

Smoking within the Household: Spousal Peer Effects and Children's Health Implications¹

Chiara Canta² and Pierre Dubois³

Abstract

We study spousal peer effects on smoking and their implication for the health of children. Smoking decisions are modeled as equilibrium strategies of an incomplete information game within the couple. Using French data, we identify two peer effects: a smoking enhancing effect of smoking partners and a smoking deterring effect of non smoking partners. An implication of these findings is that the smoking behavior could differ qualitatively in couples where both partners smoke and where only one partner smokes. This interpretation is supported by our finding that, controlling for total tobacco consumption of parents, the respiratory health of children is negatively affected only if both parents smoke.

JEL Code: C31, D10, I12.

Keywords: Smoking, Social Interaction, Simultaneous Game Model, Health.

¹We thank Jerome Adda, Magnus Johannesson, and workshop and conference participants at Toulouse School of Economics, Stockholm School of Economics, the ASHE Conference 2010 and the European Economic Association Congress 2010 for useful comments.

²NHH, Department of Economics. e-mail: Chiara.Canta@nhh.no

³Toulouse School of Economics. e-mail: Pierre.Dubois@tse-fr.eu

1 Introduction

Smoking imposes huge costs on societies, so that understanding of the determinants of smoking behavior is an important policy concern. An important question in this context is how social interactions affect smoking decisions. The behavior of peers might convey information about the costs and benefits of smoking, or affect directly the benefit that an individual derives from tobacco consumption. A number of studies provides evidence of peer effects on teenagers' tobacco consumption (Powell et al., 2005, Lundborg, 2006, Harris and González López-Valcárcel, 2008). Harris and González López-Valcárcel (2008), for instance, separately identify the impacts of peer-group smoking prevalence and peer-group non smoking prevalence using a cross-section of US young people. A peer group is constituted by young people living in the same household ("siblings"). The main finding of the paper is that the pro-smoking influence of a smoking sibling is twice as big as the deterrence effect exerted by a non smoking sibling. The authors interpret these effects in terms of learning: siblings' smoking and non smoking behavior do not convey the same kind of information about how "cool" smoking is. Concerning intra-household peer effects, Jones (1994) finds a significant impact of the presence of other smokers in the household on individuals' probability of succeeding in giving up. Cutler and Glaeser (2007) study spousal peer effects on smoking decisions, as well as the impact of other kinds of social interactions. Using instruments in order to tackle selection problems, they show the existence of such effects. Khwaja et al. (2006) and McGeary (2013), estimating a health capital model, also find that spousal health affects the individuals probability of smoking. These findings have important implications: they suggest that measuring the impact of smoking containment policies at the individual level is not appropriate, because of the multiplier effect due to social interactions.

This paper contributes to the literature by disaggregating the spousal peer effect on individual smoking into a smoking enhancing and a smoking deterring effect, in the lines

of Harris and González López-Valcárcel (2008). In our theoretical model, spousal smoking behaviors directly affect the individual utility from smoking. However, we allow this effect to be different depending on whether the spouse smokes or not. On the one hand, having a smoking partner could make smoking more valuable because of the possibility of smoking together or because of it reinforces beliefs that smoking is not too unhealthy. On the other hand, having a non smoking partner could reduce the utility from smoking, because of the (partial) internalization of some of the nuisance imposed on the partner. The main empirical result of the paper is that both a smoking enhancing and a smoking deterring effect are at work in the spousal smoking game.

As highlighted in Khwaja et al. (2006), these effects might be due to learning, altruism, or to consumption externalities within the household. In our case, both kinds of peer effects might be related to consumption externalities or learning, while the smoking deterring effect could be due to altruism. If consumption externalities or altruism were prevalent, the presence of a smoking deterring peer effect would suggest that individuals consume cigarettes in a different way, depending on the partner's smoking status. For instance, smokers could be forbidden to smoke at home by non-smoking partners, or get continual complaints from them, suffering thus a consumption externality. Alternatively, smokers might refrain from smoking at home, or smoke less than they would like to out of altruism towards non-smoking partners. In both cases, this should have an impact on the exposure to passive smoking of other family members, and ultimately on their health status. Using data on the respiratory diseases of children, we find evidence supporting the presence of consumption externalities or altruism. Interestingly, this suggests that spousal social interactions do not only affect the probability of smoking, but also the nature of the smoking behavior.

We model smoking decisions as the equilibrium strategies of an incomplete information game within the couple, while singles take smoking decisions individually. Each individual

observes a number of characteristics of her partner (if any), and she forms a prior on the probability that the partner smokes, which can be seen as the probability that the partner quits smoking or starts smoking, given her past behavior. The assumptions that the smoking game is non cooperative and that information is incomplete within the couple can be justified by the fact that smoking is addictive and result from an instantaneous decision. Thus, it might be difficult for an individual to commit to a particular behavior. In addition to that, adults are not likely to start smoking, but they might quit if they were smokers in the past. Succeeding in giving up smoking is not perfectly predictable *ex ante* and might depend on random individual shocks that the partner cannot observe or predict (for instance, health, taste or psychological shocks). Thus, at each instant in time, the smoking decision of the spouse can be considered as uncertain and based on private information.⁴

We use data from the French Health Survey 2002-2003 carried out by the French National Statistical Institute (INSEE). To our knowledge there is no study of intra-couple peer effects on smoking using French data. The literature deals mainly with U.S. or U.K. data. Smoking decisions in different countries might be subject to different social norms. Our results, however, are in line with the existing literature on the determinants of smoking. We use the methodology proposed by Bajari et al. (2010), permitting to estimate peer effects in incomplete information games, when the equilibrium probabilities of choosing a particular strategy are unobserved to the econometrician. We include in our sample individuals that play a game with their partner (couples) and individuals whose smoking behavior is determined individually (singles). This permits us to identify both the negative and positive peer effects that are at work in the smoking game. Identification uses mostly the decision to give up smoking, which, for adults, are much more prevalent than the decisions to start smoking.

⁴Alternative models have been developed in the context of peer-effects estimation, involving complete information games. For instance, Krauth (2006) and Soetevent and Kooreman (2007) consider complete information games with possibly multiple equilibria. In this context, they propose simulation methods to estimate peer effect.

The main problem in estimating household decision-making models is related to assortative matching (Becker, 1974), leading to the well known correlated effects (Manski, 1995). As emphasized by Manski (1993), or more recently by Blume et al. (2011), correlations between outcomes in a social group (here the household) can be produced by several mechanisms: (i) direct interdependences between the choices (endogenous interactions or peer effects); (ii) the impact of observable group-level characteristics or common shocks (contextual effects); (iii) the impact of unobserved common shocks and endogenous selection in the peer group on the basis of unobserved characteristics that also affect the outcome of interest (correlated effects). Thus, disentangling peer effects from correlated effects is a difficult task. To tackle endogenous matching, we control for a number of partners characteristics influencing the beliefs about the partner's decisions. In our model, each individual takes an instantaneous smoking decision in order to maximize her expected utility, knowing the characteristics of the partner. In addition to that, we control for past smoking behaviors. Thus, our model does not predict the probability of smoking of an individual, but the probability of smoking given her past smoking behavior and the partner she lives with. Controlling for the past smoking decision of both partners permits to take into account the correlation across partners' unobservables that could both affect matching and smoking decision, if this correlation is constant over time. Controlling for the past smoking status, we assume that future smoking decisions are not anticipated and that preference shocks for smoking in the future are independent across couple members conditionally on their past smoking status. In other words, we exclude that couples form on future propensity to quit or start smoking given their current smoking status. We also assume that contextual effects are fully captured by the observable variables we condition on. We will see in the identification section that results may be robust when we relax this seemingly strong assumption. So there are no correlated effects (unobserved heterogeneity correlated within the household) that could explain the results on the smoking

enhancing and smoking deterring peer effects.

By identifying significant spousal peer effects on smoking behavior, our results are markedly different from the ones of Clark and Etilé (2006). They consider a complete information framework and find that the correlation between spouses smoking decisions is entirely due to the correlation between individual random effects, not to peer effects. However, they do not consider the role of beliefs about the partner's probability to smoke in each period. More specifically, they rule out the possibility that individual effects include beliefs about the partner's instantaneous decision. We explicitly model and structurally estimate these beliefs in order to study their effect on smoking behavior, and we find that beliefs matter.

Concerning health and smoking, a relatively large literature has studied the effect of passive smoking on the health of children. For instance, Mannino et al. (2001) showed that an high level of exposure to tobacco smoke (measured by the cotinine levels in body fluids) increases the numbers of sick days in school for U.S. children. Adda and Cornaglia (2009), showed that smoking bans in public place increase the exposure of children with smoking parents. However, these studies do not look at the interaction of smoking and non smoking parents. According to our results, a non smoking parent could play a control role and push the smoking parent to smoke outside. This would imply that, controlling for the overall quantities smoked, the respiratory diseases affect more the children whose both parents smoke than the ones with just one smoking parent. Using the French health survey 2002-2003, we find that only the fact that both parents smoke affects negatively the health of children. This is consistent with the smoking game model. Non smoking parents seem to play a control role on their smoking spouses, possibly reducing the utility of smokers by imposing less smoke at home. This translates in the reduction of respiratory diseases of children that we observe in our data.

