimento no caso das coortes mais jovens que faziam 15 a 19 anos na década de 90. A hipótese avançada a partir da tendência histórica foi de um crescimento positivo do efeito coorte da ordem de 15% por coorte de nascimento. Com relação ao efeito período, houve um impacto negativo sobre a



taxa de atividade que girou em torno de uma queda de 20% por período observado. O cenário assume a continuidade dessa tendência. Em suma, o cenário 1 de projeção da TEA é marcado pelo aumento do efeito coorte da ordem de 15% por coorte e uma queda do efeito período da ordem de 20% por período.

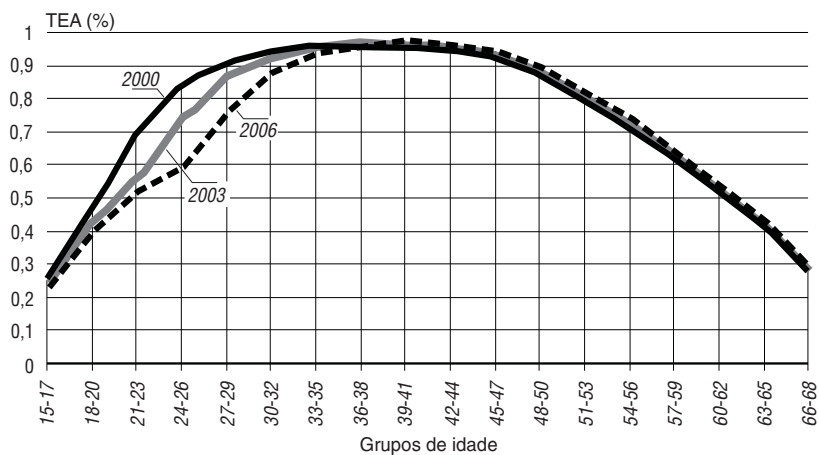
O cenário 2 apresenta pequena modificação em relação ao cenário 1. O efeito coorte continua exercendo o mesmo impacto positivo no nível de atividade, qual seja, um crescimento positivo da ordem de 15% por coorte. A diferença entre os dois cenários está no efeito de período, admitindo-se um ligeiro arrefecimento do impacto negativo do efeito de período. O último triênio da década de 90 seria marcado pela continuidade da queda de 20% que caracteriza o cenário 1. Entretanto, já entre os anos 2000 e 2003 haveria um arrefecimento do impacto negativo do efeito de período para 18% decorrente da hipótese de uma melhoria nas condições econômicas. Essa tendência persistiria entre os anos 2003 e 2006, com um outro arrefecimento do impacto negativo do efeito de período para 15%. Em suma, o cenário 2 é marcado por uma diminuição no impacto negativo do efeito de período — de 20% para 1997/2000, 18% para 2000/2003 e 15% para 2003/2006.

Os resultados da projeção da TEA masculina para o cenário 1 são apresentados no Gráfico 7 e, para o cenário 2, no Gráfico 8. Os Gráficos 7 e 8 são bastante ilustrativos, apresentando as projeções da TEA masculina metropolitana para os anos 2000, 2003 e 2006, de acordo com os cenários 1 e 2. No cenário 1, a manutenção das tendências de crescimento do nível pelo efeito coorte e de decréscimo desse nível pelo efeito período faz com que a TEA masculina seja absolutamente estável a partir do grupo etário de 33 a 35 anos. Entretanto, observa-se para os anos projetados um importante declínio das taxas de atividade dos grupos etários com menos de 33 anos. A queda é bastante pronunciada nos grupos de 24 a 26, 27 a 29 e 30 a 32 anos. Se no ano 2000 94,2% dos homens de 30 a 32 anos participarem na PEA, esse percentual cairá para 88,1% em 2006. No cenário 2, a manutenção do efeito coorte, conjugada com a contínua redução do impacto negativo do efeito período, faz com que o declínio das TEAs nas idades abaixo dos 30 anos seja ligeiramente arrefecido. Mais importante que isso, a taxa de atividade das pessoas com mais de 30 e menos de 47 anos apresenta um crescimento relativamente importante. Cabe agora analisar as tendências da TEA feminina nos cenários 1 e 2.

