Contextualizing smoking: the influence of household factors on smoking habits

Céline Goffette

EQUALSOC Summer School, 2008

Abstract

Objectives Traditional approaches studying smoking focused primarily on individuals. Nowadays there is a shift from individual-level models towards models that incorporate contextual effects. This research project takes place in this trend. The purpose of this specific analysis is to investigate the influence of household on smoking practices, in order to determine if the household is a relevant unit to study smoking. Is there evidence for an household effect on smoking? In other words, do household factors (both observable and unobservable) affect individual probabilities of smoking, all other individual characteristics being equal? Is the household effect sensitive to the national context?

Materials and methods Data from the European Community Household Panel (ECHP) were used. All individuals aged more than 16 years are interviewed within the selected households. The probability to be a daily smoker is explained by variables at the individual level and at the household level. We take into account a possible specific effect of the household by implementing a random effects probit model.

Results This study brings to light a phenomenon of clustering of smoking habits within households, which is modulated by the composition of the household. Preliminary results seem to show that the longer the sitory of tobacco consumption in the country, the more sensitive the concordance.

Conclusions This research gives evidence for clustering of smoking practices in household, the effect being different according to the context of the household and the country. Research about health concordance overwhelmingly suggests evidence for clustering of health status and health behavior. This offers room for a deeper understanding for the causes of health concordance. Next step would be to determine the reason for concordance (is it due to *ex-ante* correlation or *ex-post* convergence?).

Keywords: smoking, health concordance, household influence.

1 Introduction

1.1 General orientation

The starting points of this study were an observation and an hypothesis. The observation is that individual dimensions of smoking have already been abundantly described, but the smoker in these studies is isolated, extracted from its social environment. However it seems that the question of interactions, norms transmission and social control within the family is an important question for what concerns tobacco consumption. Therefore, the aim of this analysis is to study the social context of smoking. The social context can be studied at different levels.

- micro: individual characteristics (age, sex, social position)
- meso: close environment (family, peer groups, school, occupational environment, neighborhood)
- macro: national context (public policies, messages of the media, social acceptability of smoking)

This work focuses primarly on the meso level.

1.2 Background

The individual dimensions of smoking have already been abundantly described. Two of the most important results found in this area concern the existence of a strong sex differential that has been gradually narrowing and of a social gradient that has been progressively inverted. A theoretical model explains both sex and social differences in smoking and temporal changes in these differences: this is the diffusion model (Pampel, [8] and [7]). This model sates that there is a general tendency of the cigarette epidemic to begin with high status groups, diffuse to lower status groups and then recede among high status groups. It also states that there is a lag in the process of cigarette adoption for women. Due to a variety of historical and cultural backgrounds, countries differ in their current stage of cigarette diffusion. Therefore the same pattern is not observed at the same time in all countries. The unit of analysis of this model is the individual. But nowadays there is a shift from individual models towards models that incorporate contextual effects. This research project takes place in this trend.

Evidence of concordance in smoking and other risky practices have been found in previous studies. For example, a study conducted on the British Household Panel Survey proved that members of the same household tend to quit smoking simultaneously (Chandola, [1]). Another analysis on the same data concluded that there was a "matching" of smokers on the marriage market (Clark and Etilé, [2]). Concerning alcohol drinking, a study based on a French epidemiological cohort revealed that getting married was accompanied by an increased level of drinking (Zins et al., [5]). These studies tend to prove that it is important to look not only at individual-level predictors, but also at household-level predictors.

1.3 Research question

The influence of social context is being studied at two levels: micro level (influence of several individual characteristics such as age, gender, education, income, occupation, etc.) for variables control purposes, and meso level (influence of household). More precisely, the main objective is to take advantage of data with a hierarchical structure in order to identify and quantify the sources of the variation of smoking practices. The research question is the following: are household factors predictors of smoking, all other individual characteristics being equal? This question is split into a series of objectives:

(1) The first objective is to point out the existence of an household specific effect affecting smoking practices

(2) The second objective is to measure the portion of total variability explained by the household effect

(3) The third objective is to identify household characteristics related to tobacco use

(4) The last objective is to determine if effects vary from one country to the next by studying some particular cases. The idea here consists in taking countries that are at different moments of the diffusion process, to determine if this process affects household patterns.

