

## 4

### **A Produção do Indicador Sócio Econômico via *Teoria da Resposta ao item* (TRI).**

Já sabendo previamente que o modelo de respostas graduadas da TRI era o mais adequado para implementação, o próximo passo é analisar a natureza de cada opção de item, bem como a transformação das que se faziam necessárias. Como já foi descrito anteriormente, o MRG necessita para sua implementação que haja uma ordem graduada crescente entre as opções. Os itens que apresentavam formato dicotômico não precisaram ser modificados, pois o modelo logístico de dois parâmetros é um caso particular do MRG. Uma vez realizada a transformação das opções, o passo seguinte foi implementar uma análise de fatores para os itens intervalares. A análise fatorial (anexo 1) neste momento é importante para possibilitar a verificação de possíveis correlações espúrias.

A primeira análise fatorial teve como objetivo diminuir o número de variáveis.

Apresenta-se no anexo 4 o resultado da primeira análise fatorial realizada com auxílio do *software* SPSS<sup>®</sup>. Utilizou-se o método de extração por componentes principais, seguida da aplicação do método de rotação VARIMAX. O resultado obtido para a matriz de cargas nos fatores associadas às variáveis é apresentado, também, no anexo 4.

Percebe-se que existem algumas variáveis a priori que estavam no questionário com exatamente o mesmo objetivo de pergunta. É o caso das variáveis: **Máquina de Lavar e Lava Roupa; Rádio e Rádio Elétrico**. Uma vez percebido tal fato o procedimento foi escolher apenas uma delas. Uma outra correlação muito alta e óbvia ocorre, por exemplo, com as variáveis: **Computador e Impressora**. Também por razões óbvias retiramos a variável impressora. Pode-se dizer que os itens que apresentam esse tipo de correlação estão localmente correlacionados (na linguagem da TRI, são localmente dependentes entre si). Existem algumas correlações não interpretáveis ou espúrias na análise. É o caso da observada para **Exaustor, Cafeteira Elétrica e Forno Elétrico**.

O próximo passo foi estimar quais itens eram irrelevantes para a análise construtiva do índice, pois não se associavam adequadamente ao restante dos itens. Tal procedimento foi realizado pelo *software Parscale*<sup>®</sup>, utilizando-se as correlações polisseriais.

## 4.1

### Índices Para o Padrão de Vida

#### 4.1.1

##### Introdução

O PROCEL - Programa Nacional de Conservação de Energia Elétrica, é um programa do governo voltado para a conservação de energia elétrica. Instituído em dezembro de 1985 e implantado no ano seguinte, o PROCEL é coordenado pelo Ministério de Minas e Energia, cabendo à Eletrobrás o controle de sua execução.

Seu principal objetivo é a conservação da energia elétrica, tanto no lado da produção como no do consumo, concorrendo para a melhoria da qualidade de produtos e serviços, reduzindo os impactos ambientais e fomentando a criação de empregos. Para isso, o PROCEL desenvolve projetos nas mais diversas áreas.

O Programa utiliza recursos da Eletrobrás e da Reserva Global de Reversão - fundo federal constituído com recursos das concessionárias, proporcionais ao investimento de cada uma. Utiliza, também, recursos de entidades internacionais.

Uma pesquisa sobre posse e hábitos de consumo de eletrodomésticos e energia elétrica foi conduzida na região urbana do estado de Mato Grosso do Sul. Tal pesquisa foi encomendada pela concessionária Enersul com o objetivo de entender de que forma o consumidor está usando os serviços de energia do estado. Vale lembrar que essa pesquisa começou em todo o país muito antes dos problemas de racionamento (o chamado “apagão”). O questionário possui em torno de 60 questões das mais variadas possíveis com o intuito de traçar um perfil de consumo de energia e posse de aparelhos que consumam energia. Parte dessas questões é destinada à avaliação da condição sócio-econômica familiar do respondente. Ver **Tabela 10** abaixo.

Note-se que a questão Q04 refere-se à escolaridade do chefe de família. De fato, a variável resultante dessas questões pode ser uma boa *proxy* para a condição sócio-econômica da família e, portanto será utilizada na produção dos índices.

A primeira questão a ser observada é a da escolha dos níveis de medida das variáveis. Modelos diferentes são empregados quando se têm níveis diferentes de medida. Frequentemente, há necessidade de se aproximar uma variável com um determinado nível de mensuração para um outro nível. Este é o caso frequente da aproximação de variáveis com escalas ordinais para escalas intervalares. Essa aproximação ocorre, sobretudo, quando as ferramentas que estão disponíveis para a condução das análises não são totalmente adequadas ao nível de medida que foi empregado na construção original das variáveis, o que pode ser devido à experiência do analista ou devido à disponibilidade de *softwares* para análise ou, ainda, pela inexistência de métodos adequados.

