Catalogue no. 12-001-XIE

# Survey Methodology

June 2004


Statistics Canada Statistique Canada

Canada

## How to obtain more information

Specific inquiries about this product and related statistics or services should be directed to: Business Survey Methods Division, Statistics Canada, Ottawa, Ontario, K1A 0T6 (telephone: 1 800 263-1136).

For information on the wide range of data available from Statistics Canada, you can contact us by calling one of our toll-free numbers. You can also contact us by e-mail or by visiting our website.

| | |
|---|---|
| National inquiries line | 1 800 263-1136 |
| National telecommunications device for the hearing impaired | 1 800 363-7629 |
| Depository Services Program inquiries | 1 800 700-1033 |
| Fax line for Depository Services Program | 1 800 889-9734 |
| E-mail inquiries | infostats@statcan.ca |
| Website | www.statcan.ca |

## Information to access the product

This product, catalogue no. 12-001-XIE, is available for free. To obtain a single issue, visit our website at www.statcan.ca and select Our Products and Services.

## Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner and in the official language of their choice. To this end, the Agency has developed standards of service that its employees observe in serving its clients. To obtain a copy of these service standards, please contact Statistics Canada toll free at 1 800 263-1136. The service standards are also published on www.statcan.ca under About Statistics Canada > Providing services to Canadians.

Statistics Canada

Business Survey Methods Division

# Survey Methodology

June 2004

Published by authority of the Minister responsible for Statistics Canada



April 2006

Catalogue no. 12-001-XIE
ISSN 1492-0921

Frequency: semi-annual

Ottawa

Cette publication est disponible en français sur demande (nº 12-001-XIF au catalogue).

---

**Note of appreciation**

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued cooperation and goodwill.

Vol. 30, No. 1, pp. 115-124
Statistics Canada, Catalogue No. 12-001

# Contact and Cooperation in the Belgian Fertility and Family Survey

**Marc Callens and Christophe Croux** [1]

Abstract

Combining response data from the Belgian Fertility and Family Survey with individual level and municipality level data from the 1991 Census for both nonrespondents and respondents, multilevel logistic regression models for contact and cooperation propensity are estimated. The covariates introduced are a selection of indirect features, all out of the researchers' direct control. Contrary to previous research, Socio Economic Status is found to be positively related to cooperation. Another unexpected result is the absence of any considerable impact of ecological correlates such as urbanity.

Key Words: Nonresponse; Multilevel analysis; Fertility and Family Survey.

## 1. Introduction

The aim of this paper is to empirically assess the relative importance of correlates of contact and cooperation rates in the Belgian Fertility and Family Survey (FFS Belgium 1991).

The conceptual and theoretical nonresponse framework used in this paper has been proposed by Groves and Couper (G&C 1998). In their view, nonresponse arising from non-contact is directly influenced by survey design features such as the number and the timing of calls. Conditionally on these survey design features, other important features such as physical impediments of the housing units and accessible-at-home patterns of the would-be respondents, which are indirectly measured by various social environmental and socio-demographic attributes, also play an important role. The decision to cooperate or to refuse is primarily regarded as a direct function of a dynamic social communicative process between the interviewer and the interviewee. Survey design, main interviewer, sample person and social environment characteristics are considered to have only an indirect influence on cooperation rates.

We use both individual level and municipality level data from the 1991 Census data, matched to the fieldwork outcome variable for nonrespondents and respondents of the 1991 Belgian FFS. In this survey, individuals are the sampling units. It is a face-to-face survey with low non-contact (4%) and moderate refusal rates (22%). We consider our data to be hierarchically nested with sample units at the lower and municipalities at the higher level. Including covariates at both levels, multilevel logistic regression models for contact and cooperation propensity are estimated. The covariates are a selection of indirect features, all out of the researchers' direct control.

Some intriguing results are: (1) Socio Economic Status indicators like education are positively related to cooperation and (2) ecological factors including urbanicity are not correlated with nonresponse. This is in contrast with findings from previous US-based research.

## 2. A Theory for Contactability and Cooperation

The process of realising an interview consists of two major components: the process of contacting a sample person and dependent on contact, the process of cooperation with a survey request. An attractive multi-level theoretical framework for studying contactability and cooperation has been proposed by Groves and Couper (G&C 1998).

### 2.1 Contactability

Chronologically, the process of contacting a sample person comes first. Some sample persons are never contacted by interviewers and hence never make a decision about their survey cooperation. Relative to the process of cooperation, the process of contacting a sample person is quite simple.

G&C (1998) consider contactability to be a function of three factors: (1) whether there are any physical impediments that prevent interviewers to get in touch with the sample person, (2) when sample persons are at home and (3) when and how many times the interviewer tries to contact the sample person. The number and timing of calls

1. Marc Callens, Population and Family Study Centre (CBGS), Markiesstraat 1, B-1000 Brussels, Belgium & Dept. of Applied Economics, Katholieke Universiteit Leuven, Naamsestraat 69, B-3000 Leuven, Belgium. E-mail: Marc.Callens@cbgs.be; Christophe Croux, Dept. of Applied Economics, Katholieke Universiteit Leuven, Naamsestraat 69, B-3000 Leuven, Belgium. E-mail: Christophe.Croux@econ.kuleuven.ac.be.

by the interviewer and the accessible-at-home patterns of the sample persons are the proximate causes of contactability. The accessible-at-home patterns of the sample person are affected by the presence of physical impediments (*e.g.*, telephone presence), socio-demographic attributes (*e.g.*, commuting times) and social environmental attributes (*e.g.*, crime). Also survey design features such as the length of the data collection period and the interviewer workload might have an influence on contact rates.