Section 2 of this paper presents the data used. Section 3 examines the statistical methodology that has been implemented on this data. Section 4 analyses the results. Section 5 concludes and presents pistes for future work.

2 Data

2.1 The European Community Household Panel

The European Community Household Panel was designed to study household income dynamics at the European level. It consisted in an annual interview of randomly selected households. The first wave took place in 1991 and the eighth and last wave occurred in 1998. The ECHP database provides many information on household and individual incomes, but it also contains several modules related to other topics (education, labour, social relations, health, etc.). Information on smoking practices is available from the 6^{th} wave onwards for most countries. The final sample consists in x individuals in y households.

2.2 Variables

Smoking status was ascertained using answers to the following question: "Do you smoke or did you ever smoked?". The possible responses were: "I smoke daily","I smoke occasionally","I do not smoke, but I used to smoke daily", " I do not smoke, but I used to smoke occasionally" and "I never smoked". The variable of interest is a dummy equal to one if the individual is a daily smoker, and 0 otherwise.

The independent variables are of different types: some relates to the individual level, others to the household level.

Individual level characteristics	*Household level characteristics*
sex	household income (quartiles)
age	composition of the household
educational attainment	
activity status	
lives with at least one (other) smoker	

Three dummy variables have been constructed to represent "maximum level of education attained": Tertiary (Third level), Secondary (second stage of secondary level) and Primary (less than second stage of secondary education). There are six possible categories for "activity status": Higer grade non manual, Lower grade non manual, Skilled manual, Elementary occupation, Unemployed and Economically inactive. The household income variable is equivalised by the OECD-modified scale. The "composition of the household"

has six modalities: single person, couple without children, couple with children, other types of household. The dummy "lives with at least one (other) smoker" lies in-between the household and the individual level, because it takes into account the practices of other members of the household, but still it remains individual-specific. For example, in a couple comprising a smoker and two non-smokerw, this variable will be equal to 0 for the first member and to 1 for the two other members.

3 Statistical Methodology

3.1 Data structure and methodological consequence

In the ECHP, individuals are nested within households. We can assume that two members of the same household are more likely to 'resemble' each other than two individuals randomly drawn from the sample. In other words, an unobservable household effect may affect practices. For example, there might be a household ban on smoking which creates a local norm regulating smoking practices within the household.

From a technical point of view, it means that there is a potential correlation of residuals for members from the same household. The responses cannot be assumed to be independent even after conditioning on the observed variables, because of this unobserved household specific effect. There is great chance that the assumption of mutual independence across observations is violated, so traditional statistical models cannot be used. The advantage of panel data is that the structure of the data itself make it possible to solve the problem of non-independent observations.

3.2 A random intercept probit model

The outcome we seek to model is a discrete choice: what we observe is the result of the choice, i.e. whether the individual i belonging to the household j smokes daily or not.

$$\begin{cases} Y_{ij} = 1 \text{ "daily smoker"} & \text{if } Y_{ij}^* > 0 \\ Y_{ij} = 0 \text{ "no-smoker"} & \text{otherwise} \end{cases}$$

Y_{ij}^* is an unobserved continuous variable which can be interpreted as the utility the individual gets from smoking. This latent variable is modeled using a two level random intercept model.

$$\begin{cases} P(Y_{ij} = 1 | x_{ij}, x_j) = \beta + x_{ij}\beta_1 + x_j\beta_2 + v_{ij} \\ v_{ij} = c_j + u_{ij} \end{cases}$$

x_{ij} is a vector of individual characteristics and x_j a vector of household characteristics. c_j is the household specific effect term and can also be referred as an "unobserved heterogeneity" term. It represents a global effect of all omitted covariates (for instance, a smoking ban, or a general feeling concerning tobacco and tobacco users diffused in the household and shared by its members) that might affect the individual probability of smoking. In this model, parameters associated to individual and household variables are identical across observations, but the intercept is household-specific because it is modulated by the household specific effect.

