

# Tap Water Consumption

Assessing the Role of Attitudes towards Environment using  
Non-Parametric Mokken Scale Analysis

Olivier Beaumais\* and Mouez Fodha\*\*

\*CARE EA-2260, GRE, University of Rouen, France and \*\*Paris School of Economics

June 2011

## Motivation (0/3)

- Willingness-to-Pay (WTP) surveys currently include questions about attitudes towards environment
- Most of the time, questions that measure attitudes have more than two ordered categories (Likert scales)

Question 23: How concerned are you about the following environmental issues?

Not concerned      Fairly concerned      Concerned      Very concerned      No opinion

Waste generation  
Air pollution  
Global warming  
Water pollution  
Natural resources depletion  
GMO  
Endangered species  
Noise

Question 91: How often do you do the following in your daily life?

Never      Occasionally      Often      Always      Not applicable

Turn off the water while brushing teeth  
Take showers instead of bath  
Plug the sink when washing the dishes  
Water your garden in the coolest part of the day  
Collect rainwater

## Motivation (1/3)

- A set of questions can be designed to assess a single latent trait (psychological trait, state of health, specific abilities)
- Various models (Item Response Theory) allow to construct a score that gives a unidimensional measure of the latent trait
- Some are parametric
  - The Rasch (1960) model (initially developed for dichotomous item - question), or one-parameter logistic model
  - Extended to polytomous items (see Embretson and Reise, 2000)
  - The partial credit model (Zheng and Rabe-Hesketh, 2007)
  - Latent-class models (Morey *et al.*, 2008)
- Other are non-parametric (Hardouin, 2005), especially the Mokken model (see van Schuur, 2003 for a straightforward presentation)
- For items that satisfy the criteria of the Mokken model, the sum of the responses across items can be used to rank respondents on the latent trait (Hardouin *et al.*, 2010)

## Motivation (2/3)

- The Mokken model requires few assumptions regarding the relationship between the latent trait and the responses to the items, and so, generally allows to keep more items. As a consequence, the precision of the individuals ordering is higher (Hardouin *et al.*, 2010)
- Empirical analysis: water quality and tap water consumption
- Most of the papers in the field (see Schram, 2009) assess the willingness-to-pay to improve water quality using the averting behavior method and/or the contingent valuation method
- The willingness-to-pay for water quality improvement is found to be related to variables such as the presence of young children in the household, education, gender, job, etc.
- Estimates do vary across the literature mainly because of the variability in specification of the scenario presented to respondents
- To our best knowledge the effect of the water quality perception of the respondent on the decision to consume or not tap water has rarely been assessed (see Jakus *et al.*, 2009)

## Motivation (3/3): what does this paper add?

- Few studies have been conducted using data that allow to compare the developed and the developing world
- We use survey data of about 10,000 households from 10 OECD countries, including Czech Republic, Korea and Mexico
- We investigate whether the water quality perception of people influences their choice to drink tap water, checking for potential endogeneity issues
- Using the Mokken scale analysis, we construct a concernment score and a water-saving score which are both tested within our econometric work
- The preliminary results show, not surprisingly, strong country-specific effects
- Further, attitudes towards environment and water saving behaviors are found to be significant along with other attitudinal characteristic such as trust in information from national or local governments

## Methodological considerations (1/4)

- Mokken scale analysis relies on three assumptions (van Schuur, 2003; Hardouin, 2010):
  - Unidimensionality: the responses to the items are explained by one and only one latent trait
  - Monotonicity: given the scale value of individual  $j$ ,  $\theta_j$ , the probability of a positive response to an item  $i$  increases with increasing value of  $\theta$ . If  $\theta_s < \theta_t$ ,  $p_i(\theta_s) < p_i(\theta_t)$
  - Local independence: conditionally to the latent trait, the responses are independent or, stated differently, responses of individual  $s$  to two or more items are influenced only by  $\theta_s$
- Actual responses are compared to theoretical responses consistent with a perfect Guttman scale under the model of stochastic independence
- Attitudinal data form a perfect Guttman scale when an individual who gives a positive answer (assuming dichotomous items for simplicity) to the more difficult item (question) will also give a positive answer to all the items (questions) that are easier