No presente estudo, como pode ser observado pela forma como as questões foram construídas, há variáveis com nível de medida categórico que essencialmente admitem respostas dicotômicas (Se o domicílio aquece ou não à água, usando chuveiro elétrico? (0-não; 1-sim) etc), outras com nível intervalar, porém com valores inteiros (número de geladeiras, televisores, e etc), o que caracteriza variáveis típicas de contagem, e outras com nível ordinal de medida (as que medem a escolaridade do chefe de família).

Algumas perguntas originais foram transformadas (divididas) em outras para a busca de uma melhor análise. As questões transformadas foram com base em uma análise de distribuição de frequência das respostas dadas pelos respondentes. Segue abaixo a tabela com as perguntas que foram utilizadas para análise.

<i>Questão</i>	<i>Descrição</i>
Q01	Numero de banheiros na casa? (0; 1; 2; 3; 4 ou mais)
Q02	Numero de carros? (0; 1;2; 3; 4 ou mais)
Q03	Numero de empregados? (0; 1; 2 ou mais)
Q04	Nível de instrução do chefe da família?(1) até primário incompleto; (2) até ginásial incompleto; (3) até colegial incompleto; (4) até superior incompleto; (5) curso superior completo
Q05	Tipo de domicilio: (1-casa; 2-apartamento; 3-outro)
Q06	Tipo de parede: (1-alvenaria; 2-madeira;3-material aproveitado)
Q07	Seu piso é de terra? (1-sim; 0-não)
Q08	Seu piso é de cimento? (1-sim; 0-não)
Q09	Sua cobertura é de laje de concreto? (1-sim; 0-não)
Q10	Sua cobertura é de telha de barro? (1-sim; 0-não)
Q11	Sua cobertura é de telha de amianto? (1-sim; 0-não)
Q12	Sua cobertura é de zinco? (1-sim; 0-não)
Q13	Área construída da residência em m <sup>2</sup> : (1) ate 50; (2) de 51 a 75; (3) de 76 a 100; (4) de 101 a 150; (5) de 151 a 200; (6) acima de 200
Q14	Condição de ocupação do domicilio: (1-próprio; 2- não próprio)
Q15	O domicilio possui abastecimento de água? (1-não é rede geral; 2- é rede geral)
Q16	Número de geladeiras? (0; 1; 2; 3 ou mais)
Q17	Número de freezers? (0; 1; 2 ou mais)
Q18	Número de aparelhos de ar condicionado? (0; 1; 2; 3 ou mais)
Q19	Número de tvs? (0; 1; 2; 3; 4 ou mais)
Q20	Número e aparelhos de som? (0; 1; 2 ou mais)
Q21	Número de rádios? (0; 1; 2 ou mais)
Q22	Número de videocassetes? (0; 1; 2 ou mais)
Q23	Número de computadores? (0; 1; 2 ou mais)
Q24	Número de videogames? (0; 1)
Q25	Número de ferros de passar? (0; 1; 2 ou mais)
Q26	Número de máquinas de lavar roupa? (0; 1; 2 ou mais)
Q27	Número de máquinas de lavar louça? (0; 1; 2 ou mais)
Q28	Número de secadoras de roupa? (0; 1)
Q29	Número de fornos de microondas? (0; 1)
Q30	Número de fornos elétrico? (0; 1)
Q31	Número de liquidificadores? (0; 1)
Q32	Número de batedeiras? (0; 1)
Q33	Número de cafeteiras elétrica? (0; 1)
Q34	Número de panelas elétricas? (0; 1)
Q35	Número de exaustores? (0; 1)
Q36	Número de ventiladores? (0; 1; 2; 3; 4 ou mais)
Q37	Número de aquecedores? (0; 1)
Q38	Número de enceradeiras? (0; 1)
Q39	Número de aspiradores de pó? (0; 1)
Q40	Número de bombas d'água? (0; 1)
Q41	Número de chuveiros elétricos? (0; 1; 2; 3 ou mais)
Q42	Se o domicilio aquece ou não a água, usando chuveiro elétrico? (0-não; 1-sim)

Tabela 10 - Questões utilizadas na construção do índice

## 4.2

**Índice Obtido a Partir de Modelos para Respostas Graduadas**

Finalmente, com o intuito de buscar maior informação na construção do índice sócio-econômico, explorando-se ao máximo a estrutura das questões, utilizou-se modelos de respostas graduadas na construção do índice . Esse procedimento também foi estudado em Soares, 2003.