This household specific effect can be estimated using two different specifications: fixed effects (FE) or random effects (RE). The terms "fixed" and "random" do not relate to the nature of the effects, but to the way these effects are estimated. In both cases we can make the assumption that the household specific effect results from a random draw in a probability law. If we do not want to make any assumption on the nature of the law, we can estimate it point by point: this is the logic of the FE method. In the RE case, we suppose a functional form for the probability law and then only a limited number of parameters need to be estimated (for example, the variance). But for the results of the random effects models to be consistent, the unobserved heterogeneity term has to be unrelated to the explanatory covariates. As there is no restriction on the distribution of the heterogeneity, the fixed effects specification always gives unbiased estimators, but it does not allow to take into account household characteristics (invariant over individuals). Moreover, the main drawback in our case is that the "smokers only" and "non-smokers only" households do not contribute to the identification. This is problematic in the frame of this analysis, because the phenomenon of interest is the clustering of practices in itself. Therefore a random effects specification seems to be wiser. A drawback is that it provides unbiased and efficient estimates only if there is no correlation between the unobserved heterogeneity term and the observed characteristics. In our final model, the household specific effect is not significant anymore. As the estimates of the parameters remained quite stable as blocks of covariates were successively added up till the final model, we suppose that our random effects specification gives reliable results.

We assume that the household specific effect is normally distributed and estimated a random intercept probit model. The dataset is not balanced since households differ in the number of their members. Household with one observation do not contribute to the estimation of the unobserved component, but are taken into account for the estimation of all other parameters.

The intra-class correlation coefficient (ICC) is the proportion of the second level variance out of the total variance. σ is the second level (=household) variance. The individual variance is set to 1 in a random intercept probit model. ρ is then equal to:

$$\rho = Corr(v_{it}, v_{is}|x_{ij}, x_i) = \frac{\sigma_c^2}{\sigma_c^2 + \sigma_u^2} = \frac{\sigma_c^2}{\sigma_c^2 + 1}$$

The higher ρ, the more important the clustering.

All models were estimated using Stata®10.

4 Results

4.1 Evidence for a household specific effect

Before coming to the results of the model, we first give some descriptive evidence for a household specific effect. For the moment, only results from the French case are presented. It seems that there is a concordance of practices within households. We can see this both from the household and the individual point of view. More than 2 thirds of the households are totally smoker or totally non-smoker. From the individual point of view, individuals living with a smoker a more likely to be smokers themselves.

Table 1: Descriptive evidence of a clustering of practices within households

Clustering for more than two thirds of the households

(%)	All households	Households with 2 members and more
Non-smoker only	57,0	51,5
Smoker only	18,9	14,1
Mixed situation	24,1	34,3

Individuals living with a smoker are more likely to be smokers themselves

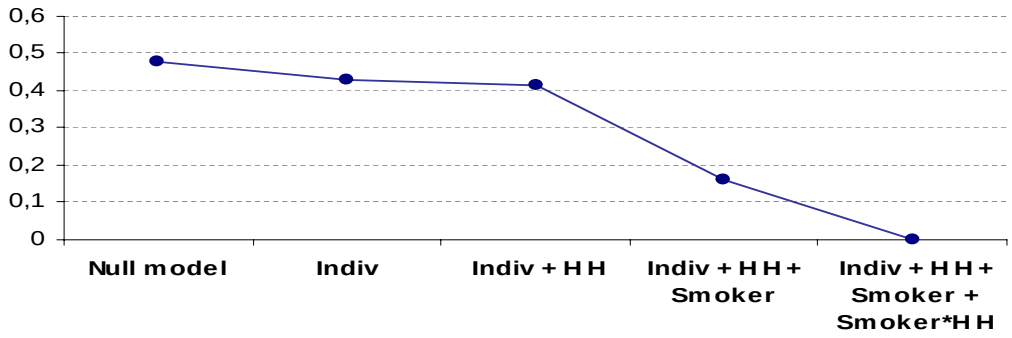
(%)	Smoker	Non-smoker
Lives with a(nother) smoker	78,5	21,5
No (other) smoker in the HH	54,4	45,6
All	*73,0*	*27,0*

4.2 Measure of the portion of variance due to the household level

Men and women have been systematically distinguished in the models to take into account the fact that the effects of the explanatory variables can differ according to the sex. In other words, all explanatory variables have been dichotomized for men and women. Blocks of variables have been included step by step in the model. This enables to observe the evolution of the proportion of total variance contributed by the household level.

