

Modelos idade-período-coorte aplicados à participação na força de trabalho: em busca de uma versão parcimoniosa

Ana Maria Hermeto Camilo de Oliveira*
Eduardo Luiz Gonçalves Rios-Neto**

Este artigo se insere na linha de abordagem que busca descrever a variação das tendências das taxas demográficas, no caso, da participação na força de trabalho, resumindo as mudanças em seus efeitos de idade, período e coorte. São utilizados dados das PNADs de 1981 a 1999 em modelos log-lineares estimados através de regressões de Poisson em que o logaritmo da PEA é a variável dependente, controlando pelo logaritmo da PIA, em função das variáveis independentes de idade, período e coorte. Ao longo deste período, a taxa de participação feminina na força de trabalho elevou-se substancialmente no Brasil, para todas as idades. O quadro não é o mesmo para os homens, que mostram um padrão de entradas e saídas do mercado de trabalho mais estável ao longo do tempo, próximo a um steady state, o que se reflete na prevalência dos efeitos de idade. Entre as mulheres, os efeitos de idade também são os mais importantes, mas os efeitos de período e de coorte se mostram relevantes. A diferenciação das coortes femininas quanto à participação na força de trabalho não pode ser atribuída à variação temporal ou etária. Buscando a parcimônia dos modelos pela utilização de medidas diretas dos efeitos de período e coorte, o nível educacional é o melhor indicador, para ambos os sexos. A experiência educacional diferenciada das coortes parece ser um determinante importante da variação da participação na força de trabalho, sobretudo entre as mulheres.

Palavras-chave: Participação na força de trabalho. Modelos idade-período-coorte.

Introdução

Eventos demográficos marcam as transições mais importantes da vida e são influenciados por vários fatores. Estas influências não são constantes no tempo e no espaço e, portanto, as taxas demográficas mostram considerável variação. A

identificação dos fatores responsáveis por esta variabilidade pode ser feita mediante o estudo das taxas em três dimensões críticas: idade à ocorrência do evento; período de ocorrência; coorte de indivíduos envolvidos. Nesta abordagem, a coorte se refere ao ano de nascimento de um indivíduo¹, a idade significa a idade cronológica de um indivíduo

* Pesquisadora do Centro de Desenvolvimento e Planejamento Regional (Cedeplar) da Universidade Federal de Minas Gerais (UFMG).

** Professor titular do Departamento de Demografia e pesquisador do Centro de Desenvolvimento e Planejamento Regional (Cedeplar) da UFMG.

¹ Poderia ser, também, o momento de alguma outra importante transição no ciclo de vida, como a entrada no mercado de trabalho, o casamento etc. Da mesma forma, a idade pode significar o tempo transcorrido desde uma importante transição – por exemplo, a duração do emprego em estudos de aposentadoria.

no momento da ocorrência do evento, e o período indica o ano de ocorrência do evento.

Este artigo se insere nessa linha de abordagem que busca descrever a variação das tendências das taxas demográficas – no caso, da participação na força de trabalho. Nosso principal objetivo é analisar os componentes das mudanças da participação na força de trabalho, concentrando a atenção no problema metodológico mais geral de resumir a variação das taxas em seus efeitos de idade, período e coorte. Para isto, é feita uma decomposição das taxas de participação específicas por sexo em componentes separados de idade, período e coorte, baseada na hipótese de que todos estes componentes são importantes na definição dos níveis e estruturas de participação na força de trabalho de homens e mulheres. Segundo Clogg (1982), as variáveis de idade e de período devem exercer uma forte influência sobre a participação na força de trabalho, e isto é importante para estimar a magnitude destes efeitos. A variável de idade, por exemplo, pode ser vista como um indicador de experiência ou como indicador de posição na estrutural social. A variação de período (efeitos de período) é usualmente tomada como indicador de forças puramente econômicas, como mudança na demanda por trabalho ou outros tipos de flutuações no mercado de trabalho. Algumas questões não respondidas por outros métodos nos motivam. Como a diferenciação das coortes condiciona as mudanças de período diretamente observadas? A participação na força de trabalho é definida por efeitos de período ou de coorte? Neste sentido, reflete ciclos ou tendências? O nível educacional é determinante da participação, como indicado pelos efeitos de coorte ou de período?

Antes dos estudos pioneiros de Winsborough (1975), Farkas (1977) e Clogg (1979 e 1982), os estudos demográficos nos quais os efeitos de coorte são estimados explicitamente estavam confinados sobretudo a análises de fecundidade, área onde esta técnica foi usada originalmente, a partir do clássico estudo de Ryder (1965). O livro

Cohort analysis in social research, editado por Fienberg e Mason (1985), sintetiza esta literatura, contendo este artigo seminal de Ryder (1965), a importante revisão feita por Hobcraft, Menken e Preston (1982) e vários outros artigos que debatem e utilizam o método de análise de idade-período-coorte (IPC). Muito pode ser apreendido pela extensão deste método aos estudos dos padrões de participação na força de trabalho.

Padrões de participação na força de trabalho foram analisados em vários trabalhos que enfatizam a participação feminina, enfocam mudanças intertemporais e tendências e fazem inferências sobre a diferenciação de coorte com relação à participação no mercado de trabalho (Farkas, 1977; Bianchi e Spain, 1986; Goldin, 1990; Hill e O'Neill, 1990 e 1992; Joshi e Paci, 1998; Blau, 1998). Wajnman e Rios-Neto (1994) apresentam projeções de taxas de participação na força de trabalho feminina com base em coortes, a partir da observação da persistência das diferenças de níveis e padrões entre as coortes. A maioria destes trabalhos conclui que as tendências de período nos padrões de participação masculina na força de trabalho não mostram o mesmo grau de mudança secular que os femininos (Clogg, 1982).

Para fazer o desenvolvimento metodológico de modelos IPC aplicados à análise das tendências de participação na força de trabalho urbana brasileira nas duas últimas décadas, utilizamos os dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) do IBGE. Nosso objetivo é descrever tais tendências da forma mais parcimoniosa possível. De acordo com Wilmoth (1990), uma característica do arcabouço IPC é fornecer descrições parcimoniosas, porque minimizam o número de parâmetros ajustados sem sacrificar a exatidão da informação.

Na primeira seção deste artigo ilustramos a motivação para a aplicação do modelo IPC no estudo da participação na força de trabalho no Brasil por meio de uma análise exploratória dos dados. Em seguida, discutimos aspectos conceituais e metodológicos do arcabouço IPC. Partimos, então, para a aplicação do modelo e discussão dos resultados.

A força de trabalho brasileira

O modelo IPC é tipicamente usado para bases de dados na forma de uma classificação de idade por período, ou coorte por período, ou coorte por idade. A mesma parametrização IPC pode ser aplicada de maneira intercambiável para cada uma destas formas, todas contendo a mesma informação. No nosso estudo, os dados obtidos por meio das múltiplas pesquisas *cross-section* são resumidos em tabelas de idade X período (ano da pesquisa): cada entrada da tabela é a taxa observada² de atividade dos indivíduos observados em um período específico (ver Tabelas 1 e 2). O espaçamento das categorias etárias (*I*) é igual ao espaçamento entre os períodos (*P*); assim, as $C = I + P - 1$ diagonais das tabelas correspondem às coortes de nascimento. Neste modelo de múltipla classificação, idade, período e coorte são tratadas como variáveis nominais.

A interpretação dos componentes das tabelas é ambígua porque a coorte está associada às duas outras dimensões temporais. Se, para uma dada coorte, a taxa

de atividade aumenta a cada período, há um efeito de período; contudo, simultaneamente, esta taxa está aumentando a cada idade, havendo um efeito de idade. Os dados são os mesmos, mas a interpretação é diferente e é impossível determinar com base puramente estatística os efeitos de coorte. Dado que há efeitos de idade sobre a participação na força de trabalho, as diferenças intracoortes confundem os efeitos de idade e de período e diferenças intercoortes confundem os efeitos de idade e de coorte. Da mesma forma ocorre com os efeitos relacionados ao período. É possível que o padrão das taxas seja inteiramente explicado pelos efeitos de idade e período e que não haja efeitos de coorte. De maneira geral, todos os efeitos podem existir e as diferenças intra e intercoortes confundem os efeitos de duas destas dimensões temporais.

Análises informais das Tabelas 1 e 2 mostram a diferenciação de coorte na participação na força de trabalho e que esta diferenciação de coorte tem diferente caráter para cada grupo de sexo. Estas comparações informais podem ser equivocadas devido à dependência funcional

TABELA 1
Taxas específicas de participação na força de trabalho por idade – homens, Brasil urbano, 1981-1999

| Idade | 1981 | 1984 | 1987 | 1990 | 1993 | 1996 | 1999 |
|-------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| 15-17 | 58,0% | 58,6% | 63,4% | 62,0% | 59,9% | 51,4% | 49,6% |
| 18-20 | 79,6% | 80,5% | 83,8% | 81,6% | 81,0% | 76,1% | 76,0% |
| 21-23 | 91,4% | 92,0% | 92,7% | 91,9% | 90,4% | 87,5% | 88,2% |
| 24-26 | 95,4% | 95,6% | 96,3% | 95,3% | 93,8% | 92,0% | 93,0% |
| 27-29 | 97,1% | 97,4% | 97,3% | 96,2% | 95,4% | 93,8% | 94,8% |
| 30-32 | 97,6% | 97,5% | 97,7% | 96,9% | 96,7% | 94,6% | 95,3% |
| 33-35 | 97,0% | 97,5% | 97,6% | 97,2% | 96,3% | 95,0% | 95,5% |
| 36-38 | 96,9% | 97,4% | 97,3% | 97,0% | 95,8% | 95,0% | 95,3% |
| 39-41 | 96,5% | 96,5% | 97,1% | 96,5% | 96,0% | 94,5% | 94,6% |
| 42-44 | 95,1% | 95,3% | 96,4% | 95,6% | 95,1% | 94,4% | 94,0% |
| 45-47 | 92,7% | 92,0% | 94,0% | 94,3% | 93,4% | 91,5% | 92,0% |
| 48-50 | 89,3% | 88,6% | 92,0% | 90,7% | 90,7% | 88,7% | 88,9% |
| 51-53 | 85,1% | 84,6% | 85,2% | 85,5% | 85,7% | 83,8% | 83,4% |
| 54-56 | 80,4% | 78,2% | 80,1% | 81,2% | 78,7% | 78,8% | 78,4% |
| 57-59 | 74,4% | 73,0% | 77,2% | 76,3% | 72,5% | 73,3% | 73,8% |
| 60-62 | 68,4% | 66,2% | 70,1% | 70,3% | 67,0% | 65,0% | 65,0% |
| 63-65 | 61,1% | 58,9% | 62,2% | 60,3% | 60,3% | 55,2% | 55,6% |
| 66-68 | 44,8% | 45,3% | 50,3% | 48,4% | 46,3% | 46,0% | 46,9% |

Fonte: Elaboração própria, a partir dos microdados das PNADs 1961-1999.

² A variável de participação na força de trabalho é dicotômica, assumindo o valor 1 se o indivíduo é economicamente ativo, e 0 caso contrário. Variáveis contínuas também podem ser analisadas dentro deste arcabouço, constando nas células da tabela seus valores médios para as categorias.

TABELA 2
Taxas específicas de participação na força de trabalho por idade – mulheres
Brasil urbano, 1981-1999

| Idade | 1981 | 1984 | 1987 | 1990 | 1993 | 1996 | 1999 |
|-------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| 15-17 | 28,7% | 31,1% | 32,8% | 31,4% | 35,0% | 30,7% | 29,9% |
| 18-20 | 40,3% | 42,8% | 45,6% | 45,2% | 48,9% | 49,7% | 51,7% |
| 21-23 | 42,9% | 45,7% | 49,7% | 50,5% | 54,4% | 55,8% | 60,0% |
| 24-26 | 42,5% | 45,4% | 49,0% | 50,7% | 55,8% | 57,3% | 62,2% |
| 27-29 | 40,9% | 45,7% | 48,3% | 51,4% | 56,3% | 58,6% | 61,7% |
| 30-32 | 42,3% | 46,9% | 52,3% | 54,3% | 57,1% | 59,8% | 63,9% |
| 33-35 | 43,8% | 47,6% | 52,3% | 54,6% | 60,3% | 60,9% | 65,6% |
| 36-38 | 43,2% | 46,9% | 53,5% | 56,6% | 61,1% | 61,7% | 65,2% |
| 39-41 | 42,5% | 47,9% | 51,6% | 54,9% | 58,4% | 61,3% | 65,3% |
| 42-44 | 38,9% | 45,1% | 49,3% | 53,1% | 58,1% | 59,8% | 62,2% |
| 45-47 | 38,2% | 41,0% | 48,0% | 48,7% | 52,8% | 55,5% | 59,9% |
| 48-50 | 34,3% | 36,9% | 42,2% | 43,9% | 49,0% | 48,7% | 53,4% |
| 51-53 | 31,4% | 33,0% | 36,8% | 39,1% | 42,8% | 43,6% | 48,8% |
| 54-56 | 26,7% | 29,1% | 32,8% | 33,7% | 35,9% | 38,1% | 42,6% |
| 57-59 | 22,5% | 25,3% | 26,7% | 28,2% | 30,5% | 31,9% | 33,4% |
| 60-62 | 18,9% | 18,2% | 21,5% | 22,1% | 23,5% | 22,6% | 24,5% |
| 63-65 | 14,5% | 14,9% | 16,9% | 17,7% | 18,6% | 17,0% | 18,1% |
| 66-68 | 10,8% | 10,1% | 12,3% | 12,0% | 14,4% | 11,8% | 14,6% |

