
O PERFIL DA PESSOA INFLUENCIA EM SUA AVALIAÇÃO DO GOVERNO E EM SUAS EXPECTATIVAS? UMA APLICAÇÃO DO MODELO DE RELAÇÕES CAUSAIS

ARTIGO

Maria Cecília Prates Rodrigues¹

Doutoranda em Administração pela Fundação Getúlio Vargas/Escola Brasileira de Administração Pública e Empresas - FGV/EBAPE. Mestre em Economia pela Universidade Federal de Minas Gerais/Centro de Desenvolvimento e Planejamento Regional - UFMG/CEDEPLAR. Economista pela UFMG. Pesquisadora e consultora em Avaliação de Projetos Sociais.
E-mail: mceciliaprates@uol.com.br

RESUMO

Tomando por base pesquisa por amostragem disponibilizada pelo Centro de Estudos de Opinião Pública da UNICAMP – CESOP (DAT/00942), realizada em dezembro/1998, buscou-se aplicar o modelo de relações causais para testar as relações de causalidade entre os conceitos (1) perfil da pessoa, (2) avaliação do 1º governo Fernando Henrique Cardoso, e (3) expectativas futuras quanto ao 2º mandato de FHC.

Os resultados evidenciaram que o “perfil da pessoa”, definido segundo critérios objetivos, não influenciou em sua “avaliação do 1º governo FHC” e nem nas “expectativas futuras”. Já o constructo “perfil da pessoa”, definido de acordo com critérios subjetivos, não se mostrou suficientemente robusto para o teste da causalidade. Mas, foi a “confiança no presidente FHC” que conseguiu, de fato, influenciar nas “expectativas futuras” – ou seja, foi o grau de confiança que as pessoas têm no presidente FHC que determinou as suas expectativas quanto às condições socioeconômicas do país no mandato seguinte.

ABSTRACT

Using data from a survey conducted in December 1998 and available at CESOP / UNICAMP (DAT / 00942), we applied a causal relation model in order to test causality relations among the following concepts: (1) person profile; (2) evaluation of Fernando Henrique Cardoso first term; and (3) future expectation about FHC second term.

The results pointed out that a “person profile”, defined according objective criteria, did not influence either his / her “evaluation of FHC first term” or his / her future expectation. Even if the construct “person profile” is defined through subjective criteria, it did not support this causality test. What we found was that it was “confidence on President FHC” that really influenced “future expectation” – in other words, the degree of confidence people had on President FHC was the decisive factor to mould their expectation related to social and economic conditions of this country for the second FHC term.

¹ A autora agradece ao Prof. Moisés Balassiano, da Fundação Getúlio Vargas/Escola Brasileira de Administração Pública e Empresas, pelos valiosos comentários.

1. INTRODUÇÃO

Com base em técnicas de interdependência das variáveis observacionais (matriz Σ de variâncias e co-variâncias), os modelos de relações causais buscam identificar a existência de causalidade entre estas variáveis e/ou entre os conceitos constituídos a partir delas.

Este artigo tem por objetivo justamente investigar a aplicação do modelo de relações causais em uma situação concreta de interpretação da relação de causa e efeito entre os dados. Como substrato para este estudo, foi selecionada uma pesquisa por amostragem (DAT/00942), disponibilizada pelo banco de dados do Centro de Estudos de Opinião Pública da UNICAMP – CESOP, intitulada *Avaliação do Governo Fernando Henrique Cardoso*. A partir das questões colocadas nessa pesquisa, buscou-se explorar as relações de causalidade entre os conceitos de (1) perfil da pessoa, (2) avaliação do governo FHC 1995-98 e (3) expectativas futuras quanto à situação socioeconômica do país, uma vez que o presidente Cardoso fora reeleito.

É evidente a relevância desse modelo no campo da Administração. A argumentação aqui é a de que a possibilidade de análise e solução de questões, como a que será abordada neste artigo, cumpre importante papel no diagnóstico, ou no entendimento lógico, das muitas variáveis envolvidas com a realidade cada vez mais complexa das organizações, sobretudo no campo do *Marketing* e da Estratégia.

2. O MODELO DE RELAÇÕES CAUSAIS: ASPECTOS TEÓRICOS

Sem dúvida, existe um paralelismo grande entre as análises de regressão e os modelos de relações causais, também conhecidos como modelos de equações estruturais. O que se pretende, nestas análises, é a identificação das relações de dependência entre variáveis, através de métodos de minimização dos resíduos. A diferença está, porém, na forma de aplicar o método.

Assim, como observa BOLLEN (1989: 1), nas análises de regressão – simples ou múltipla – o foco do analista é a minimização da soma dos quadrados dos resíduos entre os valores previstos pelo modelo e aqueles observados para a variável dependente em

cada caso. Ou seja, a modelagem caso a caso. Já nos modelos de relações causais, o que se busca é a minimização das diferenças, ou resíduos, entre a matriz S das variâncias-co-variâncias obtidas da amostra considerada e a matriz Σ estimada pelos parâmetros do modelo (θ), construído justamente para identificar as relações de dependência entre as variáveis. Isto significa que, neste segundo caso, a modelagem se dá para os casos observacionais, considerados em conjunto.

Se o desenho do modelo de relações causais se mostra adequado, os resíduos entre a matriz estimada $\Sigma(\theta)$ e a matriz S são significativamente nulos; portanto, a matriz estimada $\Sigma(\theta)$ tende a reproduzir a matriz S . A matriz estimada $\Sigma(\theta)$ se constitui justamente na hipótese central do modelo, retratada na igualdade abaixo.

$$S = \Sigma(\hat{\theta}) \quad (\text{eq.1})$$

Por implicar cálculos matriciais complexos, o estudo dos modelos causais só ganhou fôlego com o avanço e difusão dos recursos de informática. BOLLEN (1989: VII e 4) mostra que estes modelos foram inicialmente aplicados nas áreas da sociologia, psicologia e economia. O *software* LISREL (Linear Structural Relations), de JORESOG e SORBOM (1996), vem sendo um dos mais utilizados para esse tipo de análise, e foi o adotado no âmbito deste artigo².

São três os componentes definidores da causalidade, a saber: isolamento, associação e direção da influência, tomados nesta ordem. Em primeiro lugar, causalidade pressupõe que se possa isolar a variável y em questão da influência de todas as demais, exceto de x – assim, se uma mudança em y acompanha uma mudança em x , pode-se dizer que x é causa de y . Em segundo lugar, deve existir associação entre estas variáveis, muito embora seja mais do que sabido que a mera existência de correlação não é condição suficiente para a existência de causalidade. Por último, é preciso estabelecer que a associação se deve ao fato de que x influencia y , e não o contrário; detecta-se, desta forma, a direção da ação e, conseqüentemente, o sentido de temporalidade, ou seja, x preceder a y no tempo.

² Ao longo deste artigo, os termos LISREL e modelo de relações causais serão usados indistintamente.

Os submodelos e o modelo completo - Dito de maneira resumida, o modelo LISREL busca capturar relações de dependência entre variáveis observacionais e/ou entre variáveis latentes. A título de facilitar a compreensão do LISREL, são identificados aqui dois submodelos e, a seguir, o modelo completo.