The null model - that is to say a model that does not contain any explanatory variable but just take into account the household clustering - confirms the existence of an household effect, because almost half of the variance is due to the household level. Adding individual and the household variables induced a very small decrease, whereas there is a substantial decrease after the inclusion of the variable presence of a(nother) smoker in the household. The household effect disappears totally in the final model which contains individuals and household covariates, the dummy 'presence of a(nother) smoker' and an interaction term between this dummy and the household structure.

Figure 1: Evolution of the intra-class correlation coefficients during the step-by-step estimation



4.3 Identification of important household characteristics

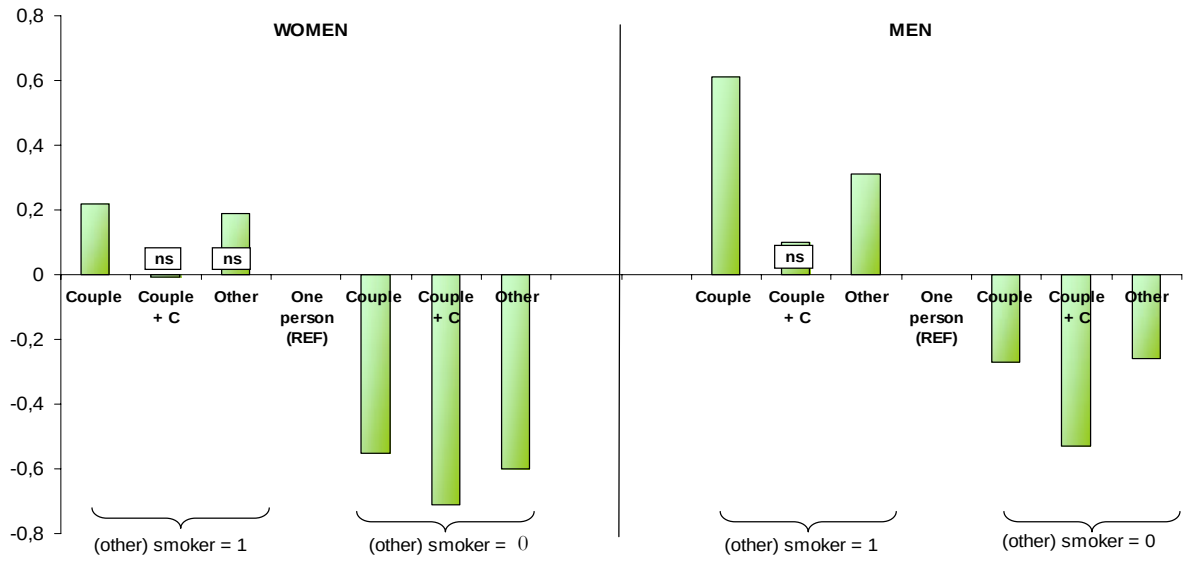
We can then conclude that the household effect is divided into:

- a clustering effect
- an effect of the household structure
- an interaction between the two effects

The results of the model show that the preliminary descriptive results remain valid after controlling for characteristics of the individual and of the household: living with a smoker increases the probability of smoking. Moreover, the effect is of the same magnitude for men and women. The effect of the household income is not significant. The household structure has a significant effect on probabilities and there is an interaction between the presence of a smoker in the household and the household structure: living with a smoker affects the probability of smoking differently according to the household structure.

Let now describe the household effect more precisely. The figure 2 shows the effect on the individual probability of smoking. For women, to live with a smoker increases the probability of smoking only in the case of the couple without children. On the contrary, to live with no-smokers is associated with a negative effect on the probability to smoke in all familial contexts, the greater effect being observed for couples with children. For men, to have at least one smoker in the household increases the probability, except for couples with children. The protective effect of a non-smoker is weaker than this observed for women.