No caso das mulheres, a projeção da TEA de acordo com o cenário 1 indica tendências muito parecidas com as estabelecidas por projeções anteriores [ver Wajnman e Rios-Neto (1994)]. De fato, no cenário 1 as tendências históricas dos efeitos coorte e período são parecidas com o desenho dos cenários daquelas projeções, que incorporam as tendências dos anos 90. Como mostra o Gráfico 9, as taxas de atividade dos grupos etários de 15 a 17 e 18 a 20 anos não variam muito quando se comparam os três anos projetados. Nos grupos etários de 21 a 23 e 24 a 26 anos, as taxas de atividade entre 2000 e 2006 apresentam pronunciada queda,

GRÁFICO 7

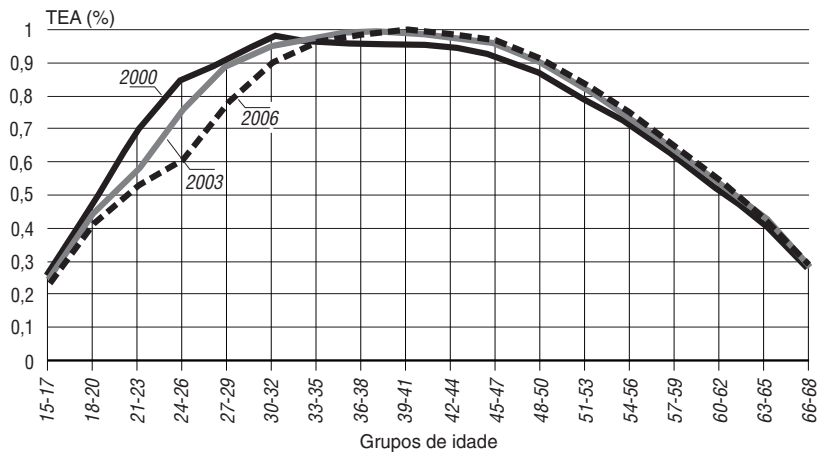
Taxas específicas de participação na PEA: homens — cenário 1



Fonte: Elaboração própria a partir de tabulações especiais da PME.

GRÁFICO 8

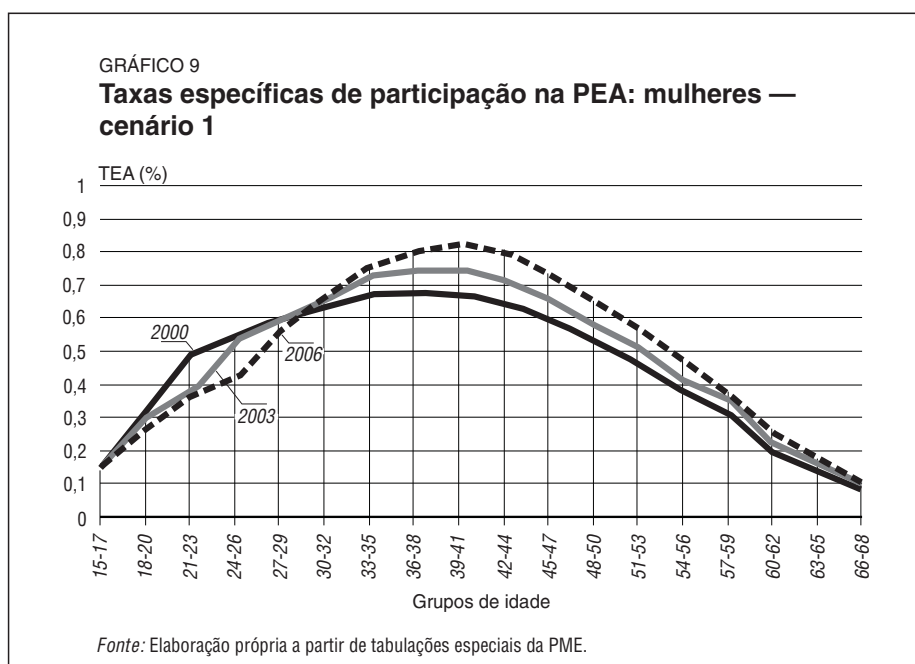
Taxas específicas de participação na PEA: homens — cenário 2