1º submodelo: Modelos de relações estruturais com variáveis observacionais

Segundo BOLLEN (1989: 80-81), este é o tipo mais comum de modelo com relações estruturais. Entre as variáveis observacionais a ser analisadas, o pesquisador deve discriminar aquelas que são as explicativas (x), ou exógenas, e aquelas que são as explicadas (y), ou endógenas. Nesta modelagem não são considerados os erros de medida, mas apenas os erros de equação (ζ), advindos da não inclusão de variáveis relevantes para a explicação da variável endógena, por razões as mais diversas (como inexistência dos dados, dificuldades de coleta e o próprio desconhecimento do pesquisador). São também consideradas nesta modelagem as associações entre os erros de equação (ψ) e as associações entre as variáveis exógenas (ϕ). A equação básica deste submodelo é a seguinte:

$y = \beta y + \Gamma x + \zeta$
Onde:
β (beta) = matriz dos coeficientes dos efeitos entre as variáveis endógenas
Γ (gama) = matriz dos coeficientes dos efeitos das variáveis exógenas x nas variáveis endógenas y
ζ (zeta) = vetor dos erros de equação

(eq.2)

2º submodelo: Análise Fatorial Confirmatória

O principal objetivo deste submodelo é a confirmação, por meio da matriz Σ , da existência de uns poucos fatores latentes (ξ ou η) a partir das muitas variáveis observacionais (x ou y). Na análise fatorial confirmatória (AFC) são levados em consideração os erros de medidas (δ ou ϵ), as associações entre estes erros, e as associações (ϕ) entre as variáveis latentes (entre os ξ). A equação básica deste submodelo tem o seu desdobramento

nas variáveis observacionais exógenas (x) e nas endógenas (y), como se vê a seguir:

$x = \Lambda x \xi + \delta$
$y = \Lambda y \eta + \epsilon$
Onde:
Λ (lâmbda) = matriz dos coeficientes das cargas fatoriais
ξ e η (Ksi e Eta) = vetores das variáveis latentes
δ e ϵ (delta e epsolon) = vetores dos erros de medida
$\theta\delta$ e $\theta\epsilon$ (theta-delta e theta-epsolon) = matriz das associações entre os erros de medidas, respectivamente entre os δ e os ϵ
Obs.: os erros de medidas não podem ser correlacionados com as variáveis latentes

(eq.3)

Na realidade, a AFC não deixa de ser uma técnica multivariada de redução da dimensionalidade dos dados, como a análise fatorial exploratória (AFE). BOLLEN (1989: 227-228) sintetiza as diferenças entre ambas. Na AFE, em que normalmente se utiliza o *software* SPSS, não se tem, de antemão, um modelo teórico relacionando as variáveis latentes às variáveis observacionais. Em princípio, todas as variáveis latentes podem estar influenciando todas as variáveis observacionais, pois não se conhece quantos e quais são os fatores latentes antes de processada a análise. Como se vê, a análise é eminentemente exploratória.

No caso da análise confirmatória, como o próprio nome indica, já se tem *a priori* um modelo teórico de análise e quer-se confirmá-lo. Portanto, já se conhece quantos são os fatores latentes e quais são as variáveis observacionais que cada um deles influencia. Além disto, alguns parâmetros ou efeitos diretos podem ser fixados em zero ou em outro valor qualquer, e os erros de medidas podem ser correlacionados ou não – na AFE os erros de medidas não podiam estar associados.

Modelo completo, com equações estruturais e equações de medida

Este modelo é constituído dos dois submodelos: o estrutural e o de medidas. As equações de medidas especificam a composição dos fatores latentes a partir das variáveis observacionais – objeto do 2º

submodelo. Já as equações estruturais especificam as relações de causalidade entre as variáveis latentes, apontando a variância explicada (pelo modelo) e a não explicada – objeto do 1º submodelo, só que este envolvia apenas variáveis observacionais. As equações do modelo, em suas vertentes estrutural e de medidas, são apresentadas a seguir:

Equação estrutural: $\eta = \beta\eta + \Gamma\xi + \zeta$ Equações de medidas: $x = \Lambda x\xi + \delta$ $y = \Lambda y\eta + \varepsilon$

(eq.4)

São as seguintes as premissas básicas do modelo LISREL completo, segundo JORESKOG e SORBON (1996: 2):

- ε não é correlacionado com η
- δ não é correlacionado com ξ
- ζ não é correlacionado com ξ
- ζ, ε e δ são mutuamente não correlacionados

Importante destacar que no estudo de caso contemplado neste artigo (item 4) foi aplicado um modelo LISREL completo, segundo uma estrutura bastante simplificada. Antes, porém, os dados foram tratados e testados pela análise fatorial exploratória (AFE) e pela análise fatorial confirmatória (AFC). Deve-se lembrar que não se dispunha aqui, *a priori*, de um modelo teórico estabelecendo a relação entre perfil da pessoa, avaliação do governo e expectativas.

Identificabilidade e avaliação do modelo: A identificabilidade do modelo, entendida como a possibilidade de solução única para cada parâmetro a ser estimado, representa condição necessária para a sua resolução. Fazendo um paralelo com os sistemas de equações simples (de 1º ou 2º grau), em que o número de equações deve ser pelo menos igual ao número de incógnitas para que os sistemas tenham solução, também no caso dos modelos de relações causais esta regra se aplica. Em outras palavras, para que um modelo LISREL seja identificável é condição necessária que o número dos parâmetros (t) a ser estimados pelo modelo seja, no máximo, igual ao número das informações conhecidas (I), fornecidas pela matriz S das variâncias-co-variâncias. Como se sabe, “ I ” corresponde aos elementos da diagonal inferior (ou superior) da matriz S . A fórmula a seguir explicita

esta condição de identificabilidade, conhecida como “regra t ”:

$t \leq n(I) = (1/2)(p+q)(p+q+1)$ onde: $p = n^\circ$ de variáveis observacionais endógenas $q = n^\circ$ de variáveis observacionais exógenas $t = n(\theta) = n$ (espaço paramétrico) = $\{\Lambda y, \Lambda x, \beta, \Gamma, \phi, \Psi, \theta\delta, \theta\varepsilon, \theta\delta, \varepsilon, \}$ sendo: $\Lambda y, \Lambda x = n^\circ$ de elementos das matrizes das cargas fatoriais de η em y , e de ξ em x $\beta = n^\circ$ de elementos da matriz dos coeficientes entre os η $\Gamma = n^\circ$ de elementos da matriz dos coeficientes dos efeitos de ξ em η ϕ (phi) = n° de elementos da matriz das associações entre os ξ Ψ (psi) = n° de elementos da matriz das associações entre os ζ (erros de equação) $\theta\delta$ e $\theta\varepsilon = n^\circ$ de elementos da matriz das associações dos erros de medida, respectivamente entre os δ e entre os ε $\theta\delta, \varepsilon = n^\circ$ de elementos da matriz das associações dos erros de medida entre os δ e os ε Nota: Se a matriz S utilizada for a de correlação (KM) e se não forem fornecidos os valores do desvio-padrão para cada variável observacional, deve-se excluir o n° de elementos da diagonal principal no cômputo de I , já que estes são iguais à unidade.