The paper is structured as follows: in Section 2 we describe the theoretical model and in

Section 3 we discuss the estimation strategy. In Section 4 we present the data and we describe the results concerning both the equilibrium smoking behavior and the effect of smoking on the respiratory diseases of children. Section 5 concludes.

2 Explaining Smoking Behavior: A Theoretical Model

As Khwaja et al. (2006), we assume that partners behave strategically with respect to smoking behavior. The assumption that the game is non-cooperative can be justified by the fact that commitment to a smoking behavior might be impossible, due to addiction, for instance.⁵ We consider a model in which individuals living in a couple choose their smoking behavior simultaneously, taking into account the expected decision of the partner. One interpretation of simultaneity is that it is a way of modelling the fact than none of them would like to move first and commit to some choice. Singles take their decision individually.

Let us first consider the utility of an individual i living with a partner j . The benefit that this individual obtains from smoking depends on her taste for smoking f_i and on the smoking behavior of her partner. Let $a_i \in \{0, 1\}$ be the smoking decision of individual i , where $a_i = 1$ means that the individual smokes. The partner makes the same dichotomous decision denoted by a_j . The utility of individual i if she smokes is:

$$f_i + sa_j - p(1 - a_j),$$

where f_i is the private benefit from smoking, p is the disutility due to the presence of a non smoking partner and s is the extra utility from smoking if the partner smokes as well.

Both parameters are assumed to be non negative in the theoretical model. They capture,

⁵Also the impossibility to write down legally binding contracts between partners might justify strategic behavior. On the alternative models of health-decisions taking within the household, see Bolin et al. (2001, 2002) .

respectively, a smoking enhancing and a smoking deterring effect related to the partner's behavior.

The utility of a non smoker is normalized to zero up to the disutility linked to passive smoking when the partner smokes. It is thus equal to:

$$-ra_j,$$

where r is a positive parameter quantifying the disutility imposed by a smoking partner on a non smoker. It is important to highlight that positive values of s , p and r may be motivated by consumption externalities, altruism or learning.

The utility from smoking of a single is simply f_i , while her utility from non smoking is normalized to zero. Using the dummy variable D_i taking value one if the individual lives with a partner and zero otherwise, we can write the utility from smoking in a more general form as:

$$f_i + sa_jD_i - p(1 - a_j)D_i.$$

In the same way, the utility of a non smoker is thus equal to:

$$-ra_jD_i.$$

We assume that the private benefit from smoking is a linear function of a vector of observed variables \tilde{x}_i , and of an individual shock, unobservable by the econometrician, ϵ_i :

$$f_i = \lambda\tilde{x}_i + \epsilon_i, .$$

The past smoking behavior is an important characteristic to be included in \tilde{x}_i , given the addictive nature of tobacco consumption.

The utility U_i of individual i can be written as a function of the individual's decision to smoke a_i , the partner's decision to smoke a_j and the individual characteristics:

$$\begin{aligned}
U_i(a_i, a_j, x_i, \epsilon_i) &= a_i [\lambda \tilde{x}_i + \epsilon_i + s a_j D_i - p(1 - a_j) D_i] - [1 - a_i] r a_j D_i \\
&= a_i [\lambda \tilde{x}_i - p D_i + \epsilon_i] - r a_j D_i + a_i a_j [s + p + r] D_i \\
&= \lambda_1 a_i x_i + \lambda_2 a_j D_i + \lambda_3 a_i a_j D_i + a_i \epsilon_i,
\end{aligned}$$

where $x_i = (\tilde{x}_i, D_i)$, $\lambda_1 = (\lambda, -p)$, $\lambda_2 = -r$, $\lambda_3 = p + r + s$. As the parameters s , r and p should be non negative, we expect in particular that $\lambda_3 \geq 0$.

We assume that both partners take their decision simultaneously. The decision is taken under incomplete information because the attitude of both partners towards smoking might depend on idiosyncratic shocks on their preferences that are private information. Think about smoking cessation (which is empirically very relevant in our sample as will be illustrated in the following): this choice might depend on some health or taste or psychological shock. Thus, the individual shock, ϵ_j , is assumed to be unobservable by the partner and the expected payoff of individual i depends on the probability that the other partner smokes which we denote $\sigma_j(x_i, x_j) = Pr(a_j = 1 | x_i, x_j)$. The belief of each individual about the partner's probability of smoking is constructed using all the available information. In other words, beliefs are contingent to (x_i, x_j) , the set of both the individual's and the partner's characteristics, which is publicly known within the household. We thus look at a Bayesian Nash equilibrium.

An individual i does smoke ($a_i = 1$) if and only if

$$\Delta EU_i \equiv E_{a_j} [U_i(1, a_j, x_i, \epsilon_i)] - E_{a_j} [U_i(0, a_j, x_i, \epsilon_i)] > 0. \tag{1}$$

Since $U_i(a_i, a_j, x_i, \epsilon_i)$ is linear in the unknown action a_j , in equilibrium the expected payoff of strategy a_i can be written as

$$E_{a_j} [U_i(a_i, a_j, x_i, \epsilon_i)] = \lambda_1 a_i x_i + \lambda_2 \sigma_j(x_j, x_i) D_i + \lambda_3 a_i \sigma_j(x_j, x_i) D_i + a_i \epsilon_i,$$

where $\sigma_j(x_j, x_i)$ is the equilibrium probability that the partner smokes. Substituting this expression in (1) one can rewrite the smoking condition for individual i as

$$\Delta EU_i = \lambda_1 x_i + \lambda_3 \sigma_j(x_j, x_i) D_i + \epsilon_i > 0.$$

This condition shows that it will be impossible to identify the parameter $\lambda_2 = -r$. The model only permits to identify $\lambda_1 = (\lambda, -p)$, and $\lambda_3 = p + r + s$. Our estimation will thus permit to recover the structural parameters p and $r + s$.

3 Identification and Estimation strategy

Given the availability of data on both the smoking behavior within households and the characteristics of the members of the households, we first study how to identify and estimate the model parameters.

Our theoretical model shows that the decision to smoke of individual i depends on whether $\lambda_1 x_i + \lambda_3 \sigma_j(x_j, x_i) D_i + \epsilon_i$ is positive or not. In our application, we will always condition on past smoking behavior (which is included in x_i) such that decision should be interpreted as the decision to quit smoking (for smokers) or to start smoking (for non smokers - even if much less frequent at adult age). Under the assumption that the preference shocks ϵ follow a logistic distribution, we obtain the following expression for the probability that individual

i smokes conditional on the individual's and the partner's characteristics, $\sigma_i(x_i, x_j)$:

$$\begin{aligned}\sigma_i(x_i, x_j) &= Pr(a_i = 1 | x_i, x_j) = Pr(\lambda_1 x_i + \lambda_3 \sigma_j(x_i, x_j) D_i + \epsilon_i > 0) \\ &= \frac{\exp(\lambda_1 x_i + \lambda_3 \sigma_j(x_i, x_j) D_i)}{1 + \exp(\lambda_1 x_i + \lambda_3 \sigma_j(x_i, x_j) D_i)} \\ &= \frac{\exp \Delta EU_i(x_i, x_j)}{1 + \exp \Delta EU_i(x_i, x_j)}\end{aligned}$$

and the same holds for the partner j .

We thus obtain the following system of equations for couples

$$\begin{aligned}\sigma_i(x_i, x_j) &= \frac{\exp(\lambda_1 x_i + \lambda_3 \sigma_j(x_i, x_j) D_i)}{1 + \exp(\lambda_1 x_i + \lambda_3 \sigma_j(x_i, x_j) D_i)} \\ \sigma_j(x_i, x_j) &= \frac{\exp(\lambda_1 x_j + \lambda_3 \sigma_i(x_i, x_j) D_j)}{1 + \exp(\lambda_1 x_j + \lambda_3 \sigma_i(x_i, x_j) D_j)}\end{aligned}$$

We can prove that this system admits at least one solution $(\sigma_i^*(x_i, x_j), \sigma_j^*(x_i, x_j))$ such that $\sigma_i^*(x_i, x_j) \in (0, 1)$ and $\sigma_j^*(x_i, x_j) \in (0, 1)$. Furthermore, a sufficient condition for this system to have a unique solution is that $|\lambda_3| < 4$. This solution thus corresponds to a Bayesian Nash equilibrium of our model.