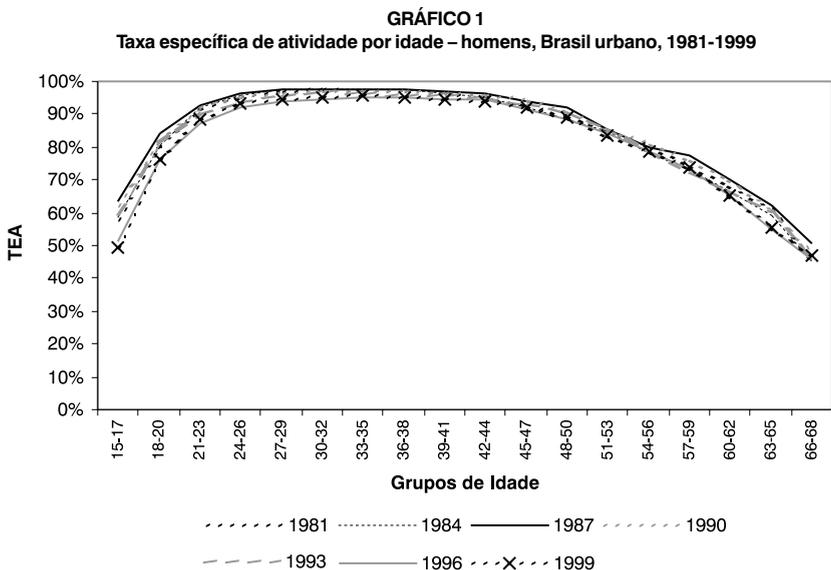
Fonte: Elaboração própria, a partir dos microdados das PNADs 1961-1999.

entre a variável de coorte e as variáveis de idade e de período. Em primeiro lugar, o que pode ser destacado é o grande diferencial de nível da participação na força de trabalho entre os sexos. Quanto à participação masculina na força de trabalho, o padrão etário na forma de U invertido é claro, com elevadas taxas entre as idades de 21 e 50 anos, e as variações de nível e estrutura entre os períodos são pequenas, localizadas nos extremos etários jovens e idosos, como mostram a Tabela 1 e o Gráfico 1. As diferentes coortes de homens parecem ser essencialmente similares em termos de seu nível e estrutura de participação na força de trabalho, como visto na Tabela 1 e no Gráfico 3. No que se refere à participação feminina na força de trabalho, as taxas específicas por idade, mostradas na Tabela 2 e no Gráfico 2, apontam para mudanças significativas tanto de nível quanto de estrutura entre os períodos. A tendência histórica de elevação do nível geral de participação feminina é evidenciada, bem como é possível visualizar a maior elevação das taxas nas idades de 18 a 54 anos. A evolução das taxas das coortes mostradas nas diagonais da Tabela 1 e no Gráfico 4 também indica uma elevação dos níveis das taxas de participação entre as coortes. As estruturas

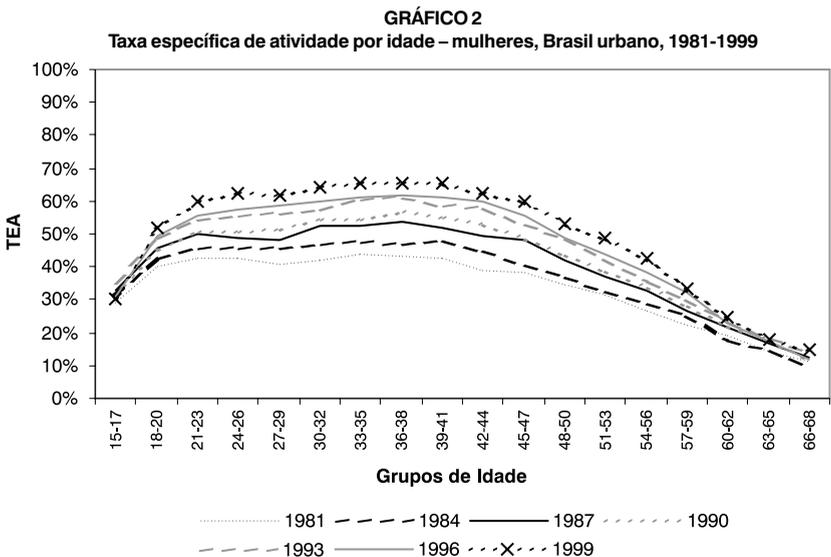
etárias da participação feminina parecem ser mais estáveis entre as coortes do que as evidenciadas entre os períodos.

O ponto é que as mudanças de período apresentadas pelas mulheres, ou a falta delas no caso dos homens, constituem evidências insuficientes, a partir das quais não é possível concluir definitivamente sobre tendências de coorte. O componente de idade é claramente importante para ambos os sexos, e os componentes de período e coorte parecem ser muito mais importantes para as mulheres. Dadas estas evidências, justifica-se a discussão dos padrões femininos e masculinos de participação na força de trabalho no Brasil dentro do arcabouço dos modelos IPC, buscando decompor a importância de cada um destes componentes.

Além disto, a exploração destes modelos permite verificar a possibilidade de encontrar bons marcadores ou medidas diretas de período e coorte. Para se especular sobre o comportamento futuro dos efeitos de período e de coorte, é necessário entender melhor que fatores o determinam. Uma hipótese é que a experiência educacional diferenciada das coortes é um destes determinantes. Nesta mesma linha, a variável de escolaridade pode ser testada como marcador dos



Fonte: Elaboração própria, a partir dos microdados das PNADs 1961-1999.



Fonte: Elaboração própria, a partir dos microdados das PNADs 1961-1999.

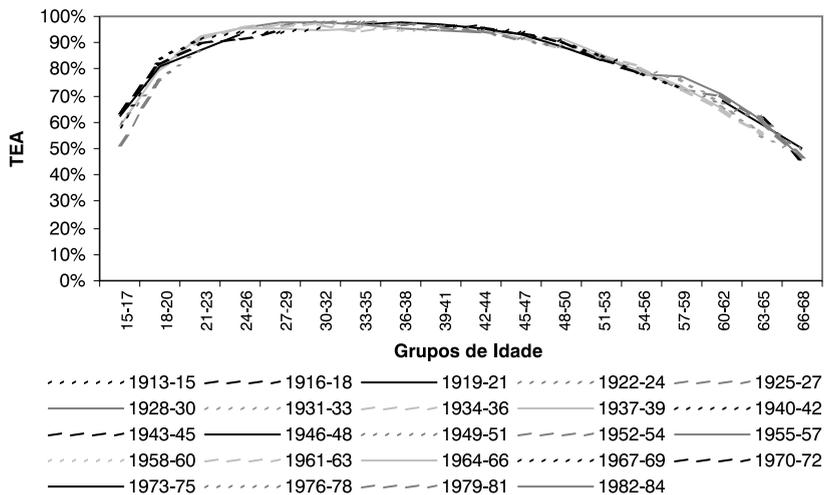
períodos (médias de anos de estudo nos períodos), assim como as médias da renda *per capita* e do salário/hora³ nos períodos. Os Gráficos de 5 a 8 ilustram a evolução

destas variáveis em relação aos períodos, nos casos da educação, renda *per capita* e salário/hora⁴, e às coortes, no caso da educação. Este último mostra mudanças

³ Não negligenciando a possível simultaneidade na determinação da participação na força de trabalho e dos salários.

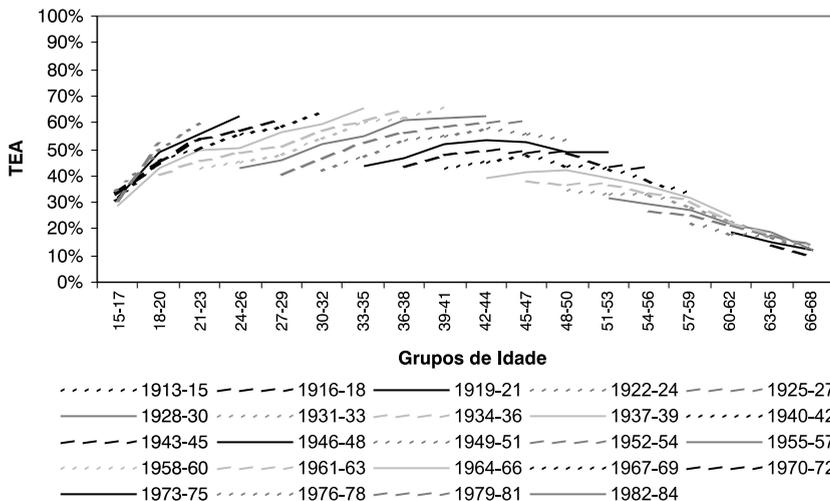
⁴ As linhas pontilhadas nos Gráficos 7 e 8 mostram as tendências estáveis, ao longo do tempo, das variáveis de renda *per capita* e salário/hora.

GRÁFICO 3
Taxa específica de atividade por idade e coorte – homens, Brasil urbano



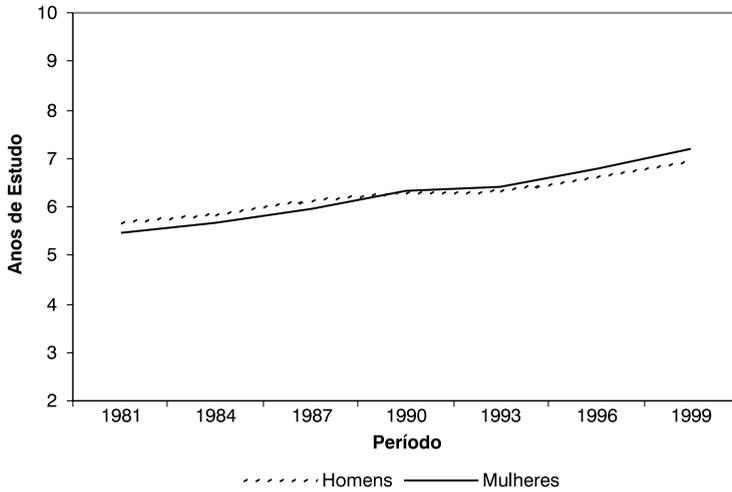
Fonte: Elaboração própria, a partir dos microdados das PNADs 1961-1999.

GRÁFICO 4
Taxa específica de atividade por idade e coorte – mulheres, Brasil urbano



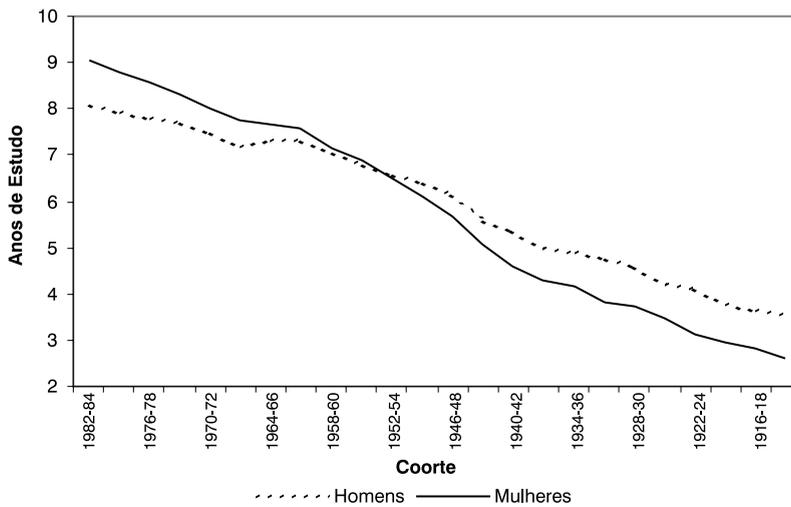
Fonte: Elaboração própria, a partir dos microdados das PNADs 1961-1999.

GRÁFICO 5
Média de anos de estudo da população de 15 a 68 anos de idade – Brasil, 1981-1999



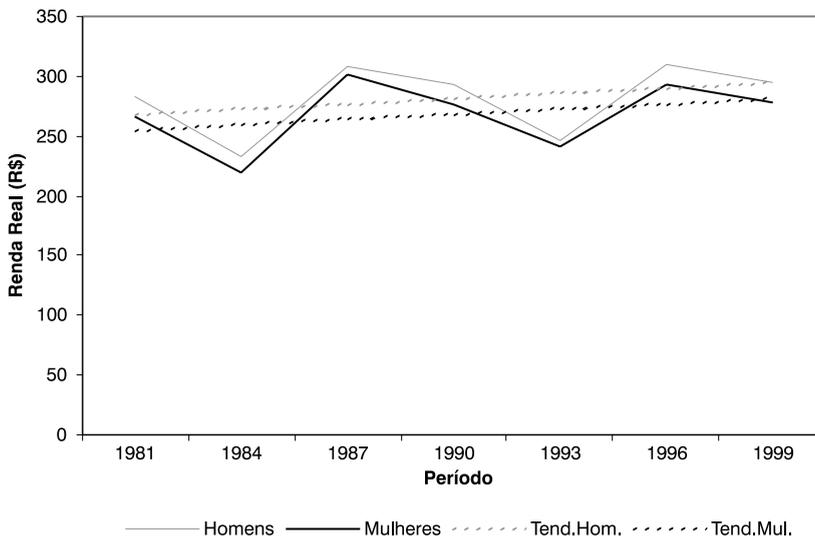
Fonte: Elaboração própria, a partir dos microdados das PNADs 1961-1999.