(eq.5)

Se $t > n(I)$, o modelo não tem solução, ou seja, não é identificável; se $t = n(I)$, o modelo é exatamente identificável; se $t < n(I)$, o modelo apresenta $[n(I) - t]$ graus de liberdade (g.l.) para processar as estimativas, podendo ocorrer sobreidentificação, ou seja, redundância de valores para algum parâmetro estimado.

Uma vez rodado o modelo LISREL, cabe avaliar as soluções encontradas, de modo que se julgue a adequação do modelo. Pode-se dizer que esta avaliação pode ser subdividida em duas partes: local e global. Na avaliação local, destacam-se como critérios mínimos para o julgamento quanto à adequação do modelo: (1) se os parâmetros

estimados são significativamente diferentes de zero, evidenciado pelo teste “t” de Student³; (2) se é satisfatório o ajustamento linear propiciado pelas equações estruturais, evidenciado pelos coeficientes de determinação R^2 para as variáveis latentes endógenas, no caso do modelo completo; e (3) se também é satisfatório o ajustamento linear propiciado pelas equações de medida, evidenciado pelos coeficientes de determinação R^2 para as variáveis observacionais.

Na avaliação global, o que está em jogo é o ajustamento do modelo como um todo, onde a hipótese nula (H_0) a ser testada, pelo teste do Qui-quadrado (χ^2), é a de que os resíduos entre a matriz S e a matriz estimada $\Sigma(\theta)$ são significativamente iguais a zero, sob a premissa de normalidade dos erros. Dentre os muitos indicadores processados pelo LISREL para avaliar este ajustamento, destacam-se aqui quatro deles: (1) o valor do qui-quadrado (χ^2) é considerado como medida da falta de ajustamento do modelo e, portanto, deve ser o menor possível, tendo em vista o número de graus de liberdade do modelo. A estatística “p” demonstra se o valor do χ^2 encontrado está na área de rejeição / aceitação de H_0 (se, por exemplo, para $\alpha = 0,05$, $p < 0,05$, rejeita-se H_0); (2) o *Goodness of Fit Index* (GFI), que varia entre zero e um, e mede a proporção das variâncias e co-variâncias de S explicada pelo modelo. Equivale ao coeficiente de determinação (R^2) utilizado em regressão; (3) o *Adjusted Goodness of Fit Index* (AGFI), que varia também entre zero e um, e corresponde ao valor do GFI ajustado para o número de graus de liberdade. Corresponde ao coeficiente de determinação ajustado (R^2 ajustado) da regressão; e (4) a Raiz Média dos Resíduos ao Quadrado Padronizados (RMRQP), que mede a dispersão média padronizada entre os elementos de S e de $\Sigma(\theta)$. Quanto menor, melhor: serve para comparar o ajustamento de modelos diferentes para um mesmo conjunto de variáveis.

³ Vale como regra básica, quando não se pode ter acesso aos valores da tabela “t”, que, se o valor do parâmetro estimado dividido pelo respectivo desvio-padrão da amostra for maior do que 2, rejeita-se a hipótese nula (H_0) de que o valor do parâmetro é zero, supondo-se intervalo de confiança uni-caudal de 95%.

3. A PESQUISA AVALIAÇÃO FHC E AS QUESTÕES INVESTIGADAS

A pesquisa: A pesquisa de opinião aqui utilizada, arquivada sob o nº 00942 no banco de dados do CESOP/UNICAMP, tem um título bastante longo, mas vale a pena transcrevê-lo pois explicita as questões nela tratadas: *Avaliação Fernando Henrique Cardoso / Expectativa Fernando Henrique Cardoso / Denúncias contra o governo / Reforma da Previdência / Avaliação do Plano Real / Avaliação Congresso Nacional / Expectativa Congresso Nacional / Avaliação Governador do Rio de Janeiro / Expectativa Governador do Rio de Janeiro*. Além das 33 questões contempladas pela pesquisa, são também levantadas as chamadas variáveis de identificação do entrevistado, tais como: idade, sexo, cidade, renda familiar, grau de escolaridade, forma de inserção no mercado de trabalho e partido político de preferência.

Esta pesquisa foi realizada pelo DATAFOLHA nos dias 10 e 11 de dezembro de 1998, período de transição entre o 1º e o 2º mandato do presidente Fernando Henrique Cardoso (FHC). O universo pesquisado foi a população adulta do Estado do Rio de Janeiro, tendo sido coletada uma amostra de 1.068 pessoas.

Questões investigadas: A partir da aplicação da análise fatorial exploratória e do modelo LISREL (2º submodelo e modelo completo), buscou-se investigar se havia (ou não) relação de causalidade entre o perfil da pessoa, sua avaliação do 1º governo Fernando Henrique Cardoso e as suas expectativas futuras. Com base nos resultados obtidos, esta questão inicial foi desdobrada em duas outras, com os devidos ajustamentos nos conceitos (redefinição e eliminação).

É importante entender que, nessas três questões aqui tratadas, utilizou-se primeiro a análise exploratória, para explorar quais indicadores da base CESOP deveriam compor os conceitos em questão. Só depois foi utilizado o modelo LISREL, para confirmar a adequação destes conceitos e testar a relação de causalidade entre eles.

As questões investigadas foram, pois, as seguintes:

- 1) O perfil das pessoas, definido por critérios objetivos, influencia em sua avaliação do 1º

governo FHC e de suas expectativas futuras quanto à situação socioeconômica do país?

- 2) Por outro lado, se o perfil das pessoas for definido por critérios subjetivos, qual é a sua influência sobre os conceitos, confiança no presidente FHC e expectativas futuras?
- 3) Será que o que conta mesmo para a formação das expectativas é o grau de confiança das pessoas no presidente FHC?

Tratamento dos dados: Como esta pesquisa não foi especialmente desenhada para investigar as questões acima colocadas, torna-se importante fazer as seguintes considerações.

Primeiro, a pesquisa não pôde ser utilizada na íntegra; foram, portanto, selecionadas apenas aquelas perguntas do questionário que, à primeira vista, pudessem fornecer algum tipo de contribuição para a análise do tema. Não apenas selecionadas, mas sobretudo tratadas: os itens de cada questão tiveram que ser padronizados, isto é, recodificados segundo uma lógica comum a todas as questões.

Esta lógica comum a todas as questões foi a ordenação dos itens de cada pergunta no sentido de uma situação/atitude, em tese, mais favorável ao governo (definido como item 1), até aquela menos favorável. Os itens que não expressaram esta escala lógica foram considerados como *missing*. Esta *rationale* foi válida para as questões que conceitualizavam tanto “avaliação”, “confiança” e “expectativas”, como também “perfil”.

No caso das questões referentes aos três primeiros conceitos, a ordenação dos itens foi bastante óbvia; porém, tratando-se de “perfil”, essa ordenação não foi tão simples. Assim, nas questões relacionadas ao “perfil”, algumas hipóteses tiveram que ser formuladas de antemão e, posteriormente, testada a sua coerência *vis-à-vis* aos demais dados. Isto porque, se a ordenação proposta não fosse adequada, os coeficientes de correlação desta questão (ou variável) com as demais ficariam negativos em sua maioria. Exemplificando, veja-se que inicialmente a variável “idade” foi ordenada no sentido dos mais idosos (categoria 1 atribuída aos maiores de 60 anos) para os menos idosos (categoria 6 atribuída aos jovens de 16 e 17 anos), supondo-se que as pessoas mais velhas tenderiam a assumir postura favorável ao governo. No entanto, esta suposição não se verificou, haja vista que os coeficientes de correlação da variável idade com as

demais foram negativos em sua grande maioria, além de muito baixos.

