



CIÊNCIAS SOCIAIS, CENSO E INFORMAÇÃO QUANTITATIVA NO BRASIL

Entrevista com Elza Berquó e Nelson do Valle Silva

ROGÉRIO JERÔNIMO BARBOSA, MURILLO ALVES
MARSCHNER DE BRITO, DIOGO FERRARI,
PATRICK SILVA, IAN ANDRADE PRATES E
LEONARDO SANGALI BARONE

RESUMO

Nesta entrevista, os pesquisadores Elza Berquó e Nelson do Valle Silva relembram suas trajetórias profissionais e a consolidação da pesquisa quantitativa nas ciências sociais no Brasil.

PALAVRAS-CHAVE: *Ciências sociais; informação quantitativa; IBGE; censos.*

ABSTRACT

In this interview, Brazilian demographer Elza Berquó and statistician Nelson do Valle Silva consider their professional careers and the consolidation of quantitative research in the social sciences in Brazil.

KEYWORDS: *Social sciences; quantitative research; Brazilian Institute of Geography and Statistics; census.*

Por sua regularidade, capilaridade, relativa padronização e abrangência temática, os censos demográficos estão entre as principais matérias-primas dos estudos sobre as mudanças de longo prazo na sociedade brasileira. No entanto, analisar dados dos censos implica também conhecer seus processos de produção e, consequentemente, a própria história dos levantamentos de informação social quantitativa no Brasil.

A entrevista a seguir foi realizada no contexto do primeiro workshop do Projeto Censo¹. Idealizado e coordenado pela professora Marta Arretche (CEM-Cebrap e USP), o Projeto Censo, iniciado em meados de 2012 no âmbito do Centro de Estudos da Metrópole (CEM), objetiva examinar as mudanças sociais, políticas e demográficas ocorridas nos últimos cinquenta anos, tomando como eixo a análise das desigualdades. No percurso de nossa pesquisa, diversas inquietações surgiram durante a utilização dos censos. Nós, entrevistadores, éramos parte da equipe de suporte metodológico do projeto e nosso objetivo ao entrevistar dois dos maiores pesquisadores brasileiros que já trabalharam com essas informações era o de (re)descobrir as histórias que envolvem a produção e o uso dos censos demográficos². A trajetória profissional dos professores Elza Berquó e Nelson do Valle Silva diz muito sobre a consolidação do campo das ciências sociais no Brasil e ambos são parte da história dos levantamentos e da análise dos dados censitários no país.

Nelson é um dos principais responsáveis por desenvolver o campo de estudos de estratificação social no Brasil. Tem inúmeros trabalhos sobre desigualdades educacionais, raciais, ocupacionais, estruturação de classes e mercado de trabalho. Atualmente professor visitante do Iesp-Uerj (anteriormente professor titular do Iuperj), foi também pesquisador titular do Laboratório Nacional de Computação Científica do então CNPq (LNCC) nas décadas de 1980 e 1990, além de ter trabalhado no IBGE nos anos 1970, inicialmente como analista de sistemas do Instituto Brasileiro de Informática (IBI-IBGE) e, posteriormente, como chefe do Departamento de Estudos da População.

Elza é professora titular da Faculdade de Saúde Pública da Universidade de São Paulo. Tem graduação em matemática, mestrado em estatística e especialização em bioestatística pela Universidade de Columbia. Foi aposentada compulsoriamente pelo AI-5, e então compôs o corpo de fundadores do Cebrap, onde deu seguimento às suas pesquisas na área de demografia. Liderou a criação do Núcleo de Estudos de População (Nepo) da Unicamp, em 1982, e da Associação Brasileira de Estudos Populacionais (Abep), em 1976. Recebeu, em 1998, a Grã-Cruz da Ordem Nacional do Mérito Científico e é membro titular da Academia Brasileira de Ciências.

Professor Nelson, apesar de suas contribuições para a sociologia, sua formação inicial não é de sociólogo. O senhor graduou-se em economia e é mestre em informática pela PUC-RJ. Somente depois é que foi fazer doutorado em sociologia na Universidade de Michigan, na área de estratificação social, defendendo uma tese sobre desigualdades de renda e de raça. Como foi esse processo de decisão que o levou até a sociologia, tendo saído de uma área tão “hard” como a computação?

[1] Cebrap, 26 e 27 de setembro de 2012.