## Methodological considerations (2/4)

- Example: van Schuur, 2003, Political science

| Response type | V1 | V2 | V3 | V4 | V5 | V6 | Frequency of response type |
|---------------|----|----|----|----|----|----|----------------------------|
| 1             | 0  | 0  | 0  | 0  | 0  | 0  | 50                         |
| 2             | 0  | 0  | 0  | 0  | 0  | 1  | 3                          |
| 3             | 0  | 0  | 0  | 0  | 1  | 1  | 2                          |
| 4             | 0  | 0  | 0  | 1  | 1  | 1  | 35                         |
| 5             | 0  | 0  | 1  | 1  | 1  | 1  | 5                          |
| 6             | 0  | 1  | 1  | 1  | 1  | 1  | 3                          |
| 7             | 1  | 1  | 1  | 1  | 1  | 1  | 2                          |
| Total         | 2  | 5  | 10 | 45 | 47 | 50 | 100                        |

- V1, V2, V3 are three indicators of party political participation (V1, Run for office, V2, is active in a campaign, V3 goes to a party meeting). V1 is the more difficult question ( $p_{V1} = 0.02$ ), V4, V5, V6 are variables that tap voting behavior for three types of office (president, representative, and sheriff)
- This table forms a perfect Guttman scale: an individual who gives a positive answer to the more difficult question will also give a positive answer to all the easier questions

## Methodological considerations (3/4)

- However, empirical data sets show model (perfect Guttman scale) violations: an individual who gives a positive response to a difficult item can give a negative response to an easier item. Such a violation is a Guttman error
- Model violations are assessed through the comparison of actual Guttman errors ( $e_{jk}$ ) and theoretical Guttman errors ( $e_{jk}^0$ ) obtained under the assumption of independence between the responses to two items  $j$  and  $k$
- Loevinger H coefficients are defined as a function of the Guttman errors

- Loevinger H coefficient between two items:  $H_{jk} = 1 - \frac{e_{jk}}{e_{jk}^0}$ .  $H_{jk} = 1$

when there is no Guttman errors

- Loevinger H coefficient to measure the integration of one item to a scale  $S$ :  $H_j^S = 1 - \frac{\sum_{k \in S, k \neq j} e_{jk}}{\sum_{k \in S, k \neq j} e_{jk}^0}$ .  $H_j^S$  is near one if the item  $j$  is well integrated in the scale  $S$

- Loevinger H coefficient of scalability:  $H^S = 1 - \frac{\sum_{j \in S} \sum_{k \in S, k > j} e_{jk}}{\sum_{j \in S} \sum_{k \in S, k > j} e_{jk}^0}$



## Methodological considerations (4/4)

- Rule of thumb:  $H^S < 0.3$ , poor scalability properties,  $0.3 \leq H^S < 0.4$ , weak scale,  $0.4 \leq H^S < 0.5$ , medium scale,  $0.5 \leq H^S$ , strong scale
- Bottom-up algorithm (best smallest scale = pair of items with the highest  $H_{jk}$  coefficient), find the next best item in the scale and iterate on the basis of a user-specified boundary
- Fortunately Hardouin, 2010 has written a stata routine (MSP, Mokken Scale Procedure) which implements this algorithm in order to find whether a set of items (questions) forms a Mokken scale
- If a set of questions forms a Mokken scale, we can define a score by adding the value taken by each response included in the scale

## Data (1)

- The data come from an environmentally-related survey implemented in 10 OECD countries (Australia, Canada, Czech Republic, France, Italy, Korea, Mexico, Netherlands, Norway and Sweden) in 2008. Data are fully described in Beaumais-Briand-Millock-Nauges, 2009, and Millock and Nauges, 2010
- About 10,000 respondents have been surveyed using a web-based access panel, regarding a set of environmentally relevant activities including use of water and energy, recycling, transportation mode.
- Respondents were also asked a series of questions regarding characteristics of their household (age, income, composition, education, ownership status), housing characteristics, and behavioural attitudes or opinions regarding the environment in general
- Specific questions on water use (water saving behavior, investment in water-saving devices)