### 2.2 Cooperation

The central question in the survey stage following contact is why sample persons do or do not cooperate with the interviewer request. In the Groves-Couper model to study cooperation, the proximate causes of the decision to cooperate or to refuse lie at the level of the householder and his or her interaction with the interviewer. Another component in the theoretical framework of G&C (1998) is the set of survey design features, such as: the agency of data collection, advance warning of the survey request, topic saliency, *etc.*

G&C (1998) consider also two factors that are out of the control of the survey designer: influences of the sample person and social environmental influences. These variables are not considered to be direct causal influences on cooperation, but indirect measures of what are essentially social psychological constructs. Important theoretical constructs in this respect are: opportunity costs, social exchange and social isolation.

#### 2.2.1 Opportunity Costs

The notion of opportunity costs implies that sample persons weigh the opportunity costs in agreeing to spend their time responding to a survey interview. An important ingredient in the opportunity costs theory is the amount of discretionary time for the sample person available to complete the survey. Those with less discretionary time are less likely to feel free to participate in a survey. Some indirect indicators for the amount of discretionary time are: the inverse of the number of adults in a household and (the amount) of labour force participation. Of course, there are also obligations away from employment tasks such as commitments to friends and relatives that also might raise the opportunity costs of a survey.

#### 2.2.2 Social Exchange

Social exchange theory considers the perceived value of equity of long-term associations between persons or between a person and societal institutions (Blau 1964). Central to all conceptualisations of social exchange is the notion that, unlike economic exchange, all social commodities are part of an intuitive bookkeeping system in which debts (*e.g.*, obligations) and credits (*e.g.*, expectations) are taken into account (G&C 1998). The social exchange perspective can be applied whenever there is an ongoing relationship between the survey organisation and the sample person (*e.g.*, government surveys).

Those receiving fewer services from the government may – in considering the cumulative effect of multiple government contacts – feel less need to cooperate. Since government services are disproportional across socio-economic strata, indicators of Socio-Economic Status (SES) should reflect exchange influences on survey participation. However, a major problem with social exchange theory is that two alternative hypotheses between SES and cooperation might be deduced from it (G&C 1998). First, one can argue that lower SES groups may have the greatest indebtedness to the government for the public assistance they may receive. Higher SES groups feel far less that they owe any sort of repayment. In this perspective, the relationship between socio-economic status and cooperation propensity is a negative one. Alternatively, a curvilinear relationship between SES and cooperation may be hypothesised. The lowest SES groups may believe that they are disadvantaged routinely compared to more fortunate people. The highest SES groups feel themselves repeatedly targeted in terms of time and money but receive little in return. In such a hypothesis, both the highest and the lowest SES feel relatively deprived in the relationship with large-scale social institutions and tend to refuse survey cooperation.

#### 2.2.3 Social Isolation

Closely related to the social exchange hypothesis is the social isolation hypothesis. Social isolates are out of touch with the mainstream culture of a society: they tend to behave in accordance with subcultural norms or in explicit rejection of those of the dominant culture. They are believed to be less likely to participate in a variety of social and political activities, including responding to surveys (Couper, Singer and Kulka 1997). In terms of SES, social isolation theory implies a positive relationship between SES and cooperation: lower SES groups are resentful of their dependence on the government, whereas higher SES groups have a greater sense of civic obligation. Such a positive relationship between SES and social isolation is opposite to the relationships predicted by social exchange theory.

Demographic indicators of social isolation are race, ethnicity, age and gender; with minorities, elderly and men in the role of the relatively isolated. Indicators of social isolation at the micro-level include whether the sample person lives in a single-person household, whether the sample person has any children, whether the sample person has moved recently and whether the sample person lives in a large multiunit structure.

#### 2.2.4 Urbanicity

At the community level contextual factors such as urbanicity, population density, crime rates and lack of social cohesion are hypothesised to influence survey cooperation. Residents of rural areas tend to cooperate at a higher level compared to residents in towns. However, it is not clear which mechanism is responsible for this urbanicity effect which might be explained in terms of greater population density, higher crime rates and higher social disorganisation that are associated with life in urban areas. Population density is hypothesised to reduce cooperation through the experience of crowding. Fear of crime may produce an unwillingness to provide information to strangers. Finally, urban life is associated with social disorganisation, characterised by weakened local kinship and friendship networks and reduced participation in local affairs.

## 3. Data and Method

### 3.1 Data

In this study we make use of both aggregated and micro-level data of the Belgian 1991 Census linked to the response status for respondents and nonrespondents from the Belgian Fertility and Family Survey (FFS-Belgium 1991) held shortly after the Census operations.