Figure 2: The effect of living with smoking or no-smoking housemates on the individual probability of smoking: results for men and women



4.4 The household effect differs according to the national context

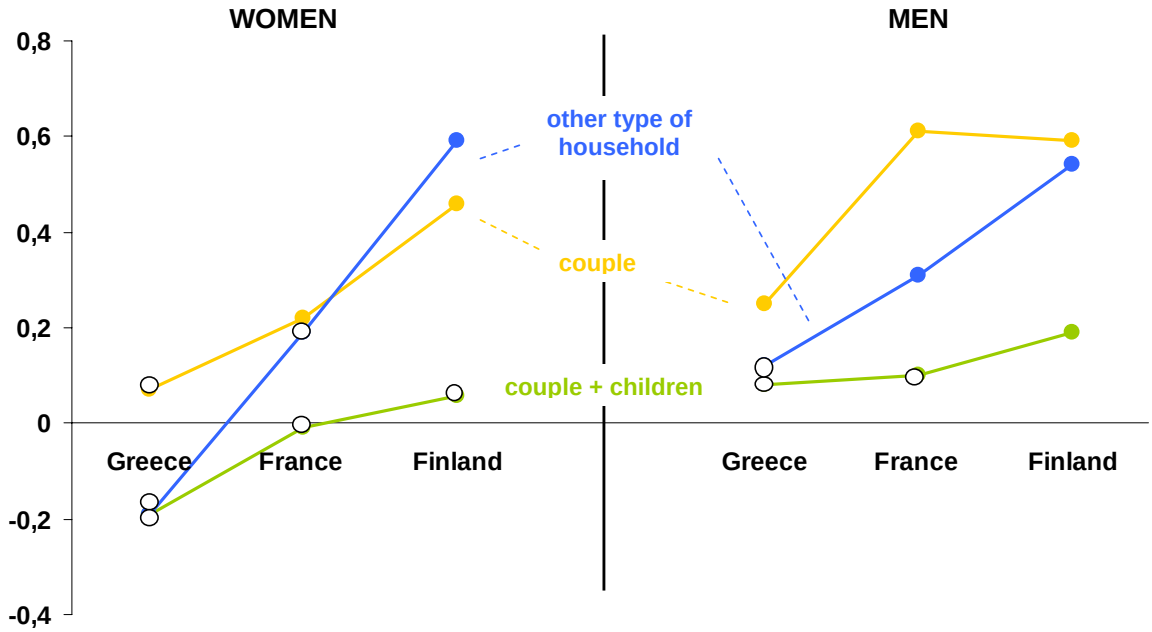
We may wonder if the concordance has been changing during the "diffusion" process of smoking. We know that men and women social gradients have known the same evolution but with a different timing. For example when smoking was receding among men of high status class, it was still rising among women for the same group. Given the fact that households are characterized by social homogamy and that this matching is stable over time, it is likely that smoking status of members of a couple have varied between concordance and discordance during the diffusion process. As we don't have data that could allow to study the concordance within French households over a long period, we tried to make use of the ECHP to replace France in the global diffusion process. Due to a variety of historical and cultural backgrounds, European countries differ in their current stage of cigarette diffusion. Greece and Finland have been chosen because they were in the two end-points of the diffusion scale (Forey et al., [4]). The same model was run for these countries. Results are reported figure 3. The magnitude of the concordance effect differs across countries. It seems that there is a stronger clustering effect as the diffusion process goes on. In other words, the "protective" effect of a non-smoker household becomes stronger, whereas the "incentive" effect of living with a smoker also increases. Theses are just preliminary results and a deeper analysis should be made using other countries of the ECHP.

