

# The impact of parents' health shocks on children's health behaviors

Sylvie Blasco\*   Eva Moreno - Galbis<sup>†</sup>   Jeremy Tanguy<sup>‡§</sup>

January 2023

## Abstract

In this paper we assess how two smoking-related parental health shocks, lung cancer and other smoking-related cancer diagnosis, affect the offspring smoking behavior depending on the timing of the health shock. We use two strategies to isolate the informational shock effect from the transmission effect. We first focus on individuals with parents having received a smoking-related cancer diagnosis and exploit heterogeneity in the age of the individual at the moment of the diagnosis. We then build a retrospective panel and use individual fixed effects to absorb the transmission effect. We find that receiving a parental diagnosis at the age at which the decision to smoke is about to be made reduces the long term probability of being a smoker and duration spells.

**Keywords**— smoking, family health shocks, intergenerational transmission of health behaviors

**JEL codes**— I10, I12, D8, D84

---

\*Le Mans University (France), GAINS-TEPP, IRA, IZA.

<sup>†</sup>Aix-Marseille Univ., CNRS, EHESS, Centrale Marseille, IRD, AMSE, Marseille, France. Email: eva.moreno-galbis@univ-amu.fr. The project leading to this publication has received funding from the French government under the “France 2030” investment plan managed by the French National Research Agency (reference :ANR-17-EURE-0020) and from Excellence Initiative of Aix-Marseille University - A\*MIDEX.

<sup>‡</sup>University Savoie Mont Blanc (France), IREGE.

<sup>§</sup>We thank Constances for providing data access. We thank the participants of the 2022 French Stata conference in applied econometrics, particularly, Jeffrey Wooldridge, Mathieu Lefebvre and Xavier D'Haultfoeuille, as well as Eve Caroli, the participants of the Gains internal seminar for useful comments.

# 1 Introduction

Smoking is related to serious chronic diseases and is a leading preventable cause of death.<sup>1</sup> To minimize these health hazards, public health policies raising awareness about the health risks of smoking and limiting access to tobacco-related products are implemented worldwide. If smoking rates have steadily decreased in the past 50 years, a significant share of the population, including teenagers, keeps smoking on a regular basis. In 2019, 14% of US adults and 18.4% of EU adults were smokers<sup>2</sup> and 10% of sixteen years old smoked at least one cigarette per day in Europe (OFDT, 2021). The question is therefore how to successfully keep reducing smoking prevalence and smoking initiation.

Recent general information anti-smoking campaigns were found to have limited impact on smoking (*e.g.* Byrne et al., 2015; Lillard et al., 2013; Palali and van Ours, 2019; Skurka et al., 2017). This may be due to the fact that people do not feel directly or individually concerned by the delivered message and therefore do not act in light of it.<sup>3</sup> If individuals dissociate the population risk and their own likelihood of experiencing the outcome, they may not perceive the health risk as salient or they may overestimate their own ability to avoid addiction and bad health outcomes. Information about own health risk or about the risks of relatives, friends and acquaintances, may serve as more personalized, and thereby more effective, informational shock. The impact is expected to be even stronger when the shock happens to parents or siblings, rather than to other social contacts, as shared genetics could make the offsprings view the health outcomes of their parents as foreshadowing of their own health.

Recent literature has failed though to provide evidence of a significant impact of smoking-related health events of social contacts, including parents, on the individual smoking behavior (Darden and Gilleskie, 2016; Li and Gilleskie, 2021). Our claim is that this puzzling result is due to the fact that the literature has neglected the role of the age of the children when the parent suffers the health event. We argue that it is crucial to look at the timing of these health shocks, as the associated informational shock effect may vary along the life cycle and with the individuals' stock of addictive capital, which depends on one's history of smoking. The younger the individual, the lower the addictive capital stock<sup>4</sup>, but also the fuzzier the expectations about future health. An informational shock could be more effective in changing smoking behavior when it happens at younger ages, as younger individuals have been smoking for a shorter period of time and would thereby be less addicted to cigarettes. However, an informational shock may not impact smoking behavior of adolescents and young adults because the youth would perceive the bad outcome as occurring in a distant future or would underestimate the addictiveness of cigarettes (Gerking and Khaddaria, 2012). All in all, the relative size of both the addictive capital stock effect (initiated by the parental behavior transmission effect<sup>5</sup>) and the informational effect changes along the life cycle. Both effects tend to increase

---

<sup>1</sup>Every year around the world 8 million deaths are related to tobacco (WHO, 2018, 2021).