[2] Os entrevistadores agradecem a contribuição daqueles que, no contexto do seminário, atuaram como “entrevistadores incidentais”, nos auxiliando na condução da conversa: os professores Alvaro Comin (CEM-Cebrap, USP e Kings College London), Carlos Antonio Costa Ribeiro (Iesp-Uerj), José Marcos Pinto da Cunha (Abep e Nepo-Unicamp); Argelina Cheibub Figueiredo (CEM-Cebrap e Iesp-Uerj), Maria Coleta de Oliveira (Nepo-Unicamp) e Marta Arretche (CEM-Cebrap e USP); e dos pesquisadores convidados Flavio Carvalhaes (Iesp-Uerj) e Daniel Waldvogel (CEM-Cebrap). Agradecemos também aos professores Márcia Lima (CEM-Cebrap e USP) e Adrian Gurza Lavalle (CEM-Cebrap e USP).

NELSON DO VALLE SILVA: Foi tudo meio por acaso. O acaso sempre foi muito forte na minha trajetória. Eu queria fazer literatura, mas o meu pai não achou a ideia nada boa. “Você quer ser professor secundário?”, me perguntou. Eu não queria. Um amigo daquela época me chamou então a atenção para um curso de economia que estava se iniciando na PUC, lecionado dentro da Escola de Sociologia e Política. Economia ainda era uma coisa nova naquela época; acho que as primeiras turmas datam dos anos 1960. Eu achei ótima ideia e fiz o vestibular, mas nunca cheguei a ser economista, nunca trabalhei na área. Comecei trabalhando na área de informática como programador do centro de computação da própria PUC, no Rio Datacentro, ainda na graduação.

Ocorreu-me fazer o tal mestrado em informática quando da reforma universitária: para ser professor em uma universidade — e eu era do Departamento de Informática — você tinha que ter mestrado. Mas eu não tinha muito interesse naquilo em si. Minha função dentro do departamento era fazer um meio de campo com as ciências humanas e sociais, ministrando aulas de computação para os cursos de psicologia, pedagogia, sociologia e economia. Além disso, eu dava apoio externo para o Datacentro, para usuários que quisessem analisar dados. A análise de dados foi minha porta de entrada para as ciências sociais. E meu primeiro contato não foi com a sociologia, mas com a geografia. A certa altura apareceu uma turma do IBGE que precisava de especialistas em geografia quantitativa. Tinham uma série de demandas: problemas de regionalização, hierarquia de cidades etc., mas não tinham software para trabalhar. Fui então contratado pelo IBGE e alocado na diretoria de informática, mas trabalhava com o pessoal da geografia — geografia urbana, mais particularmente.

Foi quando comecei a me interessar pela sociologia e, particularmente, pela demografia. Nesse período, existia um acordo do Brasil com o governo americano, o acordo MEC-USAID³. Muita gente foi fazer mestrado e doutorado nos Estados Unidos. Eu achei que era uma boa ideia e fiquei com vontade de entrar nessa história. Acontece que, nos Estados Unidos, a demografia normalmente é lecionada dentro da economia ou da sociologia. Eu fui aprovado para ir para Stanford, inicialmente. Mas no Departamento de Sociologia só havia um professor de demografia, o Dudley Kirk⁴. Pensei: vou para lá trabalhar com um só sujeito que eu nem sei quem é? E se a relação de aluno e professor não desse certo? Foi quando amigos que estavam na Universidade de Michigan me convidaram para ir para lá. O sistema norte-americano tem as áreas de *major* e *minor*: o meu *major* foi em demografia, em estudos de população e ecologia humana, e a minha área de *minor* foi justamente estratificação social. Acabei fazendo a minha tese de doutorado na minha área de *minor*, que era a área de estratificação. Foi como cheguei à sociologia.

[3] Acordos estabelecidos entre o Ministério da Educação (MEC) e a United States Agency for International Development (USAID) para promover a reforma do ensino brasileiro.

[4] Dudley Kirk (1913-2000), sociólogo pela Universidade de Harvard, foi professor das universidades de Princeton e Stanford.

Michigan já era uma referência em análise de dados quantitativos em sociologia?

Sim, mais ainda do que hoje. Hoje a coisa não está grande coisa. Em Michigan conheci o estatístico Leslie Kish, do Departamento de Sociologia, que ministrava o curso de amostragem. Lá estava eu, soterrado sob um metro de neve, fazendo o meu doutorado sob a orientação do Bill Mason⁵, que também orientou meu colega Robert Mare⁶, que era um ano mais velho que eu. Michigan ainda era “o” centro de análise de dados. Havia o Institute for Social Research, que era muito importante e que fazia (e ainda faz) um estudo de painel sobre *income dynamics* [dinâmicas de renda] (isso é realizado desde 1948). O treinamento quantitativo é bastante forte lá.

Professora Elza, assim como o professor Nelson em relação à sociologia, a senhora incorporou a demografia num ponto mais tardio de sua trajetória, depois de seu doutoramento. Na entrevista concedida ao livro de quarenta anos do Cebrap, a senhora mencionou que incorporou sozinho a demografia, aos poucos. Como foi esse processo?