O Anexo 1 apresenta as questões que foram selecionadas, e depois tratadas, a partir do questionário CESOP 00942.

A segunda consideração tem um caráter de ressalva à interpretação dos dados, objeto do próximo item. Muito provavelmente, se esta pesquisa de nº 00942 do CESOP houvesse sido especialmente desenhada para atender aos propósitos deste artigo, os resultados tenderiam a ser muito mais robustos do que os que foram efetivamente aqui encontrados. Isto porque, conforme visto no item 2, o modelo de relações causais tem por base a matriz Σ das associações entre as variáveis observacionais. Daí, seria de esperar que, no mínimo, um questionário que fosse especialmente desenhado para ser submetido a este tipo de modelagem deveria contemplar um número bem maior de opções (itens) significativamente lógicas⁴ por pergunta do que o que foi efetivamente considerado neste questionário do CESOP. Esta alteração, por si só, já representaria um *upgrade* considerável na estrutura dos dados a ser levantados, com efeitos positivos evidentes nos resultados a ser colhidos.

4. APLICAÇÃO DO MODELO LISREL E OS RESULTADOS ENCONTRADOS

Para cada uma das três questões mencionadas anteriormente serão apresentados os principais resultados encontrados. Volta-se a lembrar que a AFE foi aqui usada inicialmente em razão da inexistência de um modelo teórico que identificasse os fatores – o que foi feito com o auxílio do *software* SPSS.

1ª questão: O perfil das pessoas, definido por critérios objetivos, influencia em sua avaliação do 1º governo FHC e de suas expectativas futuras quanto à situação socioeconômica do país?

Inicialmente, para a conceitualização de “perfil das pessoas” (ξ1) foram consideradas praticamente todas as chamadas variáveis de identificação do

⁴ O importante não é aumentar o número absoluto apenas de opções por pergunta, mas sim de opções que possam ser classificadas dentro da escala lógica associada a cada questão.

entrevistado, à exceção de sexo⁵, a saber: cidade, renda familiar, escolaridade, partido político, idade e forma de inserção no mercado de trabalho.

Para o conceito “avaliação do 1º governo FHC” (η_1) foram selecionadas quatro variáveis: FHC1, Real, Congresso 1 e Corrupção. Já o conceito “expectativas futuras” (η_2) foi definido a partir das seguintes variáveis: FHC2, Congresso 2, “Econpaís”, “Econprópria”, Inflação, Desemprego e Salários.

Esta modelagem inicial contou ao todo com 17 variáveis observacionais⁶. Foram utilizados todos os 1.068 casos da amostra, e para os dados *missing*, adotou-se a opção do SPSS de substituí-los pela média da variável. Como mostra o Anexo 2, os dados *missing* são relativamente pouco significativos, à exceção, é claro, de partido político (onde eles assumem proporção bastante elevada na amostra, de 61%), e também de Congresso 2 e Corrupção (respectivamente 16% e 12% da amostra).

Análise fatorial exploratória: A AFE, utilizada para um primeiro contato com os dados, apontou que o fator “perfil da pessoa” não estava bem definido: na realidade, aquelas seis variáveis não definiam um fator apenas, mas dois. Decidiu-se manter apenas as variáveis que tinham as cargas fatoriais (depois de feita a rotação dos fatores pelo método Varimax) mais elevadas no primeiro fator, ou seja, grau de escolaridade ($\lambda=0,848$), renda familiar ($\lambda=0,834$) e cidade ($\lambda=0,346$). Com efeito, estas três variáveis guardam entre si uma certa lógica: é na capital do Estado onde a renda familiar e o grau de escolaridade das pessoas tendem a ser mais elevados; e vice-versa no que se refere à categoria “demais cidades” do Estado.

Por outro lado, as 4 variáveis que compunham o fator “avaliação do 1º governo FHC” acusaram, efetivamente, a existência de apenas um fator; porém, a variável “Congresso 1” destoava, em função de sua baixa carga fatorial *rotada* relativamente às demais ($\lambda=0,338$, sendo de 0,838

para FHC1, 0,808 para Real e 0,538 para Corrupção). O mesmo se observou em relação ao fator “expectativas futuras”, em que a variável “Congresso 2” apresentou baixa carga fatorial ($\lambda=0,397$, sendo que as seis outras tiveram carga fatorial acima de 0,6).

Uma vez eliminadas estas cinco variáveis (partido político, idade, forma de inserção no mercado de trabalho, Congresso 1 e Congresso 2), rodou-se novamente a análise fatorial exploratória para as 12 variáveis restantes. Ao invés dos 5 fatores, pouco definidos, que haviam sido identificados a partir daquelas 17 variáveis iniciais, foram agora identificados nitidamente os 3 fatores⁷ (a partir da matriz dos fatores rotados): “perfil”, “avaliação” e “expectativas”.

Análise fatorial confirmatória: Feita a AFE, o segundo passo foi confirmar esta conceitualização dos três fatores com base na AFC.

Como se pretendia identificar, além da dimensionalidade das variáveis observacionais e das associações entre os fatores latentes, optou-se aqui por considerar os três fatores latentes como exógenos (ξ). Caso contrário, se se rodasse em separado a AFC para as variáveis y (estrutura de η_1 e η_2) e depois para as variáveis x (estrutura de ξ_1), a matriz das associações entre os fatores latentes (ϕ) não existiria nos dois casos: no primeiro, porque não existe matriz PHI entre fatores latentes endógenos, e no segundo, porque só existe um fator latente exógeno.

A matriz de associação utilizada foi a de correlação (KM) e não a de variância-co-variância (CM), muito embora se tenha entrado também no modelo com os dados do desvio-padrão para cada variável observacional (o que possibilitou, dessa forma, a utilização alternativa da matriz CM). A preferência pela matriz de correlação deve-se à interpretabilidade imediata dos resultados apresentados pelo modelo, já que eles se encontram na escala padronizada, ou seja, variando entre -1 e $+1$.

⁵ A variável sexo apresentou correlação nula com todas as demais variáveis de identificação.

⁶ Daqui até o final do artigo, a referência às variáveis observacionais (questões do questionário) se dará pelo seu nome abreviado. Para melhor entendimento do significado das variáveis, ver o Anexo 1.

⁷ Nesta nova rodada da AFE notou-se que a variável FHC2 seria melhor classificada no fator “Avaliação” do que no fator “Expectativas”. Apesar disto, por razões de interpretação do fator (já que FHC2 se refere às expectativas quanto ao 2º governo FHC) a variável foi mantida no fator “Expectativas”.

Este 2º submodelo mostrou-se identificável [n(I)=78; t=27; g.l.=51]. Porém, não passou no teste da avaliação local. Assim, o teste “t” acusou que alguns parâmetros estimados não eram significativamente diferentes de zero, como as associações entre os conceitos “Perfil” e “Avaliação” ($\phi = 0,03$, sendo valor-t=0,88) e entre “Perfil” e “Expectativas” ($\phi = 0,01$, sendo valor-t=0,28). Assim, já se podia antever a inexistência de causalidade entre o perfil da pessoa *vis-à-vis* a sua avaliação do 1º governo FHC e as suas expectativas futuras, quando se rodasse o modelo completo.