Para calibração dos modelos foi utilizado o *software Parscale®*. Inicialmente produziu-se as correlações de Pearson e Polisserial de cada variável:

<i>Questão</i>	<i>Correlação Pearson</i>	<i>Correlação poliserial</i>	<i>Questão</i>	<i>Correlação Pearson</i>	<i>Correlação poliserial</i>
Q01	0.674	0.860	Q22	0.567	0.722
Q02	0.660	0.828	Q23	0.408	0.766
Q03	0.543	0.888	Q24	0.337	0.578
Q04	0.631	0.694	Q25	0.298	0.498
<b>Q05</b>	<b>0.085</b>	<b>0.214</b>	Q26	0.403	0.536
Q06	0.311	0.447	Q27	0.212	0.486
Q07	0.181	0.526	Q28	0.153	0.438
Q08	0.354	0.465	Q29	0.468	0.756
Q09	0.321	0.472	Q30	0.292	0.595
<b>Q10</b>	<b>0.145</b>	<b>0.184</b>	Q31	0.411	0.569
Q11	0.424	0.541	Q32	0.437	0.550
<b>Q12</b>	<b>-0.005</b>	<b>-0.031</b>	Q33	0.208	0.375
Q13	0.567	0.600	Q34	0.145	0.421
<b>Q14</b>	<b>0.059</b>	<b>0.088</b>	Q35	0.252	0.504
<b>Q15</b>	<b>-0.159</b>	<b>-0.343</b>	Q36	0.486	0.527
Q16	0.370	0.473	<b>Q37</b>	<b>0.076</b>	<b>0.271</b>
Q17	0.460	0.620	Q38	0.261	0.496
Q18	0.484	0.827	Q39	0.301	0.603
Q19	0.568	0.644	Q40	0.203	0.398
Q20	0.395	0.477	Q41	0.562	0.642
<b>Q21</b>	<b>0.064</b>	<b>0.082</b>	Q42	0.406	0.590

Tabela 11 - Resultados das Correlações de Pearson e Polisserial

O primeiro passo na análise é o da avaliação da correlação polisserial, para verificar se uma determinada variável apresenta correlação significativa com o escore bruto produzido com o conjunto das variáveis. Esse passo é crucial para a escolha de itens que de fato apresentam consistência interna e se associam bem ao escore que será produzido. Normalmente, aceita-se valores para a correlação polisserial superior a 0.3.

Esses resultados permitem concluir que as variáveis Q05; Q10; Q12; Q14; Q15; Q21 e Q37 não apresentam uma boa correlação com o escore bruto produzido pelo conjunto das variáveis, portanto, não apresentam boa discriminação para o índice que se deseja produzir. Essas variáveis, então, são excluídas de qualquer análise posterior.

Os parâmetros do modelo de resposta graduada, com seus respectivos erros-padrão, são apresentados abaixo.

Questão	Parâmetro de Discriminação (a)	Parâmetro de Dificuldade (b)	Parâmetro de Dificuldade de Cada item
Q01	1,511 (0,070)	0,592 (0,037)	$\infty$ 3.81 -0.191 -1.183 -1.807
Q02	1,405 (0,075)	2,239 (0,041)	$\infty$ 1.771 0.273 -0.784 -1.260
Q03	1,516 (0,120)	2,059 (0,065)	$\infty$ 0.699 -0.699
Q04	0,718 (0,030)	0,987 (0,058)	$\infty$ 1.501 0.515 -0.168 -1.847
Q06	0,626 (0,041)	-2,746 (0,117)	$\infty$ 0.996 -0.996
Q07	1,960 (0,309)	-2,437 (0,119)	*
Q08	0,636 (0,051)	-0,959 (0,085)	*
Q09	0,342 (0,052)	2,762 (0,405)	*
Q10	0,144 (0,039)	2,015 (0,599)	*
Q11	0,401 (0,043)	-0,875 (0,124)	*
Q13	0,829 (0,025)	0,147 (0,042)	$\infty$ 2.676 0.927 -1.009 -2.593
Q16	0,758 (0,028)	1,879 (0,063)	$\infty$ 3.711 -0.720 -2.991
Q17	0,862 (0,053)	2,521 (0,068)	$\infty$ 1.464 -1.464
Q18	1,349 (0,092)	2,103 (0,066)	$\infty$ 0.568 -0.025 -0.543
Q19	0,953 (0,032)	1,150 (0,041)	$\infty$ 3.010 0.267 -1.150 -2.127
Q20	0,643 (0,028)	1,574 (0,063)	$\infty$ 2.254 -2.254
Q22	1,309 (0,071)	1,721 (0,041)	$\infty$ 1.102 -1.102
Q23	1,682 (0,146)	2,399 (0,068)	$\infty$ 0.700 -0.700