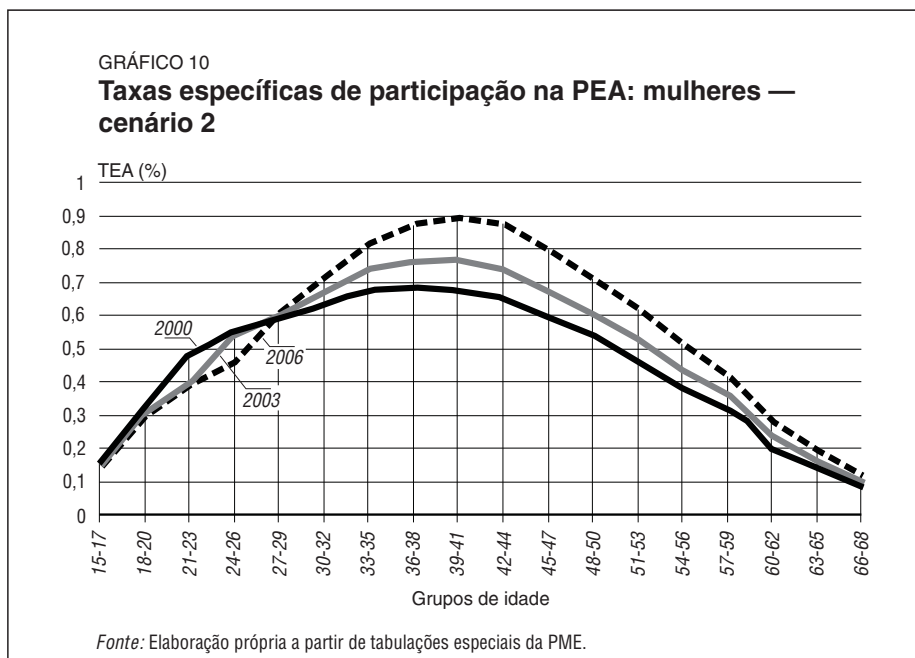


Fonte: Elaboração própria a partir de tabulações especiais da PME.

arrefecida no grupo de 27 a 29 anos. A partir dos 30 anos, a tendência entre os anos projetados é de crescimento no perfil da TEA de período. O pico de participação na TEA feminina de período varia entre os grupos de 36 a 38 e 29 a 41 anos, saindo de 68,1% em 2000 para 81,9% em 2006.

No caso do cenário 2 para as mulheres, a manutenção das tendências históricas do efeito de coorte, conjugada com a redução gradual do impacto negativo causado pelo efeito de período, causa uma pequena mudança no padrão, associada a alterações mais substanciais no nível de atividade. De acordo com o Gráfico 10, o declínio nas TEAs entre 20 e 29 anos de idade é menos pronunciado do que no cenário 1. Por outro lado, o crescimento nas TEAs é mais pronunciado a partir dos 30 anos. O pico na TEA passa de 68,1% no ano 2000 para 89,2% em 2006.





7 - Conclusão: explorando cenários a partir do modelo

Os exercícios de estimativa realizados a partir dos dados da PME nos anos 80 e 90 permitiram uma série de conclusões. Em primeiro lugar, eles confirmaram que a estimativa de modelos de IPC a partir da família de modelos lineares generalizados pode ser de grande valia para a projeção de taxas de atividade. Essa formalização estatística da projeção de TEAs parece ser extremamente compatível com o método empírico de projeção das TEAs utilizado por Wajnman e Rios-Neto (1994). Não é possível fazer uma comparação dos resultados aqui obtidos com aqueles apresentados por esses autores, pois eles se baseiam nas PNADs do IBGE e tratam do Brasil urbano, enquanto nossos resultados referem-se ao Brasil metropolitano retratado nas PMEs. Ainda assim, cumpre destacar que a vantagem desse método é sua melhor formalização, permitindo um maior detalhamento no desenho dos cenários simulados.

Em segundo lugar, este trabalho sugere a possibilidade de utilização desse método para o desenho informado de cenários futuros para a TEA. Foi realizada uma primeira exploração de cenários usando as estimativas do modelo, sendo que um cenário representou a tendência histórica dos efeitos de coorte e de período, com a constância do efeito de idade; e o outro cenário representou a manutenção da tendência histórica dos efeitos de coorte, com uma redução no impacto

negativo do efeito de período. Outros cenários podem ser desenvolvidos, sobretudo no tocante ao comportamento futuro do efeito de coorte. Para tanto, torna-se necessário entender melhor que fatores determinam o efeito de coorte. Uma hipótese é que a experiência educacional diferenciada das coortes seria um desses determinantes. Tal hipótese e outras deverão ser exploradas em trabalhos futuros.