#### 3.1.1 The FFS Survey (1991)

The Fertility and Family Survey in Belgium was organised by the Population and Family Study Centre (CBGS), a Scientific Institute from the Flemish Government. This survey was carried out between April and October 1991, which is very close to the decennial census date: April 1 in the same year. The main focus of the FFS-project is on reproductive behaviour, to be seen however in the broader context of partnership and family history, and the interaction between employment and reproduction (Cliquet and Callens 1993; Callens 1995). The target population consists of men and women of Belgian nationality, born in the period 1951–1970 and with main residence in the Flemish Region of Belgium.

A two-stage cluster sampling design was used for men and women separately. In a first stage, municipalities were selected from various socio-economic strata (Vanneste 1989). In each selected municipality, individuals were selected at random. In this way 2,975 women and 1,989 men were selected to take part in the survey. A fieldwork method was used to compensate for non-response: stratified random substitution of nonrespondents of the target sample by persons selected from a reserve sample (Chapman 1983; Vehovar 1999).

The final sample size, *i.e.*, including the substitution operation, equals 4,776 persons (2,897 women and 1,879 men). In this study we make use of respondents and nonrespondent cases of both the initial target sample and the fieldwork substitution operation ( $N = 6{,}847$).

Among both men and women, the nonresponse can be ascribed in 7 out of 10 cases to a refusal to participate in the survey. In 2 out of 10 cases, nonresponse is due to the fact that the persons selected could not be contacted, and in 1 out of 10 cases, an interview was impossible because of sickness, language difficulties or some other reason.

#### 3.1.2 Matching 1991 Census Person-Level Data (1991)

Our primary source of information on both respondent and nonrespondent cases is provided by the 1991 Census.

In an effort to reconcile privacy concerns and scientific interests, we used a simple technique to make the matching of person-level Census data and survey data anonymous. We provided a dataset to the National Institute of Statistics (NIS) containing only the national identification number and the response status for each respondent and nonrespondent case. As a result of the matching operation by the NIS, we received a selection of the 1991 Census data enriched with only two survey variables: the response status variable and an indicator whether a sample person belongs to the base or substitute sample.

The 1991 Census individual level data we have at our disposal are: the individual form and the house unit form. The individual form contains information about: the place of residence, the nationality, the labour force activity status, the first marriage, the birth year of the children, education and professional activities. The house unit form includes information on the housing unit of the household such as: the type of housing unit, the number of housing units in the building, ownership, building period, the number of rooms and corresponding squared meters, the presence of a telephone and comfort indicators such as the number of bath rooms.

#### 3.1.3 Contactability and its Determinants

To study the process of contactability, we ideally need data on the outcomes of all successive attempts to contact sample persons. In this study however, we do not have such detailed information at our disposal: we only know the final outcome of each survey request. Therefore, we can only study the probability of ever making contact with the sample person (coded $1 =$ contact and coded $0 =$ noncontact) and not whether it was easy or difficult to make contact. Sample persons that are known not to reside (anymore) on the sample address we do consider contacted. At 241 out of 6,847 sample units (3.52%), all contact attempts failed.

The data we use are measured at two levels: the individual level ($n = 6{,}847$) and the municipality level ($n = 123$). At the sample person level, we consider three types of variables: physical impediments to contact sample persons, reasons for sample persons to be present in their homes and control variables.

As there are no direct interviewer observations of physical impediments available to us, we have to rely solely on indicators for physical impediments available in the Census data. Three variables are used: whether the housing unit is a single-family structure or not, whether the housing unit is large (more than 10 units) or not and whether the sample person has a telephone or not.

Determinants of at-home patterns in this study are: civil status (unmarried, married and divorced), age (20–24, 25–29, 30–34 and 35–39 years) and activity status (inactive *vs*. other). For women only, we also consider the number of children (0, 1, 2 and 3+). For those in the labour force we have also detailed information about: working part-time *vs*. working full-time, the number of weekly working hours (<21, 21-35, 36–42, >42 hours), employment status (employee *vs*. own-account), having a second job or not and working at home or not.

We also use two control variables: substitution (whether a sample person originates from the base target sample or from the substitution sample) and gender (whether a sample person comes from the female sample or from the male sample).

At the municipality level ($n = 123$), we use five variables: population density (persons per square km for the residence of the sample person), urban status (the cities of Antwerp and Gent *vs.* other municipalities), percentage multi-unit structures (in quartile format: <7.13, 7.13–15.14, 15.14–27 and >27), percentage homes owner-occupied (in quartile format: <64.5, 64.5–71, 71–77.7 and >77.7) and percentage persons of minority race (in quartile format: <0.90, 0.9–2.22, 2.22–5.29 and >5.29).

#### 3.1.4 Cooperation and its Determinants

We are interested in the probability of ever getting cooperation (coded $1 = \text{cooperation}$ and coded $0 = \text{non-cooperation}$) conditionally on contact; not whether it was easy or difficult to get cooperation from the sample person. For 1,399 out of 6,606 contacted sample persons (21.18%), all attempts to get cooperation failed.

Again, the data we use are measured at two levels: the individual level and the municipality level. At the sample person level, we have indicators for the opportunity costs hypothesis, the exchange hypothesis and the isolation hypothesis. Substitution is used as a control variable.

Indicators for the opportunity costs hypothesis are: activity status (inactive *vs*. other), working part-time *vs*. working full-time, the number of weekly working hours (<21, 21–35, 36–42, >42 hours) and employment status (employee *vs*. own-account).