To prove this result, let us denote by F the logistic function ($F(Z) = \exp(Z)/(1 + \exp(Z))$). Note that $F' = F(1 - F)$ and define $G(z) = z - F(\lambda_1 x_i + \lambda_3 F(\lambda_1 x_j + \lambda_3 z D_j) D_i)$ for all z and $H(z) = z - F(\lambda_1 x_j + \lambda_3 F(\lambda_1 x_i + \lambda_3 z D_i) D_j)$. The two equations of our system can be rewritten as $G(\sigma_i(x_i, x_j)) = 0$ and $H(\sigma_j(x_i, x_j)) = 0$. Both $G(\cdot)$ and $H(\cdot)$ are continuous functions; furthermore, $G(0) < 0$, $G(1) > 0$. Then, there exists at least one $\sigma_i^*(x_i, x_j) \in (0, 1)$ such that $G(\sigma_i^*(x_i, x_j)) = 0$. Similarly, it can be shown that there exists at least one $\sigma_j^*(x_i, x_j)$ such that $H(\sigma_j^*(x_i, x_j)) = 0$. Consequently, there exist at least one Bayesian Nash equilibrium such that $\sigma_i^*(x_i, x_j) \in (0, 1)$ and $\sigma_j^*(x_i, x_j) \in (0, 1)$. To show that

the system has a unique solution, it is sufficient to prove that both $G(\cdot)$ and $H(\cdot)$ are strictly increasing. The first derivative of $G(z)$ is

$$G'(z) = 1 - \lambda_3^2 D_i D_j [F'(\lambda_1 + \lambda_3 F(\lambda_1 x_j + \lambda_3 z D_j) D_i)] F'(\lambda_1 x_j + \lambda_3 z D_j) \geq 1 - D_i D_j \left(\frac{\lambda_3}{4}\right)^2,$$

since $0 \leq F'(Z) \leq \frac{1}{4}$ for all z . Then, $G'(\cdot)$ is strictly positive if $|\lambda_3| < 4$. In this range of parameters, the equation $G(\sigma_i(x_i, x_j)) = 0$ has a unique solution. In the same way, it can be shown that $H'(\cdot)$ is strictly positive whenever $|\lambda_3| < 4$, so that the system admits a unique solution $(\sigma_i^*(x_i, x_j), \sigma_j^*(x_i, x_j))$.

The problem in estimating such a model is that the equilibrium probabilities $\sigma_i^*(x_i, x_j)$ and $\sigma_j^*(x_i, x_j)$ are unknown. In order to estimate the model parameters, we use a two-step procedure as in Bajari et al. (2010). In this paper the authors estimate a similar model of simultaneous decisions under incomplete information. We first estimate $\hat{\sigma}_i(x_i, x_j)$ and $\hat{\sigma}_j(x_i, x_j)$ using a logit regression. Under the assumption that individuals do not have better information than the econometrician (that is to say, the partner's ϵ is unknown to each individual), $\hat{\sigma}_j$ corresponds to the belief that individual i holds about the partner's smoking decision. Since at equilibrium the believed probabilities are equal to the equilibrium ones, $\hat{\sigma}_i(x_i, x_j)$ and $\hat{\sigma}_j(x_i, x_j)$ are unbiased estimates of the probability of smoking. Then we use the "Hotz-Miller"(1993) inversion to obtain expected utilities given observables:

$$\begin{aligned} \Delta \widehat{EU}_i(x_i, x_j) &= \ln \hat{\sigma}_i(x_i, x_j) - \ln(1 - \hat{\sigma}_i(x_i, x_j)) \\ \Delta \widehat{EU}_j(x_i, x_j) &= \ln \hat{\sigma}_j(x_i, x_j) - \ln(1 - \hat{\sigma}_j(x_i, x_j)). \end{aligned}$$

These are the empirical counterparts for the expected utility from smoking of individuals i and j . In equilibrium, these estimated expected utilities must coincide with the true expected

utility of smoking $\lambda_1 x_i + \lambda_3 \sigma_j(x_i, x_j) D_i + \epsilon_i$ where σ_j is the equilibrium belief σ_j^* for the rest of the paper. To identify the structural parameters we can then use the following moment conditions

$$E \left[\left(\Delta \widehat{EU}_i(x_i, x_j) - \lambda_1 x_i - \lambda_3 \widehat{\sigma}_j(x_i, x_j) D_i \right) \right] = 0,$$

and minimize in λ_1 and λ_3 the empirical counterpart of this moment condition across individuals.

Note that the structural model parameters are identified if and only if some variables affect only the individual smoking behavior. In particular, in the utility function of individual i , we need that x_j enters only through $\widehat{\sigma}_j(x_i, x_j)$ and is excluded from the direct utility function of individual i .

In this kind of models, assortative matching might lead to biased estimates of the peer effects parameters. The bias might come from two channels: individuals in the same household might behave similarly because of endogenous selection into the couple, or because of some unobserved common shocks.

If tobacco consumption was a relevant characteristic in the choice of a partner, a correlation between smoking behaviors might be explained by correlated effects. To tackle this problem, we control for the past smoking behavior of each partner. Under the assumption that the couple was already formed in the previous period, and that the matching did not depend on the future propensity to smoke, controlling for past smoking decisions permits to take into account the endogeneity due to assortative matching. The resulting peer effect parameter can thus be interpreted as the effect of the expected smoking behavior of the partner on the own decision to smoke today, *given* the individual past smoking status and *given* the characteristics of the partner the individual is assigned to. So the decision to smoke will be akin to a decision to quit smoking or start smoking.

Assortative matching would also imply that any correlation of spouses smoking behaviors

can be explained by the correlation of unobservable characteristics of the partners. Clark and Etilizce (2006), for instance, study how the present smoking status depends on the past smoking status of the partner and they conclude that the correlation between smoking behaviors is explained solely by the correlation of unobservables across spouses. Our strategy is to control for a number of characteristics of the partner, which enter the belief of the probability of smoking. Our approach can be motivated as follows. Suppose for instance that young people tend to smoke more. Since young people have a good chance to match, regressing the smoking behavior of an individual on the smoking behavior of her partner would lead to biased results since the correlation between behaviors could be driven by the simple fact that young people are more likely to be matched with young people. However, controlling for the age of the partner takes away this bias.

Of course, some unobservable characteristics might in fact enter the beliefs of each partner, and affect the smoking behavior. As mentioned above, a crucial assumption in the model is that either both the partners and the econometrician share the same information on the characteristics of each individual or that this information is irrelevant to predict the belief about the partner smoking decision given the variables already observed by the econometrician. This ensures that the true belief $\sigma_i^*(x_i, x_j)$ depends on observables x_i, x_j only. Since $\Delta EU_i = \ln(\sigma_i^*(x_i, x_j)) - \ln(1 - \sigma_i^*(x_i, x_j))$ and $\hat{\sigma}_i(x_i, x_j)$ is an unbiased estimate of the true individual beliefs, then we can conclude that $\Delta \widehat{EU}_i(x_i, x_j) = \ln \hat{\sigma}_i(x_i, x_j) - \ln(1 - \hat{\sigma}_i(x_i, x_j))$.

However, this assumption can somehow be relaxed. For our approach to be valid, we only need the distribution of the belief σ_i given x_i, x_j to be symmetric around 1/2. In fact, if individuals beliefs conditionally on x_i and x_j depend on some unobservable, $\hat{\sigma}_i(x_i, x_j)$ would be the conditional expectation of σ_i given x_i and x_j . In this case, $\Delta EU_i = \ln(\sigma_i) - \ln(1 - \sigma_i)$ implies that $E(\Delta EU_i | x_i, x_j) = E(\ln \sigma_i - \ln(1 - \sigma_i) | x_i, x_j)$, but not necessarily that $\Delta \widehat{EU}_i(x_i, x_j) = \ln \hat{\sigma}_i(x_i, x_j) - \ln(1 - \hat{\sigma}_i(x_i, x_j))$, due to the non linearity of the Hotz-

Miller inversion. However, the Hotz-Miller inversion is a symmetric transformation around $1/2$.⁶ Thus, if the unobservable were such that the distribution σ_i given (x_i, x_j) is distributed symmetrically around $1/2$, we would have

$$E(\ln \sigma_i - \ln(1 - \sigma_i) | x_i, x_j) = \ln(E(\sigma_i | x_i, x_j)) - \ln(1 - E(\sigma_i | x_i, x_j)) = \ln \hat{\sigma}_i(x_i, x_j) - \ln(1 - \hat{\sigma}_i(x_i, x_j)).$$

This would allow us to use the Hotz-Miller inversion using the estimated $\hat{\sigma}_i(x_i, x_j)$.

4 Empirical Estimation and Results

4.1 Data description

The French health survey 2002-2003 includes data on the demography, socioeconomic status, health status, and health consumption of 25,000 households in France. Data were collected on about 40,797 individuals, interviewed three successive times. Adult individuals (over eighteen) were also required to fill an auto-evaluation, in which they were asked to report their perceived health and their prevention behavior, including alcohol and tobacco consumption. A similar but adapted questionnaire was proposed to kids aged between 11 and 17. Out of 30,997 adults in the survey, 25,931 complied and returned an auto-evaluation. We dropped the observations containing missing values either on the reported smoking habit or on the socio-demographic variables of interest. We also limit the analysis to adults being the household reference person or the partner of the reference person and in this subgroup, we consider only individuals with an age between 25 and 60. In the case of couples we drop the households in which the husband is more than 60 years old. Our sample reduces this way to 12,770 adult individuals, 10,540 of which live in a couple.

⁶The function $f(x) = \ln(\frac{x}{1-x})$ is symmetric around $1/2$, because $f(1-x) = -f(x)$.

Our endogenous variable is the smoking behavior. We construct a dummy equal to one if the individual reports being a smoker. An individual is considered a smoker if she reports to be smoking every day.