Apêndice

TABELA A.1

TEAs observadas e estimadas, modelo taxa-Poisson — homens

Grupo etário	1982 Observada	1982 Estimada	1997 Observada	1997 Estimada
15-17	51,5	48,2	29,2	29,2
18-20	77,6	77,7	63,5	59,6
21-23	90,3	90,9	82,5	80,1
24-26	94,6	94,9	90,1	88,8
27-29	96,3	96,4	92,8	92,7
30-32	96,6	96,8	94,1	95,2
33-35	96,3	96,6	95,1	95,1
36-38	96,1	96,2	95,0	95,0
39-41	94,9	95,3	94,6	94,7
42-44	93,2	93,3	92,9	93,5
45-47	90,8	90,6	89,9	91,4
48-50	85,9	86,2	85,5	85,8
51-53	79,8	79,3	77,1	78,5
54-56	73,1	71,7	69,5	70,3
57-59	66,4	66,0	57,4	60,8
60-62	58,3	57,0	49,4	50,6
63-65	50,4	47,4	42,2	40,7
66-68	37,0	36,9	29,8	28,0

FONTE: Elaboração própria a partir de tabulações especiais da PME.

TABELA A.2

TEAs observadas e estimadas, modelo taxa-Poisson — mulheres

Grupo etário	1982	1982	1997	1997
	Observada	Estimada	Observada	Estimada
15-17	31,8	30,1	17,4	17,2
18-20	51,1	50,7	42,4	39,6
21-23	54,8	55,7	55,5	51,4
24-26	54,0	54,3	58,0	55,6
27-29	51,3	52,3	58,3	57,2
30-32	52,4	52,9	58,6	59,2
33-35	53,7	53,8	61,0	61,5
36-38	53,6	54,0	60,4	61,1
39-41	53,9	54,0	58,6	60,6
42-44	51,2	50,9	58,2	58,8
45-47	47,0	46,5	53,9	54,3
48-50	43,4	42,2	46,5	48,1
51-53	38,1	37,0	40,6	42,0
54-56	32,0	31,8	34,5	34,7
57-59	28,2	28,2	26,5	27,5
60-62	22,0	21,1	18,4	18,5
63-65	15,6	15,0	13,4	12,8
66-68	11,8	11,8	7,2	8,1

FONTE: Elaboração própria a partir de tabulações especiais da PME.

TABELA A.3
TEAs projetadas — homens

Grupos etários	Cenário 1			Cenário 2		
	2000	2003	2006	2000	2003	2006
15-17	0,268	0,247	0,227	0,268	0,253	0,232
18-20	0,477	0,438	0,403	0,477	0,449	0,413
21-23	0,703	0,562	0,517	0,703	0,576	0,529
24-26	0,844	0,741	0,593	0,844	0,760	0,607
27-29	0,912	0,867	0,761	0,912	0,889	0,778
30-32	0,942	0,927	0,881	0,981	0,950	0,901
33-35	0,964	0,954	0,939	0,964	0,978	0,961
36-38	0,958	0,971	0,961	0,958	0,996	0,984
39-41	0,956	0,963	0,977	0,956	0,988	1,000
42-44	0,945	0,954	0,961	0,945	0,977	0,984
45-47	0,924	0,934	0,943	0,924	0,958	0,965
48-50	0,872	0,882	0,892	0,872	0,904	0,913
51-53	0,794	0,807	0,816	0,794	0,827	0,836
54-56	0,714	0,722	0,737	0,714	0,740	0,751
57-59	0,619	0,629	0,636	0,619	0,644	0,651
60-62	0,516	0,526	0,533	0,516	0,539	0,546
63-65	0,409	0,416	0,425	0,409	0,427	0,434
66-68	0,282	0,283	0,288	0,282	0,290	0,295

FONTE: Elaboração própria a partir de tabulações especiais da PME.