Indicators for Socio-Economic Status in our study are: the surface of the living rooms (in squared meters: <65, 65–84, 85–104, 105–124 and >125), the number of bathrooms (0, 1 and 2+) and educational level (primary, secondary – first stage, secondary – second stage, high – nonuniversity and high – university level). Other exchange hypothesis indicators are: whether one receives a replacement income from the government or not and whether the house is owner-occupied or not.

Indicators for the social isolation hypothesis are: gender, civil status (unmarried, married and divorced), age (20–24, 25–29, 30–34 and 35–39 years), single-family structure of the housing unit and for women only: the number of children (0, 1, 2 and 3+) and the presence of children under the age of five years. Finally, substitution is included as a control variable.

At the municipality level, we use the same five variables as in section 3.1.3: urban status, population density, percentage multiunit structures, percentage owner-occupied and percentage persons of minority race.

### 3.2 Method of Analysis

#### 3.2.1 Bivariate $\chi^2$ – Test

In a first exploratory series of analyses of the correlates of contactability and cooperation, we calculate percentages for two-way contingency tables and include the results for the $\chi^2$ – test of independence against association. Such a $\chi^2$ – test, like any significance test, indicates the degree of evidence for the existence of an association, not the strength of an association. When at least one variable is ordinal, more powerful tests of independence than the $\chi^2$ – test such as the linear trend test do exist, but for reasons of simplicity of presentation, we do not use them in this paper.

#### 3.2.2 Multilevel Logistic Regression

In a second series of analyses, we use multilevel logistic regression to simultaneously estimate the impact of the various determinants (Snijders and Boskers 1999). We opt for the use of a multilevel method, because we regard our data as hierarchically nested with individuals at the lower level (level 1) and municipalities at the higher level (level 2).

Let $p_{ij}$ be the probability that an individual $i$ belonging to municipality $j$ is contacted (or cooperates). We will consider four different models for explaining this probability: the null random model, two versions of the random intercept model and the standard logistic regression model.

The empty or unconditional model does not take explanatory variables into account. We specify the model such that

logit transformed probabilities $p_{ij}$ have a normal distribution:

$$\text{logit}(p_{ij}) = 1/(1+\exp(p_{ij})) = \gamma_0 + \boldsymbol{u_{0j}}$$

where $\gamma_0$ is the population average and $\boldsymbol{u_{0j}}$ the random deviation from this average for group $j$. These deviations $\boldsymbol{u_{0j}}$ are assumed to be independent normally distributed random variables with mean zero and variance $\tau_0^2$.

When there are $r$ variables at the individual level that are potentially explicative for the observed outcomes, then they are incorporated as a linear function in the random intercept model:

$$\text{logit}(p_{ij}) = \gamma_0 + \sum_{h=1}^{r} \gamma_h x_{hij} + \boldsymbol{u_{0j}}$$

where $\gamma_1, ..., \gamma_r$ are the slope parameters measuring the effect of the explicative variables.

If we would drop the random effects $\boldsymbol{u_{0j}}$ then we obtain a standard logistic regression model:

$$\text{logit}(p_{ij}) = \gamma_0 + \sum_{h=1}^{r} \gamma_h x_{hij} .$$

By also including $s$ variables at the community level, we get an intercept model with both level–1 and level–2 covariates:

$$\text{logit}(p_{ij}) = \gamma_0 + \sum_{h=1}^{r} \gamma_h x_{hij} + \sum_{k=1}^{s} \gamma_k x_{kij} + \boldsymbol{u_{0j}} .$$

We use SAS Proc Nlmixed (SAS Institute 1999) to actually estimate the parameters. In SAS Proc Nlmixed an adaptive version of Gauss-Hermite Quadrature (numerical integration) is used to solve the maximum likelihood estimation problem. To test if a specific parameter equals zero, a Likelihood Ratio $\chi^2$ – test is used.

## 4. Results

### 4.1 Contactability

Table 1 presents the bivariate results by the $\chi^2$ – test of the percentage never contacted by various indicators of physical impediments. One strong correlate is whether the housing unit is a single-family structure or not, the latter having much higher noncontact rates (8.1%) than other units (2.4%). Also, sample persons living in large multiunit housing structures tend to have higher noncontact rates (11%) than those not living in large multiunit housing structures (3.1%). Another strong correlate is the presence of a telephone: 9.7% of those with no telephone were never contacted.

**Table 1**
Percentage Never-Contacted by 'Physical Impediments' Attributes

| Physical impediments attributes | Percentage never contacted | $\chi^2$ | *df* | *p* |
|---|---|---|---|---|
| Single-Family Structure | | 97.6 | 1 | <0.0001 |
| No | 8.1 | | | |
| Yes | 2.4 | | | |
| Large multi-unit structure (>10) | | 38.4 | 1 | <0.0001 |
| No | 3.1 | | | |
| Yes | 11.0 | | | |
| Telephone | | 88.9 | 1 | <0.0001 |
| No | 9.7 | | | |
| Yes | 2.7 | | | |

Table 2 shows the bivariate results for contactability by 'reasons to be present at home' attributes. Relatively more unmarried (4.4%) and divorced (6.9%) sample persons than married (2.9%) sample persons are never contacted. There are much lower rates of noncontacts among those that are inactive (0.9%) compared to other persons (3.5%). Having at least 3 or more children (0.9%) leads to low noncontact rates, compared to having two children (2.6%) or at most 1 child (4%). Those working at home (1.5%) and those being an independent worker (1.9%) show modestly lower noncontact rates than those working elsewhere (3.6%) or those working as an employee respectively (3.6%). Age, the number of weekly working hours, working part-time *vs.* full-time and having a second job or not have no significant influence on contactability.