GRÁFICO 6
Anos de estudo da população de 25 a 29 anos de idade, por coorte – Brasil, 1981-1999



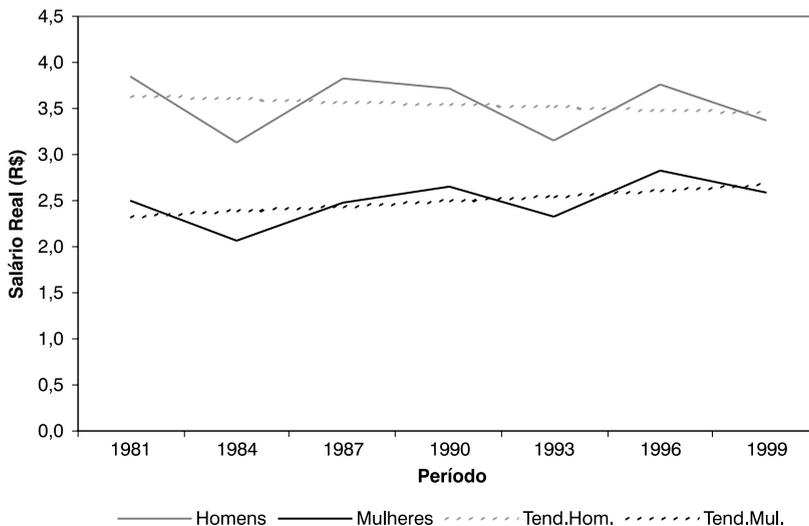
Fonte: Elaboração própria, a partir dos microdados das PNADs 1961-1999.

GRÁFICO 7
Média de renda real familiar per capita da população de 15 a 68 anos de idade – Brasil, 1981-1999



Fonte: Elaboração própria, a partir dos microdados das PNADs 1961-1999.

GRÁFICO 8
Média de salário real/hora da população de 15 a 68 anos de idade – Brasil 1981-1999



Fonte: Elaboração própria, a partir dos microdados das PNADs 1961-1999.

muito mais evidentes do que os outros casos. O nível de escolaridade das coortes mais jovens é muito mais elevado do que o das coortes mais velhas, e entre as mulheres os diferenciais intercoortes são ainda mais substantivos (Gráfico 6).

Modelos idade-período-coorte: aspectos conceituais e metodológicos

Estrutura dos dados: acumulando informações ao longo do tempo

A técnica IPC é um método de análise de dados longitudinais. A estrutura de dados longitudinais necessária para a aplicação desta técnica é obtida por meio de uma seqüência de pesquisas *cross-section* independentes repetidas, conduzidas e processadas de maneira homogênea para minimizar a variância dos métodos entre elas, nas quais as mesmas questões são aplicadas a diferentes amostras de pessoas ao longo do tempo, tal que informações sobre a mesma população agregada – no caso, a população em idade ativa no Brasil – são acumuladas ao longo do tempo. Os dados resultantes do encadeamento destas pesquisas são tipicamente mostrados na forma de uma tabela de idade por período, como visto nas Tabelas 1 e 2, ou de idade por coorte ou período por coorte. Tendo os grupos de idade e os períodos o mesmo espaçamento, subamostras na mesma coorte são encadeadas entre as pesquisas (períodos), tal como nas diagonais das Tabelas 1 e 2. Este tipo de dados possibilita o cálculo dos efeitos de idade ajustados por diferenças de coorte e de período, dos efeitos de período ajustados por diferenças de idade e de coorte, e dos efeitos de coorte ajustados por diferenças de período e de idade. Neste caso, não é preciso assumir que algum efeito potencial seja nulo.

Segundo Fienberg e Mason (1985), há problemas a serem resolvidos com a utilização desta estrutura. Primeiro, há um problema de identificação nos modelos IPC, que surge da dependência entre idade, período e coorte, o que vai ser discutido adiante. Sua solução depende do conhe-

cimento anterior sobre o processo e da teoria. Segundo, há um problema na estimação ou cálculo dos efeitos, dependendo da forma da estrutura de dados, uma vez que os dados são múltiplas *cross-sections*. Em particular, os dados são inerentemente desequilibrados, já que, apesar de serem retangulares com respeito às idades e aos períodos, não são retangulares em relação às coortes. Se a estrutura dos dados é modificada para tornar equilibradas as coortes, torna-se desequilibrada com relação à idade e ao período; não há solução para este problema, o que pode influenciar nos efeitos calculados. Terceiro, dados na forma de *cross-sections* repetidas fornecem estoques e não fluxos, ou seja, não há informações sobre os encadeamentos dos indivíduos ao longo do tempo. Neste caso, como o interesse da análise de coorte é na descrição quantitativa agregada das variações temporais no comportamento de populações, não é relevante a mudança do comportamento de cada indivíduo separadamente ao longo do tempo. Os dados das coortes não são meramente uma soma de um conjunto de histórias individuais; cada coorte tem composição e caráter distintos, que refletem as circunstâncias de sua origem e história.

Especificação de idade, período e coorte

Como mencionado anteriormente, dados de coorte são organizados seqüencialmente a partir de observações do tempo de ocorrência do evento e do intervalo desde a ocorrência do evento definidor da coorte. Para coortes de nascimento, o intervalo é a idade; se t é o momento de ocorrência e i é a idade naquele momento, as observações para a idade i e o tempo t se aplicam à coorte nascida no ano $t - i$, assim como observações para a idade $i - 1$ e tempo $i - 1$ etc. Sob este enfoque, a coorte é definida pelo período e idade em que uma observação individual entra no conjunto de dados de idade por período. O pressuposto da análise IPC é que todos os fatores que influenciam a taxa de ocorrência de um evento podem ser agrupados significativamente nas

categorias de idade, período e coorte. Pode haver interações ou sobreposições entre os conjuntos de fatores, mas assume-se implicitamente que estas influências são menos importantes do que os efeitos principais associados a cada dimensão. A estratégia analítica típica dos estudos IPC é identificar os padrões de mudança nas taxas demográficas associados a estas dimensões e, em seguida, as influências específicas que determinam os padrões observados de idade, período e coorte (Fienberg e Mason, 1985).

Baseada nesta descrição, a lógica da análise IPC parece simples e relativamente direta. A exigência mínima de dados para a análise de coorte é que mais de uma coorte seja seguida ao longo do tempo. Uma "tabela padrão de coorte" usa os mesmos intervalos de idade e período, tal que as coortes são seguidas nas diagonais, como nas Tabelas 1 e 2⁵. Hipóteses são formuladas sobre padrões de variação (ou falta de variação) na tabela de coorte. A inspeção visual da tabela permite detectar o padrão de variação que ocorreria se toda a variação se devesse a efeitos de idade, período ou coorte: os efeitos puros. Nos dados reais, os padrões previstos por estes efeitos puros raramente ocorrem.

Estudos pioneiros como Mason *et al.* (1973) tentaram discutir a análise de dados agrupados na qual efeitos de idade, período e coorte se apresentam simultaneamente. Quando os dados são agrupados de acordo com a idade do indivíduo e o intervalo de tempo no qual a observação ocorreu, medidos em unidades de extensão iguais, isto é um claro exemplo de diagonais com importância analítica. Neste estudo, a frequência de eventos é aferida em relação ao tamanho da população em risco, configurando uma taxa; no caso, a população economicamente ativa (PEA) sobre a população em idade ativa (PIA).

Há um problema fundamental na abordagem IPC relacionado ao fato de que, matematicamente, coorte + idade = período.

Do ponto de vista de alguns autores (Glenn, 1976; Rodgers, 1982), a conexão matemática exata entre idade, período e coorte torna todas as formas de análise IPC sem significado, dado que mudanças em um processo demográfico em uma das dimensões não podem ser distinguidas estatisticamente de mudanças nas outras dimensões. Outros autores (Fienberg e Mason, 1985; Clogg, 1979 e 1982; Wilmoth, 1985, 1990 e 1997) propuseram soluções para este problema e argumentam que resultados úteis podem ser derivados em um arcabouço IPC. Não obstante, todos os autores concordam que o problema de identificação que atormenta a análise IPC é de difícil resolução.

O problema de identificação e estratégias de estimação

Uma vez que há uma grande controvérsia sobre a metodologia de análise de coorte, é necessário justificar os métodos específicos a serem usados. Os modelos IPC de contagem são objeto de atenção há muito tempo, em parte devido ao desafio colocado pelo problema de identificação, o qual é inexorável, não importa com qual representação se escolha trabalhar. O valor de qualquer uma das variáveis é completamente determinado pelos valores das outras variáveis. Entretanto, cada variável pode ser tomada como um indicador de um conceito importante com um significado conceitualmente distinto, e ignorar um deles poderia ser contraproducente para a análise. Vários estudos ilustram que resultados válidos podem ser derivados dentro de um arcabouço IPC, aceitando-se que não há uma solução mágica para o problema de identificação (Fienberg e Mason, 1985; Wilmoth, 1997). A abordagem para estimar os efeitos de idade, período e coorte sobre uma taxa demográfica começa pela especificação de um modelo geral (Bloom, 1984; Wilmoth, 1997):

$$f(r_{ijk}) = \gamma + I\alpha + P\phi + C'\varphi + \varepsilon_{ijk}$$

⁵ Tal tabela deve ser constituída por categorias de coortes de tamanho equilibrado: suficientemente estreitas para serem úteis e amplas para conterem amostras de tamanho suficiente para que sejam obtidas estimativas confiáveis da variável dependente.

onde r_{ijk} é uma taxa demográfica observada para algum evento que ocorra na idade i no ano j para a coorte k e a função $f(.)$ é uma transformação aplicada às taxas observadas. α , ϕ e φ são vetores de parâmetros de efeitos de idade, período e coorte. O parâmetro γ é uma constante de escala, estabelecendo um nível global para $f(r_{ijk})$, e ε_{ijk} é um termo de erro na especificação do modelo ou nos dados originais e flutuações aleatórias. I' , P' e C' são vetores de variáveis indicadoras que têm um elemento correspondente igual a 1 para cada observação, com todos os outros elementos iguais a 0⁶.

Apesar de o modelo IPC padrão ser simples na forma, seus parâmetros não são facilmente estimados. À primeira vista, parece que estes parâmetros poderiam ser estimados por regressão de mínimos quadrados ordinários; entretanto, devido ao fato de que as variáveis de idade, período e coorte são relacionadas pela identidade $k + i = j$, os parâmetros não podem ser identificados separadamente, ou seja, este é um caso de multicolinearidade perfeita. Isto significa que não há um meio óbvio de se identificar uma única série de estimativas de parâmetros que forneça um ajuste aos dados observados⁷.

O esquema de classificação de idade e coorte nos dados de múltiplas *cross-sections* e a restrição linear nas variáveis independentes levam ao problema de identificação, restrição que leva ao processo de "tradução demográfica" (Ryder, 1965) de taxas de período em taxas de coorte e vice-versa⁸. Poder-se-ia pensar que uma solução seria refinar uma das classificações dos

dados, enquanto se deixa a outra classificação como está; por exemplo, um gradiente de idade mais refinado, enquanto se mantém o gradiente de período constante e menos refinado. No entanto, neste caso, o problema de identificação original permanece e um adicional é criado. Se supusermos que a classificação de período é mais refinada, sem fazermos nada com a classificação de idade, é possível que o problema de identificação original seja resolvido, mas isso só pode ser feito devido à perda de capacidade de localizar os indivíduos ao longo do tempo corretamente. Assim, o problema de identificação não é resolvido. Outra pretensa solução seria reformular o modelo em termos de mensurações contínuas de idade, período e coorte; desse modo, os parâmetros de efeito seriam interpretados como numa regressão. Esta reformulação, no entanto, não resolve o problema de identificação, porque o problema básico permanece. Este problema de identificação afeta a estimação estatística dos parâmetros e sua interpretação subsequente. Assim, é necessário explorar abordagens alternativas de estimação. Algumas estratégias para estimar e/ou modificar o modelo padrão IPC à luz do problema de identificação são: imposição de restrições a partir de pressupostos arbitrários e mensuração direta dos componentes IPC⁹.

A abordagem do problema de identificação que envolve a imposição de restrições nos vetores de parâmetros a serem estimados é discutida amplamente por Mason *et al.* (1973) e Fienberg e Mason

⁶ Do ponto de vista de uma regressão, ϕ é interpretado como a mudança esperada em r quando a coorte muda uma unidade, mantendo constantes idade e período. Entretanto, esta interpretação não é possível, posto que, quando a idade e o período são mantidos constantes, a coorte não pode mudar porque, por definição, há uma dependência linear do tipo $I_i = t - C_i$. Assume que os efeitos de idade, período e coorte são aditivos e lineares. O modelo indica que a idade, o período e a coorte têm efeitos sobre r .

⁷ Tipicamente, as discussões dos modelos IPC enfatizaram o problema da identificação da tendência linear dos parâmetros de idade, período e coorte. Componentes de ordem mais alta destes parâmetros (quadrática, cúbica etc.) e interações entre as três séries de fatores foram geralmente ignorados porque tornam a identificação do modelo mais difícil.

⁸ A noção de tradução demográfica pode ser entendida a partir da redução do termo de coorte a um componente residual, quando a tendência linear ou de longo prazo da coorte se traduz nas dimensões de idade e período.