## Data (2)

- Attitudinal variables (environment):

Question 23: How concerned are you about the following environmental issues?

|                             | Not concerned | Fairly concerned | Concerned | Very concerned | No opinion |
|-----------------------------|---------------|------------------|-----------|----------------|------------|
| Waste generation            |               |                  |           |                |            |
| Air pollution               |               |                  |           |                |            |
| Global warming              |               |                  |           |                |            |
| Water pollution             |               |                  |           |                |            |
| Natural resources depletion |               |                  |           |                |            |
| GMO                         |               |                  |           |                |            |
| Endangered species          |               |                  |           |                |            |
| Noise                       |               |                  |           |                |            |

- Attitudinal variables (water saving behavior):

Question 91: How often do you do the following in your daily life?

|  | Never | Occasionally | Often | Always | Not applicable |
|--|-------|--------------|-------|--------|----------------|
| Turn off the water while brushing teeth          |       |              |       |        |                |
| Take showers instead of bath                     |       |              |       |        |                |
| Plug the sink when washing the dishes            |       |              |       |        |                |
| Water your garden in the coolest part of the day |       |              |       |        |                |
| Collect rainwater                                |       |              |       |        |                |

## Data (3)

Table: Respondents' opinion about tap water

| Country                                   | % satisfied<br>with tap water | % of dissatisfied<br>having<br>taste concern | % of dissatisfied<br>having<br>health concern | % drinking<br>tap water |
|---|-------------------------------|--|---|-------------------------|
| <b>High quality tap water countries</b>   |                               |  |   |                         |
| Netherlands                               | 95                            | 63   | 31  | 91                      |
| Sweden                                    | 92                            | 68   | 24  | 95                      |
| Norway                                    | 90                            | 67   | 29  | 97                      |
| <b>Medium quality tap water countries</b> |                               |  |   |                         |
| Czech Republic                            | 72                            | 52   | 39  | 75                      |
| Australia                                 | 71                            | 55   | 42  | 83                      |
| France                                    | 70                            | 59   | 37  | 63                      |
| Canada                                    | 67                            | 43   | 56  | 67                      |
| Italy                                     | 56                            | 33   | 61  | 39                      |
| <b>Low quality tap water countries</b>    |                               |  |   |                         |
| Korea                                     | 30                            | 11   | 86  | 39                      |
| Mexico                                    | 21                            | 5  | 92  | 19                      |

- Question 95a: Do you drink tap water for your normal household consumption?

## Mokken scale analysis

- Question 23 and question 91 are investigated in search of a Mokken scale
- The eight items of question 23 refer to an environmentally concerned attitude
- The five items of question 91 refer to a water-saving attitude
- All the eight items of question 23 are found to form a Mokken scale. Thus we create a score (`score_env`) by adding the value taken by each item of question 23
- Only items 4 and 5 of question 91 are found to form a Mokken scale. These two items are related to garden work (use of water tank, timing of garden watering). Thus we create a score (`score_wat`) by adding the value taken by items 4 and 5 of question 91

## Mokken scale analysis: concernment scale

Table: Concernment score

| Country                                   | Mean  | Std. Dev. | Min | Max |
|---|-------|-----------|-----|-----|
| <b>High quality tap water countries</b>   |       |           |     |     |
| Netherlands                               | 12.47 | 5.69      | 0   | 24  |
| Sweden                                    | 13.56 | 5.77      | 0   | 24  |
| Norway                                    | 13.84 | 5.14      | 0   | 24  |
| <b>Medium quality tap water countries</b> |       |           |     |     |
| Czech Republic                            | 15.22 | 5.12      | 0   | 24  |
| Australia                                 | 16.13 | 5.03      | 0   | 24  |
| France                                    | 16    | 5.21      | 0   | 24  |
| Canada                                    | 16.14 | 4.87      | 0   | 24  |
| Italy                                     | 17.34 | 4.20      | 0   | 24  |
| <b>Low quality tap water countries</b>    |       |           |     |     |
| Korea                                     | 18.06 | 4.44      | 0   | 24  |
| Mexico                                    | 20.09 | 3.07      | 4   | 24  |

