

Roger Keller Celeste¹João Luiz Bastos^{II}

Mid-point for open-ended income category and the effect of equivalence scales on the income-health relationship

Ponto médio em categoria aberta de renda e o efeito de escalas de equivalência nas relações renda-saúde

ABSTRACT

To estimate the mid-point of an open-ended income category and to assess the impact of two equivalence scales on income-health associations. Data were obtained from the 2010 Brazilian Oral Health Survey (*Pesquisa Nacional de Saúde Bucal – SBBrazil 2010*). Income was converted from categorical to two continuous variables (*per capita* and equalized) for each mid-point. The median mid-point was R\$ 14,523.50 and the mean, R\$ 24,507.10. When *per capita* income was applied, 53% of the population were below the poverty line, compared with 15% with equalized income. The magnitude of income-health associations was similar for continuous income, but categorized equalized income tended to decrease the strength of association.

DESCRIPTORS: Dental Health Surveys, Statistics as Topic, Income, Economics, Oral Health.

RESUMO

Objetivou-se estimar o ponto médio de categoria aberta de renda e o efeito de escalas de equivalência nas associações renda-saúde. Dados foram obtidos da Pesquisa Nacional de Saúde Bucal 2010 (SBBrazil). A renda foi convertida da variável categórica para contínua (*per capita* e equalizada) para cada ponto médio. A mediana da categoria aberta de renda foi R\$ 14.523,50 e a média, R\$ 24.507,10. O percentual da população abaixo da linha de pobreza foi de 53% com renda *per capita* e 15% com renda equalizada. A magnitude das associações foi similar na escala contínua, mas, quando categorizada, a renda equalizada reduziu sua força.

DESCRIPTORES: Inquéritos de Saúde Bucal, Estatística como Assunto, Renda, Economia, Saúde Bucal.

¹ Departamento de Odontologia Preventiva e Social. Faculdade de Odontologia. Universidade Federal do Rio Grande do Sul. Porto Alegre, RS, Brasil

^{II} Departamento de Saúde Pública. Centro de Ciências da Saúde. Universidade Federal de Santa Catarina. Florianópolis, SC, Brasil

Correspondence:

Roger Keller Celeste
Faculdade de Odontologia
Universidade Federal do Rio Grande do Sul
R. Ramiro Barcelos 2492, 3º andar
90035-003 Porto Alegre, RS, Brasil
E-mail: roger.keller@ufrgs.br

Received: 05/24/2012
Approved: 03/19/2013

INTRODUCTION

Income is commonly used in health research and has been operationalized in different ways by distinct researchers across studies. For instance, it may take into account the respondent's own wage or her/his family disposable income. Producing valid data on income poses many difficulties to interested researchers. There may be diverse sources of income and the respondent may not know other family members' earnings; income questions can be intrusive, making respondents skip them, and increasing the rate of non-response.³ Richer individuals may be tempted not to inform the interviewer of their actual incomes, underestimating the real value of their revenues.³

To make income data more valid, some researchers have developed and used income questions in a pre-coded format, with an upper open-ended category. Nonetheless, categorical measurement does not allow for the calculation of *per capita* income. To overcome such issues, all individuals within a category are usually assigned the mid-point income of their corresponding category. For the upper open-ended category, however, there is no mid-point. So, three common options to overcome this problem are: a) to adopt the lowest value of the upper category; b) to use an arbitrary income value for that category, or; c) to estimate a mid-point for the category, based on the Pareto Curve.⁴ The last option is believed to be better, as it is based on data, and not on an arbitrary value defined by the researcher.

Disposable household income, measured in many surveys, is a better option for grasping an individual's material wellbeing than his/her own income. However, there is the need to take into account the number of household members, a concept known as economies of scales. As the family size increases, the costs associated with an additional member are smaller than the costs of previous ones. Scales account for this issue using different formulas and the main impact described is on poverty levels.¹ To the authors' knowledge, the impact of different equivalence scales on the associations between income and other health-related variables has not yet been investigated.

In Brazil, the 2010 Brazilian Oral Health Survey (*Pesquisa Nacional de Saúde Bucal – SBBrazil 2010*) produced income data according to seven categories, with an upper open-ended category; the number of residents living on this income was also reported. Therefore, the aims of the present short-report are twofold: to estimate the upper mid-point of the open-ended income category, as well as to assess the impact of different equivalence scales on income-health associations.

METHODS

We used data from the SBBrazil 2010. Its sampling design has been reported by local official authorities. Data, collected between March and November 2010, on total monthly disposable household income and the number of people living in the household were obtained by means of a self-administered questionnaire.