Individual variables include the age, a dummy equal to one if the level of education is above high school (HS), and the body mass index (BMI). Descriptive statistics on the whole adult population of the survey point to a negative correlation between smoking and BMI. Another variable susceptible to influence smoking behavior is the exposure to smoke at the workplace. This variable is self reported by each individual and takes the form of a dummy equal to one if the individual is currently exposed to smoke at work. It is worth noticing that in France smoking bans in public places were not as strict in 2002 as they are today. Consequently the exposure to work variable exhibits some variability, and almost 15 percent of the individuals in our sample reported to be exposed to smoke at the workplace. Finally, the past smoking behavior is included in the analysis: in particular, we use the two years lagged smoking behavior. A two years period is short enough to reasonably assume that most couples did not change their partner; at the same time it is a period large enough to have some variation between past and current smoking behaviors. Taking longer lags would raise the problem of having to deal with individuals changing partner or moving from single to couple. We do not include in the explanatory variables the price of cigarettes for the simple reason that this price in France is regulated and homogeneous at the national level. Another variable that is often used in smoking models is the price (or the price variation) of cigarettes during the adolescence of the individual, that is to say when the individual was more likely to start smoking.⁷ We refrain from using this variable, since it would be the same for all individuals in the same age cohort, and thus reduce to a mere proxy for age.

Household specific variables include the family income adjusted for the number of people

⁷See for instance Douglas and Hariharan (1994).

in the household, according to the OECD-modified scale (assigning a weight of 1 to the household head, of 0.5 to each additional adult, and of 0.3 to each child), and a dummy variable taking value 1 if at least one child under the age of 15 belongs to the household. Note that the presence of children does not mean that these are the children of the reference individuals (they could be the children of just one of the two partners). For our purposes, this is not problematic, since we want to estimate the impact of the presence of children in the same household on the smoking behavior. Finally, we control for the number of children under age 15 living in the household.

4.2 Reduced Form Results on Smoking Behavior

In the first stage of the empirical analysis we estimate a reduced form model in order to recover the estimated probability of smoking of each individual and the empirical analogue of her expected utility.

We run a logit regression where the endogenous variable is the smoking dummy. We include in the explanatory variables individual characteristics of both the individual and her partner, if any. In particular we include the past smoking behavior, in the form of a dummy taking value one if the individual smoked two years ago. We use a two years lag to have more variation across present and past smoking status without having to deal with changing partners. Remark that the correlation across past and present behavior is very high, probably because of addiction. In particular, in our sample, the probability of being a smoker is equal to 25.9%. However this probability is equal to 85.6% for past smokers and to just 1.17% for people who did not smoked in the past. This piece of evidence shows that among adults quitting smoking is much more frequent than starting smoking. Since we control for the past smoking behavior, we will be able to look at the impact of spousal smoking on the probability of switching behaviors and in particular on the probability of smoking cessation.

We also include in the analysis the gender, the age, the education level and the current work exposure to smoke of the individual. Some household characteristics are also taken into account: for instance, we distinguish among individuals being singles and the ones living in couple. The number of children under 15 and the presence of at least one child under 15 are variables that are susceptible to influence smoking behavior. Finally, we include the family income, adjusted for the size of the household. Descriptive statistics for people living in couple and singles are reported in Tables 1 and 2.

[INSERT TABLE 1 HERE]

[INSERT TABLE 2 HERE]

The results of the reduced form logit regression are reported in Table 3. Heteroscedasticity-robust standard errors allowing for clustering at the household level are reported in parentheses.

[INSERT TABLE 3 HERE]

The current smoking behavior is strongly correlated with the past smoking behavior, as expected. The level of education correlates negatively with the probability to smoke, as well as the BMI for people living in a couple. The presence of children does not display any correlation with the probability of smoking. For couples, the past smoking behavior of the partner has a positive and highly significant effect on the probability of smoking, while the partner's level of education affects negatively the probability of smoking. These parameters, however do not correspond to the structural model parameters. As pointed out by Clark and Etilé (2006), the effect of the partner's past smoking behavior could be overestimated due to the exclusion of individual effects. The first step estimation is thus only instrumental to getting the estimated probability $\hat{\sigma}_i(x_i, x_j)$ that individual i smokes conditional on the

observables. For our purposes, it does not matter whether the partner's smoking behavior affects $\hat{\sigma}_i(x_i, x_j)$ directly or through the correlation of individual effects. All that matters is that $\hat{\sigma}_i(x_i, x_j)$ is an unbiased estimate of the equilibrium individual beliefs, whose effect on smoking decisions we want to test. In order to recover the structural parameters, we use the Hotz-Miller inversion and we obtain an estimate of the expected utility from smoking for each of individual i :

$$\Delta \widehat{EU}_i(x_i, x_j) = \ln \hat{\sigma}_i(x_i, x_j) - \ln(1 - \hat{\sigma}_i(x_i, x_j))$$

Remind that, under our assumptions, $\hat{\sigma}_j(x_i, x_j)$ is a consistent estimate of the beliefs that individual i holds about the probability of smoking of her partner. In equilibrium the beliefs coincide with the true probabilities. Thus, $\Delta \widehat{EU}_i(x_i, x_j)$ is the empirical analogue of the expected utility of individual i . We can then proceed with the second step of our analysis.

4.3 Structural Estimation of Smoking Behavior

Once an unbiased estimation of the expected utility from smoking has been recovered, it is possible to recover the parameters of the structural model under the sole moment condition discussed above, $E \left[\left(\Delta \widehat{EU}_i(x_i, x_j) - \lambda_1 x_i - \lambda_3 \hat{\sigma}_j(x_i, x_j) D_i \right) \right] = 0$. In particular, we obtain estimates of the coefficients $\lambda_1 = (\lambda, -p)$ and $\lambda_3 = p + r + s$. Both p and $r + s$ are identified.

Remark that the assumption that allows us to recover these parameters is that partner's characteristics x_j enter the individual expected utility function only through the beliefs about the partner's smoking behavior $\hat{\sigma}_j(x_i, x_j)$ and thus are excluded from the direct utility of the individual.

The results of the described regression are reported in Table 4. Bootstrap standard errors

are reported in brackets.⁸

[INSERT TABLE 4 HERE]

The past smoking behavior affects positively the utility of smoking. The corresponding parameter is strongly significant and quantitatively important. This is not surprising, since the smoking behavior is persistent over time. Remember that, on the one hand, only about 15% of past smokers succeeded in giving up smoking by the time of the interview. On the other hand, individuals that were non smokers two years before are very unlikely to start smoking. This accounts for the explanatory power of the past smoking behavior.

However, given the past smoking behavior, other variables significantly affect the present utility from smoking. Males seem to be more prone to smoke than women. The BMI is negatively correlated with the utility of smoking. This piece of evidence suggests that overweight people might be more concerned about the health damages of smoking; the fact that obesity and tobacco consumption could be substitute and not complements is not new to the literature (see for instance, Gruber and Frakes, 2006, and Flegal et al., 1995). Of course, the BMI is potentially an endogenous variable. Smoking might substitute excessive eating and affect individual weight. However, the causal relationship between BMI and tobacco consumption goes beyond the scope of this paper, and the cross-sectional nature of our data does not permit us to trace back the evolution of the BMI, which is the result of long term decisions. The level of education is negatively correlated with the expected utility of smoking. This is an intuitive result in line with previous findings from the literature (see for instance Kenkel et al., 2006). The exposure to smoke at work has no significant effect on the utility of smoking. The results also suggest that the utility from smoking decreases as family income increases, which is in line with previous findings. The presence of at least one child and the

⁸A bootstrap procedure is motivated by the fact that in the second stage we include estimated variables. The bootstrapping procedure was performed at the household level, in order to preserve the information concerning both partners in each bootstrap sample.

number of children in the household do not have any significant effect on individual smoking. This result is counter intuitive but quite robust to different specifications. Parents do not seem to perceive an extra cost of smoking with respect to non parents.

Let us now discuss the parameters corresponding to peer effects. On the one hand, the impact of the probability of smoking of the partner is positive and strongly significant. This effect corresponds in the theoretical model to $p + r + s$ and is estimated to be 0.55. On the other hand, living in a couple has a negative and significant (at the 95% confidence level) effect estimated to be -0.34. The absolute value of this parameter corresponds to the cost of smoking when the partner does not, p . We can thus identify $r + s$. If the partner has a probability to smoke close to one, the individual expected utility function increases of around 0.55 in absolute value with respect to the case in which it is very unlikely that the partner will smoke. In the model, the presence of single individuals permits to decompose this effect. With respect to a single, an individual with a partner who smokes gets an extra utility from smoking. This extra utility is modeled through the parameter $r + s$ and is estimated to be about 0.21. Conversely, with respect to a single, an individual with a non smoking partner gets a loss of utility if she smokes. This loss, corresponding to the parameter p in the model, is estimated to be equal to -0.34.

These findings suggest that the smoking behavior of the partner influences the utility from smoking in two distinct ways. First, if both partners smoke, this increases the benefit that they can both extract from tobacco consumption. This is a smoking enhancing effect linked to the presence of a smoker in the household. Second, if just one partner smokes, she will have to bear a cost due to the fact that she lives with a non smoker. This extra cost might be linked to the partner complaints, or to the internalization of some of the passive smoke externality. Furthermore, it could be due to the fact that a smoker matched with a non smoker might have to smoke outside the house or reduce her tobacco consumption.

Finally, note that the estimated $\lambda_3 = r + p + s$ is smaller than four in absolute value. As shown in Section 3, this ensures that the Bayesian Nash equilibrium that we have empirically characterized is indeed unique.

The coefficients listed in Table 4 give us a qualitative idea of spousal peer effects. However, the results refer to the impact on the utility from smoking which is an ordinal quantity. Thus, they do not allow us to quantify the magnitude of the peer effects. As pointed out before, these magnitudes are important in order to measure the impact of smoking containment measures. In the next section, we present the marginal effects of spousal smoking behavior on the individual probability of smoking.