## Mokken scale analysis: water-saving scale

Table: Water-saving score

| Country                                   | Mean | Std. Dev. | Min | Max |
|---|------|-----------|-----|-----|
| <b>High quality tap water countries</b>   |      |           |     |     |
| Netherlands                               | 1.92 | 1.94      | 0   | 6   |
| Sweden                                    | 1.56 | 1.99      | 0   | 6   |
| Norway                                    | 1.19 | 1.43      | 0   | 6   |
| <b>Medium quality tap water countries</b> |      |           |     |     |
| Czech Republic                            | 2.30 | 2.29      | 0   | 6   |
| Australia                                 | 3.39 | 1.99      | 0   | 6   |
| France                                    | 2.36 | 2.28      | 0   | 6   |
| Canada                                    | 2.10 | 1.90      | 0   | 6   |
| Italy                                     | 3.87 | 1.86      | 0   | 6   |
| <b>Low quality tap water countries</b>    |      |           |     |     |
| Korea                                     | 1.15 | 1.58      | 0   | 6   |
| Mexico                                    | 2.64 | 1.90      | 0   | 6   |

## Specification

- Two variables of interest: *tap* (1 drink tap water, 0 do not drink tap water) and *sat* (1 satisfied, 0 not satisfied)
- Question 95: Are you satisfied with the quality of your tap water for drinking?
- Cross table:

|       |       | Sat   |       |  |
|-------|-------|-------|-------|--|
| Tap   | 1     | 0     | Total |  |
| 0     | 878   | 2500  | 3,378 |  |
| 1     | 5,610 | 865   | 6,475 |  |
| Total | 6,488 | 3,365 | 9,853 |  |

Independence of *sat* and *tap* is clearly rejected ( $\chi^2(1) = 3600$ )

- We seek to estimate a probit model:

$$\Pr(\text{tap} = 1) = \Phi(\beta'x + \delta\text{sat} + \sum_{i=1}^9 \text{country}_i)$$

- Where  $x$  is a column vector of exogenous variables and  $\text{country}_i$  are binaries for each country (Australia is the reference)



## Econometrics issues

- Are the scores (score\_env and score\_wat) independent? **No**  
(Spearman test, clear rejection of the null hypothesis of independence)
- Endogeneity concerns: we must check whether *sat* is exogenous or not
- As *sat* is dichotomous variables, we must estimate a recursive bivariate probit model. See Greene (1998, 2008) and more recently, Wilde (2000), Monfardini and Radice (2008)
- Marginal effects are more involved than in a basic bivariate probit model (again, see Greene, 1998 or Kassouf and Hoffmann (2006))

## The model (1/2)

- The recursive bivariate probit model is as follows:

$$\Pr(\text{sat} = 1, \text{tap} = 1 \mid w, x) = BVN(\alpha'w, \beta'x + \delta, \rho)$$

$$\Pr(\text{sat} = 1, \text{tap} = 0 \mid w, x) = BVN(-\alpha'w, \beta'x, -\rho)$$

$$\Pr(\text{sat} = 0, \text{tap} = 1 \mid w, x) = BVN(\alpha'w, -\beta'x - \delta, -\rho)$$

$$\Pr(\text{sat} = 0, \text{tap} = 0 \mid w, x) = BVN(-\alpha'w, -\beta'x, \rho)$$

where  $BVN$  indicates the cumulative distribution function of the bivariate standard normal distribution with correlation  $\rho$ .  $w$  is a column vector of exogenous variables. Note that the same exogenous regressors can appear in both equation, without raising identification issues (Wilde, 2000)

- As Greene (1998) recalls, the model can be estimated by maximum likelihood as the usual bivariate probit model

## The model (2/2)

- The expected value for  $tap$  is given by:

$$\begin{aligned} E(tap | w, x) &= \Pr(sat = 1)E(tap | sat = 1, w, x) + \Pr(sat = 0)E(tap | sat = 0, w, x) \\ &= \Pr(sat = 1) \Pr(tap = 1 | sat = 1, w, x) + \Pr(sat = 0) \Pr(tap = 1 | sat = 0, w, x) \\ &= \Pr(tap = 1, sat = 1) + \Pr(tap = 1, sat = 0) \end{aligned}$$