ELZA BERQUÓ: Formada em matemática pela PUC de Campinas, tornei-me assistente do prof. Pedro Egydio de Oliveira Carvalho, chefe do Departamento de Bioestatística da Faculdade de Saúde Pública da USP. Pedro Egydio era médico com grande conhecimento de matemática. Ofereceu-me a chance de fazer um curso de especialização em estatística matemática na USP, proferido pelo professor americano William Meadow, baseado no livro recém-publicado de Harold Cramer — *Mathematical methods of statistics*.

Depois disso, com bolsa da Oficina Sanitária Panamericana, realizei na Universidade de Columbia um programa que constava de uma pós-graduação em bioestatística na Escola de Saúde Pública e de diversos cursos de teoria das probabilidades e inferência estatística no Departamento de Estatística.

De volta ao Brasil, pude colaborar na aprimoração dos cursos de pós-graduação em saúde pública, contando com o intercâmbio de professores da Universidade de Columbia. Em seguida, preparei e realizei o concurso de livre-docente.

O falecimento repentino de Pedro Egydio, aos 58 anos, colocou para mim a obrigação de concorrer ao concurso de professor titular. Volto a Columbia para discutir com meus colegas o tema de minha tese. Sob a orientação da saudosa Agnes Berger, decidimos pela aplicação da análise sequencial à área da saúde pública⁷.

Na chefia do departamento, senti a necessidade de incorporar ao programa de ensino e pesquisa a disciplina de demografia. Mas, além do IBGE, não se dispunha de recursos humanos nesse campo. Dada a interdisciplinaridade da demografia, era preciso preparar um grupo

[5] William Maxwell Mason é professor emérito do Departamento de Sociologia da Universidade da Califórnia, Los Angeles. Trabalha com demografia social e metodologia quantitativa.

[6] Robert Denis Mare é sociólogo e professor do Departamento de Sociologia da Universidade da Califórnia, Los Angeles.

[7] Berquó, Elza. “Sobre a análise sequencial para testes de hipóteses relativas a proporções. Aplicação a problemas de Medicina e de Saúde Pública” (1960).

de pesquisadores de formações diversas para preencher essa lacuna do conhecimento entre nós. Foi nesse contexto que solicitamos apoio da Oficina Sanitária Panamericana para uma consultoria no sentido de estabelecer e organizar um núcleo de demografia. Contamos para essa tarefa com a consultoria da eminente demógrafa Irene Tauber, da Universidade de Princeton. Com ela, elaboramos um plano para criar um centro de estudos de dinâmica demográfica que resultou no Centro de Estudos de Dinâmica Populacional (Cedip), em 1967.

Conseguimos quatro bolsas e convidamos profissionais para se prepararem no exterior, com o compromisso de, ao voltar, estarem preparados para desenvolver pesquisa e ensino na área de demografia. Nesse grupo estavam Paul Singer, da área de economia; Jair Lício Ferreira Santos, da área de estatística; Neide Patarra, da área de sociologia; e João Yunes, médico. Os quatro receberam bolsas e cada um foi para uma universidade americana. Foram para Princeton, Chicago e Michigan. Candido Procópio Ferreira de Camargo, primeiro presidente do Cebrap, de notório saber, fez uma série de viagens pelo mundo mais desenvolvido, visitando centros de demografia. Eu mesma sou autodidata em demografia.

Tudo ia muito bem: esse pessoal havia voltado do exterior e trabalhava no Cedip. Foi quando o AI-5 nos tirou da Universidade de São Paulo e nos aposentou compulsoriamente. Fomos aposentados, eu e Paul Singer, no final de 1968, e em 1969, sob a liderança de Fernando Henrique Cardoso, o Cebrap foi criado. Viemos para cá e trouxemos as pesquisas que estavam em andamento no Cedip para serem desenvolvidas com os pesquisadores daqui. Nós continuamos nossos projetos e os ampliamos. Estávamos desenvolvendo, na época, uma pesquisa nacional sobre reprodução humana. Dialogávamos com aquela visão “controlista” que pensava que o país só se desenvolveria se a população diminuísse.

Quando veio a abertura política, todo mundo pôde voltar à universidade. Eu tinha dois lugares para ir, tanto para a Matemática como voltar para a Faculdade de Saúde Pública. Até hoje me emociono com essa história. O coração me levava para a Faculdade de Saúde Pública, mas a razão me levava para a Matemática. Decidi que voltaria para a Faculdade de Saúde Pública. Comuniquei primeiro ao Instituto de Matemática (IME) da USP. Procurei em seguida o diretor da Faculdade de Saúde Pública, o Forattini⁸, e dei a notícia de que queria voltar. No entanto, metade da congregação votou contra a minha volta. Eu ainda era uma “comunista muito perigosa” e a faculdade não me queria lá. Forattini deu o voto de minerva para que eu voltasse. Diante disso, preferi recusar. Tampouco iria ficar com a minha segunda opção, o IME. Fiquei no Cebrap. Em 1982, o reitor da Unicamp, o dr. José Aristodemo Pinnotti, me convidou para criar o Núcleo de Estudos da População (Nepo).