4.4 Marginal effects

Given the estimated parameters, we now analyze the marginal effects of x_i or x_j (the characteristics of the partner of i) on the probability of smoking of an individual i . This probability is equal to

$$\sigma_i(x_i, x_j) = F(\Delta EU_i(x_i, x_j)),$$

where F is the logistic function satisfying $F'(u) = F(u)(1 - F(u))$ for all u .

Thus, we have

$$\frac{\partial}{\partial x_i} \sigma_i(x_i, x_j) = \sigma_i(x_i, x_j) [1 - \sigma_i(x_i, x_j)] \left[\lambda_1 + \lambda_3 \frac{\partial}{\partial x_i} \sigma_j(x_i, x_j) D_i \right]$$

for the marginal effect of an individual characteristic on its own likelihood to smoke, and

$$\frac{\partial}{\partial x_j} \sigma_i(x_i, x_j) = \sigma_i(x_i, x_j) [1 - \sigma_i(x_i, x_j)] \left[\lambda_3 \frac{\partial}{\partial x_j} \sigma_j(x_i, x_j) D_i \right]$$

for the marginal effect of a partner characteristic on the own likelihood to smoke.

Note that each individual characteristic has both a direct effect on the probability of smoking and an indirect effect through the partner's probability of smoking. The direct effect of x_i on the probability of smoking of i is $\lambda_1 \sigma_i(x_i, x_j) [1 - \sigma_i(x_i, x_j)]$. The indirect effect due to the implied change in the probability of smoking of the partner is equal to $\lambda_3 \sigma_i(x_i, x_j) [1 - \sigma_i(x_i, x_j)] \frac{\partial}{\partial x_i} \sigma_j(x_i, x_j)$. This indirect effect can either reinforce or reduce the effect of x_i on the probability to smoke.

Substituting the same expressions for $\frac{\partial}{\partial x_i} \sigma_j(x_i, x_j)$ and $\frac{\partial}{\partial x_j} \sigma_j(x_i, x_j)$ in the previous equations, after some rearrangements we get:

$$\frac{\partial}{\partial x_i} \sigma_i(x_i, x_j) = \frac{\lambda_1 \sigma_i(x_i, x_j) [1 - \sigma_i(x_i, x_j)]}{1 - \lambda_3^2 \sigma_j(x_i, x_j) [1 - \sigma_j(x_i, x_j)] \sigma_i(x_i, x_j) [1 - \sigma_i(x_i, x_j)] D_i},$$

and

$$\begin{aligned} \frac{\partial}{\partial x_j} \sigma_i(x_i, x_j) &= \frac{\lambda_1 \lambda_3 \sigma_i(x_i, x_j) [1 - \sigma_i(x_i, x_j)] \sigma_j(x_i, x_j) [1 - \sigma_j(x_i, x_j)] D_i}{1 - \lambda_3^2 \sigma_i(x_i, x_j) [1 - \sigma_i(x_i, x_j)] \sigma_j(x_i, x_j) [1 - \sigma_j(x_i, x_j)] D_i} \\ &= \frac{\partial}{\partial x_i} \sigma_i(x_i, x_j) [\lambda_3 \sigma_j(x_i, x_j) [1 - \sigma_j(x_i, x_j)] D_i]. \end{aligned}$$

The interaction effect increases the marginal effect of the own individual characteristics on the probability to smoke since

$$1 - \lambda_3^2 \sigma_j(x_i, x_j) [1 - \sigma_j(x_i, x_j)] \sigma_i(x_i, x_j) [1 - \sigma_i(x_i, x_j)] D_i < 1$$

whenever $|\lambda_3| < 4$, which is the case empirically.

We estimated the marginal effect of continuous variables for each individual, and we take the averages on the whole sample. Concerning the BMI, for a person measuring 1m70 the effect of a 3 kilos weight increase reduces the probability of smoking by a factor equal to 0.20 percentage points on average (ranging from -2% to 0). An increase in annual income of

10,000 euros corresponds to a decrease in the probability of smoking of 3.7 percentage points on average.

In Table 5 we report the average marginal effect for the peer effect parameters. The standard deviations are obtained by bootstrap. For individuals living in couple, the effect of the probability that the partner smokes $s + r + p$ is equal to

$$\frac{\partial}{\partial \sigma_j} \sigma_i(x_i, x_j) = \lambda_3 \sigma_i(x_i, x_j) [1 - \sigma_i(x_i, x_j)],$$

and is estimated to be equal to 0.02 for individuals living in couple. This effect ranges from 0 to 0.13 depending on the value of σ_i . This effect seems low, compared to the existing empirical literature; for instance, Cutler and Glaeser (2007) find that a smoking partner increases the individual probability of smoking of 40 percent. However, we are already controlling for the past smoking behavior for which we know that there is a strong path dependence. Thus, marginal effects on transitions from smoking to non smoking are generally low. In addition to that, the average marginal effect for individuals that smoked in the past (t-2) goes up to 6.4 percent. The interpretation of this number is that a smoker is about 6.4 percentage points less likely to give up smoking if her partner smoking probability moves from zero to one. Remember that the unconditional probability to quit smoking is equal to 15 percent. The estimated peer effect is therefore relatively important.

The marginal effect p of being in a couple on the probability of smoking is calculated as the difference in the probability of smoking of individual i when the couple dummy, D_i , passes from zero to one. Since for singles $\hat{\sigma}_j = 0$, in order to evaluate this marginal effect separately from the effect of $\hat{\sigma}_j$, one has to keep $\hat{\sigma}_j$ fixed and equal for all the individuals in the sample. First, we impose $\hat{\sigma}_j = 0$. The average marginal effect over the full sample is equal to -1.5 percent. This figure goes up to -4.1 percent for individuals who smoked two

years before. This result suggests that smokers living in a couple and expecting the partner not to smoke are 4.1 percentage points more likely to give up smoking than singles.

We also look at the marginal effect of being in couple imposing $\hat{\sigma}_j = 1$. The effect is much smaller than the previous one both for smokers and the full sample. As one would expect, the smoking deterring effect of living in a couple is smaller if the partner is expected to smoke with probability one. Living with a smoking partner increases the probability of smoking by a number (corresponding to $s+r$ in the theoretical model) equal to the difference between the marginal effect of $\hat{\sigma}_j$ (corresponding to $s+r+p$) and the marginal effect of the couple dummy calculated at $\hat{\sigma}_j = 1$ (corresponding to $-p$). For past smokers, this increase is estimated to be equal to 1.7 percentage points.

[INSERT TABLE 5 HERE]

The results from this section suggest that a non smoking partner might reduce the utility from smoking. This effect might work through three channels. First, the non smoker could impose a cost on the smoker, sometimes succeeding in having the latter quit. For instance, he could require the smoker to smoke less or to leave the house in order to smoke (consumption externality). Second, the individual could be altruistic towards a non smoking partner, and voluntarily impose on himself the transportation costs of smoking outside or the costs related to guilt. Finally, the non-smoking status of the partner might convey information on the harmful effects of smoking, reducing the pleasure of tobacco consumption. If we were able to observe the qualitative aspects of the smoking behavior (smoke in/out, for instance), it would be possible to disentangle learning effects from consumption externalities and altruism, since the latter affect the behavior of the smoker. We do not observe these qualitative variables. However, in the next section, we consider partners with children and analyze the effects of parental smoking on the health of their children. If the deterring effect identified in this

section reduces the individual utility through an obligation for the smoker not to smoke at home, we should find that, everything else equal, children whose both parents smoke are more exposed to passive smoking than children with a non smoking parent. Thus, smoking externality and altruism seem to play a role in smoking spousal peer effects.

4.5 Gender specific peer effects

In a couple, male and female partners might be affected in different ways by the behavior of the partner. In this section we consider a different setting where the structural parameters p , r and s depend on gender. We denote the gender by $g \in \{f, m\}$ where m refers to male individuals and f to female ones. The utility of individual i of gender g equals:

$$\begin{aligned}
 U_{ig}(a_i, a_j, x_i, \epsilon_i) &= a_i [\lambda_g \tilde{x}_i + \epsilon_i + s_g a_j D_i - p_g (1 - a_j) D_i] - [1 - a_i] r_g a_j D_i \\
 &= a_i [\lambda_g \tilde{x}_i - p_g D_i + \epsilon_i] - r_g a_j D_i + a_i a_j [s_g + p_g + r_g] D_i \\
 &= \lambda_{1g} a_i x_i + \lambda_{2g} a_j D_i + \lambda_{3g} a_i a_j D_i + a_i \epsilon_i,
 \end{aligned}$$

where $x_i = (\tilde{x}_i, D_i)$, $\lambda_1 = (\lambda_g, -p_g)$, $\lambda_2 = -r_g$, $\lambda_3 = p_g + r_g + s_g$. The parameters s_g , r_g and p_g are expected to be non negative for $g = f, m$.

We apply the same estimation method as in the previous sections. In the first stage regression (the results are omitted), we allow the parameters to be different for men and women. In the second stage we estimate the gender specific peer effects, by interacting the variables of interest, $\hat{\sigma}_j$ and Couple, with a dummy taking value one if the individual is male. The results of the analysis are reported in Table 6.