⁹ Há outras soluções, que não serão exploradas neste artigo, como, por exemplo, a inclusão de termos de interação. Para uma discussão sobre esta estratégia ver Glenn (1976), Fienberg e Mason (1985), Clogg (1982) e Wilmoth (1990 e 1997). Um argumento para não utilizarmos esta estratégia, mesmo reconhecendo sua validade, é que um dos critérios no desenvolvimento da compreensão científica é a parcimônia no modelo explicativo, e a aplicação deste critério envolve simplificações. Segundo Glenn (1976), haveria uma inadequação empírica em modelo que não permite termos de interação, mas a inclusão de termos de interação complicaria tanto o problema de identificação que tem sido geralmente evitada (Fienberg e Mason, 1985; Clogg, 1982).

(1985), que sugeriram um meio de estimar os componentes lineares, com a imposição arbitrária de uma ou mais restrições lineares sobre a relação entre qualquer uma das variáveis preditoras e a variável dependente, idealmente a partir do conhecimento prévio. No contexto de um modelo de múltipla classificação com medidas categóricas de idade, período e coorte, o modelo padrão IPC – com três efeitos – pode ser identificado e os parâmetros estimados a partir do pressuposto de que os efeitos de algum par de coortes (ou idades ou períodos) são iguais. Esta restrição permite a estimação do modelo, apesar de ter o custo de assumir que os efeitos específicos de coorte para as coortes deste par são iguais¹⁰. Ao invés de impor uma única restrição, tal como a igualdade de duas coortes, várias restrições deste tipo podem ser impostas simultaneamente. Por exemplo, várias coortes podem ser restritas a serem iguais entre si, e mesmo todas as coortes (ou grupos de idades ou períodos) podem ser tomadas como iguais. Neste caso, um dos efeitos seria inteiramente eliminado do modelo. Este é o caso especial de solução com dois fatores (ver Pullum, 1980; Rodgers, 1982; Wilmoth, 1990). Assim, diferentes modelos, com diferentes conjuntos de restrições, podem ser então comparados em termos da qualidade do ajuste aos dados, tendo em mente que a validade dos modelos é condicional à validade destas restrições¹¹. Não há como testar os pressupostos, mas é possível tentar diferentes pressupostos de identificação para testar a robustez dos resultados.

Possuindo informações suficientes para legitimar as restrições de identificação, seria possível avançar e abandonar este esquema descritivo, definindo variáveis para as quais idade, período e coorte são somente *proxies*. Idade, período e coorte não são conceitos teóricos, mas indicadores de conceitos como processos biológicos de envelhecimento, condições em períodos críticos nos ciclos de vida das pessoas nascidas em anos específicos, e condições econômicas e políticas. Portanto, idade, período e coorte geralmente não são de interesse intrínseco, mas são usados como indicadores convenientes diretamente mensuráveis de conceitos mais básicos. Sendo possível obter medidas mais diretas destes conceitos, sua mensuração seria mais válida e resultaria no desaparecimento do problema de parâmetros não estimáveis. Se for possível substituir ao menos um deles por um conjunto de variáveis mais diretamente relevantes, todos os parâmetros se tornam estimáveis. Neste sentido, outra solução para o problema de identificação que afeta a análise IPC é a especificação e mensuração de variáveis teóricas para as quais idade, período e coorte são indicadores indiretos, ou seja, mensurar diretamente os fatores cujos efeitos são resumidos pelos coeficientes do modelo padrão IPC. A dificuldade óbvia desta estratégia é o problema de definir e mensurar estas variáveis teóricas; há sempre o risco de haver efeitos relacionados à idade, ao período ou à coorte não captados por estas medidas¹². De acordo

¹⁰ Aplicações desta estratégia são encontradas em Pullum (1980), Clogg (1982), Wilmoth (1990), Halli e Rao (1992) e Rios-Neto e Oliveira (1999).

¹¹ De acordo com Wilmoth (1997), em geral, mesmo se o conhecimento prévio sustenta a escolha de uma restrição identificadora específica, é importante reconhecer a sensibilidade dos resultados a pequenas violações do pressuposto escolhido. Se os efeitos de duas coortes são próximos, mas não iguais, o erro dos resultados derivados de igualá-los seria insignificante para coortes adjacentes, mas bem maior para coortes mais distantes. A importância desta imprecisão depende da aplicação e da sensibilidade dos resultados a qualquer pressuposto e pode ser verificada facilmente.

¹² Por exemplo, se o período é relacionado com uma variável dependente específica devido a um conjunto de condições econômicas especificáveis – taxa de inflação, de desemprego etc. – e se medidas diretas destas condições são disponíveis, o período por si pode ser eliminado do modelo explicativo. Os parâmetros para as variáveis remanescentes – idade, coorte e medidas econômicas – são todos estimáveis. Contudo, há o perigo de haver efeitos relacionados ao período mas não devido às condições econômicas especificadas no modelo; ou seja, estas medidas não captam todos os aspectos, as possíveis influências, de cada dimensão. Isto ocorrendo, as estimativas dos efeitos lineares de idade e coorte seriam enviesadas por estarem contaminadas pelos efeitos lineares residuais do período. Além disto, estimativas dos parâmetros para os termos não-lineares e lineares estariam enviesadas se há interações envolvendo o termo residual de período.

com Clogg (1982), a seleção das variáveis causais apropriadas a serem consideradas na modelagem é uma tarefa teórica tão difícil quanto a aplicação do arcabouço IPC. Ou seja, embora pareça, a princípio, que a mensuração direta seja uma solução óbvia para o problema de identificação IPC, a aplicação prática desta técnica é mais repleta de dificuldades do que o problema inicial. Escolhe-se trabalhar no arcabouço IPC exatamente porque não se conhece suficientemente as variáveis explicativas relevantes que poderiam substituir as *proxies* idade, período e coorte.

O modelo IPC aplicado à participação na força de trabalho

A importância da idade e do período sobre a participação na força de trabalho é bastante documentada. No caso da idade, seus efeitos podem ser pensados na perspectiva da teoria do capital humano, como um indicador de experiência, ou como indicador de posição na estrutura social. No caso do período, a variação da participação na força de trabalho é usualmente tomada como resultado de forças puramente econômicas – por exemplo, a mudança na demanda por trabalho ou outros tipos de flutuações no mercado de trabalho. A estimação de efeitos de período e idade é uma tentativa de resumir os termos empíricos do papel da estrutura etária e da mudança no mercado de trabalho na caracterização da força de trabalho agregada (Clogg, 1979).

O conceito de uma variável de coorte e os efeitos de tal variável sobre a participação na força de trabalho são mais sutis e se justificam no contexto explicativo de variações na distribuição da população em idade ativa entre os estados na força de trabalho. A persistência de características da força de trabalho ao longo do tempo para uma dada coorte é um efeito de coorte, indicando uma mudança social indiretamente observável se as coortes são estatisticamente distintas umas das outras entre diferentes padrões de participação, sendo cada padrão característico de uma coorte particular, e se o efeito de coorte não

pode ser atribuído aos efeitos do envelhecimento e/ou das flutuações de período.

Farkas (1977) faz uma primeira tentativa na direção de incorporar a análise de coorte nos estudos sobre os padrões de participação na força de trabalho. Este autor estima efeitos líquidos de idade, período e coorte sobre o emprego de mulheres brancas usando uma análise de regressão de dados de um arquivo de história contínua de trabalho. Ao conceituar efeitos específicos de período como devidos às flutuações cíclicas da economia, o autor evita a multicolinearidade entre os efeitos de coorte, idade e período e consegue realizar uma análise confiável com noções de coorte usuais em análises da fecundidade e noções macroeconômicas usuais em estudos de emprego. Ele foi então capaz de mensurar diferenças de emprego intercoortes, após considerar o padrão etário do emprego e flutuações de curto prazo na economia. A tendência de coorte de longo prazo de aumento do emprego feminino é comparada com os efeitos estimados de curto prazo associados ao ciclo econômico, e os resultados indicam que ambos os efeitos são significantes, sendo que a tendência de coorte tem um efeito maior.

Clogg (1979 e 1982) aplicou o modelo IPC para analisar taxas de participação na força de trabalho. Segundo ele, há fortes razões para esperar que a variável dependente seja influenciada por fatores relacionados à idade, período e coorte, mas há uma base fraca para supor, *a priori*, uma relação específica entre algum par de efeitos. O autor restringe os dois últimos efeitos de coorte a diferir em 1 ou 2%, como meio de obter um conjunto único de estimativas dos parâmetros. Ele apresenta um modelo com um tipo especial de interação idade-período, em adição aos efeitos principais das três variáveis indexadoras, como meio de captar choques transitórios de período, que influenciam diferencialmente a chance de participação dos grupos etários jovens e idosos. Seus resultados indicam que os efeitos de coorte são muito influentes na conformação da tendência observada. Especificamente, seus resultados indicam que coortes mais

jovens de homens e mulheres brancos e mulheres negras apresentam maiores tendências a participar da força de trabalho do que coortes mais avançadas, enquanto coortes mais jovens de homens negros têm menos tendência a participar do que as mais avançadas. Estes resultados são usados para decompor a mudança ao longo do tempo em uma parte devida aos efeitos de coorte e outra devida aos efeitos de período.

Halli e Rao (1992) aplicam a análise de idade, período e coorte para a participação na força de trabalho usando estimativas de modelos lineares por mínimos quadrados, com a variável dependente sendo obtida a partir de uma transformação logit da variável dependente participação na força de trabalho. Seus modelos de efeitos principais são obtidos a partir das estimativas de regressões específicas para as variáveis *dummies* de idade, período e coorte.

Rios-Neto e Oliveira (1999) aplicam um modelo IPC para a atividade econômica no Brasil metropolitano, usando dados da Pesquisa Mensal de Emprego, com o objetivo de analisar as tendências e propor um método para projeção das taxas específicas de participação na força de trabalho. Este trabalho formaliza a aplicação mediante a implementação de modelos log-lineares topológicos de taxa (*rate models*), que confirmam a viabilidade de se desenhar cenários de projeções mais rigorosos.

Aplicação à participação na força de trabalho no Brasil

A variável de participação na força de trabalho no Brasil é analisada a partir do arcabouço IPC. A motivação para esta análise parte da constatação de que, dividindo a população em idade ativa em distintos estados de participação na força de trabalho, a distribuição das pessoas entre estes estados depende da idade, do período e da coorte. Devido às dificuldades em se especificar o que as variáveis de

idade, período e coorte realmente representam em termos causais, nossa análise faz o mínimo possível de pressupostos. Os modelos usados para estimar os efeitos de idade, período e coorte sobre a participação na força de trabalho são análogos aos usados em contexto similar por Clogg (1979 e 1982) e similares aos discutidos por Pullum (1978 e 1980) e Fienberg e Mason (1985). A partir da observação da experiência parcial de várias coortes, introduzir a variável de coorte e o conceito de coorte na análise da dinâmica da força de trabalho enriquece a compreensão da participação na força de trabalho.

Nossos procedimentos de estimação partem da observação, para o Brasil, da relação da taxa de participação com as variáveis demográficas – idade, período e coorte; ou seja, é possível inferir potenciais efeitos destas variáveis. As taxas mostram grandes flutuações nas três dimensões para as mulheres, na direção de níveis ascendentes; sutis flutuações para os homens, no sentido de um pequeno declínio dos níveis, e mudanças nas estruturas para ambos os sexos. Neste contexto, um ponto importante é a diferença entre as taxas de participação femininas e masculinas. Algumas hipóteses para investigação usando o método IPC para o caso brasileiro são: (a) os efeitos de idade são pronunciados na participação dos homens e das mulheres na força de trabalho; (b) dado que as taxas de participação das mulheres cresceram ao longo do período enquanto ficaram aproximadamente constantes para os homens, os efeitos de período parecem ter afetado as taxas de mulheres mais do que as dos homens; (c) os efeitos de coorte devem ser mais significativos para as mulheres¹³. Diferentes coortes que vivam em momentos de mudanças sociais rápidas na força de trabalho têm diferentes padrões de participação na força de trabalho e estas diferenças de coorte são diferentes entre os grupos de sexo. A abordagem de coorte fornece indicações sobre a variação

¹³ Como os dados estão restritos a um período relativamente curto (18 anos), portanto, com histórias incompletas de vida, a generalização dos resultados é limitada, não havendo como conhecer as diferenças intercoortes antes da data inicial.

verdadeira de período da participação da força de trabalho e sua distribuição etária (Clogg, 1979).

A distribuição entre os estados de participação na força de trabalho depende da idade, porque a força de trabalho e os retornos do mercado de trabalho são organizados de maneiras distintas em relação à idade do indivíduo. Depende do período no qual a distribuição é observada, porque as flutuações do mercado de trabalho mudam o risco da participação na força de trabalho. As taxas de período podem variar devido a genuínos efeitos de período ou devido a mudanças intercoortes no nível ou *timing* do evento demográfico. E depende da coorte, devido a um composto de efeitos: tamanho de coorte – como discutido em Leveson e Newitt (1980), Easterlin (1978) e Welch (1979) –, influências de socialização específicas por coorte, experiências peculiares de coortes e dotações de capital humano, tal como o nível educacional (Weiss e Lillard, 1978), são alguns atributos de coorte que podem ser relevantes. O método IPC decompõe tais efeitos sobre as taxas de participação na força de trabalho; esperam-se diferenças na capacidade de cada fator explicar a tendência destas taxas. Um dos objetivos deste método aplicado à força de trabalho é associar os efeitos de coorte com as tendências de participação de longo prazo, e contrastar seus efeitos com os efeitos de período, no curto prazo.