[8] Oswaldo Forattini (1924-2007), epidemiologista, foi pioneiro na produção de metodologias de coleta e identificação de vetores e agentes etiológicos de doenças. Foi fundador da *Revista de Saúde Pública*.

Os senhores usaram muito os censos demográficos do IBGE, que são uma fonte importantíssima de dados sociais, mas cujo volume dificulta muito o manuseio e a utilização. Ainda hoje é difícil lidar com essa quantia de informações. Como era possível usar os censos e desenvolver tantas análises substantivas e importantes com recursos tecnológicos limitados, na década de 1970? Como era trabalhar com o censo naquele período? Quais eram as dificuldades e quais soluções vocês desenvolveram?

NELSON DO VALLE SILVA: Para mim foi fácil, porque eu trabalhava no instituto de informática do IBGE. Eu tinha um *mainframe* para “brincar”. Mas mesmo assim havia problemas sérios. Para trabalhar com o Censo de 1970, que estava em final de processamento quando entrei para o IBGE, havia problemas de software. Nós formamos um grupo para produzir um software de processamento de *big data*, que batizei de Sistema de Análise de Dados Estatísticos, ou SADE, em homenagem ao marquês [risos]. Os softwares existentes até então não se apoiavam em algoritmos adequados para cálculos estatísticos em grandes bancos. Por exemplo, não havia registro que aguentasse o acúmulo do produto de vinte milhões de informações; isso estourava a capacidade dos computadores e dos softwares. Os erros nos resultados chegavam a ser ridículos. Se você fizesse uma análise fatorial ou algo do tipo, o resultado podia ser um *nonsense* completo. É importante tomar cuidado e checar os algoritmos. Não sei como as coisas estão atualmente, estou falando de quarenta anos atrás. Mas é importante verificar se os algoritmos resistem a essa quantidade de dados. Você tem que ter certos truques para fazer operações simples, como para calcular uma correlação, por exemplo.

Em 1979 ou 1980, o senhor começou a fazer parte do Laboratório Nacional de Computação Científica (LNCC) do CNPq. Creio que a partir desse período a infraestrutura e as formas de acesso aos recursos tecnológicos existentes para análise de dados mudaram. Como foi a experiência no LNCC?

Era o mesmo grupo que estava no IBGE elaborando o SADE. Houve uma briga danada entre o grupo do IBI e a direção do IBGE por causa de desperdício de recursos públicos. Esse pessoal do IBI acabou pedindo demissão e foi trabalhar no Centro Brasileiro de Pesquisa Física (CBPF) — mais especificamente no Laboratório de Cálculo. Nessa época, eu estava nos Estados Unidos, fazendo doutorado. Aquele grupo tinha como líder o Marco Antonio Raupp⁹, que também liderou e criou dentro do CBPF o Laboratório de Cálculo. Esse grupo foi se fortalecendo dentro do CBPF. Surgiu então a ideia de criar um centro de informática, nas áreas de ciência da computação e matemática aplicada, com foco na segunda. O LNCC foi criado no segundo semestre de 1980, sob a direção de Antonio César Olinto. Nesse momento,

[9] Marco Antonio Raupp, atual ministro de Ciência, Tecnologia e Inovação, foi diretor geral do Inpe e do LNCC e professor no IME-USP e da UnB.

eu estava no IBGE, mas já me aborrecendo de novo com as brigas todas com a nova diretoria, que havia mudado em 1979. O Marco Antonio me convidou para me juntar ao Laboratório de Computação, recém-criado no âmbito do CNPq e logo transformado em instituto de pesquisa. Lá eu fiz minhas pesquisas de matemática aplicada, na área de ciências sociais. Depois de um tempo, outras pessoas foram se juntando. Nós formamos um núcleo dentro do laboratório, composto de pesquisadores da área de ciências sociais, incluindo o Sérgio Costa Ribeiro (tio de Carlos Antônio Costa Ribeiro), que vinha da física, mas estava interessado em problemas educacionais.

Nessa época, o que levava cientistas sociais a direcionar suas demandas e convergir seus interesses com os de um laboratório de computação científica?