The results for the individual characteristics are very similar to the ones reported in Table 5, with the important exception of the gender dummy. Concerning the peer effects, however there are some noticeable differences. Women seem to feel a higher peer pressure.

In particular, the so-called smoking enhancing effect is greater for women than for men. The difference is statistically significant. In order to have a sense of the magnitude of these different effects, we calculated the average marginal effect of a smoking partner on the own probability of smoking. For both men and women, the probability of smoking is 2.1 percentage points higher if their partner smokes (the standard deviation is equal to 0.0060 and 0.0068 , respectively).

The difference between men and women appears when computing the average marginal effect of σ_j for the subsample of past smokers. This is very high for women: a woman living in a couple is 6.8 percentage points less likely to give up smoking if her partner's smoking probability goes from zero to one (with standard deviation equal to 0.0192). For past smoking men, this figure is only of 6.2 percentage points (standard deviation of 0.0183). Under our interpretation, these results suggest that women get more pleasure from smoking with their partner, or internalize more the externality imposed on their partner. This results are in line with the findings of McGeary (2003), which shows that women are more likely to quit smoking if their partner suffers a negative health shock. Concerning the smoking deterring effect, represented by the parameter p , the results suggest that there are no significant differences between men and women.

The gender dummy has no significant explanatory power in this specification, while in the previous one it was significant. This suggests that gender affects the smoking behavior only through the interaction with the expectations on the partner's smoking behavior.

[INSERT TABLE 6 HERE]

4.6 Children's health and parents' behavior

In this section, we study the relationship between the smoking behavior of parents and the number of non chronic respiratory diseases of their children. We only include households in

which the reference person lives in a couple. We also limit the analysis to the case in which the reference person is a man (only 10 households are removed) and is between the age of 25 and 60. There are 5,274 such households. For simplicity, in the following we will define the reference person as the husband and his partner as the wife, even though we do not control for their effective marital status. Overall, 2,791 such households have at least one child under the age of 15; 4853 children live in these families. In these household, 670 husbands and 831 wives reported being smokers. In 374 cases both parents smoke.

[INSERT TABLE 7 HERE]

We regress the number of non chronic respiratory diseases on a set of individual variables, such as the age, the gender, the BMI of the child and a dummy equal to one if she is affected by chronic respiratory condition; we also include in the regressors the number of children under 15 present in the household. The education level of both the mother and the father and their BMI are also taken into account, since these characteristics may influence the way parents care about the health of their children. Finally, we include two dummy variables summarizing parents' smoking behavior: we distinguish the case in which both parents smoke and the one in which just one parent smokes.

In our data set the number of sicknesses and the level of health care consumption of children is reported by parents. This can lead to measurement problems if parents tend to misreport the health status of their children. In the case of respiratory diseases, this difficulty seems to be particularly strong since some parents might report to the interviewer problems such as a light flu, while others would not. If this reporting bias is correlated with the variables of interest, this can lead to a misinterpretation of the coefficients. For instance, if smoking parents tend to underestimate the health problems of their children, we might find no effect of smoking on the non chronic respiratory diseases of children, even though

passive tobacco exposure actually affects their health. In order to reduce this bias, we control for the total number of non respiratory diseases and the total number of doctor visits for health matters different from the respiratory ones. These variables are useful to control the global health of the child and might capture the reporting bias of parents, if the latter is systematically the same for any health condition.

Another problem to take into account is endogeneity. On the one hand if a parent observes that her children are often sick, this may affect her smoking behavior. On the other hand, even though we control for chronic respiratory conditions, which have a high chance to affect both the parents' behavior and the occurrence of non chronic respiratory diseases, there could still be some correlation between our measures of parents' tobacco consumption and some unobservables affecting the occurrence of non chronic respiratory diseases. For instance, smoking parents might be less concerned with the health status of their children as they are less by their own health. This would lead to an overestimation of the effect of smoking on children's health. Regressing naively the health status of children on the smoking behavior of parents might thus lead to biased results. In particular, the impact of smoking on children respiratory health could be underestimated. We try to overcome this problem using instruments. We use as instrument the age of both parents and their exposure to smoke at the workplace. The age of parents does not appear to be correlated with the error term once one controls for the age of children. We think that it is quite reasonable to consider the exposure to smoke at the workplace as exogenous to the health care of children by parents. Evans *et al.* (1999) using U.S. data, showed that the sorting across jobs based on smoking status seems to be relatively weak. We perform an OLS and 2SLS regression. We run a Sargan test suggesting that the instruments are indeed valid (see Table 8).

[INSERT TABLE 8 HERE]

Since we are dealing with count data (the number of respiratory diseases in the period of interest ranges from 0 to 4), we also run a Poisson regression and a Poisson regression with control functions in order to account for endogeneity (using the same instrumental variables as in the 2SLS). The latter estimates are obtained by a technique suggested by Blundell and Powell (2003). We perform a first stage regression of the endogenous variables on all exogenous variables and instruments and we use the residuals and polynomials of those residuals as additional control variables in the main regression. This seems to be a more appropriate estimation strategy, because of the zero inflated discrete distribution of the dependent variable.

Results of both the OLS and the IV regressions are reported in the first two columns of Table 8. Clustered standard errors at the household level are reported in brackets. The presence of chronic conditions seems not to influence the number of non chronic respiratory diseases. However, the total number of doctor visits and the number of diseases different than respiratory both present positive and highly significant effect on the endogenous variable. As pointed out before, this could be due to the fact that these variables are proxies for the general health of the child. However, their positive impact might be linked to a systematic reporting bias of the parents. The age of the child affects positively the respiratory health of the child, as expected. The square of the age, the gender and the BMI have a small or insignificant effect in both specifications. The variable controlling for the number of children in the household presents a negative estimated parameter. This result seems counter intuitive since one could expect the presence of siblings to increase the risk to fall sick. The father's education level seems to have a positive correlation with the respiratory diseases of children. The mother's education and BMI have a small effect which is significant effect only under certain specifications. The fact that the father's education increases the number of non chronic respiratory condition is at first sight counter intuitive too. One explanation could be

that more educated households tend to report more respiratory diseases. The family income has a positive and significant effect on the number of sicknesses.

The results on smoking variables are the ones presenting a particular interest in order to test our smoking model. In the OLS specification, the estimates suggest no significant effect of the smoking behavior of parents. In the IV regression, however, the fact that both parents smoke has a significant (at the 95% confidence interval) and positive effect on the respiratory diseases of children. Having just one parent smoking at home does not affect significantly the children's health. The impact of the fact that both parents smoke is quantitatively very high (0.668), considering that in the period of observation, the average number of non chronic respiratory diseases for children is equal to 0.3. If just one parent smokes, the effect of smoking on children seems to be null. This evidence supports the hypothesis that non smoking parents might have a control role inside the household, protecting children from passive tobacco exposure.

In the third and fourth columns of Table 8, we report the results of the Poisson regressions. For most variables the results of the least square regression are confirmed. However, the parameters corresponding to the smoking behavior of parents are not significant in this specification.

These results weakly point towards a control role of a non smoking parent. However, this first analysis does not take into account the number of cigarettes effectively smoked by each parents and thus lacks precision. We refine the analysis by looking at the effect of each cigarette smoked per day by household members. In particular, we constructed a variable corresponding to the daily tobacco consumption of the smoking parent when just one parent smokes, and a variable corresponding to the daily aggregate consumption when both parents smoke. No information concerning where these cigarettes are consumed is available. In particular, the available data does not allow us to know whether the parents smoke at home

or not. The estimates of the smoking game suggest that when an individual is in couple with a non smoker, this reduces its utility from smoking. Our interpretation is that this disutility might come from the fact that smokers living with non smokers and having kids may be forced to smoke less at home or leave the house in order to smoke. Thus, we expect to see a smaller effect of each cigarette on the respiratory health of children when just one parent smokes instead of both.

In columns (a) and (b) of Table 8 we report the results of the OLS and the IV regressions. In the OLS regression, no effect of the smoking behavior is found. In the IV specification, the respiratory health of children seems to be affected by each extra cigarette that a parent smokes only if the other parent is also a smoker. The parameter is significant at the 95% confidence level. In column (3) of Table 8, we also include among the instrumented variables the square of the total number of cigarettes smoked by parents (if one or both smoke). This variable is meant to control for non linear effects of the number of cigarettes smoked. More precisely, we want to test for the hypothesis that the number of cigarettes matters only if a critical number of cigarettes is consumed (or that the negative effect on health is convex). The results seem to reject this hypothesis. Controlling for this variable, the marginal cigarette smoked leads to an increase of 0.06 in the number of respiratory diseases when both parents smoke. Again, no effect is detected in households where just one parent smokes. In columns (c) and (d), we report the results of a Poisson regression and of a Poisson with control functions (for endogeneity). The results are qualitatively similar to the previous ones. The number of cigarettes smoked in households where both parents smoke has a positive impact on the number of diseases.

Summarizing, there is some evidence that non smoking parents exert a control role over smoking ones, in particular in protecting children from passive smoking. This points towards an interpretation of the empirical results of the smoking game in terms of consumption

externalities or altruism. For instance, having one parent smoking 20 cigarettes a day may be less detrimental than having two parents smoking 10 cigarettes a day because of the unobserved change in behavior concerning smoking inside or outside, the smoker with a non smoking partner being more likely to smoke outside.