- That is:

$$E(tap | w, x) = BVN(\alpha'w, \beta'x + \delta, \rho) + BVN(-\alpha'w, \beta'x, -\rho)$$

- Which allows to compute the marginal effects

## Marginal effects (1/3, Kassouf and Hoffmann, 2006)

- Marginal effect of being satisfied with the water quality on the probability to drink tap water:

$$\begin{aligned} Mef(sat) &= \Pr(tap = 1 \mid sat = 1, w, x) - \Pr(tap = 1 \mid sat = 0, w, x) \\ &= \frac{BVN(\alpha'w, \beta'x + \delta, \rho)}{\Phi(\alpha'w)} - \frac{BVN(-\alpha'w, \beta'x, -\rho)}{1 - \Phi(\alpha'w)} \end{aligned}$$

where  $\Phi$  is the cumulative distribution function of the standard univariate normal distribution

- When  $\rho = 0$ , the joint probability is the product of the marginals:

$$BVN(\alpha'w, \beta'x + \delta, \rho) = \Phi(\alpha'w)\Phi(\beta'x + \delta)$$

- Thus, in this case,  $Mef(sat) = \Phi(\beta'x + \delta) - \Phi(\beta'x)$

## Marginal effects (2/3, Kassouf and Hoffmann, 2006)

- Marginal effect of a binary variable  $y$  that belongs to  $x$  and/or  $w$  on the probability to drink tap water:

$$Mef(y) = BVN(\alpha' w_1, \beta' x_1 + \delta, \rho) + BVN(-\alpha' w_1, \beta' x_1, -\rho) \\ - BVN(\alpha' w_0, \beta' x_0 + \delta, \rho) - BVN(-\alpha' w_0, \beta' x_0, -\rho)$$

where  $w_1, x_1, w_0, x_0$  are the vectors  $w$  and  $x$  in which  $y$  equals 1 and 0 respectively.

- A convenient way to interpret this marginal effect is to split it in two parts:

$$Mef(y) = Mef_1(y) + Mef_0(y)$$

$$Mef_1(y) = BVN(\alpha' w_1, \beta' x_1 + \delta, \rho) - BVN(\alpha' w_0, \beta' x_0 + \delta, \rho)$$

$$Mef_0(y) = BVN(-\alpha' w_1, \beta' x_1, -\rho) - BVN(-\alpha' w_0, \beta' x_0, -\rho)$$

- $Mef_1(y)$  is the marginal effect on the probability to drink tap water for individuals being satisfied
- $Mef_0(y)$  is the marginal effect on the probability to drink tap water for individuals not being satisfied

## Marginal effects (3/3, Kassouf and Hoffmann, 2006)

- Marginal effect of a continuous variable  $z$  that belongs to  $x$  and/or  $w$  on the probability to drink tap water:

$$Mef(z) = \frac{\partial E(\text{tap}|w,x)}{\partial z} = Mef_1(z) + Mef_0(z)$$

- With:

$$Mef_1(z) =$$

$$\phi(\alpha'w)\Phi(\beta'x + \delta - \frac{\rho\alpha'w}{\sqrt{1-\rho^2}})\alpha_z + \phi(\beta'x + \delta)\Phi(\alpha'w - \frac{\rho(\beta'x + \delta)}{\sqrt{1-\rho^2}})\beta_z$$

$$Mef_0(z) = \phi(-\alpha'w)\Phi(\beta'x - \frac{\rho\alpha'w}{\sqrt{1-\rho^2}})\alpha_z + \phi(\beta'x)\Phi(-\alpha'w - \frac{\rho(\beta'x)}{\sqrt{1-\rho^2}})\beta_z$$

where  $\phi$  is the density function of the standard univariate normal distribution.

- $Mef_1(z)$  and  $Mef_0(z)$  receive the same interpretation as the previous ones. The expressions simplify when  $\rho = 0$ .