Modelagem das taxas

Assim como em Clogg (1979) e Rios-Neto e Oliveira (1999), são usados neste artigo modelos multiplicativos para análise de coorte baseados em pesquisas amostrais sucessivas. Considerando conjuntamente as variáveis de idade, período e coorte para o caso da variável dependente dicotômica de interesse (participação na força de trabalho), propõe-se a utilização de modelos log-lineares. No caso, os modelos das taxas de participação se baseiam em modelos de contagem e são estimados através de regressões de Poisson em que o logaritmo da população economicamente

ativa é a variável dependente, controlando pelo logaritmo da população em idade ativa. Utilizamos a regressão de Poisson para modelar a taxa de participação na força de trabalho como uma função de algumas variáveis independentes: idade, período e coorte. Na regressão de Poisson, assume-se que a variável dependente Y , número de ocorrências de um evento, tem uma distribuição de Poisson, dadas as variáveis independentes X_1, X_2, \dots, X_m (Cameron e Trivedi, 1998):

$$P(Y = k \mid x_1, x_2, \dots, x_m) = \frac{e^{-\mu} \mu^k}{k!}$$

onde o logaritmo da média μ é pressuposto como uma função linear das variáveis independentes:

$$\log(\mu) = \text{intercepto} + b_1 X_1 + b_2 X_2 + \dots + b_m X_m$$

que implica que μ é uma função exponencial de variáveis independentes

$$\mu = e^{(\text{intercepto} + b_1 X_1 + b_2 X_2 + \dots + b_m X_m)}$$

Em nossa aplicação, é necessário modelar a taxa de incidência do evento ao invés do número de ocorrências. A incidência é o número de ocorrências dividido pelo número total de pessoas em risco; assim, um maior número de pessoas expostas resulta em maior número de ocorrências, devido à taxa de ocorrência e à composição da população, sendo necessário ajustar pelo número de pessoas em risco. O logaritmo da incidência,

$\log\left(\frac{\mu}{N}\right)$ é modelado, portanto, como uma função linear das variáveis independentes:

$$\log\left(\frac{\mu}{N}\right) = \text{intercepto} + b_1 X_1 + b_2 X_2 + \dots + b_m X_m$$

O modelo de regressão de Poisson então é expresso como:

$$\log(\mu) = \log(N) + \text{intercepto} + b_1 X_1 + b_2 X_2 + \dots + b_m X_m$$

onde N é o número total de pessoas em risco (população em idade ativa). A não ser pelo termo de $\log(N)$, este modelo seria o

mesmo usado para modelar o número de ocorrências. O logaritmo da variável N é usado como *offset*, isto é, uma variável da regressão com um coeficiente constante de 1 para cada observação. O modelo IPC com efeitos principais teria a especificação:

$$\log\left(\frac{PEA}{PIA}\right) = \text{intercepto} + b^I X^I + b^P X^P + b^C X^C$$

Os parâmetros no lado direito desta equação representam os principais efeitos das três variáveis independentes, idade, período e coorte, sobre a variável dependente, taxa de participação na força de trabalho. Os parâmetros de idade denotam os efeitos das categorias etárias sobre a taxa, e os outros parâmetros são interpretados da mesma forma. Os efeitos de idade indicam o caráter de evolução etária da participação na força de trabalho, ou efeitos de ciclo de vida; os efeitos de período se referem aos choques transitórios no mercado de trabalho e a outras influências específicas por período; os efeitos de coorte se referem às tendências das coortes a participar. O método de máxima verossimilhança é usado para estimar os parâmetros dos modelos de regressão de Poisson, gerando estatísticas de razão de verossimilhança, χ^2 , dos modelos e identificando as estimativas dos parâmetros. Como já mencionado, há um problema de identificação que deve ser encarado para se obter estimativas de parâmetros não enviesadas.

Este modelo IPC com três fatores está sujeito a restrições de identificação discutidas anteriormente¹⁴. Para satisfazer as restrições da regressão com variáveis *dummy* de idade, período e coorte, restringi-las implica o estabelecimento de um coeficiente igual a zero em duas variáveis e dois coeficientes iguais a zero na outra variável. A solução clássica para identificação dos parâmetros é restringir como

iguais, na maioria das vezes nulos, os coeficientes estimados para duas coortes adjacentes. Partindo da hipótese de que, com o envelhecimento, a identidade de coorte se turva, a idade se torna progressivamente menos precisa como indicadora das características sociais de uma pessoa e coortes adjacentes tendem a se permear, restringimos os coeficientes das duas coortes mais velhas ao valor nulo¹⁵. Em nossa aplicação, há *dummies* para 18 grupos de idades, 7 períodos e 24 coortes; no modelo com três fatores, I_{18} , P_7 , C_{23} e C_{24} foram restritos a zero. No contexto do modelo log-linear, esta saída de identificação dos parâmetros permite estimar efeitos de idade, período e coorte.

Modelos com dois fatores também são aplicados e seus ajustes são comparados: idade e período, idade e coorte, período e coorte. Como dito anteriormente, isto equivale a restringir os efeitos de um fator a serem nulos, ou todos os parâmetros deste fator em 1. Por exemplo, no modelo que exclui coorte, os efeitos de coorte estão implicitamente presentes, mas são iguais a 1; no modelo completo, os efeitos de coorte estão explicitamente presentes e podem variar. Neste sentido, os modelos de dois fatores são casos especiais do modelo de três fatores. Hierarquicamente, todos os modelos de dois fatores estão no mesmo nível e são inferiores aos de três fatores. É contrastada com estes modelos com dois e três fatores uma série de modelos que relacionam as taxas aos parâmetros que denotam os principais efeitos de idade, período e coorte. Cada modelo a ser considerado é um caso especial do modelo completo. O modelo de idade expressa a taxa somente em termos dos efeitos principais de idade, fixando os parâmetros de período e coorte em 1. O mesmo ocorre com os modelos de período e de coorte.

Outra maneira de evitar a multicolinearidade, já discutida anteriormente, é tratar a

¹⁴ Estimativas de modelos lineares por mínimos quadrados com a transformação logital da variável dependente, para solucionar o problema de identificação, foram feitas por Halli e Rao (1992).

¹⁵ Mason *et al.* (1973) e Glenn (1976), entre outros, ressaltam o conhecimento da situação substantiva e seu uso na escolha das restrições, já que a exigência de que dois efeitos sejam iguais tem repercussões em todas as estimativas, e diferentes parametrizações são derivadas dependendo de que par de efeitos é fixado como igual.

idade e a coorte como variáveis categóricas e o período como uma co-variável contínua representada por medidas diretas, tais como médias de anos de estudo, de renda familiar *per capita* e de salário/hora. Da mesma forma, é possível tratar a idade e o período como variáveis categóricas e a coorte como uma co-variável contínua representada por medidas diretas, tais como médias de anos de estudo. Também é importante incluir medidas de mudanças de longo prazo além dos indicadores de ciclo econômico, usualmente incluindo variáveis de tendência temporal, t e t^2 , nas regressões, partindo da premissa de que as forças operam contínua e suavemente de forma linear ou curvilínea ao longo do tempo.

Dados e variáveis

As tabelas para mulheres e homens separadamente foram organizadas de modo que as três variáveis ficassem explícitas (Tabelas 1 e 2). Estas são a variável de linha (l) com categorias $i = 18, \dots, 1$, correspondentes às idades 15-68 em grupos trianuais; a variável de coluna de período (P) com classes $j = 7, \dots, 1$, correspondentes aos anos 1981-1999; e a variável dependente da participação na força de trabalho, cujas taxas foram construídas a partir da classificação de 0 e 1, correspondentes à inatividade e à atividade. As duas categorias de uma variável dependente são definidas pela contagem do número de pessoas ativas e pela contagem do número de pessoas inativas. A passagem destes dados para taxas é feita pela divisão da contagem de ativos por um denominador que é a soma da contagem de ativos e de inativos correspondente. Tal soma configura o chamado *offset*, mencionado anteriormente.

Implícita na tabela 18 X 5 está a variável de coorte (C), marcando a categoria de coorte cronológica na qual um dado indivíduo se localiza¹⁶. Se a idade i e o período j são conhecidos, então a categoria

k de coorte também é conhecida e aqui é dada por $k = i - j + 7$. Uma relação linear exata existe entre l , P e C , da forma $C = l - P + 7$. A variável C está implícita nestas tabelas, mas os dados podem ser reorganizados para torná-la explícita e tornar implícita uma das outras variáveis de tempo ou idade. Há $k = 1, \dots, 24$ categorias de coorte incluídas no estudo por pelo menos um período, correspondentes ao número de diagonais na tabela com 18 categorias de idade e 7 períodos. O grupo de pessoas mais recentemente entrando na força de trabalho denota a coorte 1 e assim sucessivamente. Dado que a aplicação do método IPC requer intervalos etários iguais, o Quadro 1 apresenta a configuração do esquema IPC (como em Halli e Rao, 1992).

Tais dados apresentam limitações, conforme afirma Pullum (1978). Analiticamente, a identificação de coorte tem a mesma importância que identificações de idade e período. Entretanto, as diagonais de uma tabela não chamam a atenção da mesma forma que as linhas e colunas. Não é convencional mostrar taxas para diagonais, e apesar de as linhas terem o mesmo tamanho, assim como as colunas, as diagonais variam de tamanho. Uma coorte pode, no máximo, ser observada por 18 anos, como as coortes 7 a 18; algumas coortes são observadas por poucos anos, tal que pouco se espera conhecer sobre suas experiências particulares. O cruzamento de idade X período foi escolhido sobretudo por conveniência; as linhas e colunas deste cruzamento demonstram um grau muito elevado de regularidade e é possível explorar esta regularidade de estrutura por idade e período para apreender algo sobre a natureza da atividade de coorte. As variáveis que representam as medidas diretas de período e coorte foram calculadas por médias, nos casos de anos de estudo (para coorte, foi feita uma projeção para as que não tinham ainda completado 25 anos de idade), de renda familiar *per capita* e de salário/hora.

¹⁶ As pessoas em um grupo de idade x em um ano t podem ter nascido no ano $t-x$, ou $t-x-1$, ou $t-x-2$; assim, três anos de nascimento estão associados a cada coorte.

QUADRO 1
Configuração do Esquema IPC

| Idade | Período | 1981 | 1984 | 1987 | 1990 | 1993 | 1996 | 1999 |
|-------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| | | P ₇ | P ₆ | P ₅ | P ₄ | P ₃ | P ₂ | P ₁ |
| 15-17 | A ₁₈ | C ₇ | C ₆ | C ₅ | C ₄ | C ₃ | C ₂ | C ₁ |
| 18-20 | A ₁₇ | C ₈ | C ₇ | C ₆ | C ₅ | C ₄ | C ₃ | C ₂ |
| 21-23 | A ₁₆ | C ₉ | C ₈ | C ₇ | C ₆ | C ₅ | C ₄ | C ₃ |
| 24-26 | A ₁₅ | C ₁₀ | C ₉ | C ₈ | C ₇ | C ₆ | C ₅ | C ₄ |
| 27-29 | A ₁₄ | C ₁₁ | C ₁₀ | C ₉ | C ₈ | C ₇ | C ₆ | C ₅ |
| 30-32 | A ₁₃ | C ₁₂ | C ₁₁ | C ₁₀ | C ₉ | C ₈ | C ₇ | C ₆ |
| 33-35 | A ₁₂ | C ₁₃ | C ₁₂ | C ₁₁ | C ₁₀ | C ₉ | C ₈ | C ₇ |
| 36-38 | A ₁₁ | C ₁₄ | C ₁₃ | C ₁₂ | C ₁₁ | C ₁₀ | C ₉ | C ₈ |
| 39-41 | A ₁₀ | C ₁₅ | C ₁₄ | C ₁₃ | C ₁₂ | C ₁₁ | C ₁₀ | C ₉ |
| 42-44 | A ₉ | C ₁₆ | C ₁₅ | C ₁₄ | C ₁₃ | C ₁₂ | C ₁₁ | C ₁₀ |
| 45-47 | A ₈ | C ₁₇ | C ₁₆ | C ₁₅ | C ₁₄ | C ₁₃ | C ₁₂ | C ₁₁ |
| 48-50 | A ₇ | C ₁₈ | C ₁₇ | C ₁₆ | C ₁₅ | C ₁₄ | C ₁₃ | C ₁₂ |
| 51-53 | A ₆ | C ₁₉ | C ₁₈ | C ₁₇ | C ₁₆ | C ₁₅ | C ₁₄ | C ₁₃ |
| 54-56 | A ₅ | C ₂₀ | C ₁₉ | C ₁₈ | C ₁₇ | C ₁₆ | C ₁₅ | C ₁₄ |
| 57-59 | A ₄ | C ₂₁ | C ₂₀ | C ₁₉ | C ₁₈ | C ₁₇ | C ₁₆ | C ₁₅ |
| 60-62 | A ₃ | C ₂₂ | C ₂₁ | C ₂₀ | C ₁₉ | C ₁₈ | C ₁₇ | C ₁₆ |
| 63-65 | A ₂ | C ₂₃ | C ₂₂ | C ₂₁ | C ₂₀ | C ₁₉ | C ₁₈ | C ₁₇ |
| 66-68 | A ₁ | C ₂₄ | C ₂₃ | C ₂₂ | C ₂₁ | C ₂₀ | C ₁₉ | C ₁₈ |

Resultados

A exploração preliminar dos dados indicou uma regularidade do padrão etário para coortes ao longo do tempo entre os homens e uma mudança entre as mulheres. Estas constatações não respondem à questão de quais efeitos – de período ou de coorte – predominam na determinação das mudanças na participação na força de trabalho, sobretudo entre as mulheres. Outra questão a ser respondida é se o nível educacional é determinante da participação, como indicado pelos efeitos de coorte ou de período. A discussão sobre os fatores que potencialmente determinam os efeitos de coorte e de período é um ponto fundamental desta análise. A hipótese mais forte é que a experiência educacional diferenciada das coortes seria um determinante importante da variação da participação na força de trabalho, principalmente entre as mulheres. Este efeito de educação de coorte seria mais forte do que a medida de educação de período, o que leva à hipótese de que a evolução dos níveis de escolaridade ao longo do tempo determina, de alguma forma, os efeitos de período, assim como teriam alguma influência as tendências de variação da renda *per capita* e do salário/hora. Os efeitos

destes últimos indicadores de período seriam também mais fortes do que o efeito de educação de período.