TABELA A.4
TEAs projetadas — mulheres

Grupos etários	Cenário 1			Cenário 2		
	2000	2003	2006	2000	2003	2006
15-17	0,159	0,146	0,134	0,159	0,150	0,146
18-20	0,317	0,292	0,268	0,317	0,299	0,292
21-23	0,485	0,388	0,357	0,485	0,398	0,890
24-26	0,557	0,525	0,420	0,557	0,598	0,457
27-29	0,591	0,592	0,558	0,591	0,607	0,608
30-32	0,636	0,658	0,658	0,636	0,674	0,717
33-35	0,672	0,722	0,746	0,672	0,740	0,813
36-38	0,681	0,745	0,800	0,681	0,763	0,871
39-41	0,672	0,749	0,819	0,672	0,768	0,892
42-44	0,649	0,720	0,802	0,649	0,738	0,874
45-47	0,597	0,659	0,731	0,597	0,675	0,796
48-50	0,537	0,589	0,651	0,537	0,605	0,709
51-53	0,464	0,518	0,569	0,464	0,531	0,619
54-56	0,381	0,422	0,470	0,381	0,432	0,512
57-59	0,313	0,344	0,380	0,313	0,353	0,414
60-62	0,205	0,233	0,257	0,205	0,239	0,280
63-65	0,139	0,155	0,176	0,139	0,159	0,191
66-68	0,089	0,097	0,108	0,089	0,100	0,118

FONTE: Elaboração própria a partir de tabulações especiais da PME.

Abstract

The integration of period and cohort rates is an important method for the projection of economically active population. This paper tries to formalize models of age-period-cohort (APC) by means of log-linear (rate) models. Topological models are applied to index cohorts. This work starts with a review and methodological session focused on the identification problem in APC models. In the empirical part of the paper data from the Monthly Employment Survey in the eighties and nineties are applied to test the method's fit. The results are similar to less formal cohort/period models. The method proposed in the paper has the advantage of formalization, which improves one's capability to design different scenarios for economically active population projections.

Bibliografia

- AGRESTI, A. *Categorical data analysis*. New York: John Wiley & Sons, 1991.
- BLOOM, D. E. *On the nature and estimation of age, period, and cohort effects in demographic data*. Center for Population Studies, Harvard University, 1984 (Discussion Paper, 84-87).
- CLOGG, C. C. *Measuring underemployment: demographic indicators for the United States*. New York: Academic Press, 1979.
- FARKAS, G. Cohort, age, and period effects upon the employment of white females. *Demography*, v.14, n.1, p.33-35, 1977.
- FIENBERG, S. E., MASON, W. M. Specification and implementation of age, period and cohort models. In: MASON, W. M., FIENBERG, S. E. (eds.). *Cohort analysis in social research*. New York: Springer Verlag, 1985.
- GLENN, N. D. *Cohort analysis*. Beverly Hills, CA: Sage, 1977.
- HALLI, S. S., RAO, K. V. *Advanced techniques of population analysis*. New York: Plenum Press, 1992.
- HANOCH, G., HONIG, M. "True" age profiles of earnings: adjusting for censoring and for period and cohort effects. *Review of Economics and Statistics*, v. 67, n. 3, p. 383-394, 1985.
- HECKMAN, J., ROBB, R. Using longitudinal data to estimate age, period and cohort effects in earnings equations. In: MASON, W. M., FIENBERG, S. E. (eds.). *Cohort analysis in social research*. New York: Springer Verlag, 1985.
- HOBcraft, J., MENKEN, J., PRESTON, S. Age, period, and cohort effects in demography: a review. *Population Index*, v. 48, p. 4-43, 1982.
- JUHN, C., MURPHY, K. M., PIERCE, B. Wage inequality and the rise in returns to skill. *Journal of Political Economy*, v. 101, n. 3, p. 410-442, 1993.
- MASON, K. O., MASON, W. M., WINSBOROUGH, H. H., POOLE, W. K. Some methodological issues in cohort analysis of archive data. *American Sociological Review*, v. 38, p. 242-258, 1973.
- RODGERS, W. L. Estimable functions of age, period, and cohort effects. *American Sociological Review*, v. 47, p. 774-787, 1982.
- RYDER, N. B. The cohort as a concept in the study of social change. *American Sociological Review*, v. 30, n. 6, p. 843-861, 1965.
- WAJNMAN, S., RIOS-NETO, E. L. G. Participação feminina na população economicamente ativa no Brasil: alternativas para projeções de níveis e padrões. *Pesquisa e Planejamento Econômico*, Rio de Janeiro, v. 24, n. 2, p. 203-233, 1994.

WINSBOROUGH, H. H. Age, period, cohort and education effects on earnings by race. In:
LAND, K., SPILLERMAN, S. (eds.) *Social indicators models*. New York: Russell
Sage Foundation, 1975.

(Originais recebidos em junho de 1999. Revisos em fevereiro de 2000.)