**Table 2**
Percentage Never-Contacted by 'Reasons to be Present at Home' Attributes

| Reasons to be present at home | Percentage never contacted | $\chi^2$ | *df* | *p* |
|---|---|---|---|---|
| Civil status | | 19.4 | 2 | <0.0001 |
| Unmarried | 4.4 | | | |
| Married | 2.9 | | | |
| Divorced | 6.9 | | | |
| Inactive *vs.* other | | 4.0 | 1 | 0.04 |
| Inactive | 0.9 | | | |
| Other | 3.5 | | | |
| Number of children[a] | | 14.5 | 3 | 0.0023 |
| 0 | 4.3 | | | |
| 1 | 4.0 | | | |
| 2 | 2.6 | | | |
| 3+ | 0.9 | | | |
| Employment place[b] | | 4.6 | 1 | 0.03 |
| At home | 1.5 | | | |
| Elsewhere | 3.6 | | | |
| Employment status[b] | | 4.0 | 1 | 0.05 |
| Employee | 3.6 | | | |
| Own-account | 1.9 | | | |

[a] subsample of women only ( $n$ = 4,098)
[b] subsample of active persons only ( $n$ = 5,368)

In addition, substitution is associated with higher noncontact rates (5.9%) compared to the base sample (2.6%). No significant difference has been found for the male and the female subsample.

In a multiple logistic regression model of the combined effects of those individual-level indicators that have some marginal bivariate effect on contactability only single-family structure ($\chi^2 = 35.75$, $p = <0.0001$), telephone ($\chi^2 = 52.63$, $p = <0.0001$) and substitution ($\chi^2 = 28.59$, $p = <0.0001$) remain significant.

In Table 3, noncontact rates for various environmental attributes are presented. Cities (6.6%) have higher non-contact rates compared to nonurban areas (3.1%). The percentage never contacted is higher for high-density areas (5.4%) than low-density areas (1.7%). The presence of multiunit structures and the presence of persons of other nationalities tend to increase non-contact rates. Finally, the percentage of owner-occupied houses shows a negative association with noncontact rates.

**Table 3**
Percentage Never-Contacted by 'Environmental' Attributes

| Environmental attribute | Percentage never contacted | $\chi^2$ | *df* | *p* |
|---|---|---|---|---|
| Urban status | | 24.0 | 1 | <0.0001 |
| Cities | 6.6 | | | |
| Other | 3.1 | | | |
| Population density | | 34.4 | 3 | <0.0001 |
| Lowest quartile | 1.7 | | | |
| Second quartile | 3.2 | | | |
| Third quartile | 3.8 | | | |
| Highest quartile | 5.4 | | | |
| % Multi-unit structures | | 50.4 | 3 | <0.0001 |
| Lowest quartile | 2.0 | | | |
| Second quartile | 2.2 | | | |
| Third quartile | 4.0 | | | |
| Highest quartile | 5.9 | | | |
| % Persons of other nationalities | | 23.1 | 3 | <0.0001 |
| Lowest quartile | 2.5 | | | |
| Second quartile | 2.3 | | | |
| Third quartile | 4.3 | | | |
| Highest quartile | 4.8 | | | |
| % Homes owner-occupied | | 64.4 | 3 | <0.0001 |
| Lowest quartile | 6.4 | | | |
| Second quartile | 3.6 | | | |
| Third quartile | 1.6 | | | |
| Highest quartile | 2.7 | | | |

We complement now the bivariate analysis with a multivariate analysis. In Table 4 four models for modelling contact relative to noncontact are presented. Model 1 is the null random model at the municipality level. Model 2 is a multiple logistic regression model. In this model, we have included the person-level effects that remained significant in a multivariate context (*i.e.*, single-family structure, telephone and substitution) and the variable activity status because of its theoretical importance. Model 3 is a random intercept version of model 2. In Model 4, we have extended Model 3 with the municipality level variable multi-units structures (in %) only.

**Table 4**
Results of (Multilevel) Logistic Regression Models of Contactability

| Results | Model 1: Null Random | Model 2: Logistic Regression | Model 3: Random Intercept Level 1 | Model 4: Random Intercept Level 1&2 |
|---|---|---|---|---|
| *Intercept* | 4.01*** | 3.08*** | 3.68*** | 4.15*** |
| | (0.16) | (0.73) | (0.77) | (0.79) |
| *Individual Characteristics* | | | | |
| Single-family structure | | 1.16*** | 1.02*** | 0.92*** |
| | | (0.15) | (0.17) | (0.17) |
| Telephone | | 1.19*** | 1.25*** | 1.26*** |
| | | (0.16) | (0.17) | (0.17) |
| Inactive *vs.* other | | −1.23 | −1.34 | −1.33 |
| | | (0.72) | (0.75) | (0.74) |
| Substitution sample | | −0.78*** | −0.64*** | −0.62*** |
| | | (0.14) | (0.15) | (0.15) |
| *Municipality Characteristics* | | | | −0.02* |
| Multi-unit structures (%) | | | | (0.01) |
| *Estimated variances* | | | | |
| Var(Intercept) | 1.03 | | 0.82 | 0.79 |
| *Goodness of fit* | | | | |
| Deviance | 1,720 | 1,658 | 1,606 | 1,599 |

Notes: Standard errors in parentheses. *$p$<0.05, ** $p$<0.01, ***$p$<0.001, one-tailed tests.