Uma direção de interpretação dos modelos estimados é a qualidade do ajuste, ou seja, a extensão em que as expectativas modeladas concordam com as observações. Esta avaliação do ajuste é feita por meio da comparação entre os valores preditos e observados em cada célula da tabela de contingência, ou por meio de uma medida global de adequação do modelo. Mudanças nas estatísticas de qualidade do ajuste são usadas para avaliar a contribuição de subconjuntos de variáveis explicativas para um modelo. Uma estratégia para seleção das variáveis é ajustar uma seqüência de modelos, começando com um modelo simples, no caso o nulo, com somente um termo de intercepto, e depois incluir uma variável explicativa adicional em cada modelo sucessivamente.

Uma medida de qualidade de ajuste é G^2 , a estatística χ^2 da razão de verossimilhança ou desvio em relação ao modelo saturado (*deviance*). Esta estatística de teste compara o modelo restrito vs. o irrestrito e tem distribuição assintótica χ^2 sob o pressuposto de que o modelo restrito é verdadeiro, sendo que para o modelo

saturado $G^2 = 0$. Os graus de liberdade são a diferença entre o número de células iniciais e o número de parâmetros ajustados (Powers e Xie, 2000). Tomando como modelo de referência o modelo nulo, a significância de cada modelo estimado é medida pela diferença entre os desvios ou log-verossimilhanças ajustados entre os modelos sucessivos e o nulo, dG^2 , a qual permite mensurar a importância da variável explicativa adicional. R^2 é a razão entre esta diferença e o valor de G^2 do modelo nulo, indicando em que medida o modelo se aproxima mais do modelo saturado, o que representa a melhoria da qualidade do ajuste em relação ao modelo nulo. Ou seja, a inspeção dos valores de R^2 indica os melhores modelos.

Neste ponto, vale mencionar que, nos modelos estimados, foram inicialmente detectados problemas de sobredispersão. Para a distribuição de Poisson, a média e a variância são iguais, o que implica que G^2 dividida pelos graus de liberdade deve ser aproximadamente 1. Os valores encontrados foram maiores que 1, indicando sobredispersão, ou seja, a variância verdadeira é maior que a média, o que sugeriria um ajuste inadequado do modelo de Poisson. Foi necessário, então, estimar modelos de regressão binomial negativa para testar a razão de verossimilhança baseado nas distribuições Poisson e binomial negativa¹⁷. Este procedimento permite testar a igualdade da média e da variância imposta pela distribuição de Poisson vs. a alternativa de que a variância é maior que a média. Os valores da log-verossimilhança destes modelos binomiais negativos foram então comparados aos valores dos modelos Poisson, mediante o teste da razão de verossimilhança, para

verificar a significância da sobredispersão. Em todos os casos, não se rejeitou a hipótese nula de que o parâmetro de dispersão fosse igual a 0, ou seja, sempre se chegou à conclusão de que a média e a variância seriam iguais. Assim, o pressuposto da distribuição de Poisson não teve de ser abandonado, mas somente corrigido, fixando-se o valor da razão entre G^2 e os graus de liberdade em 1.

Dado que o tamanho das amostras difere de amostra para amostra, e dado que a variação total não explicada nos dados (dada por G^2) também é diferente de amostra para amostra, esta estatística, sozinha, é difícil de interpretar. Há na literatura (Powers e Xie, 2000) uma crítica à estatística G^2 como procedimento para rejeitar um modelo em favor de outro em grandes amostras. Quando o tamanho da amostra é grande, é mais fácil aceitar modelos mais complexos porque G^2 é planejado para detectar qualquer desvio entre um modelo e os dados observados; neste sentido, adicionar mais termos a um modelo sempre melhoraria o ajuste. Uma solução é usar a estatística Critério de Informação Bayesiano (BIC) para buscar modelos mais parcimoniosos que forneçam ajuste adequado aos dados, o que é um dos principais objetivos de nossa análise. O BIC fornece uma aproximação à razão de verossimilhança entre um modelo (M_0) e outro (M_1), a partir da idéia de comparar a plausibilidade relativa dos dois modelos e não achar o desvio absoluto dos dados observados em relação a um modelo específico. Sua fórmula é: $BIC = G^2 - \text{graus de liberdade} \cdot \log n$. Esta expressão mostra que o BIC utiliza os graus de liberdade e o tamanho da amostra (n) para avaliar a qualidade do ajuste. Um BIC negativo

¹⁷ Para a distribuição binomial negativa, a variância = média + $k \cdot \text{média}^2$ ($k > 0$; a distribuição binomial negativa se reduz à Poisson quando $k = 0$). A hipótese nula é: $H_0: k = 0$ e a hipótese alternativa é: $H_a: k > 0$. Ajustar uma regressão binomial negativa é similar ao ajuste da regressão de Poisson. Assume-se que o modelo é o mesmo, ou seja, que o log da média, μ , é uma função linear de variáveis independentes, $\log(\mu) = \alpha + b_1X_1 + b_2X_2 + \dots + b_mX_m$, que implica que μ é uma função exponencial de variáveis independentes $\mu = e^{\alpha + b_1X_1 + b_2X_2 + \dots + b_mX_m}$. Ao invés de assumir que a distribuição é Poisson, assume-se uma distribuição binomial negativa da variável dependente. Isto significa relaxar o pressuposto de igualdade entre a média e a variância, dado que a variância da binomial negativa é igual a $\mu + k \mu^2$, onde $k >= 0$ é o parâmetro de dispersão. O método de máxima verossimilhança é usado para estimar k e os parâmetros do modelo de regressão para $\log(\mu)$.

significa que o modelo M_0 é mais provável que M_j , tal que deva ser escolhido. Comparando múltiplos modelos, um valor mais baixo de BIC significa um modelo com melhor ajuste.

Uma medida alternativa da adequação dos modelos é fornecida pela estatística de razão de verossimilhança, G^2 , dividida por seus graus de liberdade, que permite verificar a progressão desta quantidade, dos modelos mais simples para os mais complexos, para cada grupo. Se um aumento no grau de ajuste do modelo com três fatores é insignificante, este é descartado em favor do modelo mais simples com dois fatores. A medida em que o modelo com três fatores representa uma melhoria em relação aos modelos com dois fatores determina a relativa importância das distâncias. Outra medida da qualidade do modelo que não é afetada pelo tamanho da amostra é Δ , o índice de dissimilaridade:

$$\Delta = \left(\frac{1}{2} n\right) \sum |n_{ijm} - \hat{n}_{ijm}|$$

onde n é o tamanho da amostra total. Este índice informa a proporção de casos que necessitariam mudar na tabela observada (ou esperada) para alcançar correspondência exata com a tabela esperada (ou observada). Para procedermos à escolha do modelo com melhor ajuste, os dados apresentados nas Tabelas 3 e 4 representam as tentativas de ajuste dos modelos para as taxas de participação na força de trabalho das mulheres e dos homens, respectivamente.

As informações da Tabela 3 indicam que a dimensão de idade (modelo 1) apresenta o maior poder preditivo da participação feminina: 83%. O modelo 3, somente com a dimensão de coorte, apresenta o segundo maior poder preditivo (58%), ao passo que a dimensão temporal do modelo de período (modelo 2) explica um componente mais baixo das taxas femininas. Os modelos 4 e 5, com dois fatores, de idade e período e de idade e coorte, explicam uma grande proporção da dispersão da tabela de contingência (cerca de 99%). Quando o

TABELA 3
Ajuste dos modelos – mulheres, Brasil urbano

| Modelos | G ² | g.l. | DG ² | dg.l. | R ² (%) | BIC | G ² /g.l. | Δ |
|--------------------------------------|----------------|------|-----------------|-------|--------------------|----------|----------------------|----------|
| 0 – Nulo | 27425.48 | 125 | | | | | 219.40 | 11.59 |
| 1 – Idade | 4668.75 | 108 | 22756.73 | 17 | 82.98 | 4146.43 | 43.23 | 5.40 |
| 2 – Período | 23282.00 | 119 | 4143.48 | 6 | 15.11 | 22706.48 | 195.65 | 10.56 |
| 3 – Coorte | 11568.09 | 102 | 15857.38 | 23 | 57.82 | 11074.79 | 113.41 | 7.81 |
| 4 – Idade + Período | 360.77 | 102 | 27064.70 | 23 | 98.68 | -132.53 | 3.54 | 1.22 |
| 5 – Idade + Coorte | 282.84 | 85 | 27142.64 | 40 | 98.97 | -128.24 | 3.33 | 1.15 |
| 6 – Período + Coorte | 8451.38 | 96 | 18974.10 | 29 | 69.18 | 7987.10 | 88.04 | 8.20 |
| 7 – Idade + Período + Coorte | 115.71 | 80 | 27309.77 | 45 | 99.58 | -271.20 | 1.45 | 0.72 |
| 8 - I + M.Ed.C | 647.33 | 107 | 26778.15 | 18 | 97.64 | 129.85 | 6.05 | 1.84 |
| 9 - I + P + M.Ed.C | 302.48 | 101 | 27122.99 | 24 | 98.90 | -185.98 | 2.99 | 1.11 |
| 10 - I + M.Ed.P | 685.53 | 107 | 26739.95 | 18 | 97.50 | 168.04 | 6.41 | 1.74 |
| 11 - I + C + M.Ed.P | 280.71 | 84 | 27144.76 | 41 | 98.98 | -125.53 | 3.34 | 1.14 |
| 12 - I + C + M.Ed.P + t2 | 243.14 | 83 | 27182.33 | 42 | 99.11 | -158.27 | 2.93 | 1.03 |
| 13 - I + C + M.R.Pc.P | 264.57 | 84 | 27160.90 | 41 | 99.04 | -141.67 | 3.15 | 1.16 |
| 14 - I + C + M.R.Pc.P + t2 | 223.91 | 83 | 27201.56 | 42 | 99.18 | -177.50 | 2.70 | 0.99 |
| 15 - I + C + M.Sal/H.P | 280.16 | 84 | 27145.31 | 41 | 98.98 | -126.09 | 3.34 | 1.15 |
| 16 - I + C + M.Sal/H.P + t2 | 245.28 | 83 | 27180.20 | 42 | 99.11 | -156.13 | 2.96 | 1.06 |
| 17 - I + MedC + M.R.Pc.P + t | 488.45 | 105 | 26937.02 | 20 | 98.22 | -19.36 | 4.65 | 1.43 |
| 18 - I + MedC + M.R.Pc.P + t + t2 | 401.30 | 104 | 27024.18 | 21 | 98.54 | -101.67 | 3.86 | 1.30 |
| 19 - I + MedC + M.Sal/H.P + t | 503.85 | 105 | 26921.62 | 20 | 98.16 | -3.96 | 4.80 | 1.42 |
| 20 - I + MedC + M.S.al/H.P + t + t2 | 424.70 | 104 | 27000.78 | 21 | 98.45 | -78.28 | 4.08 | 1.33 |
| 21 - I + MedC + Tend.Rpc + Ciclo Rpc | 488.54 | 105 | 26936.94 | 20 | 98.22 | -19.27 | 4.65 | 1.43 |
| 22 - I + MedC + Tend.S/h + Ciclo S/h | 495.96 | 105 | 26929.52 | 20 | 98.19 | -11.85 | 4.72 | 1.40 |
| 23 I + MedC + t + t2 | 425.10 | 105 | 27000.38 | 20 | 98.45 | -82.71 | 4.05 | 1.33 |

Fonte: Elaboração própria, a partir dos microdados das PNADs 1961-1999.

Nota: G² = desvio; g.l. = graus de liberdade; dG²=diferença entre os desvios (modelo – modelo nulo); dg.l. = diferença entre os graus de liberdade; R²= dG²/G² modelo nulo; BIC = G² – g.l. log n; Δ = índice de dissimilaridade.

Nota: I = idade; P = período; C = coorte; M.Ed.C = média de educação da coorte; M.Ed.P = média de educação do período; M.R.Pc.P = média de renda *per capita* do período; M.Sal/H.P = média de salário/hora do período; TendRpc = medida de tendência da renda *per capita*; Ciclo Rpc = medida de ciclo da renda *per capita*; Tend.S/h. = medida de tendência do salário/hora; Ciclo S/h. = medida de ciclo do salário/hora; t e t2 = medidas de tendência temporal.

efeito de idade é eliminado, o montante da variação das taxas de participação na força de trabalho explicada para as mulheres é somente de 70%. O modelo 7, de independência condicional ou de efeitos principais, explica 99,6% da dispersão observada na tabela de contingência, sendo o mais forte preditor dentre todos os modelos.