5 Discussion

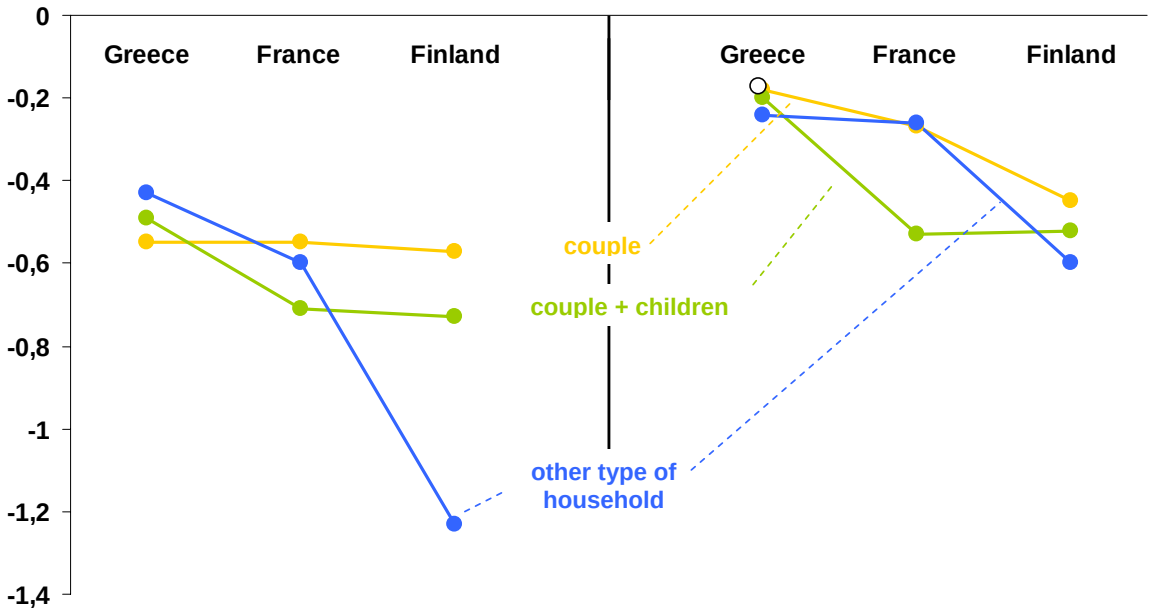
This research gives evidence for clustering of smoking practices in household, the effect being different according to the contexts of both the household and the country. The effect of living with smoking or no-smoking housemates indeed depends on the structure of the household, but this effect also differs according to the length of smoking history in the country. It seems that the later the stage of the smoking diffusion process, the more sensitive the concordance. In a given country, the diffusion process shapes the smoking patterns observed across cohorts and social classes. It probably also shapes the distribution of smoking within households. The social determinants of smoking shift across cohorts as the diffusion process goes on. For example, the social class, which was no longer a predictor of smoking status for men some decades ago, has regained its predictive power (but in the inverse way). The household effect also changed over time. The stronger clustering effect that is observed as the diffusion process goes on tends to show that there are more and more "all no-smokers" and "all smokers" households,

Figure 3: Evaluation of concordance in three national contexts: the cases of Greece, France and Finland

The influence of living with a smoker on the individual probability to smoke



The influence of living in a no-smoker household on the individual probability to smoke



Lecture : Results come from random intercept probit models allowing for different influences of living with a smoker or not on the individual probability to smoke, according to the household composition. Single adults are the reference category. Estimates that are not statistically significant are denoted by the sign ○.

maybe because of a growing stigmatization of smoking practices?

Research about health concordance overwhelmingly suggests evidence for clustering of health status and health behavior (Falba et al., [3]; Meyler et al., [6]; Wilson, [9]. This offers room for a deeper understanding for the causes of health concordance: is it due to *ex ante* correlation (which might be caused by assortive mating, social homogamy) or to *ex post* convergence (which might occured because of social control, contagion and imitation, influence of shared environment and ressources)?

Table 2: Comparison between a single probit and the random intercept probit model

	Probit				Random intercept probit model			
	Female		Male		Female		Male	
	Coef.	Sig.	Coef.	Sig.	Coef.	Sig.	Coef.	Sig.
Constant	-0,3073	***	0,0632		-0,4400	***	0,0492	
Age group								
Under 30	0,0046		-0,0577		-0,0010		-0,0411	
30-39 (ref.)	-	-	-	-	-	-	-	-
40-49	-0,2806	***	-0,2434	***	-0,3082	***	-0,2731	***
50-59	-0,8284	***	-0,6151	***	-1,0766	***	-0,7223	***
More than 60	-1,4136	***	-1,1349	***	-1,8480	***	-1,4142	***
Educational attainment								
Third level	-0,3088	***	-0,3647	***	-0,3907	***	-0,4780	***
Secondary level	-0,0373		-0,1833	***	-0,0648		-0,2362	***
Less than secondary level (ref.)	-	-	-	-	-	-	-	-
Activity status								
Higher grade non manual	0,2336	**	-0,1156	*	0,3610	**	-0,1345	*
Lower grade non manual	0,1262		0,0340		0,1828		0,1013	
Skilled Manual (ref.)	-	-	-	-	-	-	-	-
Elementary occupation	0,1520		0,0914		0,2800		0,0706	
Unemployed	0,2981	**	0,2850	***	0,4011	**	0,3917	***
Economically inactive	0,1070		-0,0872		0,1863		-0,1193	
Household income								
1 Highest quartile	0,0305		0,0862		0,0423		0,0828	
2	0,0019		-0,0452		0,0055		-0,0748	
3 (ref.)	-	-	-	-	-	-	-	-
4 Lowest quartile	-0,0034		0,0757		0,0054		0,0993	
Composition of the household								
Single person (ref.)	-	-	-	-	-	-	-	-
Couple without children	-0,1884	***	0,0069		-0,2872	***	-0,0165	
Couple with children	-0,3168	***	-0,2118	***	-0,4294	***	-0,2702	***
Other type of household	-0,1576	*	0,0312		-0,2160	*	0,0362	
Random part								
HH level random intercept					0,4305 (***)			