## Results (1/3): same exogenous regressors in both equations, Wilde (2000)

| Variable      | sat equation |                          | tap equation   |         |
|---------------|--------------|--------------------------|--|---------|
|               | Coefficient  | p-value                  | Coefficient  | p-value |
| income cat. 2 | -0.023       | 0.610                    | -0.025   | 0.605   |
| income cat. 3 | 0.035        | 0.431                    | 0.024  | 0.635   |
| income cat. 4 | 0.036        | 0.415                    | -0.157   | 0.002   |
| income cat. 5 | -0.007       | 0.941                    | -0.236   | 0.037   |
| score_wat     | 0.021        | 0.006                    | 0.003  | 0.717   |
| score_env     | -0.005       | 0.067                    | 0.010  | 0.006   |
| age           | -0.011       | 0.088                    | -0.010   | 0.171   |
| age*age       | 0.0001       | 0.024                    | 0.0001   | 0.209   |
| gender (male) | 0.112        | 0.243                    | 0.134  | 0.220   |
| age*gender    | 0.0009       | 0.649                    | -0.004   | 0.086   |
| sat           | -            | -                        | 1.226  | 0.022   |
| i_Canada      | -0.115       | 0.068                    | -0.554   | 0.000   |
| i_Netherlands | 1.108        | 0.000                    | -0.033   | 0.848   |
| i_France      | -0.064       | 0.312                    | -0.759   | 0.000   |
| i_Mexico      | -1.342       | 0.000                    | -1.621   | 0.000   |
| i_Italy       | -0.441       | 0.000                    | -1.354   | 0.000   |
| i_Czech       | -0.070       | 0.325                    | -0.383   | 0.000   |
| i_Sweden      | 0.891        | 0.000                    | 0.440  | 0.008   |
| i_Norway      | 0.754        | 0.000                    | 0.754  | 0.000   |
| i_Korea       | -1.036       | 0.000                    | -0.951   | 0.000   |
| intercept     | 0.660        | 0.000                    | 0.396  | 0.323   |
| N=9212        |              | Log likelihood=-8097.308 | LR test: $\rho = 0$ not rejected ( $chi^2(1) = 0.36$ ) |         |

## Results (2/3): instruments in sat equation, Monfardini and Radice (2008), score\_wat dropped

| Variable                  | sat equation             |         | tap equation  |         |
|---------------------------|--------------------------|---------|---|---------|
|                           | Coefficient              | p-value | Coefficient   | p-value |
| income cat. 2             | -0.022                   | 0.602   | -0.025  | 0.606   |
| income cat. 3             | 0.027                    | 0.547   | 0.025   | 0.618   |
| income cat. 4             | 0.019                    | 0.672   | -0.155  | 0.002   |
| income cat. 5             | -0.014                   | 0.885   | -0.236  | 0.037   |
| score_env                 | -0.004                   | 0.166   | 0.010   | 0.004   |
| age                       | -0.008                   | 0.243   | -0.010  | 0.164   |
| age*age                   | 0.0001                   | 0.131   | 0.0001  | 0.192   |
| gender (male)             | 0.086                    | 0.369   | 0.135   | 0.213   |
| age*gender                | 0.0014                   | 0.520   | -0.004  | 0.085   |
| sat                       | -                        | -       | 1.181   | 0.005   |
| prop (owner=1)            | 0.083                    | 0.017   | -   | -       |
| nbr of yrs in home cat. 2 | 0.011                    | 0.790   | -   | -       |
| nbr of yrs in home cat. 3 | 0.045                    | 0.327   | -   | -       |
| nbr of yrs in home cat. 4 | 0.142                    | 0.004   | -   | -       |
| trust in gov (no=1)       | -0.041                   | 0.000   | -   | -       |
| i_Canada                  | -0.155                   | 0.013   | -0.560  | 0.000   |
| i_Netherlands             | 1.058                    | 0.000   | -0.026  | 0.855   |
| i_France                  | -0.096                   | 0.124   | -0.762  | 0.000   |
| i_Mexico                  | -1.419                   | 0.000   | -1.641  | 0.000   |
| i_Italy                   | -0.492                   | 0.000   | -1.355  | 0.000   |
| i_Czech                   | -0.002                   | 0.973   | -0.385  | 0.000   |
| i_Sweden                  | 0.834                    | 0.000   | 0.444   | 0.001   |
| i_Norway                  | 0.663                    | 0.000   | 0.754   | 0.000   |
| i_Korea                   | -1.107                   | 0.000   | -0.974  | 0.000   |
| intercept                 | 0.759                    | 0.000   | 0.433   | 0.192   |
| N=9212                    | Log likelihood=-8084.137 |         | LR test: $\rho = 0$ not rejected ( $chi2(1) = 0.72$ ) |         |