A ideia compreendida pela expressão “computação científica” naquela época hoje poderia ser chamada de matemática aplicada. O propósito era desenvolver aplicações da matemática dentro das diversas áreas. Havia um pessoal da engenharia mecânica, outro pessoal da engenharia elétrica, por exemplo. O LNCC era um grupo totalmente heterogêneo. Nem sei como a gente conseguia se entender. Os temas eram os mais diversos. A informática era instrumental. A ideia era desenvolver aplicações da matemática dentro de outras áreas para resolver os seus problemas particulares. Mas não havia mais ninguém das ciências sociais propriamente. Eu era o único, os outros eram matemáticos, engenheiros etc. Por isso eu sempre mantive um contato com o Iuperj. Esse contato me permitia ter problemas para resolver e foi onde começou, em 1981, acho, minha cooperação mais estreita com o Carlos Hasenbalg e com o Amaury de Souza.

A senhora participou de diversas edições do censo como membro do Conselho Consultivo. Poderia falar um pouco para nós sobre esse trabalho de bastidores?

ELZA BERQUÓ: No começo, minhas pesquisas não eram baseadas em censos, eram pesquisas de campo. Por exemplo, a Pesquisa Nacional de Reprodução Humana (realizada entre 1973 e 1978) foi uma pesquisa com design de amostragem, com representatividade nacional. A análise estatística do banco de dados precisou ser feita em Chicago, uma vez que no Brasil não se dispunha de computador de grande porte que pudesse realizar o trabalho.

No início do seminário se falou muito sobre as dificuldades de comunicação com o IBGE. Só a título de curiosidade, de 1936, quando foi criado o instituto¹⁰, até hoje, foram ao todo 31 presidentes. De 1936 até 1967 foram quinze; de 1967 a 2011 foram dezesseis. Alguns presidentes ficaram menos de um ano no cargo. Houve muita rotatividade, por tratar-se de indicação política.

[10] O IBGE foi instalado em 1936, com o nome de Instituto Nacional de Estatística. Em 1938, mudou seu nome para Instituto Brasileiro de Geografia e Estatística.

Nos últimos quarenta anos o IBGE contou com presidentes que deixaram marcas positivas por seu desempenho e descortino: Isaac Kerstenetzky (1970-1979), Jessé de Souza Montello (1979-1985), Simon Schwartzman (1994-1989), Sergio Besserman Vianna (1999-2003) e Eduardo Pereira Nunes (2003-2011). A partir de 2011, conta com a competência de Wasmália Bivar.

Eu estava presente quando tivemos que tomar a decisão histórica de não fazer o censo em 1990, no governo Collor. O presidente do IBGE era Eduardo Augusto Guimarães (1990-1992). Nós do comitê do censo percebemos que a data programada para o censo estava chegando perto e ainda não tínhamos os recursos financeiros necessários no IBGE para começar a deslanchar a operação censitária.

Tentamos até o último momento. Cada um nós ofereceu os contatos que tinha para que, por outras vias, ele fosse recebido no Ministério da Fazenda. Mas ele afirmava não ser necessário: “Eu vou conseguir”, dizia. Conclusão: não foi recebido. Chegou a hora do Censo de 1990 e não tínhamos os recursos. Uma das coisas mais tristes nesse comitê foi o momento em que votamos, por unanimidade, que o censo não seria realizado em 1990, quebrando a série histórica. Então ele passou para 1991. Isso foi realmente muito sofrido.

A preparação de um censo enfrenta necessariamente questões de custo e benefício: incorporar quesitos que respondam às mudanças ocorridas na sociedade nos últimos dez anos e não perder totalmente a comparabilidade com o censo anterior.

Há, aliás, uma história interessante sobre o Censo de 1960. Ele foi realizado antes do golpe militar de 1964. Acontece que os dados sumiram. Durante aquela confusão de 1964, uma parte do governo achava que o IBGE não tinha capacidade para processar aquilo tudo dentro do tempo planejado. Então parece que parte desses dados viajou do Brasil para os Estados Unidos, para que pudessem ser processados lá. Esta é a explicação que chegou até nós, que justifica o fato de que nós só pudemos ter acesso aos dados do Censo de 1960 em 1978. Foram dezoito anos sem informação censitária. Essa é também uma das explicações para o fato de até hoje não termos acesso aos microdados completos do censo: faltam alguns estados do Norte e Nordeste. Teriam se perdido durante o transporte, um processo de desmagnetização da informação durante o voo. Há um mistério cercando isso tudo. Essa era a história contada para nós.

Essa perda foi um marco extremamente negativo na história dos nossos censos. Mas não foi a única. Nós não tivemos os censos de 1900 e 1910. O de 1900 foi feito, mas com resultados questionáveis. Nem para o Rio de Janeiro, que era a capital do Brasil, foi possível aproveitar os resultados. Quanto ao de 1910, Nilo Peçanha não assinou a Lei do Censo por questões políticas. O Censo de 1930 também não

aconteceu, por causa da revolução. Questões políticas afetaram muito a história dos censos e nos impediram de ter uma série completa. A ditadura tirou a informação de cor/raça do Censo de 1970, afirmando que fazer essa pergunta era, na verdade, racismo. Tiraram uma informação fundamental. Nós, como membros do Conselho Consultivo, sempre fizemos frente a essas dificuldades de decidir mudanças e incorporações. Por exemplo, na questão sobre situação conjugal, a adição da alternativa “união consensual” também não foi simples, não foi trivial decidir pela sua inclusão. Mas felizmente foi incluída. As outras categorias foram mantidas, o que permitiu a continuidade da análise.