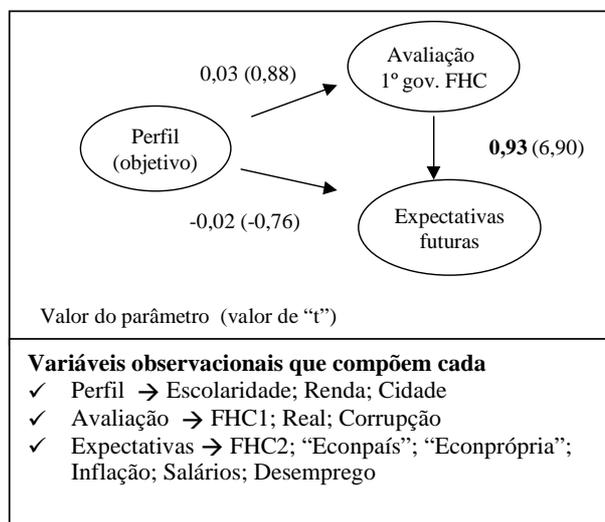
Ainda no âmbito da avaliação local, os valores de R^2 para algumas regressões foram muito baixos: de 0,06 para a variável cidade, em razão do fator latente perfil; de 0,11 para corrupção em razão do fator avaliação; e de 0,16 para desemprego, em razão de expectativas. Considerando que esta pesquisa do CESOP não foi especialmente desenhada para os propósitos deste artigo, e ainda tendo em vista que, nesta modelagem, cada uma das variáveis observacionais está explicada apenas por um fator latente, admite-se que um valor de R^2 local de 0,20, embora bastante baixo, possa ser tido como aceitável.

Também no que se refere à avaliação global do modelo, os resultados foram insatisfatórios. Vale notar que o valor do χ^2 chegou a 402,01, com $p=0,0$, ou seja, rejeitando-se a hipótese nula, a um nível de significância (α) de 0,05, de que os resíduos entre a matriz S e a matriz estimada Σ (θ) sejam significativamente nulos. Por sua vez, o AGFI foi de apenas 0,89, valor geralmente considerado baixo (inferior a 0,90).

Modelo completo: Além das equações de medidas (testadas no submodelo anterior), foram também testadas no modelo completo as relações estruturais de causalidade entre os fatores latentes. Como havia sido antecipado pela análise fatorial confirmatória, os efeitos de “perfil das pessoas” em “avaliação do 1º governo FHC” e em “expectativas futuras” foram significativamente nulos, sendo “ γ ” respectivamente de 0,03 e de -0,02 (Figura 1). Isto quer dizer que o perfil das pessoas, definido segundo variáveis objetivas, praticamente não influenciou na avaliação que elas faziam do 1º governo Cardoso e nas expectativas que elas tinham quanto à evolução socioeconômica do país no mandato seguinte. Por outro lado, cabe ressaltar o

efeito bastante expressivo da variável endógena “Avaliação” em “Expectativas” ($\beta_{21} = 0,93$).

Figura 1: Questão 1 – As relações estruturais do modelo completo



2ª questão: Por outro lado, se o perfil das pessoas for definido por critérios subjetivos, como se comporta a sua influência sobre os conceitos, confiança no presidente FHC e expectativas futuras?

Para investigar esta nova questão foram feitas algumas alterações na conceitualização não só do fator “perfil das pessoas” como também dos dois outros.

A caracterização do “perfil das pessoas” (ξ_1) passou a ser determinada por critérios subjetivos, ou seja, pela percepção da pessoa sobre a sua própria condição. Exemplificando, o que conta agora não é mais a renda familiar da pessoa (critério objetivo), mas a sua percepção dos rendimentos auferidos por ela e sua família, se são suficientes ou precários, o que é explicitado pela variável “dinheiro”. Além desta variável, compõem também este conceito as variáveis observacionais “econprópria”, que mede a percepção da pessoa sobre a própria situação econômica, e “partido”⁸, que denota a percepção da pessoa sobre o quadro partidário do país.

⁸ Ressalte-se aqui o caráter dúbio da variável partido: do ponto de vista do perfil da pessoa, ela foi vista tanto como uma característica objetiva quanto subjetiva.

Como visto no Anexo 2, dentre as variáveis selecionadas, a variável partido é a que apresenta, de longe, o maior número de dados *missing*. Como se tinha uma suspeita inicial de que esta variável fosse influente na avaliação do governo FHC, optou-se inicialmente por trabalhar apenas com os entrevistados que efetivamente tivessem manifestado alguma preferência partidária: com isto, a amostra teve que ser reduzida para 412 casos (foram desconsiderados todos os casos com valores *missing* para partido).

Também como resultado da análise exploratória para as 12 variáveis da questão anterior, foi identificado que o segundo fator, “avaliação”, ficaria melhor definido pelas seguintes variáveis: FHC1, FHC2 e Real. Com esta nova composição, este fator (η_1) perde a conotação de avaliação do 1º governo FHC e ganha o sentido mais amplo de “confiança no presidente FHC”.

Tendo perdido uma variável observacional para “perfil das pessoas” (“econprópria”) e outra para “confiança no governo FHC” (FHC2), a nova configuração do fator “expectativas futuras” (η_2) passou a ser composta das seguintes variáveis: “econpaís”, inflação, salários e desemprego.

Análise fatorial exploratória: Rodou-se a AFE para estas 10 variáveis observacionais selecionadas. O interessante a destacar é que, quando se consideram os autovalores maiores do que 1, são gerados apenas 2 fatores, que explicam 50% da variância total das variáveis. Observando-se a matriz dos fatores rotados, pode-se identificar os fatores “confiança no presidente FHC” (carregado nas variáveis FHC1, FHC2, Real e Partido) e “expectativas futuras” (carregado nas variáveis “econpaís”, inflação, salários, desemprego e “econprópria”). Já a variável dinheiro apresenta carga fatorial baixa e semelhante para os dois fatores, em torno de 0,2.

Quando se passa a considerar a geração de 3 fatores, agora explicando 60% da variância total das variáveis, a interpretação não fica muito clara, com exceção do primeiro fator (este define bem a “confiança no presidente FHC”, incorporando a variável partido). A idéia de “expectativas futuras” parece ficar diluída nos dois outros fatores: um carrega nas variáveis “econprópria” e “econpaís” (esta associação é explicável, uma vez que as pessoas tendem a fazer projeções semelhantes para

a própria situação econômica e a do país); e o outro fator carrega nas variáveis inflação e desemprego.

Análise fatorial confirmatória: Mesmo com o resultado um pouco dúbio à luz da AFE, resolveu-se testar esta segunda hipótese de conceitualização dos 3 fatores latentes com base na AFC.

O modelo mostrou-se identificável, com 32 graus de liberdade. À primeira vista, esta nova forma de conceitualização mostrou-se melhor do que na questão anterior, tanto no que se refere à avaliação global como à local. O coeficiente de determinação ajustado do modelo como um todo (AGFI) foi de 0,93. O teste “t” evidenciou que todos os parâmetros estimados foram significativamente diferentes de zero.