<sup>2</sup>See <https://www.cdc.gov/tobacco/campaign/tips/resources/data/cigarette-smoking-in-united-states.html> and <https://ec.europa.eu/eurostat/fr/web/products-eurostat-news/-/edn-20211112-1>

<sup>3</sup>Another explanation would be that people are well aware of the risks, but that the perception of health risk is not the main driver in smoking decisions (Cutler and Glaeser, 2005; Khwaja et al., 2009). A number of studies show though that people have biased perception of the health risks associated with smoking (Krosnick et al., 2017; Slovic, 2001; Viscusi, 1990).

<sup>4</sup>Most smokers take up the habit before reaching the age of 20 (Holford et al., 2014; Lillard et al., 2013)

<sup>5</sup>There is considerable evidence that children of smokers are more likely to become smokers themselves (see

with age and we are unable to anticipate which one will be dominant.

We use the French population-based epidemiological cohort Constances to solve this indetermination and gain insights on the relative size of each effect at each stage of the offspring life. Our final objective is to isolate the informational shock effect to identify the impact of parental health events on the offspring smoking behavior. Identifying such an informational effect requires dealing with confounding factors that may explain both the parent’s diagnosis and the individual health-related behaviors. These confounding factors are likely to include in particular unobserved parental risky behaviors, given the intergenerational transmission of risky behaviors. Since we cannot account for selection into the parent’s diagnosis, due to lack of information at the time of – or before – the diagnosis, we first restrict the analysis to individuals whose parents experienced a health shock and thereby assess the impact of the timing of the shock. the identifying assumption here is that, conditional to the occurrence of the health shock and to observables, the timing of the parent’s health shock is exogenous to the individual. This identifying assumption may be invalid if the timing of the parents’ health shock is endogenously determined by the intensity of their smoking behavior. Therefore, we then control for time-invariant unobserved factors by constructing a retrospective panel from the individual smoking history and modeling the smoking probability at each age using individual fixed effects. Doing so, we aim to neutralize the transmission effect and to identify the informational shock effect of the parents’ diagnosis.

We find that, among individuals whose parents were diagnosed with a lung cancer, there is no significant effect of the age at diagnosis on the probability of having ever smoked. The informational shock effect seems to be offset by the addictive capital stock initiated with the behavioral transmission effect coming from parents. When looking at smoking duration among those who have ever smoked though, receiving a parental diagnosis between 15 and 18 years old, rather than later, significantly reduces the duration until smoking cessation. This is consistent with findings from the retrospective panel which confirm that individuals having received the parental diagnosis at the age when the decision to smoke is about to be made (15-18 years old) are the least likely to be smoker. Among individuals whose parents were diagnosed with another smoking-related cancer, the relative size of the addictive capital stock outweighs the informational shock effect, when the diagnosis is made before, rather than after, the age of 40. When focusing on those who have ever smoked, the timing of the diagnosis has no effect on the smoking spell duration. Put together, these results suggest that the informational shock of a lung cancer diagnosis is greater than that of another smoking-related cancer diagnosis. Results from the retrospective panel confirm a lower negative impact of other smoking related cancer diagnosis on the probability to be a smoker.

We contribute to the literature that investigates the determinants of adopting or quitting unhealthy behaviors. The bulk of this research is about risk perception and about intergenerational transmission of behaviors. Papers about intergenerational transmission aim at determining to which extent the observed positive correlation between parents’ and children’s health behaviors results from a causal impact of parents’ habits on the children or from confounding factors (see Darden and Gilleskie 2016; Göhlmann et al. 2010; Loureiro et al. 2010; Pan and Han 2017 for smoking). Here we are interested in the impact of the parent’s outcome (not behavior) that serves as an informational shock about the harmful consequences of smoking.

---

Jackson and Henriksen, 1997; de Vries et al., 2003; Göhlmann et al., 2010; Melchior et al., 2010 for evidence on smoking initiation among teens as a function of parental smoking behavior or Darden and Gilleskie, 2016 for evidence on the impact of parental smoking on adults’ offspring smoking behavior).