To estimate the mid-point of the upper income category, we used formulas described by Parker & Fenwick⁴ (1983). These formulas are based on Pareto's law of income distribution, which states that the logarithm of the percentage of units with an income in excess of some value is a negatively sloped linear function of the logarithm of that value. According to this theorem, the mean income for the top-coded category = $X(v/v-1)$, and the median income for the top-coded category = $10^{(0.301/v)}*(X)$, where: X = lower value of the top-coded/open-ended category and $v = c-d/b-a$, and where: a = log of lower limit of interval preceding top-coded/open-ended category; b = log of lower limit of top-coded/open-ended category; c = log of the sum of the frequencies in the top-coded category and the category preceding it; d = log of the frequencies in the top-coded category.

We equalized household income utilizing two parametric equivalence scales.^{1,a} Such scales can be written generically as y/n^θ , where y is the total disposable household income, n is the number of members in the household and θ is the parameter for economies of scales (or elasticity). In the first scale $\theta = 1$, which corresponds to *per capita* family income, one of the most common used parameters. The second scale, $\theta = 0.5$, is also commonly used as it is also easy to implement. When $\theta = 0.5$, n^θ represents the square root of household members. Other common scales, the scale of the Organization for Economic Co-operation and Development and Canada's Market Basket Measure require data on the number of residents according to their age, as the costs of a child are lower than the costs of an adult. These latter two scales were not assessed in the present study, but should be considered, when data are available.

Outcome variables comprised two oral health variables: Decayed, Missing and Filled Teeth (DMFT) index and the Oral Impact on Daily Performances (OIDP) score. They were used both as continuous and dichotomized variables (OIDP > 0, DMFT > 0). When used as continuous variables, partial correlations with income variables were produced; when used as dichotomous variables, logistic regressions were fitted to produce odds ratios (OR) as a measure of their associations

^a Expert Group on Poverty Statistics, Rio Group. Compendium of best practices in poverty measurement. Rio de Janeiro; 2006.

with income variables. Sex and age (numerical discrete variable) were used as control variables in all analyses.

When income was used as continuous variables, we worked with log-income, as it showed a skewed distribution. The symmetry of log-income was checked after transformation and was deemed acceptable. We also dichotomized income (*per capita* and equivalized by squared root) at the poverty threshold (R\$255.00) and included them in logistic regression models and compared with the original categories as collected (not equivalized). All analyses were carried out in Stata v.11.2, using survey commands with the sampling weights and cluster structure provided within the data.

The SBBrasil 2010 was approved by the Brazilian National Council for Research Ethics under protocol number 15,498 on January 7th, 2010. Subjects' participation was voluntary and they signed the Free Prior and Informed Consent.

RESULTS

Data were obtained from all 35,929 individuals of all ages, who responded to the income question. Correlation coefficients and ORs were calculated for those older than 14 years of age ($n = 21,623$); all other analyses were carried out with the whole sample.

As seen in the Table, applying the formulas previously described, the estimated median mid-point income for the upper open-ended category was R\$14,522.50 and the mean mid-point was R\$24,507.10. The percentage of individuals below the Brazilian poverty line (1/2 of Minimum Wage, MW=R\$510.00 at the time of data collection) was 52.6% (95%CI 52.1;53.1), when using *per capita* income, and 14.6% (95%CI 14.5;15.2), when using equivalent income. The average *per capita* income

using the median and mean mid-point were, respectively, R\$ 475.20 (95%CI 429.40;521.00) and R\$ 507.40 (95%CI 446.20;568.60). The average equivalent income calculated with the median and mean mid-points were, respectively, R\$ 850.60 (95%CI 769.50;931.60) and R\$ 909.90 (95%CI 799.30;1020.60).

The partial Pearson's correlations showed no expressive differences in coefficients, as they only appeared in the third decimal place. The partial correlation between ODP and all four income variables was $r = -0.17$ ($p < 0.01$), and the partial correlation between DMFT and all four income variables was $r = -0.12$ ($p < 0.01$). The correlation between *per capita* and equivalent income was $r = 0.97$.