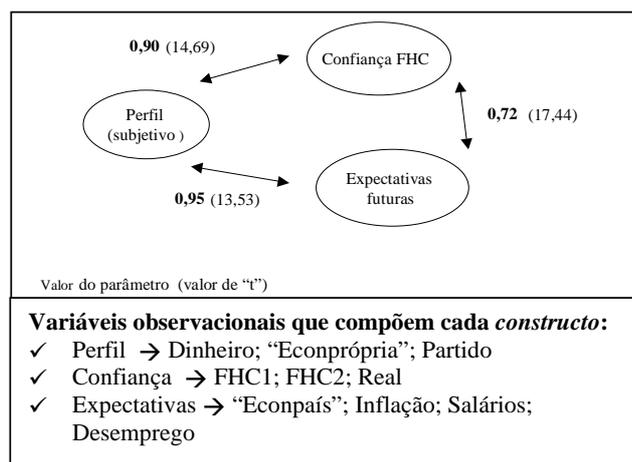
Porém, os coeficientes locais de determinação (R2) mostraram-se muito baixos para as variáveis dinheiro (0,10) e partido (0,18) – neste último caso, contrariando, até mesmo, a hipótese inicial assumida de que partido pudesse ser uma variável relevante para a caracterização do perfil subjetivo da pessoa. Como ambas as variáveis vêm compor, juntamente com uma terceira, um mesmo fator, o “perfil subjetivo”, esta constatação parece sugerir que a conceitualização deste fator ainda se encontra frágil no modelo, ou seja, ele ainda não está bem definido.

Além disso, note-se que, se no modelo de medidas anterior (questão 1) a associação de “perfil objetivo” com os dois outros fatores latentes era praticamente nula, a situação agora se inverte: os valores de ϕ para a associação entre perfil/confiança e entre perfil/expectativas tornam-se extremamente elevados, respectivamente de 0,90 e de 0,95 (Figura 2). Uma possível explicação, baseada na observação da matriz de correlação (KM) dos indicadores, é a de que os indicadores que compõem o “perfil subjetivo” estão mais associados com os indicadores que compõem os outros dois constructos do que com os indicadores do próprio constructo – tal é o caso da variável partido *vis-à-vis* ao constructo “confiança no presidente FHC”, e da variável “econprópria” em relação ao conceito “expectativas futuras”.

Particularmente no que se refere a essa elevada associação entre “perfil subjetivo” e “confiança”, isto não é bom, já que ambos viriam a ser, no modelo completo, as variáveis independentes da equação de regressão, em que “expectativas

futuras” seria a variável dependente. Com isso, o risco de multicolinearidade torna-se muito grande. E, como se sabe, a presença de multicolinearidade costuma gerar instabilidade nos modelos, produzindo resultados “inesperados” e não condizentes com a análise conduzida *a priori*.

Figura 2: Questão 2 – As associações entre os constructos na AFC



De fato, ao se rodar o modelo completo para estes três fatores latentes, o efeito obtido (γ_{21}) de “perfil subjetivo” em “expectativas” foi de 1,56, o que não poderia ocorrer já que os valores estavam padronizados e, portanto, deveriam estar contidos no intervalo entre -1 e $+1$. Por esta mesma razão, o efeito encontrado (β_{21}) de “confiança” em “expectativas” foi negativo ($-0,68$), o que não condiz obviamente com o resultado esperado (no sentido de que quanto maior a confiança no presidente FHC, melhores são as expectativas) e com a *rationale* da construção dos indicadores aqui utilizada.

Em situações de multicolinearidade, recomenda-se a remoção de uma variável do par das variáveis preditoras altamente correlacionadas (JOHNSON e WICHERN, 1998: 409-410). Uma vez que, conforme visto, a conceitualização do fator “perfil subjetivo” encontrava-se bastante frágil, optou-se aqui por excluir este fator. Passa-se, pois, a seguir, à

análise da relação entre os dois outros constructos, “confiança no presidente FHC” e “expectativas futuras”.

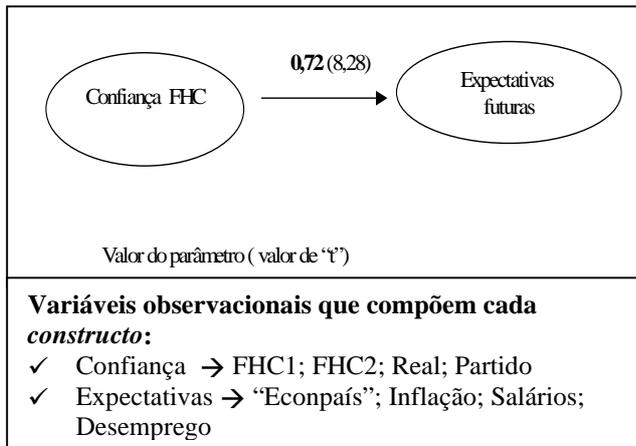
3ª questão: Será que o que é relevante para a formação das expectativas é o grau de confiança da pessoa no presidente FHC?

O fator latente exógeno (ξ_1) “confiança no presidente FHC” passou a ser definido pelas seguintes variáveis observacionais: FHC1, FHC2, Real, Partido – lembrar que na AFE da questão 2, quando foram gerados os 2 fatores, a variável partido compunha com estes três outros um mesmo fator. O fator latente endógeno (η_1) “expectativas futuras” foi conceitualizado por meio das variáveis: “econpaís”, inflação, salário e desemprego. Como se vê, o modelo tornou-se agora bem mais parcimonioso, com apenas 2 variáveis latentes e 8 variáveis observacionais. Como a variável partido continuou incluída no modelo, e já que seria interessante mantê-la com todo o seu poder discriminatório, a amostra trabalhada seguiu sendo a dos 412 entrevistados que manifestaram preferência partidária.

Análise fatorial exploratória: A AFE para as 8 variáveis gerou 2 fatores, e a interpretação para a matriz *rotada* coincidiu exatamente com a conceitualização aqui proposta para “confiança no presidente FHC” e “expectativas futuras”. Apenas estes dois fatores explicaram 56% da variação total das variáveis observacionais.

Modelo completo: Esta versão final do modelo completo ficou bastante simples e os seus resultados foram praticamente os mesmos da análise fatorial confirmatória. Também não poderia ser diferente, já que, no âmbito dessa modelagem, a diferença básica entre os dois submodelos foi que, no modelo completo, houve a inclusão do efeito direto (γ_{11}) de “confiança” em “expectativas”. Note-se que o valor encontrado para este efeito foi de 0,72, demonstrando que confiança no presidente FHC tem um efeito positivo elevado nas expectativas futuras das pessoas (Figura 3).

Figura 3: Questão 3 – relação de causalidade de “confiança” em “expectativas”



O modelo se mostrou identificável com 19 graus de liberdade. Em termos da avaliação local, os resultados foram relativamente satisfatórios. Com base no teste “ t ”, os parâmetros estimados foram todos significativamente diferentes de zero. O coeficiente R^2 para a variável latente endógena “expectativas futuras” foi de 0,51, mostrando que o modelo explica 51% da variação total desta variável. Já os valores de R^2 estimados para as variáveis observacionais foram bons em alguns casos (ou seja, assumindo valores de 0,73 para FHC2, 0,67 para FHC1 e 0,46 para Real) e apenas aceitáveis nos demais (0,39 para inflação, 0,36 para desemprego, 0,33 para salários, 0,28 para “econpaís” e 0,20 para partido).