The literature has analyzed the impact of parents’ diseases or death on the offspring’s health and well-being (Cas et al., 2014; Johnston et al., 2013; Le and Nguyen, 2018), education and skills (Adda et al., 2011; Alam, 2015; Case and Ardington, 2006; Chen et al., 2009; Senne, 2014), cognitive and non-cognitive development and wealth (Arora, 2016; Le and Nguyen, 2017, 2018), but not so much on their health behaviors.

The existing literature on the impact of informational shocks about health risk on health risky behaviors has looked at the impact of informational shocks due to the outbreak of a disease (Adda, 2007; Jin et al., 2022), to health shocks of social contacts (Khwaja et al., 2006; Li and Gilleskie, 2021; Lin and Sloan, 2015; Sloan et al., 2003) and to own diagnosis (Darden, 2017; Smith et al., 2001). Overall, the literature finds limited impact of health shocks of social contacts, but significant decrease in smoking prevalence and intensity following one’s own diagnosis (Arcidiacono et al., 2007; Bünnings, 2017; Sloan et al., 2003; Smith et al., 2001; Sundmacher, 2012). To the best of our knowledge, only Darden and Gilleskie (2016) and Li and Gilleskie (2021) look at the impact of an informational shock induced by parents’ smoking-related illness on the children’s smoking behaviors. Both studies show limited impact.<sup>6</sup>

Our work extends these studies by considering past shocks instead of contemporaneous shocks (note that in Darden and Gilleskie, 2016 the authors consider both past and contemporaneous shocks). This extension is of importance here: as it may be difficult to quit smoking due to addiction, individuals may take time to act upon the informational shock and effectively quit smoking. Therefore, the impact of the informational shock could not be fully captured with a short window of observation and one need to consider a longer observation period as we do here. Moreover, by looking at the timing of the parents’ health shock, we allow the associated informational shock effect to vary over the life cycle. This is of importance for the anti-smoking public health policy. The literature that evaluates the anti-smoking public health policy has looked at the age-dependent effects of cigarette taxes but not so much of informational campaigns. Moreover, most studies about the impact of information shocks and risk perception about the risks of smoking use samples containing either only young or only old individuals, which gives limited insights about the differences in the effects across ages. Our results reveal significant effect heterogeneity with respect to age. They confirm that interventions at the time of smoking initiation are crucial and that information campaigns need to find a better way to communicate about the risks of smoking at younger ages. Our findings indicates that campaigns should be impactful and, for example, feature smoker teenagers who receive the news that their parents have been diagnosed with lung cancer.

The paper is organized as follows. Section 2 presents the databases, the variables used as well as descriptive statistics. The econometric strategy and results are presented in section 3. Section 4 concludes.

## 2 Data and descriptive statistics

**Constances cohort.** We use the Constances cohort which is a large population-based and general-purpose epidemiological cohort (see Zins et al. (2015) for a detailed description). Started at the end of 2012, it includes about 200,000 individuals. It was designed to be representative of the general French adult population aged between 18 and 69, excluding agricultural and self-employed workers.<sup>7</sup> Participation

---

<sup>6</sup>Hillebrandt (2022) evaluates among others how the death of a partner or a parent affects smoking behavior and body weight. He does not restrict the analysis to smoking-related deaths.

<sup>7</sup>Representativeness was in terms of age and gender.

was voluntary upon invitation sent to randomly selected individuals. In case they agreed to participate in the cohort, individuals were invited to go to a Health Screening Center (HSC) to undergo a health examination. They also had to complete at home a questionnaire about their health, lifestyle and parents' health. Participants were then asked to fill in self-administered follow-up questionnaires once a year and they had to update their health clinic examination every 5 years.

All randomly selected individuals invited to participate in the Constances project did not eventually participate and some participants dropped out, which challenges the representativeness of the sample. To deal with non random participation, individuals aged between 18 and 69 living nearby a partner HSC were drawn randomly by stratified sampling with unequal probabilities, over-representing individuals with a higher probability of non-volunteering according to age.

We compared the age distribution of Constances sample with the age distribution of the sample provided by the French Labor Force Survey (see the Online Appendix). Both the youngest and the oldest age categories are slightly under-represented in Constances, while intermediate age categories are over-represented even if the relative composition of these intermediate categories corresponds well to the one found in the representative French Labor Force Survey.<sup>8</sup>

The socio-demographic composition of the Constances sample is summarized in the Online Appendix.