O Censo de 2010 incluiu a informação sobre casais do mesmo sexo. Eu estava lutando para isso desde o Censo de 1991, sem sucesso. Não foi trivial. Também hoje já não se usa mais o conceito de “chefia do domicílio”. Trocou-se o “chefe” pela “pessoa de referência”. Os movimentos feministas não queriam que se falasse em chafia. A consulta aos usuários foi sempre muito importante. A contribuição dos movimentos negros no Brasil foi fundamental para que o quesito raça continuasse lá. Avançou-se na questão da população indígena também. Esse Censo de 2010 é mais rico com relação a essa informação.

Quando o Censo de 1970 ficou disponível para consulta?

O Censo de 1970 foi uma experiência extremamente positiva. Os resultados saíram rápido, demorou de três a quatro anos. Os demógrafos da época já estavam sem fôlego de tanto usar dados antigos: tínhamos parado em 1950, porque o Censo de 1960 não havia saído. Nessa época, a Fundação Ford funcionava na base de balcão: você mandava o projeto para pedir financiamento e eles faziam a análise. A Ford agiu muito bem. Reuniu todos os demógrafos e não demógrafos que queriam trabalhar com o Censo de 1970 e sugeriu que trabalhássemos em conjunto, cada um fazendo uma parte. Lembro-me que a experiência nossa com o Cedeplar¹¹ se deu nesse contexto. Nós definimos um projeto conjunto para a análise do censo na área de imigração, fecundidade e assim por diante. O Censo de 1970 foi um bálsamo naquele vazio de informações. E foi tudo muito bem conduzido. Houve inclusive um financiamento substantivo da Fundação Ford para que os pesquisadores pudessem prosseguir suas análises.

Gostaríamos de ouvir um pouco mais sobre essa “caixa-preta” do Censo de 1960. Sabemos que o IBGE não comercializa os dados de uma certa amostra de 1,25% (que possui todos os estados). Algumas instituições, como a Universidade de Minnesota, têm versões da amostra oficial de 10% (à qual a professora Elza se referiu), mas sempre faltam alguns estados (UFs) e outras informações importantes de documentação. A base de dados que temos utilizado no Projeto Censo nos foi

[11] Cedeplar – Centro de Desenvolvimento e Planejamento Regional da Universidade Federal de Minas Gerais.

informalmente disponibilizada pelo professor Carlos Antonio Costa Ribeiro. O professor Nelson do Valle, aliás, utiliza o Censo de 1960 em sua tese de doutorado. Como o senhor obteve esses dados?

NELSON DO VALLE SILVA: A minha versão da história do Censo de 1960 é um pouquinho diferente daquela contada pela Elza. O IBGE tinha comprado o maior computador do mundo, naquela época. Uma coisa fantástica, de primeiríssima. Era um computador para cálculo de trajetória de satélites. Possuía uma velocidade fantástica de processamento. Só que o output era uma máquina de escrever. Pelo que entendi, o presidente do IBGE na época foi processado e a justiça embargou o computador. A digitalização do Censo de 1960, que estava ocorrendo aos pouquinhos, parou. Foram digitalizados apenas alguns estados. Então colocaram tudo lá, em um armazém do IBGE, onde eu o encontrei dez anos depois, coberto de poeira, de terra. Tinha uma camada de terra mesmo! Fisicamente, o censo estava armazenado naquelas pastas. Então o Albert Fishlow veio para o Brasil fazer um estudo sobre distribuição de renda e, por alguma razão, não sei exatamente por quê, ele foi o primeiro a usar. Para isso, resolveram tirar uma amostra desse censo, com uma proporção de 1,25%. Isso foi feito fisicamente mesmo: pegaram os questionários, fizeram o sorteio aleatório. Agora atenção: a minha tese não foi a única que usou essa versão não. O artigo do Fishlow e o livro do Langoni também usaram. Todo o debate sobre distribuição de renda no Brasil nas décadas de 1960 e 1970 se baseou nessa amostra.