O interessante a destacar é que, ao se rodar novamente o modelo excluindo a variável partido e incluindo todos os 1.068 casos da amostra, estes coeficientes locais de R^2 não apresentaram melhora generalizada, como seria de esperar. Alguns deles aumentaram um pouco, como o R^2 para “econpaís”, que passou de 0,28 para 0,38, mas a maioria piorou, como o R^2 do fator endógeno “expectativas”, que caiu de 0,51 para 0,46, e do desemprego, que reduziu de 0,36 para 0,28.

Em termos da avaliação global, apesar do tamanho da amostra (412 casos) ser menor do que o tamanho mínimo crítico⁹, definido pelo modelo em

⁹ Abaixo deste tamanho de amostra, o teste de χ^2 perde o poder de discriminar o ajuste global do modelo; com isto, podem ser obtidos valores pouco significativos do χ^2 , apesar de existirem erros substanciais na especificação do modelo (BOLLEN, 1989: 338).

436 casos, o teste de χ^2 continuou a rejeitar a hipótese nula de que a matriz estimada $\Sigma(\theta)$ tende a reproduzir a matriz observada S ($p = 0,017 < \alpha = 0,05$). As demais medidas de avaliação global foram favoráveis: AGF de 0,98, AGFI de 0,96 e RMRQP de 0,031.

Pode-se concluir, portanto, à luz dessa última modelagem do LISREL, que o fator “confiança no presidente FHC” consegue, de fato, influenciar as expectativas futuras da população quanto ao desempenho socioeconômico do país.

5. CONSIDERAÇÕES FINAIS

Do que foi visto, ficou evidenciado o “casamento” feliz e útil da análise fatorial exploratória com o modelo LISREL, quando não se tem *a priori* um modelo teórico de análise dos dados. Enquanto a primeira busca explicitar a dimensionalidade dos dados, ou seja, capturar os fatores latentes subjacentes às relações entre as variáveis, o segundo a complementa, uma vez que detecta as relações lógicas de causa e efeito entre os fatores.

Recapitulando, veja-se que a AFE teve aqui papel importante na (1) identificação, entre as 17 variáveis selecionadas intuitivamente no início, das mais indicadas, que foram 12, para compor os conceitos iniciais de “perfil objetivo da pessoa”, “avaliação do 1º governo FHC” e “expectativas”; na (2) alteração do sentido deste segundo conceito para “confiança no presidente FHC”, com base na interpretação das variáveis que se lhe mostraram mais ajustadas; na (3) sugestão da existência de 2 fatores latentes, e não 3, a partir das 10 variáveis analisadas na questão 2.

No âmbito do LISREL, a análise fatorial confirmatória serviu para confirmar os resultados encontrados na AFE e para antever os resultados do modelo completo. Assim, na questão 1, os coeficientes de associação não significativos entre “perfil” e “avaliação” e entre “perfil” e “expectativas” obtidos da AFC anteciparam a ausência de efeito entre estes conceitos. Na questão 2, a frágil conceitualização do fator “perfil subjetivo da pessoa”, percebida por meio da AFE, ficou evidente pela AFC, culminando com a eliminação deste constructo do modelo.

Já o modelo completo LISREL cumpriu aqui papel relevante na detecção das relações de causa e efeito entre as variáveis latentes do modelo. Assim, constatou-se que o “perfil da pessoa”, definido por critérios objetivos, não influenciou em sua “avaliação do 1º governo FHC” e nem nas “expectativas futuras”. Mas, foi a “confiança no presidente FHC” que conseguiu, de fato, influenciar as “expectativas futuras” – ou seja, foi o grau de confiança que as pessoas têm no presidente FHC que determinou as suas expectativas quanto às condições socioeconômicas do país no mandato seguinte.

Finalmente, é preciso ter claras as limitações desta aplicação do modelo LISREL. A pesquisa CESOP, embora muito rica em informações, não foi especialmente desenhada para este estudo. Como já salientado, a consequência disto foi o empobrecimento da base de dados¹⁰ efetivamente analisada, atribuído (1) às poucas variáveis relevantes ao tema; (2) às poucas possibilidades de relacionamento, ou inter cruzamento, entre estas variáveis; e (3) ao reduzido número de opções significativamente lógicas por pergunta. São estas limitações da base de dados trabalhada que explicam, de certa forma, os resultados limitados aqui encontrados. Como foi visto, em nenhum momento se conseguiu o ajuste global do modelo, medido pela estatística do χ^2 . Também os valores de R^2 para as variáveis observacionais foram, em geral, muito baixos, em parte porque elas estavam associadas a apenas um fator latente.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- BOLLEN, K. A. *Structural Equations with Latent Variables*. USA: John Wiley & Sons Inc., 1989.
- CESOP (Centro de Estudos de Opinião Pública) / UNICAMP (Universidade Estadual de Campinas). Banco de Dados de Pesquisas por Amostragem, *Pesquisa DAT/00942*. Rio de Janeiro: dezembro 1998.
- JOHNSON, R. A. e WICHERN, D. W. *Applied Multivariate Statistical Analysis*. 4. ed. USA: Prentice_Hall Inc., 1998.
- JORESOG, K. e SORBOM, D. *LISREL 8: User's Reference Guide*. USA: Scientific Software International Inc., 1996.

¹⁰ Entendida como as questões selecionadas e tratadas no âmbito da Pesquisa CESOP nº 00942.

ANEXO 1: Questões selecionadas e tratadas do Questionário CESOP para serem utilizadas no modelo de relações causais

var *label* idade “Idade” (**Idade**)

- 1 “60 anos ou mais”.
- 2 “45 a 59 anos”
- 3 “35 a 44 anos”
- 4 “25 a 34 anos
- 5 “18 a 24 anos”
- 6 “16 e 17 anos”

var *label* p1 (**FHC1**) “No final do ano termina o primeiro mandato do presidente Fernando Henrique. Na sua opinião, Fernando Henrique está fazendo um governo”:

- 1 “Ótimo”
- 2 “Bom”
- 3 “Regular”
- 4 “Ruim” ou
- 5 “Péssimo?”

Var *missing*: “Não sabe”

var *label* p3 (**FHC2**) “Em 1º de janeiro o presidente Fernando Henrique Cardoso, que foi reeleito em outubro, inicia seu segundo mandato. Na sua opinião, ele vai fazer um governo”:

- 1 “Ótimo”
- 2 “Bom”
- 3 “Regular”
- 4 “Ruim” ou
- 5 “Péssimo?”

Var *missing*: “Não sabe”

var *label* p8 (**Real**) “De um modo geral você acha que o Plano Real está sendo”:

- 1 “Ótimo”
- 2 “Bom”
- 3 “Regular”
- 4 “Ruim” ou
- 5 “Péssimo?”

Var *missing*: “Não sabe”

var *label* p9a (**Inflação**) “Na sua opinião daqui para a frente a inflação vai aumentar, diminuir ou ficar como está?”.

- 1 “Vai diminuir”
- 2 “Vai ficar como está”
- 3 “Vai aumentar”

Var *missing*: “Não sabe”

var *label* p9b (**Desemprego**) “E o desemprego, vai aumentar, diminuir ou ficar como está?”.