The effects of the person-level covariates in Models 2, 3 and 4 are in accordance with the findings of the bivariate analysis. Single-family structure and the presence of a telephone have a positive influence on contactability, while the effect of activity status is not significant. The impact of field substitution is negative. We also notice a (rather small) reduction of the regression coefficient for single-family structure and substitution in the multilevel models 3 and 4. Models 3 and 4 have one variance component for the intercept. To test the null hypothesis that the random intercept variance equals zero, we use the Likelihood Ratio test and compare the conventional logit model (Model 2) with the random intercept model (Model 3). The difference in deviance between both models is large (52). So, there might be some variance in the intercept to explain by municipality level covariates. By introducing municipality characteristics one at a time, we can test for significant effects by calculating deviance differences between Model

4 and Model 3. The only deviance difference of importance noted is the case of the variable 'multi-unit structures' (7 units difference). No differences in deviances are found for the introduction of the other level-two variables (urban status, percentage owner occupied, population density and persons of other nationalities).

We consider Model 3 and Model 4 as the better models. According to these multilevel models, noncontact rates vary considerably across municipalities. However, the municipality level covariates in our study are not able to explain much of this variation.

## 4.2 Cooperation

In Table 5, we present the bivariate results for the opportunity costs hypothesis indicators. Being inactive or not does not seem to have an effect on the cooperation rate. However, when we use indicators of discretionary time, such as working part-time versus working full-time or the weekly number of working hours, the predicted negative relationship does show up in the bivariate results. In addition, self-employed sample persons have lower cooperation rates compared to employees.

**Table 5**
Percentage Cooperation by 'Opportunity Cost Hypothesis' Indicators

| Opportunity cost indicators | Percentage cooperated | $\chi^2$ | *df* | *p* |
|---|---|---|---|---|
| Inactive *vs*. other | | 0.41 | 1 | 0.52 |
| Inactive | 77.0 | | | |
| Other | 78.9 | | | |
| Part-time *vs*. Full-time[b] | | 10.04 | 1 | 0.001 |
| Part-time | 82.3 | | | |
| Full-time | 77.4 | | | |
| Number of working hours[b] | | 15.3 | 3 | 0.0016 |
| <20 | 80.1 | | | |
| 21 – 35 | 84.7 | | | |
| 36 – 42 | 77.6 | | | |
| > 43 | 75.7 | | | |
| Employment status[b] | | 4.2 | 1 | 0.04 |
| Employee | 78.7 | | | |
| Own-account | 74.6 | | | |

[b] subsample of active persons only ( $n = 5{,}180$)

The predictions of the exchange hypothesis theory do not show up in the bivariate results presented in Table 6. SES indicators like the surface of the living room and the number of bathrooms are not negatively, but positively related to cooperation. Of course, these measures are not ideal, because we are not able to control for household size. Another indication of a positive relationship between cooperation and SES is the case of educational level. Whether one receives a replacement income or not and whether the house is owner-occupied or not has no impact on cooperation rates.

In a multiple logistic regression model of the combined effects of those social exchange indicators that have some marginal bivariate effect on cooperation, only the effects of educational level ($\chi^2 = 39.35$, $df = 4$, $p < 0.0001$) and surface of the living room ($\chi^2 = 13.4$, $df = 4$, $p = 0.0095$) remain significant.

**Table 6**
Percentage Cooperation by 'Exchange Hypothesis' Indicators

| Exchange indicators | Percentage cooperated | $\chi^2$ | *df* | *p* |
|---|---|---|---|---|
| Surface living rooms ($m^2$) | | 26.8 | 4 | <0.0001 |
| < 65 | 74.8 | | | |
| 65 – 84 | 77.6 | | | |
| 85 – 104 | 78.6 | | | |
| 105 – 124 | 79.9 | | | |
| > 125 | 83.1 | | | |
| Number of bathrooms | | 7.9 | 2 | 0.02 |
| 0 | 74.2 | | | |
| 1 | 78.6 | | | |
| 2 | 83.5 | | | |
| Educational level | | 46.7 | 4 | <0.0001 |
| Primary | 76.6 | | | |
| Secondary, first stage | 74.5 | | | |
| Secondary, second stage | 78.7 | | | |
| High, non-university | 85.1 | | | |
| High, university | 82.2 | | | |
| Replacement income | | 0.3 | 1 | 0.58 |
| No | 78.7 | | | |
| Yes | 79.5 | | | |
| Owner occupied | | 3.4 | 1 | 0.06 |
| No | 77.4 | | | |
| Yes | 79.4 | | | |

In the section for the exchange hypotheses, we have found support for the notion that those with low SES, cooperate less with surveys than those in the high SES groups. Such a positive relationship between SES is predicted by the social isolation hypothesis. Demographic indicators of social isolation theory are gender, civil status and age (See Table 7). No effects are found for gender, civil status (however, divorced sample persons are probably less cooperative) and single-family structure. Age seems to have a negative effect on cooperation. For women only, we have also data on the presence of children. We find that the number of children has a positive effect on cooperation rates. The age of the children is also important: the presence of young children is associated with higher cooperation.