Entre os homens (Tabela 4), o modelo de idade explica aproximadamente 98,5% da variância total; a alta porcentagem da variância explicada pelo modelo reflete principalmente os contrastes na participação ao longo das idades, um padrão cujos contornos gerais mudaram pouco para os homens, mas mudaram muito para as mulheres. O modelo de período praticamente não explica nada na variação masculina (0,62%). O modelo de coorte explica cerca de 52%. Assim como entre as mulheres, entre os homens os modelos 4 e 5, com dois fatores, de idade e período e de idade e coorte, explicam cerca de 99% da dispersão da tabela de contingência. Eliminar o efeito de idade implica um montante da

variação explicada para os homens de apenas 52% (modelo 6). O modelo de três fatores também é o mais forte preditor entre os homens.

Na explicação da variação das taxas de participação feminina na força de trabalho, além do modelo que inclui a média de educação de coorte (modelo 9), os modelos que incluem as médias de período, juntamente com os efeitos de idade e de coorte e a tendência temporal, fornecem ótimos ajustes (modelos 12, 14 e 16). No caso das mulheres, portanto, a educação parece se ajustar bem tanto para as coortes quanto para os períodos, mas o melhor ajuste para coortes é bem mais claro (modelo 9), segundo todas as medidas expostas na Tabela 3. Estes resultados viriam no sentido de confirmar as hipóteses de que a evolução dos níveis de escolaridade ao longo do tempo determina, de forma importante, os efeitos de período entre as mulheres, assim como teriam alguma influência as tendências de variação da renda *per capita* e do salário/hora.

TABELA 4
Ajuste dos modelos – homens, Brasil urbano

| Modelos | G ² | g.l. | dG ² | dg.l. | R ² (%) | BIC | G ² /g.l. | Δ |
|--------------------------------------|----------------|------|-----------------|-------|--------------------|----------|----------------------|------|
| 0 – Nulo | 20208.50 | 125 | | | | | 161.67 | 7.98 |
| 1 – Idade | 302.69 | 108 | 19905.80 | 17 | 98.50 | -219.62 | 2.80 | 0.77 |
| 2 – Período | 20083.95 | 119 | 124.55 | 6 | 0.62 | 19508.43 | 168.77 | 7.98 |
| 3 – Coorte | 9694.18 | 102 | 10514.32 | 23 | 52.03 | 9200.88 | 95.04 | 4.85 |
| 4 – Idade + Período | 199.84 | 102 | 20008.66 | 23 | 99.01 | -293.47 | 1.96 | 0.65 |
| 5 – Idade + Coorte | 140.54 | 85 | 20067.96 | 40 | 99.30 | -270.55 | 1.65 | 0.56 |
| 6 – Período + Coorte | 9513.83 | 96 | 10694.67 | 29 | 52.92 | 9049.55 | 99.10 | 5.56 |
| 7 – Idade + Período + Coorte | 91.94 | 80 | 20116.56 | 45 | 99.55 | -294.96 | 1.15 | 0.50 |
| 8 - I + M.Ed.C | 283.57 | 107 | 19924.93 | 18 | 98.60 | -233.91 | 2.65 | 2.62 |
| 9 I + P + M.Ed.C | 193.05 | 101 | 20015.45 | 24 | 99.04 | -295.42 | 1.91 | 0.64 |
| 10 - I + M.Ed.P | 267.30 | 107 | 19941.20 | 18 | 98.68 | -250.19 | 2.50 | 2.61 |
| 11 - I + C + M.Ed.P | 136.86 | 84 | 20071.63 | 41 | 99.32 | -269.39 | 1.63 | 0.57 |
| 12 - I + C + M.Ed.P + t2 | 110.09 | 83 | 20098.40 | 42 | 99.46 | -291.32 | 1.33 | 0.52 |
| 13 - I + C + M.R.Pc.P | 138.37 | 84 | 20070.12 | 41 | 99.32 | -267.88 | 1.65 | 0.57 |
| 14 - I + C + M.R.Pc.P + t2 | 119.92 | 83 | 20088.57 | 42 | 99.41 | -281.49 | 1.44 | 0.53 |
| 15 - I + C + M.Sal/H.P | 139.84 | 84 | 20068.66 | 41 | 99.31 | -266.41 | 1.66 | 0.57 |
| 16 - I + C + M.Sal/H.P + t2 | 122.35 | 83 | 20086.14 | 42 | 99.39 | -279.06 | 1.47 | 0.53 |
| 17 - I + MedC + M.R.Pc.P + t | 256.12 | 105 | 19952.38 | 20 | 98.73 | -251.69 | 2.44 | 0.68 |
| 18 - I + MedC + M.R.Pc.P + t + t2 | 214.83 | 104 | 19993.67 | 21 | 98.94 | -288.15 | 2.07 | 0.66 |
| 19 - I + MedC + M.S./H.P + t | 256.89 | 105 | 19951.60 | 20 | 98.73 | -250.92 | 2.45 | 0.68 |
| 20 - I + MedC + M.S./H.P + t + t2 | 216.74 | 104 | 19991.75 | 21 | 98.93 | -286.23 | 2.08 | 0.66 |
| 21 - I + MedC + Tend.Rpc + Ciclo Rpc | 256.10 | 105 | 19952.39 | 20 | 98.73 | -251.71 | 2.44 | 0.68 |
| 22 - I + MedC + Tend.S/h + Ciclo S/h | 256.95 | 105 | 19951.55 | 20 | 98.73 | -250.86 | 2.45 | 0.68 |
| 23 I + MedC + t + t2 | 219.30 | 105 | 19989.20 | 20 | 98.91 | -288.51 | 2.09 | 0.63 |

Fonte: Elaboração própria, a partir dos microdados das PNADs 1961-1999.

Nota: G² = desvio; g.l. = graus de liberdade; dG² = diferença entre os desvios (modelo – modelo nulo); dg.l. = diferença entre os graus de liberdade; R² = dG²/G² modelo nulo; BIC = G² - g.l. log n; Δ = índice de dissimilaridade.

Nota: I = idade; P = período; C = coorte; M.Ed.C = média de educação da coorte; M.Ed.P = média de educação do período; M.R.Pc.P = média de renda *per capita* do período; M.Sal/H.P = média de salário/hora do período; TendRpc = medida de tendência da renda *per capita*; Ciclo Rpc = medida de ciclo da renda *per capita*; Tend.S/h. = medida de tendência do salário/hora; Ciclo S/h. = medida de ciclo do salário/hora; t e t2 = medidas de tendência temporal.

Entre os homens (ver Tabela 4), seguindo este critério da medida BIC, outro modelo com bom ajuste é o modelo de dois fatores, com idade e período, o que confirmaria a menor importância do efeito de coorte entre eles. Além deste, outros modelos importantes na explicação da variação das taxas de participação masculinas são os modelos que incluem as médias de período (educação, renda *per capita* e salário/hora) e as variáveis de coorte juntamente com as variáveis de tendência temporal, que partem do pressuposto de que as forças operam contínua e suavemente ao longo do tempo. Vale ressaltar, contudo, que o modelo que inclui somente a média de educação da coorte e as tendências temporais, e não inclui as médias de período, fornece um ajuste ainda melhor. E o modelo 9, com a média de educação de coorte, o período e a idade, também fornece um ótimo ajuste, bastante superior ao ajuste dos modelos que incluem a média de educação de período (modelos 11 e 12).

Quando o efeito de idade é controlado, período e coorte contribuem de forma distinta para homens e mulheres (Tabela 5). Particionando a variação explicada, entre as mulheres, período, coorte, e período e coorte juntos explicam 92% ou mais quando a idade é controlada. Entre os homens, o período explica somente 34%, a coorte somente 53% e período e coorte juntos explicam 70%. Estes resultados refletem a maior importância dos efeitos de coorte e de período sobre a variação da participação feminina na força de trabalho. Como mencionado acima, a variação da participação masculina na força de trabalho explicada pelo efeito de idade reflete os contrastes ao longo do ciclo de vida dos homens, um padrão cujos contornos gerais mudaram pouco. Por outro lado, entre as mulheres este perfil etário mudou bastante ao longo do tempo. Os Gráficos 9, 10 e 11 permitem visualizar a importância destes efeitos¹⁸ para homens e mulheres, utilizando o conjunto de coeficientes dos modelos.

As curvas mostradas no Gráfico 9 indicam a média do padrão etário de atividade para as sucessivas coortes de homens e mulheres ao longo do tempo, estimadas a partir dos modelos. Em ambos os grupos, as taxas são baixas à entrada na força de trabalho, ascendem até um nível onde ficam estáveis, e depois decrescem. O declínio a partir dos 40 anos para as mulheres é bastante acentuado; entre os homens, somente é visualizado a partir dos 50 anos, e de forma suave. Este perfil mais inclinado das mulheres indica a maior saída feminina da atividade econômica ao longo da sua trajetória longitudinal de vida.

O Gráfico 10 mostra os efeitos de período, os quais representam a evolução média da participação ao longo do tempo. É bastante claro que estes efeitos são mais marcantes entre as mulheres, de forma ascendente, refletindo a tendência conhecida de elevação das taxas femininas de participação na força de trabalho ao longo do tempo. Entre os homens, os efeitos de período são estáveis, sendo possível visualizar uma ligeira influência positiva entre 1984 e 1990, e a partir daí uma pequena tendência negativa.

O diferencial de gênero dos efeitos de coorte é o mais marcante, como mostra o Gráfico 11. Os dados apresentados neste gráfico mostram o impacto das várias coortes identificadas pelos anos em que tinham entre 15 e 17 anos. O gráfico mostra que as coortes femininas apresentam níveis crescentes de participação na força de trabalho, ritmo que se acelera e se mantém elevado até meados da década de 90, quando começa a arrefecer. Este arrefecimento se deve ao fato de que nestes últimos anos da década de 90 ainda não se completou o ciclo de vida ativa destas coortes; ao contrário, este somente se iniciou em idades até 24 anos, quando o aumento do nível de escolaridade significa uma maior frequência à escola e concomitante diminuição da participação na força de trabalho, o que é positivo. Entre os homens, os efeitos de coorte sobre a participação são sempre mais suaves.

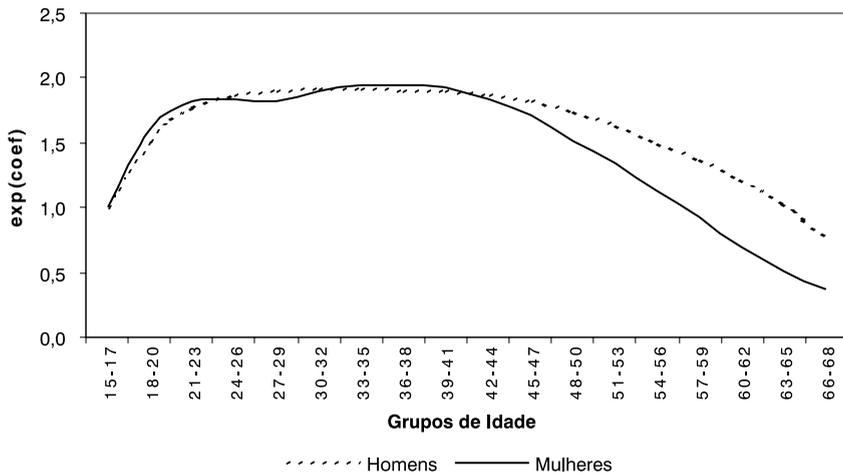
¹⁸Estes efeitos podem ser pensados como taxas médias de participação por idade (médias dos perfis etários de participação de coortes sucessivas ao longo do tempo), período ou coorte.

TABELA 5
Proporção da variação residual (controlando por idade) explicada pelos efeitos de período e coorte

| Partição de R ² | Mulheres | Homens |
|----------------------------|----------|--------|
| 1 - I | 0,1702 | 0,0150 |
| (IP-I)/(1-I) | 0,9224 | 0,3400 |
| (IC-I)/(1-I) | 0,9395 | 0,5333 |
| (IPC-I)/(1-I) | 0,9753 | 0,7000 |

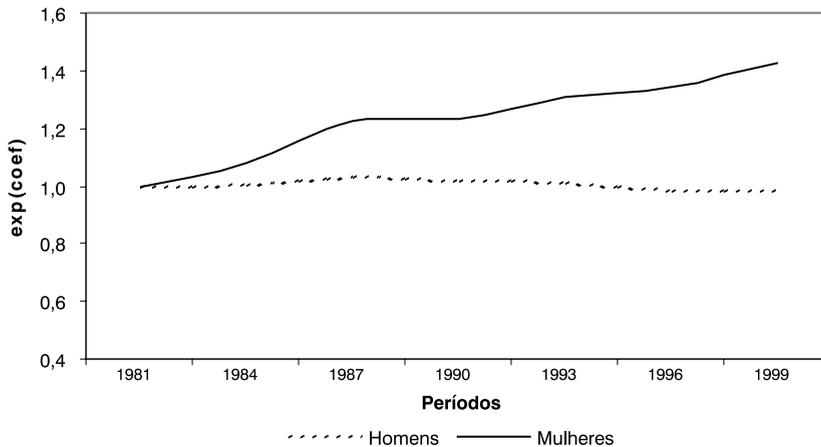
Fonte: Elaboração própria, a partir das Tabelas 3 e 4.
 I = idade, P= período, C = coorte.