## Tests

- No endogeneity issues in the bivariate probit (see Greene, 2008 for a straightforward explanation)
- This result is robust to the specification. Both equations are independent
- Income falls into five classes, class 1 [resp. class 5] gathering households with the lowest [resp. highest] income
- Class 1 is the reference. Are the income variables jointly significant? **Yes:**  $\chi^2 = 20.98[8 \text{ df}]$
- Are the age/gender variables jointly significant? **Yes:**  $\chi^2 = 46.29[8 \text{ df}]$
- Are "the number of years living in the primary residence" variables jointly significant? **Yes:**  $\chi^2 = 10.30[3 \text{ df}]$
- Is score\_env significant? **Yes:**  $\chi^2 = 10.82[2 \text{ df}]$
- What if we replace score\_env and score\_wat by naive mean scores (preoc\_env and habit\_wat)?

## Results (3/3): naive scores in both equations

| Variable                  | sat equation |                          | tap equation  |         |
|---------------------------|--------------|--------------------------|---|---------|
|                           | Coefficient  | p-value                  | Coefficient   | p-value |
| income cat. 2             | -0.024       | 0.573                    | -0.027  | 0.582   |
| income cat. 3             | 0.027        | 0.542                    | 0.026   | 0.604   |
| income cat. 4             | 0.023        | 0.602                    | -0.149  | 0.004   |
| income cat. 5             | -0.015       | 0.876                    | -0.239  | 0.035   |
| habit_wat                 | 0.074        | 0.002                    | 0.092   | 0.001   |
| preoc_env                 | -0.595       | 0.026                    | 0.055   | 0.073   |
| age                       | -0.008       | 0.206                    | -0.011  | 0.130   |
| age*age                   | 0.0001       | 0.120                    | 0.0001  | 0.171   |
| gender (male)             | 0.086        | 0.375                    | 0.132   | 0.223   |
| age*gender                | 0.0014       | 0.500                    | -0.004  | 0.095   |
| sat                       | -            | -                        | 1.160   | 0.006   |
| prop (owner=1)            | 0.083        | 0.021                    | -   | -       |
| nbr of yrs in home cat. 2 | 0.012        | 0.787                    | -   | -       |
| nbr of yrs in home cat. 3 | 0.045        | 0.328                    | -   | -       |
| nbr of yrs in home cat. 4 | 0.141        | 0.005                    | -   | -       |
| trust in gov (no=1)       | -0.041       | 0.000                    | -   | -       |
| i_Canada                  | -0.123       | 0.052                    | -0.520  | 0.000   |
| i_Netherlands             | 1.066        | 0.000                    | -0.003  | 0.982   |
| i_France                  | -0.082       | 0.192                    | -0.742  | 0.000   |
| i_Mexico                  | -1.384       | 0.000                    | -1.604  | 0.000   |
| i_Italy                   | -0.461       | 0.000                    | -1.318  | 0.000   |
| i_Czech                   | 0.033        | 0.652                    | -0.339  | 0.000   |
| i_Sweden                  | 0.867        | 0.000                    | 0.490   | 0.001   |
| i_Norway                  | 0.721        | 0.000                    | 0.835   | 0.000   |
| i_Korea                   | -1.040       | 0.000                    | -0.896  | 0.000   |
| intercept                 | 0.636        | 0.000                    | 0.156   | 0.641   |
| N=9212                    |              | Log likelihood=-8074.691 | LR test: $\rho = 0$ not rejected ( $chi2(1) = 0.81$ ) |         |