- 1 “Vai diminuir”
- 2 “Vai ficar como está”
- 3 “Vai aumentar”

Var *missing*: “Não sabe”

var *label* p9c (**Salário**) “E o poder de compra dos salários, vai aumentar, diminuir ou ficar como está?”.

1 “Vai aumentar”

2 “Vai ficar como está”

3 “Vai diminuir”

Var *missing*: “Não sabe”

var *label* p10a (**Econpaís**) “Na sua opinião, nos próximos meses, a situação econômica do país vai melhorar, vai piorar ou vai ficar como está?”.

1 “Vai melhorar”

2 “Vai ficar como está”

3 “Vai piorar”

Var *missing*: “Não sabe”

var *label* p10b (**Econprópria**) “E no seu caso pessoal, você acha que a sua situação econômica vai melhorar, vai piorar ou vai ficar como está?”.

1 “Vai melhorar”

2 “Vai ficar como está”

3 “Vai piorar”

Var *missing*: “Não sabe”

var *label* p11 (**Dinheiro**) “Você diria que o dinheiro que você e sua família ganham é”:

1 “Mais do que suficiente”

2 “É exatamente o que precisam para viver”

3 “Não é suficiente, às vezes falta” ou

4 “É muito pouco, trazendo muitas dificuldades?”

var *label* p16 (**Congresso 1**) “No final do ano termina o mandato dos atuais deputados federais e de parte dos senadores que formam o atual Congresso Nacional. Na sua opinião, os atuais senadores e deputados federais, eleitos em 1994, estão tendo um desempenho”:

1 “Ótimo”

2 “Bom”

3 “Regular”

4 “Ruim”

5 “Péssimo?”

Var *missing*: “Não sabe”

var *label* p17 (**Congresso 2**) “E os novos senadores e deputados eleitos que tomarão posse no ano que vem, terão um desempenho”:

1 “Ótimo”

2 “Bom”

3 “Regular”

4 “Ruim”

5 “Péssimo?”

var *label* p24 (**Corrupção**) “O presidente Fernando Henrique Cardoso e alguns membros do PSDB estão sendo acusados de serem donos de uma empresa, criada em Janeiro de 1994, nas Bahamas, cuja conta bancária guardaria mais de trezentos milhões de dólares, e que serviria para guardar sobras de dinheiro utilizado em campanhas políticas. Na sua opinião”:

1 “as denúncias são falsas e não devem ser investigadas”

2 “as denúncias são falsas, mas devem ser investigadas” ou

3 “não sabe se é verdadeira ou falsa, mas devem ser investigadas”

4 “as denúncias são verdadeiras e devem ser investigadas”

Var *missing*: “Outras respostas”; “Não sabe”; “Devem ser investigadas”

var *label* **pea (Inserção no mercado de trabalho)** “Atualmente você trabalha? (Se sim) Qual é a sua ocupação principal? (Se não) Você só estuda, é aposentado, é dona-de-casa ou está desempregado? (Se desempregado) Está procurando emprego ou não?”

1 “Assalariado registrado”; “Funcionário Público”; “Autônomo regular (paga ISS)”; “Profissional Liberal (Autônomo universitário)”; “Empresário”: Inserção formal no mercado de trabalho

2 “Só dona-de-casa”; “Só aposentado”; “Só vive de rendas”

3 “Só estudante”; “Estagiário / aprendiz (remunerado)”

4 “Assalariado sem registro”; “Freelance / bico”: Inserção informal no mercado de trabalho

5 “Desempregado (procura emprego) – PEA

Var *missing*: “Outros - PEA”; “Outros - Não PEA”; “Desempregado (Não procura emprego) - Não PEA”¹¹

var *label* **partido** “Qual é o partido¹² de sua preferência? (Espontânea e única)”

1 “PFL”; “PSDB”; “PPB (PDS / PPR)”: Pró FHC

2 “PMDB”; “PTB”; Posição indefinida

3 “PT”; “PDT”; “PSB”; “PL”: Contra FHC

Var *missing*: “Outro partido”; “Nomes e referências”; “Nenhum / não tem”

var *label* **escola (Escolaridade)** “Até que ano da escola você estudou?”

1 “Pós-Graduação”

2 “Superior completo”

3 “Superior incompleto”

4 “Colegial completo”

5 “Colegial incompleto”

6 “Ginasial completo”

7 “Primário completo / ginasial incompleto”

8 “Analfabeto / primário incompleto”

var *label* **rendaf (Renda familiar)** “Somando a sua renda com a renda das pessoas que moram com você, quanto é aproximadamente a renda familiar na sua casa?”¹³

1 “R\$6.501,00 ou mais”

2 “De R\$2.601,00 até R\$6.500,00”

3 “De R\$1.301,00 até R\$2.600,00”

4 “De R\$651,00 até R\$1.300,00”

5 “De R\$391,00 até R\$650,00”

6 “De R\$261,00 até R\$390,00”

7 “... até R\$260,00”

Var *missing*: “Não sabe”; “Recusa”

¹¹ Na definição desta categoria, vale registrar que os organizadores do questionário cometeram uma certa confusão conceitual. Isto porque, por definição, o desempregado faz parte da PEA, que é constituída (1) por ocupados e (2) por não ocupados procurando trabalho, que são justamente os desempregados. E o desempregado já havia sido incluído no item 5. Quem não trabalha e não procura trabalho compõe a NÃO PEA dentro da PIA (População em Idade Ativa).

¹² Embora “PRN” tenha sido uma categoria prevista no questionário, nenhum entrevistado manifestou preferência por este partido.

¹³ Para a caracterização da renda do entrevistado seria preferível que a pesquisa CESOP tivesse adotado a renda familiar *per capita*, e não simplesmente a renda familiar. Supondo-se 2 entrevistados com a mesma renda familiar, aquele que vive em uma família com 10 filhos está em situação financeira bem pior do que aquele que vive em uma família com apenas 1 filho.

var label cidade “**Cidade**”:

“São Gonçalo”; “Nova Iguaçu”; “Duque de Caxias”; “Petrópolis”; “São João de Meriti”; “Niterói”; “Magé”; “Paty do Alferes”; “Miracema”; “Macaé”; “Campos dos Goytacazes”; “Bom Jardim”; “Três Rios”; “Arraial do Cabo”; “Araruama”; “Resende”; “Volta Redonda”; “Queimados”: Demais cidades
1 “Rio de Janeiro”: Capital

Fonte: CESOP / UNICAMP, Pesquisa n° 00942

ANEXO 2: Percentual de casos *missing* por variável selecionada da Pesquisa CESOP

Variáveis	% de casos <i>missing</i> (*)	Variáveis	% de casos <i>missing</i> (*)
Cidade	0,0	Corrupção	12,5
Idade	0,0	FHC2	7,7
Escolaridade	0,0	Congresso2	15,7
Renda familiar	3,1	Econpaís	3,1
Trabalho	3,7	Econprópria	3,4
Partido	61,4	Inflação	3,7
FHC1	1,4	Desemprego	2,1
Real	0,4	Salários	2,9
Congresso1	5,9	Dinheiro	0,0

Fonte: CESOP / UNICAMP, Pesquisa n° 00942

(*) N° total de casos analisados: 1.068