The control variable substitution has a slightly negative effect on cooperation ($\chi^2 = 4.24$, $p = 0.039$) with lower

cooperation rates for the substitution sample (77.3%), compared to the base sample (79.5%).

**Table 7**
Percentage Cooperation by 'Social isolation Hypothesis' Indicators

| Social isolation indicators | Percentage cooperated | $\chi^2$ | *df* | *p* |
|---|---|---|---|---|
| Gender | | 1.56 | 1 | 0.21 |
| Male | 78.1 | | | |
| Female | 79.3 | | | |
| Civil status | | 3.11 | 2 | 0.21 |
| Unmarried | 79.8 | | | |
| Married | 78.6 | | | |
| Divorced | 75.4 | | | |
| Single-family structure | | 0.76 | 1 | 0.38 |
| No | 78.9 | | | |
| Yes | 77.7 | | | |
| Age | | 17.5 | 3 | 0.0006 |
| 20–24 | 80.8 | | | |
| 25–29 | 80.7 | | | |
| 30–34 | 78.3 | | | |
| 35–39 | 75.5 | | | |
| Number of children[a] | | 18.2 | 3 | 0.0004 |
| 0 | 77.9 | | | |
| 1 | 76.3 | | | |
| 2 | 81.7 | | | |
| 3+ | 84.9 | | | |
| Presence of young children[a] | | 12.3 | 1 | 0.0005 |
| No | 77.8 | | | |
| Yes | 82.8 | | | |

[a] subsample of women only ( $n = 3{,}955$)

Table 8 contains the bivariate results for social environmental differences in cooperation. Population density has a curvilinear effect on cooperation. Being a resident in a large metropolitan area has no effect. Thus, the evidence for the literature that crowding and high levels of stimulus input are negatively associated with cooperation is of a mixed nature.

The effect of indicators for social cohesion is not clear. Only the variable percentage owner-occupied has a (curvilinear) effect. The variables percentage persons of other nationalities and percentage multi-unit structures seem to have no effect.

Finally, we present in Table 9 a series of regression models for cooperation similar to those in section 4.1. In these models, we have included four individual level covariates: surface of the living room (<84, >84 m²), education (up to secondary-second stage *vs*. high level), age (20–29, 30–39 years) and substitution sample. Surface of the living room and education have been selected as the only significant exchange hypothesis indicators in the previously described multiple logistic regression model. Age was the only significant effect in the bivariate analysis on the social isolation hypothesis. Finally, substitution is introduced to control for possible fieldwork effects. The slightly negative effect of substitution in Model 2 might indicate that fieldwork substitution negatively influences cooperation. However, this effect disappears completely when a random intercept is introduced (Models 3 and 4). The effects of the other individual level covariates are in accordance with the findings of the bivariate analysis and do not change across Models 2 to 4. SES indicators like education and surface of the living room have a positive effect and age has a negative effect on cooperation. These effects rather confirm the social isolation hypothesis than the exchange hypothesis.

**Table 8**
Percentage Cooperation by 'Environmental' Attributes

| Environmental attribute | Percentage cooperated | $\chi^2$ | *df* | *p* |
|---|---|---|---|---|
| Urban status | | 0.84 | 1 | 0.36 |
| Cities | 80.1 | | | |
| Other | 78.7 | | | |
| Population density | | 10.7 | 3 | 0.014 |
| Lowest quartile | 80.0 | | | |
| Second quartile | 79.9 | | | |
| Third quartile | 76.0 | | | |
| Highest quartile | 79.4 | | | |
| % Multiunit structures | | 3.1 | 3 | 0.38 |
| Lowest quartile | 80.1 | | | |
| Second quartile | 79.2 | | | |
| Third quartile | 77.9 | | | |
| Highest quartile | 78.1 | | | |
| % Homes owner-occupied | | 12.3 | 3 | 0.0063 |
| Lowest quartile | 79.7 | | | |
| Second quartile | 76.2 | | | |
| Third quartile | 78.5 | | | |
| Highest quartile | 80.9 | | | |
| % Persons of other nationalities | | 5.2 | 3 | 0.16 |
| Lowest quartile | 77.7 | | | |
| Second quartile | 77.6 | | | |
| Third quartile | 79.6 | | | |
| Highest quartile | 80.2 | | | |

The only level two variable of (modest) importance is multi-unit structures (in %) and has been kept in Model 4. The Likelihood Ratio test for introducing this variable gives a difference of two units in deviance terms. The introduction of one or more other second level variables gives Likelihood Ratio tests differences close to zero in deviance terms. We consider Model 3 and 4 as the most suitable models. The difference in deviance terms between model 3 and model 2 is 8 units, which is significant. The variance for the intercept term is moderate (0.21). The introduction of second level covariates (including multi-unit structures)

leaves this variance term practically unchanged. Therefore, we may state that environmental attributes like urbanicity are not important for explaining cooperation.