Results from logistic regression showed that dichotomizing income may slightly change the association with oral health outcomes. Being below R\$255.00 of *per capita* income was associated with having a DMFT > 0 OR = 1.77 (IC95% 1.31;2.39) and ODP > 0 OR = 1.64 (95%CI 1.40;1.92). Being below R\$255.00 equivalent income was associated with having a DMFT > 0 OR = 1.52 (95%CI 1.05;2.21) and ODP > 0 OR = 1.54 (95%CI 1.24;1.93). When using the categorical income question as it was collected, without applying equivalence scales, the OR of being below R\$250.00 was associated with having a DMFT > 0 OR = 1.74 (95%CI 0.76;3.99) and ODP > 0 OR = 2.49 (95%CI 1.60;3.86).

DISCUSSION

In the present study, we showed a way to derive income as a continuous variable from its categorical form adopting either the median or the mean income as the mid-point of the upper category. Furthermore, the use of equivalized

Table. Mid-point values, based on medians and means of each income category and the impact of using equivalized income on poverty level. SBBrasil, 2010.

Household income categories (R\$)	n	%	Mid-points (R\$)	
			Median	Mean
Up to 250	1,212	3.4	125	125
251 to 500	4,857	13.5	375	375
501 to 1,500	18,910	52.6	1,000	1,000
1,501 to 2,500	6,232	17.4	2,000	2,000
2,501 to 4,500	2,984	8.3	3,500	3,500
4,501 to 9,500	1,222	3.4	7,000	7,000
More than 9,500 ^a	512	1.4	14,523	24,507
Total	35,929	100.0		
<i>Per capita</i> income	Average		477.7	520.6
	individuals < R\$ 255.00	18,898	52.6	
Equivalized income	Average		866.5	942.7
	individuals < R\$ 255.00	5,353	14.9	

^a mid-points calculated based on Pareto's curve

income had a negligible effect on income-health associations, but a relevant impact on poverty levels.

The impact of equivalence scales on poverty level has been described as dramatic,¹ and was also expressive in the present analysis. Perhaps, contrary to common sense, larger families are not often as poor as supposed, because large families tend to have many earners that share assets, so economic well-being is much a matter of total disposable income. Considering such a large change in the magnitude of poverty level, it is important to use equivalized income, which is a theoretically more appropriate choice.

To our knowledge, no study has previously shown the impact of different equivalence scales on the income-health relationship. In our study, there was no difference between scales when income was a continuous variable, but there was a difference when income was

dichotomized. This was due to the fact that income-health relationship is curvilinear² and *per capita* income classified more individuals as “poor”.

Limitations of this study include that it is unknown whether the median or mean income of the upper category would be more appropriate. However, the impact of such estimates depends on the percentage of individuals in the upper category, which was only 1.4% in this study. For example, studies with fewer categories will have a larger upper category. A strength of the study is that we used a large representative population-based, dataset.

In conclusion, transforming income from categorical to continuous is a simple task that allows researchers to apply equivalence scales. Not taking this into account may overestimate poverty rates and, to some extent, affect income-health associations when the variable is dichotomized.

REFERENCES

1. Buhmann B, Rainwater L, Schmaus G, Smeeding TM. Equivalence scales, well-being, inequality, and poverty: sensitivity estimates across 10 countries using the Luxembourg Income Study (LIS) Database. *Rev Income Wealth*. 1988;34(2):115-42. DOI:10.1111/j.1475-4991.1988.tb00564.x
2. Celeste RK, Nadanovsky P. Income and oral health relationship in Brazil: is there a threshold? *Community Dent Oral Epidemiol*. 2009;37(4):285-93. DOI:10.1111/j.1600-0528.2009.00474.x
3. Davern M, Rodin H, Beebe TJ, Call KT. The effect of income question design in health surveys on family income, poverty and eligibility estimates. *Health Serv Res*. 2005;40(5 Pt 1):1534-52. DOI:10.1111/j.1475-6773.2005.00416.x
4. Parker RN, Fenwick R. The Pareto Curve and its utility for open-ended income distributions in survey research. *Soc Forces*. 1983;61(3):872-85.

The *Pesquisa Nacional de Saúde Bucal 2010* (SBBrazil 2010, Brazilian Oral Health Survey) was financed by the General Coordination of Oral Health/Brazilian Ministry of Health (COSAB/MS), through the *Centro Colaborador do Ministério da Saúde em Vigilância da Saúde Bucal, Faculdade de Saúde Pública at Universidade de São Paulo* (CECOL/USP), process no. 750398/2010.

This article underwent the peer review process adopted for any other manuscript submitted to this journal, with anonymity guaranteed for both authors and reviewers. Editors and reviewers declare that there are no conflicts of interest that could affect their judgment with respect to this article.

The authors declare that there are no conflicts of interest.