GRÁFICO 9
Coefficientes da variável de idade – Brasil urbano

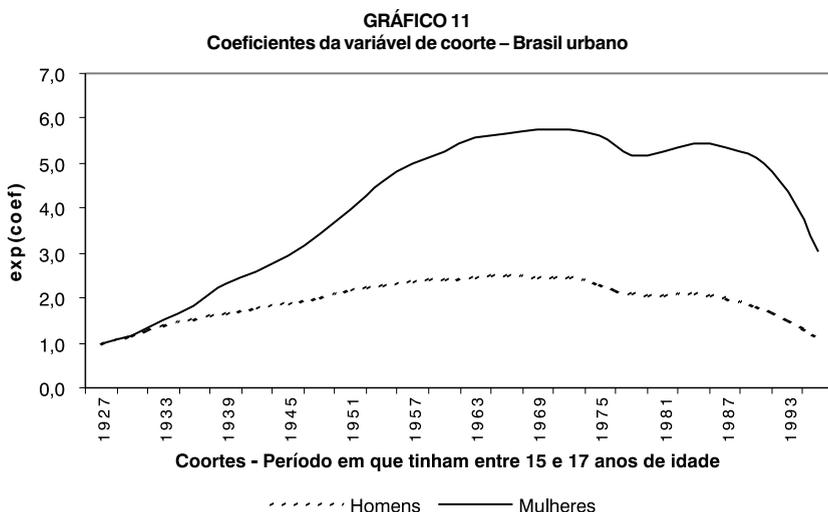


Fonte: Elaboração própria a partir dos coeficientes das regressões.

GRÁFICO 10
Coefficientes da variável de período – Brasil urbano



Fonte: Elaboração própria a partir dos coeficientes das regressões.



Fonte: Elaboração própria a partir dos coeficientes das regressões.

A partir de todas as considerações sobre a qualidade de ajuste dos modelos, poderíamos concluir que o modelo completo, de três fatores, é o mais aceitável tanto para homens quanto para mulheres. No entanto, outro objetivo da nossa análise é encontrar medidas diretas que sejam bons indicadores dos efeitos de período e coorte, buscando modelos mais parcimoniosos e passíveis de diálogo mais amplo com outros campos de análise. O nível educacional parece ser determinante da participação, principalmente como indicado pelos efeitos de coorte. Enfatizando a medida de ajuste (BIC) que leva em consideração o tamanho da amostra e os graus de liberdade, o que reforça a busca pela parcimônia, o modelo que inclui a média de educação da coorte, além dos efeitos de idade e de período, apresenta excelente ajuste, tanto para os homens quanto para as mulheres. Isto sugere a confirmação da hipótese de que a experiência educacional diferenciada das coortes é um determinante importante da variação da participação na força de trabalho, principalmente entre as mulheres.

Conclusões

O propósito deste artigo foi mostrar uma possível utilização dos modelos IPC na

descrição de eventos demográficos complexos e sua aplicação na análise da variabilidade das taxas de participação masculina e feminina no mercado de trabalho brasileiro, sugerindo caminhos metodológicos em situações nas quais se dispõe de dados longitudinais. O principal resultado desta aplicação metodológica foi a constatação da importância de levar em conta a escolaridade das coortes e as diferenças de escolaridade entre coortes na explicação das mudanças nas taxas de participação.

As mudanças nas taxas de atividade feminina ao longo das últimas décadas são marcantes. Fica claro, a partir da análise exploratória dos dados, que as coortes e os períodos têm um impacto muito mais significativo sobre a participação na força de trabalho das mulheres no Brasil nos anos 80 e 90 do que sobre a participação dos homens. Esta constatação foi confirmada e elaborada pela aplicação dos modelos IPC. De fato, ao longo do período estudado, a taxa de participação feminina na força de trabalho elevou-se substancialmente no país, para todas as idades (exceto o grupo etário mais jovem), e tende a continuar em expansão. O quadro não é o mesmo para os homens, que mostram maior consistência na participação na força de trabalho,

entrando e saindo do mercado de trabalho em um padrão estável ao longo do tempo, mais próximo de um *steady state*. Isto se reflete na prevalência dos efeitos de idade entre os homens; num *steady state* perfeito, só o efeito da idade existiria. Entre as mulheres, os efeitos de idade também são os mais importantes, mas os efeitos de período e de coorte mostram-se extremamente relevantes. A diferenciação das coortes femininas quanto à participação na força de trabalho não pode ser atribuída à variação temporal ou etária. Os diferentes resultados do modelo para homens e mulheres ilustram, de forma clara, as diferenças de comportamentos no mercado de trabalho, apontando para a importância da variável de coorte no caso feminino.

Buscando a parcimônia dos modelos através da utilização de medidas diretas que sejam bons indicadores dos efeitos de período e coorte, podemos concluir que o nível educacional parece ser o melhor indicador, para ambos os sexos. A hipótese de que a experiência educacional diferenciada das coortes é um determinante importante da variação da participação na força de trabalho, sobretudo entre as mulheres, foi confirmada. Os ajustamentos dos modelos são bons com um número

relativamente pequeno de variáveis explicativas. Especificamente para o caso das mulheres, a escolha do modelo parcimonioso que inclui a média de escolaridade das coortes foi feita levando em conta as medidas de ajuste dos modelos. Os resultados parecem também confirmar a hipótese de que a evolução dos níveis de escolaridade ao longo do tempo determina de forma importante os efeitos de período entre as mulheres, assim como têm influência as tendências de variação da renda *per capita* e do salário/hora.

A exploração destes modelos amplia a possibilidade de diálogo com outros campos de análise, facilitando a replicação da metodologia, inclusive o desenho de cenários alternativos de projeções. Para se especular sobre o comportamento futuro dos efeitos de coorte é necessário entender melhor que fatores o determinam. Por exemplo, sendo a variável de escolaridade claramente um dos determinantes pessoais mais importantes para a projeção da PEA e tendendo a variar entre as coortes, o desenho de cenários alternativos de oferta de trabalho masculina e feminina pode ser enriquecido pela indicação de que as médias de anos de estudo são um bom marcador das coortes.

Referências bibliográficas

BIANCHI, S. e SPAIN, D. **American women in transition**. Nova York: Russel Sage, 1986. 286p.

BLAU, F.D. Trends in the well-being of American women, 1970-1995. **Journal of Economic Literature**, v. 36, n. 1, 1998, p. 112-165.

BLOOM, D.E. **On the nature and estimation of age, period, and cohort effects in demographic data**. Harvard: Harvard University, Center for Population Studies, 1984 (Discussion Paper 84-7).

CAMERON, A.C. e TRIVEDI, P.K. **Regression analysis of count data**. Nova York: Cambridge University Press, 1998. 411p.

CLOGG, C.C. The dependence of labor force status on age, time-period, and cohort. In: CLOGG, C.C., **Measuring underemployment: demographic indicators for the United States**. Nova York: Academic Press, 1979, chapter 6, p. 121-148.

_____. Cohort analysis of recent trends in labor force participation. **Demography**, v. 19, n. 4, 1982, p. 459-479.

EASTERLIN, R.A. What will 1984 be like? Socioeconomic implication of recent twists in age structure. **Demography**, v. 15, n. 4, 1978, p. 397-432.

FARKAS, G. Cohort, age, and period effects upon the employment of white females. **Demography**, v. 14, n. 1, 1977, p. 33-42.

FIENBERG, S.E. e MASON, W.M. Specification and implementation of age, period and cohort models. In: MASON, W.M. e FIENBERG, S.E. (eds.), **Cohort analysis in social research: beyond the identification problem**. Nova York: Springer, 1985, p. 44-88.

FREEMAN, R. The effects of demographic factors on age-earnings profiles. **Journal of Human Resources**, v. 14, n. 3, 1979, p. 289-318.

LENN, N.D. Cohort analysts' futile quest: statistical attempts to separate age, period, and cohort effects. **American Sociological Review**, v. 41, n. 5, 1976, p. 900-904.

_____. **Cohort analysis**. Beverly Hills, CA: Sage, 1977. 72p.

GOLDIN, C. **Understanding the gender gap: an economic history of American women**. Nova York: Oxford University Press, 1990. 287p.

HALLI, S.S. e RAO, K.V. **Advanced techniques of population analysis**. Nova York: Plenum, 1992. 266p.

HECKMAN, J.J. e ROBB, R. Using longitudinal data to estimate age, period and cohort effects in earnings equations. In: MASON, W.M. e FIENBERG, S.E. (eds.), **Cohort analysis in social research**. Nova York: Springer, 1985, p. 137-150.

HILL, M.A. e O'NEILL, J. E. Intercohort change in women's labor market status. In: EHRENBERG, R.G. (ed.), **Research of Labor Economics**. Greenwich, Conn.: JAI, v. 13, 1992, p. 215-286.

_____. **A study of intercohort change in women's work patterns and earnings**. Nova York: Baruch College, Center for the Study of Business and Government, 1990. 99p.

HOBcraft, J., MENKEN, J. e PRESTON, S. Age, period, and cohort effects in demography: a review. **Population Index**, v. 48, n. 1, 1982, p. 4-43.

JOSHI, H. e PACI, P. **Unequal pay for women and men: evidence from the British cohort studies**. Cambridge, MA: MIT Press, 1998.

LEVESON, I. e NEWITT, J. **Generational crowding: economic, social and demographic effects of changes in relative cohort size**. Quebec: Hudson Institute, 1980.

MASON, K.O., MASON, W.M., WINSBOROUGH, H.H. e POOLE, W. Some methodological issues in cohort analysis of archive data. **American Sociological Review**, v. 38, n. 2, 1973, p. 242-258.

POWERS, D. e XIE, Y. **Statistical methods for categorical data analysis**. San Diego, CA: Academic Press, 2000.

PULLUM, T.W. Parametrizing age, period, and cohort effects, with application to U.S. delinquency rates, 1964-1973. **Sociological Methodology**, v. 9, 1978, p. 116-140.

_____. Separating age, period, and cohort effects in white U.S. fertility, 1920-1970. **Social Science Research**, v. 9, n. 3, 1980, p. 225-244.

RIOS-NETO, E.L.G. e OLIVEIRA, A.M.H.C. Aplicação de um modelo de idade-período-coorte para a atividade econômica no Brasil metropolitano. **Pesquisa e Planejamento Econômico**, v. 29, n. 2, 1999, p. 243-271.

RODGERS, W.L. Estimable functions of age, period, and cohort effects. **American Sociological Review**, v. 47, n. 6, 1982, p. 774-787.

_____. Interpreting the components of time trends. **Sociological Methodology**, v. 20, 1990, p. 421-446.

RYDER, N.B. The cohort as a concept in the study of social change. **American Sociological Review**, v. 30, n. 6, 1965, p. 843-861.

WAJNMAN, S. e RIOS-NETO, E.L.G. Participação feminina na população economicamente ativa no Brasil: alternativas para projeções de níveis e padrões. **Pesquisa e Planejamento Econômico**, v. 24, n. 2, 1994, p. 203-233.

WEISS, Y. e LILLARD, L.A. Experience, vintage and time effects in the growth of earnings: American scientists 1960-70. **Journal of Political Economy**, v. 86, n. 3, 1978, p. 427-448.

WELCH, F. Effects of cohort size on earnings: the baby boom babies' financial bust. **Journal of Political Economy**, v. 87, n. 5, 1979, pt2, p. S65-97.

WILMOTH, J.R. **Identifiable age, period, and cohort effects**: an exploratory approach applied to Italian female mortality. Laxenburg, Áustria: International Institute for Applied Systems Analysis, 1985 (Working Paper; 85-69).

_____. Variation in vital rates by age, period, and cohort. **Sociological Methodology**, v. 20, 1990, p. 295-335.

_____. Age-period-cohort models in demography. In: CASELLI, G., VALLIN, J. e WUNSCH, G. (eds.), **Démographie: analyse et synthèse**; Actes du Séminaire de San Miniato. Paris: Institut National d'Études Démographiques, 1997, p. 187-204.

WINSBOROUGH, H.H. Age, period, cohort and education effects on earnings by race. In: LAND, K. e SPILLERMAN, S. (eds.), **Social indicators models**. Nova York: Russell Sage Foundation, 1975, p. 201-217.

Abstract

Age-period-cohort models applied to the participation of the labor force: searching for a parsimonious version

This paper applies age, period and cohort (APC) models to the analysis of trends of the participation in the Brazilian urban labor force over the last two decades, using data from the PNAD demographic surveys. Rate models based on count models were estimated by Poisson regressions, where the logarithm of the economically active population is the dependent variable, controlled by the logarithm of the population at working ages, as a function of the independent variables of age, period and cohort. Over the period, the percentage of women on the labor market in Brazil increased substantially for all ages. Men, on the other hand, showed a stable pattern of entering and leaving the labor market, maintaining a nearly steady state. This means a prevalence of age effects among men. Among women, age effects are also the most important, but period and cohort effects are also extremely relevant. This differentiation of female cohorts regarding participation in the labor force is not attributable to temporal or age variations. Seeking parsimony of models using direct measures of period and cohort effects, educational level is the best indicator for both men and women. The differentiated educational experience of the cohorts is an important determinant of the variation in participation in the labor force, especially among women.

Key words: Participation in the labor force. Age-period-cohort models.

Recebido para publicação em 2/9/2003.

Aceito para publicação em 1/2/2004.