[INSERT TABLE 9 HERE]

5 Conclusion

In this paper we analyze intra-couple peer effects on smoking. Our empirical results suggest that the smoking behavior of the partner influences the individual probability of smoking in two ways. If the partner smokes, the individual utility from smoking is enhanced. If the partner does not smoke, the individual utility from smoking is smaller than the one of a single. We interpret this result as the effect of smoking externalities. Smoking together might be more pleasant than smoking alone. However, smoking alone as a single is better than smoking when living with a non smoker. This might be due to the fact that the non smoker imposes a cost on the smoking partner. This cost may materialize in a different smoking behavior: for instance smoking outside the house.

We quantify the impact of spousal tobacco consumption on the individual probability of smoking. More particularly, we find that smokers living in couple and expecting the partner not to smoke are 4.3 percentage points more likely to give up smoking within a period of two years with respect to singles. This might be due to the fact that individuals anticipate some extra cost from smoking if their partner does not smoke. However, if the partner is likely to smoke, this situation is reversed, and smokers living in couple are 8 percentage points less likely to give up smoking with respect to singles. Having a smoking partner enhances the utility an individual can get from smoking. This effect is higher for women than for men.

Finally, we look at the impact of smoking on children respiratory diseases. There seems to be some effect of parental smoking on children health only when both parents smoke. This result is consistent with the evidence from the smoking game we modeled above, and in particular with the fact that, if only one parent smokes, she might be obliged by the partner to protect the children from passive smoke. In terms of policy, the results show the importance of peer effects within the family and the role of partner's behavior in quitting smoking. It suggests that incentives to smoking cessation should be given in priority to couples of smokers specially if they have children and even if they jointly do not smoke as much as singles.

References

- [1] Adda J. and F. Cornaglia, 2010, "The Effect of Taxes and Bans on Passive Smoking", *American Economic Journal Applied Economics*, 2 (1), 1-32.
- [2] Bajari P., H. Hong, J. Krainer and D. Nekipelov, 2010, "Estimating Static Models of Strategic Interactions", *Journal of Business and Economic Statistics*, 28 (4), 469-482.
- [3] Bartley, M., P. Martikainen, M. Shipley and M. Marmot, 2004, "Gender differences in the relationship of partner's social class to behavioural risk factors and social support in the Whitehall II study", *Social Science and Medicine*, 59(9): 1925-1936.
- [4] Becker G.S., 1974, "A theory of marriage. Part I", *Journal of political Economy*, 81, 813-846.
- [5] Blume, L.E., W.A. Brock, S.N. Durlauf and Yannis M. Ioannides (2011), "Identification of Social Interactions". in J. Benhabib, M. O. Jackson and A. Bisin (Eds.), *Handbook of Social Economics*. The Netherlands: North-Holland, 1B: 853-964.
- [6] Blundell R. and J. Powell, 2003, "Endogeneity in nonparametric and semiparametric regression models," in *Advances in Economics and Econometrics : Theory and Applications*, Dewatripont M., L.P. Hansen and S.J. Turnovsky (eds.), Eighth World Congress, Vol II, Cambridge : Cambridge University Press.
- [7] Bolin, K., L. Jacobson and B. Lindgren (2001), "The family as the health producer - when spouses are Nash-bargainers", *Journal of Health Economics*, 20: 349-362.
- [8] Bolin, K., L. Jacobson and B. Lindgren (2002), "The family as the health producer - when spouses act strategically", *Journal of Health Economics*, 21: 475-495.

- [9] Clark A.E. and F. Ettiliçe, 2006, “Don’t give upon me baby: Spousal correlation in smoking behavior”, *Journal of Health Economics*, 25 (5), 958-978.
- [10] Cutler D.M. and E.L. Glaeser, 2007, “Social interactions and smoking”, NBER Working Paper W13477.
- [11] Douglas S. and G. Hariharan, 1994, “The hazard of starting smoking: Estimates from a split population duration model”, *Journal of Health Economics*, 13 (2), 213-230.
- [12] Evans W.N., M.C. Farrelly and E.B. Montgomery, 1999, “Do workplace smoking bans reduce smoking?”, *American Economic Review*, 85 (5), 729-747.
- [13] Flegal K.M., R.P. Troiano, E.R. Pamuk, R.G. Kuczmarski and S.M. Campbell, 1995, “The Influence of Smoking Cessation on the Prevalence of Overweight in the United States”, *The New England Journal of Medicine*, 333 (18), 1165-1170.
- [14] Gruber J. and Frakes M., 2006, “Does Falling Smoking Lead To Rising Obesity?”, *Journal of Health Economics*, 25 (2), 183-197.
- [15] Harris J.E. and B. Gonziçepez Liçepez-Valçiçercel, 2008, “Asymmetric peer effects in the analysis of cigarettes smoking among young people in the United States, 1992-1999”, *Journal of Health Economics*, 27, 249-264.
- [16] Hotz V.J. and R.A. Miller, 1993, “Conditional Choice Probabilities and the Estimation of Dynamic Models”, *The Review of Economics Studies*, 60, 497-529.
- [17] Jones A.M., 1994, “Health, addiction, social interaction and the decision to quit smoking”, *Journal of Health Economics*, 13, 93-110.
- [18] Kenkel D., Lillard D. and Mathios A., 2006, “The Roles of High School Completion and GED Receipt in Smoking and Obesity” , *Journal of Labor Economics*, 24 (3), 635-660.

- [19] Kenkel D.S. and J.V. Terza, 2001, "The effect of physician advice on alcohol consumption: count regression with an endogenous treatment effect", *Journal of Applied Econometrics*, 16(2), 165-184.
- [20] Khwaja A., Sloan F. and S. Chung, 2006, "The effects of spousal health on the decision to smoke: Evidence on consumption externalities, altruism and learning within the household", *Journal of Risk and Uncertainty*, 32, 17-35.
- [21] Krauth B.V., 2006, "Simulation-based estimation of peer effects", *Journal of Econometrics*, 133(1): 243-271.
- [22] Lundborg P., 2006, "Having the wrong friends? Peer effects in adolescent substance use", *Journal of Labor Economics*, 25 (2), 214-233.
- [23] Mannino J.S., J.E. Moorman, B. Kingsley, D. Rose and J. Repace, 2001, "Health effects related to environmental tobacco smoke exposure in children in the United States: data from the third National Health and Nutrition Examination Survey", *Archives of Pediatrics and Adolescent Medicine*, 155, 36-41.
- [24] Manski, C, 1993, "Identification of Endogenous Social Effects: The Reflection Problem", *Review of Economic Studies*, 60(3): 531-542.
- [25] Manski C., 1995, *Identification problems in the social sciences*, Harvard University Press, Cambridge, MA.
- [26] McGeary K.A., 2013, "Spousal effects in smoking cessation: matching, learning, or bargaining?" NBER Working Paper W19274.

- [27] Powell L.M., J.A. Tauras and H. Ross, 2005, "The importance of peer effects, cigarette prices and tobacco control policies for youth smoking behavior", *Journal of Health Economics*, 24 (5), 950-968.
- [28] Soetevent, A.R. and P. Kooreman (2007), "A discrete-choice model with social interactions: with an application to high school teen behavior", *Journal of Applied Econometrics*, 22(3): 599-624.

Tables

TABLE 1. Individuals in couple. Descriptive statistics

Variable	Mean	Std. Dev.
Smoking	0.242	0.4286
Smoking two years before	0.278	0.4480
Gender (1: man, 0:woman)	0.500	0.5000
Age	41.907	9.7213
Education>High School	0.380	0.4854
BMI	24.670	4.1080
Work exposure	0.143	0.3499
At least one child under 15 years	0.529	0.49962
Number of children under 15 years	0.920	1.0491
Income	17,912	11,286
Observations	10,540	

TABLE 2. Singles. Descriptive Statistics

Variable	Mean	Std. Dev.
Smoke	0.335	0.4709
Smoked two years before	0.361	0.4804
Gender (1: man, 0:woman)	0.3910	0.4896
Age	42.596	9.8718
Education higher than High School	0.434	0.4957
BMI	23.931	4.3927
Work exposure	0.153	0.3604
At least one child under 15	0.195	0.3963
Number of children under 15	0.294	0.6943
Income	17,173	11,924
Observations	2,230	

TABLE 3. Smoking Behavior: Logit Model

VARIABLES	Parameter	(Standard error)
Smoking at year-2*couple	6.097***	(0.1243)
Smoking at year-2*single	6.798***	(0.2914)
Gender*couple	0.325***	(0.1142)
Gender*single	0.082	(0.2431)
Age*couple	0.081*	(0.0442)
Age ² *couple	-0.001*	(0.0005)
Age*single	0.057	(0.1080)
Age ² *single	-0.001	(0.0013)
Education higher than High School*couple	-0.352***	(0.1114)
Education higher than High School*single	-0.6882***	(0.2540)
BMI*couple	-0.235**	(0.0928)
BMI ² *couple	0.003	(0.0017)
BMI*single	-0.2244***	(0.008)
BMI ² *single	0.003***	(0.0012)
Work exposure*couple	0.118	(0.1225)
Work exposure*single	0.562*	(0.3163)
Partner smoking at year-2*couple	0.824***	(0.0995)
Partner's age*couple	0.007	(0.0106)
Partner's education>HS*couple	-0.280**	(0.1120)
Partner's BMI*couple	0.007	(0.0123)
Partner's work exposure*couple	-0.077	(0.1300)
Couple	0.888	(2.9193)
At least one child under 15*couple	0.3899	(0.2917)
At least one child under 15*single	-4955	(0.9296)
Number of children under 15*couple	-0.437	(0.2792)
(Number of children under 15) ² *couple	0.111**	(0.0618)
Number of children under 15*single	0.267	(0.9034)
(Number of children under 15) ² *single	-0.019	(0.1795)
Income*couple	-2.42e-05***	(8.93e-06)
Income ² *couple	2.31e-10**	(9.54e-11)
Income*single	-4.93e-05**	(1.99e-05)
Income ² *single	4.84e-10**	(2.04e-10)
Constant	-1.396	(2.5248)
Observations	12770	