**Behavior and health-related outcomes of interest.** The Constances data contain rich information about health-related outcomes and behaviors. Past health-related behaviors and previous health diagnosis are collected at inclusion in the cohort. Information about self-assessed general health, own diagnosis and current health-related behaviors are collected at inclusion and is updated annually. As for the parents, we do not have information about their health-related behaviors, but we know if they developed a smoking-related cancer. Information about the parents' diagnosis is collected at inclusion and is not updated in the follow-up questionnaires. As a result, we only use the inclusion questionnaires for our analysis.

Table 1 summarizes the main smoking behaviors of individuals at inclusion (Panel A), among those who have ever smoked (Panel B) and among those who no longer smoke (Panel C).<sup>9</sup> About half of the sample has ever smoked and almost one fifth smokes at the time of inclusion. On average, smokers started around age 17 and smoke almost 12 cigarettes a day. The average quitting age is 34.<sup>10</sup> In the Online Appendix we show that, at inclusion, smokers and non-smokers have statistically different socio-demographic characteristics. This also holds when distinguishing by gender.

---

<sup>8</sup>Imputing to the Constances sample the average weights computed by gender, nativity and 5 age categories from the French LFS, leads to an over-representation of the older age category and does not improve the representation of the youngest age category of the Constances sample with respect to the LFS sample (see the Online Appendix). We thus keep the original sample without implementing any counterfactual reweighting procedure.

<sup>9</sup>See Table A.1 in appendix A for an analysis by gender.

<sup>10</sup>See the distribution of the starting and stopping ages of smoking in Figures A.1 and A.2 in Appendix A.

**Table 1:** Sample composition - Smoking Behaviors

	%	N
<b>Panel A: Smoking at inclusion</b>		
Ever smoked	0.523	96,014
Smoker	0.185	33,923
<b>Panel B: Among ever smoker</b>		
Smoker	0.353	33,923
Age at smoking onset	17.438	96,014
Number of cigarettes per day	11.975	96,014
<b>Panel C: Among no longer smokers</b>		
Age at smoking onset	17.349	62,091
Number of cigarettes per day	12.827	62,091
Age at smoking cessation	34.405	62,091

Source: Constances (2012-2018).

Concerning parents' health, we consider whether either parent has been diagnosed with lung cancer or another smoking-related cancer. The former covers lung cancer, but also bronchial and thoracic cancers.<sup>11</sup> Other smoking-related cancers, include, according to the French Institute of Cancer, tongue, throat, nasal, esophageal, laryngeal, otorhinolaryngology, ureter, bladder, breast, uterus, kidney, liver, blood, colon, ovarian, pancreatic, stomach, pharyngeal and bone marrow cancers. We distinguish between these two types of diagnosis to test whether cancers which are clearly associated with smoking in the opinion (*i.e.* lung cancers) have stronger impact than other smoking-related cancers.

4% of the sample at inclusion has parents that were diagnosed with lung cancer and 21.1% with another smoking-related cancer (Table 2).<sup>12</sup>

**Table 2:** Proportion of the population with parents having cancer

	%	N
Parents with lung cancer	0.040	7,411
Parents with other smoking-related cancer	0.211	38,634

Source: Constances (2012-2018).

The proportion of people who have ever smoked is higher among people whose parents have been diagnosed. (Table 3) In contrast, the proportion of current smokers is lower for this subgroup. Smokers whose parents have been diagnosed smoke though more cigarettes per day than smokers whose parents have not been

<sup>11</sup>Smoking causes 90% of lung cancers for men (70% for women). The remaining lung cancers are mainly due to exposure to radon, pollution and asbestos, with the risks of developing a lung cancer due to such exposures increasing in case of smoking.

<sup>12</sup>Table A.2 in Appendix A presents the analysis by gender. Figure A.3 in Appendix A displays the distribution of the age at which parents of the interviewee were diagnosed with lung cancer (Panel A) or another smoking-related cancer (Panel B).

diagnosed.<sup>13</sup>

**Table 3:** Smoking behavior by parental diagnosis (%)

	Parents with ...					
	lung cancer			other smoking-related cancer		
	No cancer	Cancer	P-value	No cancer	Cancer	P-value
Ever smoker	0.521	0.572	0.000	0.517	0.544	0.000
Current smoker	0.185	0.168	0.000	0.191	0.162	0.000
Number of cigarettes per day	11.891	13.696	0.000	11.817	12.531	0.000
Own diagnosis	0.022	0.027	0.003	0.019	0.033	0.003

Source: Constances (2012-2018).