**Table 9**
Results of (Multilevel) Logistic Regression Models of Cooperation

| Results | Model 1: Null Random | Model 2: Logistic Regression | Model 3: Random Intercept Level 1 | Model 4: Random Intercepts Level 1&2 |
|---|---|---|---|---|
| *Intercept* | 1.41*** | 1.24*** | 1.30*** | 1.39*** |
| | (0.06) | (0.06) | (0.08) | (0.10) |
| *Individual Characteristics* | | | | |
| Substitution sample | | −0.15* | −0.03 | −0.02 |
| | | (0.07) | (0.07) | (0.07) |
| Surface living rooms | | 0.23*** | 0.24*** | 0.24*** |
| | | (0.06) | (0.06) | (0.06) |
| Educational level | | 0.45*** | 0.47*** | 0.47*** |
| | | (0.08) | (0.08) | (0.08) |
| Age | | −0.23*** | −0.23*** | −0.23*** |
| | | (0.06) | (0.06) | (0.06) |
| *Municipality Characteristics* | | | | -0.006 |
| Multi-unit structures (%) | | | | (0.004) |
| *Estimated variances* | | | | |
| Var(Intercept) | 0.21 | | 0.21 | 0.21 |
| *Goodness of fit* | | | | |
| Deviance | 6,664 | 6,664 | 6,596 | 6,594 |

Notes: Standard errors in parentheses. * $p<0.05$, ** $p<0.01$, *** $p<0.001$, one-tailed tests.

## 5. Discussion

In this paper, we have used 1991 individual and municipality level Census data matched to the response status variable of the Belgian Fertility and Family Survey to analyse the relative importance of correlates of contact and cooperation.

We have organised our analysis according to the Groves-Couper conceptual framework. In the bivariate analysis stage, we have found essentially the same kind of correlates as was predicted and actually found in an US-based multi-survey analysis (G&C 1998). One important difference between the present study and the US-results seems to be the nature of the effect of SES indicators (*e.g.*, education) on cooperation. In the present study, we find a positive relationship; in the US-study the inverse relationship is found. We can imagine two alternative explanations for these conflicting findings. A first one is based on survey design effects such as topic saliency. The FFS-survey in Belgium might be atypical in being disproportionally attractive to the higher educated because of the specific content of the survey. Replicating the present analysis for surveys about varying topics can easily test such a hypothesis. Another possible hypothesis is that effects of education on survey cooperation do vary across societies. Then the challenge is to find out why this relationship varies across countries. Such a hypothesis is far less easy to test in real, as data for several countries are needed.

In the multilevel logistic regression analysis stage, the impact of all but one contextual factor completely vanished. Only the impact of the variable percentage of multi-unit structures shows, however only weakly, some resistance against ecological randomness present in the random intercept models. To us, this is a very intriguing result. Random ecological variation at the municipality level seems to dominate largely even the urban-rural dichotomy. A possible explanation is that the variation at the community level is dominated by interviewer effects, not by ecological factors.

## Acknowledgements

This paper grew out of our activities in the design and organisation of surveys at the Population and Family Study Centre (CBGS), a scientific Institute of the Flemish Government in Belgium. We are grateful to the Belgian National Statistical Institute, the local authorities and the National Register for their cooperation. We also thank the Editor and two anonymous referees for their valuable comments and suggestions. Finally, this research was supported by a FWO grant Bijzondere doctoraatsbeurs 2002–2003.

## References

Blau, P.M. (1964). *Exchange and Power in Social Life*. New York: John Wiley & Sons, Inc.

Callens, M. (1995). *De 'Fertility and Family Survey' in Vlaanderen* (*Nego V*, *1991*), *De gegevensverzameling*. Brussel: CBGS-document. 1995, 4.

Chapman, D.W. (1983). The impact of substitution on survey estimates. In *Incomplete Data in Sample Surveys*, *Proceedings of the Symposium*, (Eds. W.G. Madow and I. Olkin). New York: Academic. 2, 45-61.

Cliquet, R.L., and Callens, M. (Ed.) (1993). *Gezinsvorming in Vlaanderen. Hoe en Wanneer*? Brussel: CBGS Monografie 1.

Couper, M.P., Singer, E. and Kulka, R.A. (1997). Participation in the decennial census: Politics, privacy, pressures. *American Politics Quarterly*, 26, 59-80.

Groves, R.M., and Couper, M. (1998). *Nonresponse in Household Interview Surveys*. New York: John Wiley & Sons, Inc.

SAS Institute Inc. (1999). *SAS/STAT® User's Guide*. Version 8, Cary, NC: SAS Institute Inc.

Snijders, T.A.B., and Boskers, R.J. (1999). *Multilevel Analysis*, *an Introduction to Basic and Advanced Multilevel Modeling*. London: Sage Publications.

Vanneste, D. (1989). *Economische Typering Van de Belgische Gemeenten*. Leuven: Leuvense Geografische Papers.

Vehovar, V. (1999). Field substitution and unit nonresponse, *Journal of Official Statistics*, 15, 2, 335-350.