TABLE 4. Expected utility of smoking (OLS)

VARIABLES	Parameter	(Standard Error)
Smoked at t-2	6.256***	(0.1184)
$\hat{\sigma}_j$ (r+p+s)	0.547***	(0.1239)
Couple (-p)	-0.339*	(0.1952)
Age	0.082*	(0.0435)
(Age) ²	-0.001	(0.0005)
Gender (1: man, 0 woman)	0.186**	(0.0870)
Education higher than High School	-0.534***	(0.0967)
BMI	-0.080***	(0.0117)
Work exposure	0.171	(0.1159)
At least one child under 15	-0.085	(0.1585)
Number of children under 15	0.065	(0.0751)
Income	-1.11e-05***	(4.02e-06)
Constant	-3.379***	(0.9293)
Observations	12770	

TABLE 5. Marginal peer effects

	Marg. Eff.	Std. Dev.
<i>All individuals</i>		
$\hat{\sigma}_j$ (couples only)	0.022	0.0008
Couple (non smoking partner)	-0.015	0.0080
Couple (smoking partner)	-0.009	0.0077
<i>Past smokers</i>		
$\hat{\sigma}_j$ (couples only)	0.064	0.0026
Couple (non smoking partner)	-0.041	0.0216
Couple (smoking partner)	-0.017	0.0204

TABLE 6. Expected utility of smoking/Gender differences (OLS)

VARIABLES	Parameter	(Standard Error)
Smoked at t-2	6.283***	(0.1205)
$\hat{\sigma}_j$ (r+p+s)	0.569***	(0.1655)
$\hat{\sigma}_j$ *Gender	-0.038*	(0.2117)
Couple (-p)	-0.367	(0.2246)
Couple*Gender	0.065	(0.2635)
Age	0.080*	(0.0431)
(Age)2	-0.001	(0.0005)
Gender	0.098	(0.2394)
Education>HS	-0.541***	(0.0978)
BMI	-0.079***	(0.0116)
Work exposure	0.170	(0.1168)
At least one child under 15	-0.091	(0.1601)
Number of children under 15	0.064	(0.0757)
Income	-1.11e-05***	(4.03e-06)
Constant	-3.953***	(0.9352)
Observations	12770	

TABLE 7. Children. Descriptive statistics

Variable	Obs	Mean	Std. Dev.
Non chronic respiratory diseases	4853	0.310	0.5782
At least one chronic condition	4853	0.084	0.2773
Doctor visits-else than respiratory	4853	0.401	0.8896
Sicknesses-else than respiratory	4853	0.932	1.0878
Age	4853	7.010	4.3335
Gender	4853	0.512	0.5000
BMI	4853	17.121	3.4891
Number of children under 15	4853	2.117	0.9244
Only one parent smokes	4853	0.268	0.4430
Both parents smoke	4853	0.133	0.3393
Cigarettes if one parent smokes	1301	14.862	7.9370
Cigarettes if both parents smoke	644	30.846	11.6197

TABLE 8. Children Respiratory Diseases (a)

VARIABLES	(a) OLS	(b) 2SLS	(c) Poisson	(d) CF Poisson
One parent smokes	-0.005 (0.0190)	-0.050 (0.3390)	-0.020 (0.0623)	2.420 (1.7390)
Both parents smoke	0.040 (0.0251)	0.668** (0.3240)	0.119 (0.0776)	1.793 (1.4660)
One chronic condition	-0.002 (0.0292)	-0.029 (0.0345)	-0.018 (0.0924)	-0.083 (0.1020)
Visits-else than respiratory	0.035*** (0.0109)	0.035*** (0.012)	0.045* (0.0270)	0.043 (0.0297)
Sicknesses-else than respiratory	0.047*** (0.0090)	0.047*** (0.012)	0.160*** (0.0271)	0.163*** (0.0333)
Age	-0.043*** (0.0074)	-0.042*** (0.0084)	-0.085*** (0.0228)	-0.085*** (0.0244)
(Age)2	0.001** (0.0005)	0.001* (0.0006)	-0.001 (0.0017)	-0.0004 (0.0018)
Gender (1: male, 0:female)	0.014 (0.016)	0.012 (0.0184)	0.040 (0.0520)	0.040 (0.0553)
BMI	-0.001 (0.0026)	-0.002 (0.0029)	-0.0042 (0.0088)	-0.0070 (0.0095)
Number of children under 15	-0.029*** (0.0090)	-0.027*** (0.0098)	-0.096*** (0.0302)	-0.077** (0.0315)
Mother's education higher than HS	0.022 (0.019)	0.049** (0.0236)	0.080 (0.0613)	0.144** (0.0693)
Father's education higher than HS	0.052*** (0.0200)	0.100*** (0.0300)	0.165*** (0.0622)	0.292*** (0.0877)
Mother's BMI	0.002 (0.0024)	0.006** (0.0031)	0.007 (0.0077)	0.017* (0.0093)
Father's BMI	0.0005 (0.0020)	0.0047* (0.0028)	0.002 (0.0066)	0.012 (0.0084)
Income	-3.30e-06*** (1.05e-06)	-1.92e-06 (1.44e-06)	-1.15e-05*** (3.63e-06)	-7.93e-06* (4.60e-06)
Constant	0.504*** (0.0850)	0.179 (0.1500)	-0.708** (0.2800)	-2.002*** (0.5110)
Observations	4,853	4,853	4,853	4,853
Sargan	.	1.165 (P val. 0.5584)	.	.
R^2	0.067	.	.	.

TABLE 9. Children Respiratory Diseases (b)

VARIABLES	(a) OLS	(b) 2SLS	(c) 2SLS	(d) Poisson	(e) CF Poisson
Cigarettes if one smokes	0.0002 (0.0011)	-0.010 (0.0138)	0.0392 (0.0350)	0.003 (0.0052)	0.120 (0.1060)
Cigarettes if both smoke	0.001 (0.0007)	0.022** (0.0098)	0.059** (0.0262)	0.007 (0.0067)	0.162** (0.0806)
(Cigarettes)2			-0.001 (0.0009)	-7.65e-05 (0.0002)	-0.005* (0.0027)
One chronic condition	-0.001 (0.0292)	-0.022 (0.0344)	-0.026 (0.0336)	-0.017 (0.0924)	-0.079 (0.0996)
Visits-else than respiratory	0.035*** (0.0109)	0.039*** (0.0127)	0.029** (0.0139)	0.045* (0.0270)	0.027 (0.0354)
Sicknesses-else than respiratory	0.047*** (0.0090)	0.042*** (0.0110)	0.059*** (0.015)	0.160*** (0.0271)	0.195*** (0.0454)
Age	-0.043*** (0.0074)	-0.043*** (0.0084)	-0.042*** (0.0082)	-0.086*** (0.0228)	-0.085*** (0.0235)
(Age)2	0.001** (0.0005)	0.001** (0.0006)	0.001** (0.0006)	-0.001 (0.0017)	-0.0004 (0.0018)
Gender (1: male, 0: female)	0.014 (0.0161)	0.008 (0.0188)	0.019 (0.0197)	0.041 (0.0520)	0.056 (0.0591)
BMI	-0.001 (0.0026)	-0.001 (0.0029)	-0.002 (0.0029)	-0.004 (0.0088)	-0.008 (0.0093)
Number of children under 15	-0.029*** (0.0090)	-0.029*** (0.0099)	-0.024** (0.0101)	-0.096*** (0.0302)	-0.073** (0.0321)
Mother's education higher than HS	0.023 (0.0194)	0.049** (0.0242)	0.035 (0.025)	0.082 (0.0614)	0.111 (0.0752)
Father's Education higher than HS	0.053*** (0.0198)	0.096*** (0.0296)	0.088*** (0.0294)	0.168*** (0.0623)	0.261*** (0.0865)
Mother's BMI	0.002 (0.0024)	0.006* (0.0035)	0.007** (0.0034)	0.007 (0.0077)	0.019* (0.0102)
Father's BMI	0.0005 (0.0020)	0.0049* (0.0030)	0.003 (0.0032)	0.002 (0.0066)	0.007 (0.0097)
Income	-3.33e-06*** (1.05e-06)	-2.20e-06* (1.28e-06)	-2.08e-06* (1.25e-06)	-1.14e-05*** (3.63e-06)	-8.09e-06** (4.01e-06)
Constant	0.505*** (0.0848)	0.210 (0.1560)	0.188 (0.1530)	-0.722*** (0.2790)	-1.566*** (0.4680)
Observations	4,853	4,853	4,853	4,853	4,853
Sargan	.	2.272 (P val. 0.3211)	0.002 (P val. 0.9676)	.	.
R^2	0.675