Quando eu estava prestes a embarcar para meu doutorado em Michigan, pensei: a maneira mais fácil de produzir um *paper* é ter dados novos, dados que ninguém tem. Então eu fui lá ver o que tinha. Havia vários arquivos no IBGE. E descobri que um companheiro de IBI, estatístico da velha guarda, tinha justamente participado do processo de fazer a amostra. Ele tinha esse arquivo que levei comigo para Michigan. Tese você também faz com dados novos. Peguei os dados, fiz minha tese e deixei uma cópia do Censo de 1960 na Universidade de Michigan. Suponho que o Censo de 1960 que vocês têm seja o mesmo. Embora existissem cópias em outros lugares, como na Universidade da Flórida.

É uma história fantástica, porque depois o Censo de 1960 sumiu de novo. No IBGE era assim: as coisas sumiam. Se algo está faltando, é porque provavelmente o sujeito que armazenou acabou guardando de modo errado, e errado ficou.

A Universidade do Texas também tem uma cópia do Censo de 1960. Talvez ela tenha sido uma das primeiras universidades a ter esse banco de dados — antes mesmo do que muita gente aqui do Brasil. Eles vieram aqui e levaram as fitas magnéticas do censo “debaixo do braço”. Investiram um dinheirão na época para conseguir um equipamento que pudesse ler as fitas.

[12] O IBGE produziu, para o Censo de 1980, amostras de uso público de 0,7% e de 3%. Trata-se de bancos de dados produzidos a partir do sorteio aleatório estratificado de domicílios da amostra de 25%. A proposta era facilitar o acesso e o uso por parte de pesquisadores que não possuíam acesso à supercomputadores. Mas até 1991 as amostras oficiais dos censos demográficos ainda eram de 25%. A partir de 2000, a fração amostral passou a ser 10% — e nesse mesmo ano o IBGE produziu também uma amostra de uso público de 1%.

[13] Giorgio Mortara (1885-1967) foi um economista, demógrafo e estatístico italiano. Em 1939 foi forçado a deixar a Itália e mudou-se para o Brasil, onde contribuiu intensamente para a organização dos censos demográficos de 1940 e 1950 e para o desenvolvimento da demografia no país. No IBGE fundou o Gabinete Técnico, que posteriormente transformou-se no Laboratório de Estatística, no qual se formaram diversos técnicos e estatísticos que contribuíram para o desenvolvimento do campo. Em 2007 o IBGE organizou um documento para discussão denominado “Giorgio Mortara: ampliando os horizontes da demografia brasileira” em homenagem aos quarenta anos de seu falecimento.

Em 1979 ou 1980, o restante dos questionários foi processado. Os questionários ainda existiam no meio daquela poeira toda que cobria o computador encaixotado, as fitas e as pastas. O presidente do IBGE naquela época, o professor Isaac Kerstenetzky, mandou processar o que estava faltando. Talvez tenha sido isso.

No caso do Censo de 1980, nos Estados Unidos, por exemplo, não se usa a amostra de 25%. Os americanos achavam um absurdo o tamanho de uma amostra de 25%. Em muitos lugares de lá, quando se vai analisar os dados de 1980, usa-se uma amostra de 2% ou 3%. Hoje, depois da amostra do Censo de 2000 (de 10%)¹², vemos que de fato não são necessárias amostras grandes.

Quando estávamos na Comissão Consultiva do Censo de 1991, eu insistia para que se fizesse uma amostra de 1%. As pessoas simplesmente recusaram. O argumento era que os “novos” computadores seriam capazes de processar os dados da amostra de 25%. Todo mundo teria acesso e o processamento seria um problema do usuário. Mas vocês estão vendo aí que seria bem melhor se estivessem trabalhando com uma amostra de 1% — 99% dos problemas se resolve com uma amostra de 1%.

ELZA BERQUÓ: Mudando de assunto, acho que seria injusto se a gente falasse de censo no Brasil sem fazer uma menção a Giorgio Mortara¹³. Ele veio para o Brasil fugindo do fascismo em 1938 e foi o responsável pelo belo Censo de 1940. Tinha experiência de professor universitário, era economista, atuário, geógrafo etc. Com uma competência muito grande, ele se envolveu de corpo e alma na operação censitária de 1940. Como não houve censo em 1930, estávamos há vinte anos sem informações. Ele fez retroprojeções para tentar entender o que tinha acontecido por aqui. O registro civil ainda estava engatinhando, então o censo era um instrumento importante. Felizmente, hoje o registro civil é de outra qualidade. Giorgio Mortara é o responsável pela introdução de inúmeras questões importantes, mas não é o caso de detalhá-las aqui. Posteriormente, ele conduziu também o de Censo de 1950. Mas quando as coisas melhoraram na Europa, ele regressou à Itália.

E com respeito aos temas cobertos pelos censos? Existe ainda alguma ausência notável? Que tipo de problemas de pesquisa vocês acham que poderiam ser tratados no censo e ainda não foram?