## Comments

- Whatever the specification, habit\_wat is strongly significant
- Again, no endogeneity issues
- Coefficients other than those of preoc\_env and habit\_wat are stable
- What tells us the Cronbach's  $\alpha$  about the scalability of question 23 and question 91?
- Question 23: Cronbach'  $\alpha = 0.86$  (good scalability, not surprising because all items are also included in the Mokken scale)
- Question 91: Cronbach'  $\alpha = 0.52$  (bad scalability, but the Mokken scale analysis identifies a scale and not the Cronbach analysis)
- A naive approach (computation of scores without checking the scalability) could lead to false conclusions (influence of water saving attitudes on the probability to drink or not tap water)
- The Mokken scale analysis appears to be more precise than the Cronbach analysis (on these data) and allows for testing richer specifications

## Marginal effects (1/2)

- The delta method is used to derive a 95 % confidence interval (restricted model,  $\rho = 0$ )
- $male = 1, age = 42, prop = 1, sat = 1$ , Australia is the reference, no trust in gov, first revenue category, reference number of years in the primary residence (less than two years).
- Joint probability,  $\Pr(tap = 1, sat = 1) = 0.73$  ; marginal probability  $\Pr(tap = 1) = 0.94$

## Marginal effects (2/2)

| Variable         | Marginal Effect on Pr(tap=1) | p-value | [95% Conf. Interval] |         |
|------------------|------------------------------|---------|----------------------|---------|
| Mef(sat)         | 0.40                         | 0.000   | 0.36                 | 0.45    |
| Mef1(score_env)  | 0.0009                       | 0.003   | 0.0003               | 0.001   |
| Mef0(score_env)  | 0.0009                       | 0.003   | 0.0003               | 0.001   |
| Mef1(age)        | 0.0005                       | 0.227   | -0.0003              | 0.001   |
| Mef0(age)        | -0.001                       | 0.001   | -0.001               | -0.0004 |
| Mef1(prop)       | 0.024                        | 0.017   | 0.004                | 0.044   |
| Mef0(prop)       | -0.013                       | 0.018   | -0.025               | -0.002  |
| Mef1(notrustgov) | -0.011                       | 0.000   | -0.017               | -0.005  |
| Mef0(notrustgov) | 0.006                        | 0.000   | 0.002                | 0.009   |
| Mef1(sweden)     | 0.177                        | 0.000   | 0.137                | 0.217   |
| Mef0(sweden)     | -0.077                       | 0.000   | -0.101               | -0.053  |
| Mef1(Korea)      | -0.463                       | 0.000   | -0.503               | -0.424  |
| Mef0(Korea)      | 0.028                        | 0.066   | -0.001               | 0.058   |
| Mef1(inc_cat2)   | -0.013                       | 0.012   | -0.023               | -0.002  |
| Mef0(inc_cat2)   | 0.004                        | 0.356   | -0.004               | 0.013   |
| Mef1(inc_cat3)   | -0.009                       | 0.073   | -0.019               | -0.0008 |
| Mef0(inc_cat3)   | 0.008                        | 0.076   | -0.0008              | 0.017   |
| Mef1(inc_cat4)   | -0.026                       | 0.000   | -0.038               | -0.015  |
| Mef0(inc_cat4)   | -0.008                       | 0.112   | -0.018               | 0.001   |
| Mef1(inc_cat5)   | -0.035                       | 0.012   | -0.063               | -0.007  |
| Mef0(inc_cat5)   | -0.014                       | 0.161   | -0.035               | 0.005   |

## Conclusions

- Estimations on the pooled data show strong country specific effects
- **The water quality perception do influence** the decision to drink or not tap water (strong effect of the *sat* variable)
- Constructing Mokken scales allows us to assess accurately the effect of attitudinal variables on the individual decision to drink or not tap water. Concerned individuals do consume more tap water
- More generally, attitudinal variables have a small but significant influence on the decision to drink tap water (at least for our respondent profile)