Note : ***: significant at 1% level **: significant at 5% level *: significant at 10% level

Table 3: Final model

	Random intercept probit model			
	Female		Male	
	Coef.	Sig.	Coef.	Sig.
Constant	-0,4084	***	0,1183	
Age group				
Under 30	-0,0445		-0,1117	*
30-39 (ref.)	-	-	-	-
40-49	-0,2936	***	-0,2623	***
50-59	-0,8214	***	-0,5987	***
More than 60	-1,3596	***	-1,0173	***
Educational attainment				
Third level	-0,2881	***	-0,3677	***
Secondary level	-0,0297		-0,1912	***
Less than secondary level (ref.)	-	-	-	-
Activity status				
Higher grade non manual	0,2603	**	-0,1243	*
Lower grade non manual	0,1207		0,0615	
Skilled Manual (ref.)	-	-	-	-
Elementary occupation	0,1755		0,0707	
Unemployed	0,2985	**	0,3153	***
Economically inactive	0,1124		-0,0904	
Household income				
1 Highest quartile	0,0604		0,0800	
2	0,0024		-0,0467	
3 (ref.)	-	-	-	-
4 Lowest quartile	-0,0166		0,0483	
Composition of the household and presence of a smoker (1/0)				
Single adult (ref.)	-	-	-	-
Couple without children, 1	0,2149	**	0,6151	***
Couple with chidren, 1	-0,0063		0,1104	
Other type, 1	0,2001		0,3119	**
Couple without children, 0	-0,5630	***	-0,2693	***
Couple with chidren, 0	-0,7154	***	-0,5263	***
Other type, 0	-0,5997	***	-0,2637	**
Random part				
HH level random intercept	1.87e-06			

References

[1] T. Chandola, J. Head, and M. Bartley. Socio-demographic predictors of quitting smoking: how important are household factors? *Addiction*, 99:770–777, 2004.

[2] A.E. Clark and F. Etilé. Don't give up on me baby: spousal correlation in smoking behaviour. *DELTA Working Paper*, (2003-25), 2003.

[3] T.A. Falba and J.L. Sindelar. Spousal concordance in health behavior change. *Health Services Research*, 43(1), 2008.

[4] B. Forey, J. Hamling, P. Lee, and N. Wald (eds.). *International Smoking Statistics*. Oxford University Press, 2007.

[5] A. Guéguen, J.P. Nakache, and M. Zins. Utilisation des modèles marginaux et des modèles mixtes dans l'analyse de données longitudinales (1992-1996) concernant mariage et consommation d'alcool des femmes de la cohorte gazel. *Revue de Statistique Appliquée*, 48(3):57–73, 2000.

[6] D. Meyler, J.P. Stimpson, and M.K. Peek. Health concordance within couples: A systematic review. *Social Science Medicine*, 64:22972310, 2007.

[7] F.C. Pampel. Cigarette diffusion and sex differences in smoking. *Journal of Health and Social Behavior*, 42(4):388–404, 2001.

[8] F.C. Pampel. Inequality, diffusion, and the status gradient in smoking. *Social Problems*, 49(1):35–57, 2002.

[9] S.E. Wilson. The health capital of families: an investigation of the inter-spousal correlation in health status. *Social Science Medicine*, 55:11571172, 2002.