NELSON DO VALLE SILVA: Eu acho que o censo tem coisa demais. Eu participei de várias discussões de montagem de questionários. Sempre há uma demanda muito grande para se introduzir questões. Eu acho que o censo poderia ter um corpo menor, de modo que a maior

parte das questões poderia ser tranquilamente e até mais bem tratada por pesquisas como a PNAD¹⁴. Mas tem muita gente que não acredita em amostragem. Não estou brincando. No de 1980, por exemplo, há uma questão sobre alfabetização no questionário do universo porque o presidente do Mobral, naquela época, não acreditava em amostragem. Acho que o censo tem, sobretudo, a função de gerar parâmetros para que a gente possa fazer outras pesquisas. Essa é sua função primordial. É possível fazer pesquisas com o censo propriamente dito; mas principalmente em temas muito “macro”, como demografia.

[14] Pesquisa Nacional por Amostra de Domicílios, realizada anualmente pelo IBGE.

ELZA BERQUÓ: Eu concordo. A tendência é termos PNADs quase contínuas. Na minha área há o tema da esterilidade, que pode ser aprofundado em pesquisas complementares. Ainda haveria falta de informação, claro. Mas eu concordo que não dá para colocar tudo no censo. O ideal é melhorar a qualidade da informação mantendo a comparabilidade. Esse é o grande drama. Quando se passa de um censo para o outro, como fica? Acho que o censo vai ficar cada vez mais no questionário básico, mais enxuto. O que está lá, já está lá. Possíveis acréscimos deveriam ser feitos no máximo no questionário da amostra.

Como vocês veem a evolução dos procedimentos de coleta de dados no âmbito do censo?

ELZA BERQUÓ: Acho realmente que evoluíram bastante tanto a qualidade da informação quanto a forma como a informação é coletada. Não há nenhuma razão para pensar que as coisas possam piorar. Porque o IBGE continua sendo dirigido por pessoas de competência. A Wasmália, atual presidente, é uma funcionária antiga do IBGE, que conhece a máquina e tem uma relação com a comunidade de usuários. Acredito que é como o presidente anterior. Os últimos presidentes do IBGE têm ficado mais tempo, têm podido conhecer melhor a instituição e as comunidades de pesquisa. Respeitam o que os movimentos solicitam e discutem o que é ou não é possível. Eu não vejo no horizonte nenhuma razão para imaginar que as coisas possam piorar.

NELSON DO VALLE SILVA: Toda a tecnologia que está sendo usada hoje só melhora as coisas e diminui o ruído no processamento das informações. Mas o censo é uma operação de guerra e, como tal, depende muito do general. A qualidade do final vai depender muito do general que estiver no comando.

O uso do PDA¹⁵ melhorou muito a consistência dos dados. Não há base de dados completamente consistente, mas o PDA tem um mecanismo que detecta inconsistências internas já durante o procedimento de preenchimento do questionário. No entanto, existem outros problemas que

[15] PDA é sigla para Personal Digital Assistant, um computador de mão que, no Censo de 2010, substituiu o uso dos questionários de papel.

ocorrem na coleta de dados. Talvez o maior deles seja o de “entrar nas casas”. Quem faz pesquisa domiciliar sabe que nos últimos dez anos está cada vez mais difícil não apenas conseguir o consentimento, mas também conseguir que as pessoas respondam adequadamente.

ELZA BERQUÓ: Nos condomínios de luxo, por exemplo.

Nós programamos muito mais perguntas do que foi possível fazer hoje, com mais questões específicas sobre as obras e trajetórias de Nelson do Valle Silva e Elza Berquó. Mas infelizmente temos de deixar para outra ocasião. Gostaríamos de agradecer aos professores Nelson e Elza. Foi uma honra tê-los aqui conosco.

ELZA BERQUÓ: Eu gostaria de agradecer muito essa oportunidade. E queria citar realmente os jovens que estão aí, nesse trabalho difícil. Acho que vocês vão deixar um produto da maior importância para usuários futuros. Vocês estão fazendo um trabalho inovador, da maior importância. Se o Projeto Censo fosse só isso, conhecer mais sobre as bases de dados, já teria valido a pena. Afinal de contas, entrar nesse âmbito é difícil. O trabalho de vocês será um produto que o CEM e o Cebrap poderão divulgar, e do qual outros pesquisadores, de outras instituições, inclusive o IBGE, vão se beneficiar.

Recebido para publicação
em 4 de fevereiro de 2013.

NOVOS ESTUDOS

CEBRAP

95, março 2013

pp. 143-155

Os entrevistadores são pesquisadores do Centro de Estudos da Metrópole (CEM). ROGÉRIO e MURILLO são doutorandos em sociologia pela USP; DIOGO e PATRICK, mestrando em ciência política pela USP; IAN, mestrando em sociologia pela USP; e LEONARDO, doutorando em administração pública pela FGV.