Q24	0,769 (0,080)	2,105 (0,163)			*
Q25	0,889 (0,044)	0,631 (0,085)	$\infty$	2.920	-2.920
Q26	0,778 (0,041)	1,584 (0,067)	$\infty$	2.771	-2.771
Q27	0,759 (0,114)	3,004 (0,332)			*
Q28	0,764 (0,151)	3,555 (0,526)			*
Q29	1,493 (0,120)	1,379 (0,059)			*
Q30	0,803 (0,100)	2,568 (0,229)			*
Q31	0,967 (0,069)	-1,083 (0,065)			*
Q32	0,896 (0,060)	0,193 (0,049)			*
Q33	0,666 (0,080)	2,498 (0,235)			*
Q34	0,712 (0,152)	3,780 (0,623)			*
Q35	0,739 (0,095)	2,666 (0,255)			*
Q36	0,619 (0,020)	1,498 (0,055)	$\infty$	2.833	0.475 -0.868 -0.440
Q38	0,702 (0,087)	2,590 (0,245)			*
Q39	1,206 (0,126)	2,027 (0,118)			*
Q40	0,398 (0,080)	4,227 (0,765)			*
Q41	1,166 (0,044)	1,117 (0,035)	$\infty$	1.950	-0.449 -1.502
Q42	1,233 (0,092)	1,151 (0,060)			*

Tabela 12 - Parâmetros dos itens

Uma vez terminada análise de correlação polisserial e tendo o percentual de posses dos itens que se mostram relevantes para a aplicação do modelo apresentamos no anexo 5 os resultados dos parâmetros estimados com seus gráficos.

Uma segunda análise fatorial, cujo resultado se encontra abaixo, foi realizada com o objetivo de verificar se a hipótese da unidimensionalidade está sendo satisfeita. Ao contrário da primeira análise fatorial cujo objetivo principal era o de analisar correlações espúrias e investigar outras variáveis latentes associadas às questões. A segunda análise fatorial trará a informação sobre a violação ou não do postulado do modelo unidimensional logístico MRG. Se o pressuposto estiver sendo violado, significa apenas que o modelo não está adequado, mas não implica que a TRI não possa ser usada. Haja visto que existem modelos apropriados para casos multidimensionais (Linden, 1997).

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	16.619	59.353	59.353	16.619	59.353	59.353
2	3.499	12.495	71.848	3.499	12.495	71.848
3	1.183	4.225	76.073	1.183	4.225	76.073
4	.875	3.126	79.198			
5	.680	2.429	81.628			
6	.573	2.046	83.673			
7	.473	1.690	85.364			
8	.449	1.602	86.965			
9	.405	1.445	88.410			
10	.371	1.326	89.736			
11	.364	1.299	91.035			
12	.301	1.076	92.111			
13	.292	1.041	93.153			
14	.264	.942	94.095			
15	.200	.714	94.809			
16	.189	.676	95.485			
17	.181	.645	96.131			
18	.168	.598	96.729			
19	.162	.577	97.306			
20	.149	.534	97.840			
21	.139	.497	98.337			
22	.121	.434	98.770			
23	.104	.372	99.143			
24	.100	.358	99.500			
25	6.461E-02	.231	99.731			
26	4.964E-02	.177	99.908			
27	1.738E-02	6.208E-02	99.970			
28	8.329E-03	2.975E-02	100.000			

Extraction Method: Principal Component Analysis.

Figura 7 – Variância total explicada

Com base na segunda análise fatorial apresentada acima podemos perceber que cerca de 59% da variância total é explicada pela componente 1. Isto nos leva a conceber que a hipótese de unidimensionalidade não foi violada. Segundo Hambleton (1991) a unidimensionalidade não pode ser inspecionada diretamente, porém existem evidências que podem ser coletadas e analisadas. Foi isso que fizemos ao produzir uma segunda análise fatorial. Hambleton afirma ainda que se a unidimensionalidade é satisfeita, a investigação da dependência local não é necessária, pois esta é consequência direta da presença da primeira. Porém a independência local pode ser obtida mesmo que a unidimensionalidade não seja satisfeita (Hambleton (1991) p 10-11).