### Computing the age of the interviewee when the parents were diagnosed with cancer.

In the Constances data, we know the age of the parents at the time of the diagnosis but not the age of the interviewee at the time of the parent's diagnosis. We compute a proxy of this age by exploiting the average age of parents at the birth of their children available in different data sources. For workers, we compute by gender and by occupation the average age of parents at the birth of children in the first, second, third and higher ranks using the DADS-EDP panel.<sup>14</sup>

For not-working people, not included in the DADS-EDP panel, we use the 1993 French Labor Force Survey (LFS) to compute by gender the average age at the birth of the first child.<sup>15</sup> On the LFS data, we compute the average age of individuals not working and declaring having a child below 3 years old.<sup>16</sup> We subtract 3 years from this average age, so as to compute a lower bound of the age at the birth of the first child. We compute by gender the average age in 1993 at the birth of the first child for the whole population using the same procedure. We then compute by gender the standard deviation in the age at the birth of the first child between not-working people and the whole population. Fourth, we use DADS-EDP panel and compute for every year between 1968 and 2000 the average age at the birth of the first child by gender and apply the standard deviation computed for not-working people in 1993. This allows us to compute the approximate age at the birth of the first child for not-working men and women over the 1968-2000 period.

Since Constances provides information on the occupation of the interviewee's parents and on the rank of the interviewee within her siblings, we can impute for each interviewee the approximative age of her parents

<sup>13</sup>See the Online Appendix for the analysis stratified by gender.

<sup>14</sup>These data result from the merge between The *Déclarations Annuelles de Données Sociales* (DADS), a representative administrative panel of French salaried workers, and the *Echantillon Démographique Permanent* (EDP), a socio-demographic panel based on the population census since 1967.

<sup>15</sup>The 1993 LFS wave is the earliest wave in which we had the required information to compute the age of the parents when their first child was born.

<sup>16</sup>Information in the LFS is less detailed than in the DADS-EDP panel. We only know the total number of children and if the individual has a child below 3 years old. We consider the population sample declaring having a child below 3 and for whom the number of children below 3 equals the number of children below 18. This makes it possible to isolate the first child, except in rare cases where the age difference within the fraternity is more than 15 years.

at the time of their birth. Then, given that we have information on the age at which the parents were diagnosed with cancer, we can proxy the age of the individual at the time of the parent’s diagnosis.

### 3 Empirical strategy and estimation results

We aim at evaluating the impact of an informational health shock represented by parent’s cancer diagnosis (lung cancer or other smoking related cancer) on the offspring’s smoking behavior. Identifying such an impact requires controlling for the parents’ health-related behaviors that are correlated to both parent diagnosis and individual health-related behaviors. The probability of being diagnosed is indeed higher for parents adopting health risk behaviors. Moreover, due to the intergenerational transmission of behaviors, individuals are likely to adopt the same habits as their parents. Put differently, individuals with and without a parent diagnosed are likely to differ in family background and parental health behavior. Not accounting for this selection in the parent’s diagnosis could strongly downward bias our estimate of the impact of the informational shock (*i.e.* health shock), as the possibly limiting effect of the diagnosis on risky behaviors would be counterbalanced by the fact that people who are more likely to adopt risky behaviors are also more likely to have a diagnosed parent. Unfortunately, the data at hand do not provide information on the smoking behavior of the parents or on the individual at the time of the diagnosis.

To address this issue, we proceed in two steps. We first restrict the analysis to the sample of individuals whose parents have been diagnosed with lung cancer or another smoking-related cancer. Doing so, we perform the analysis on a more homogeneous sample. Implicitly we assume that the transmission effect will be homogeneous among people whose parents were diagnosed with cancer. This limits the selection bias. This strategy implies that we evaluate the impact of the timing of the diagnosis (given that a parent got diagnosed), instead of the impact of having a parent diagnosed. Our identifying assumption is that conditional on having a parent diagnosed, the timing of the diagnosis is exogenous to the individual.

This identifying assumption may be questionable, as the timing of the diagnosis may be influenced by the intensity of the risky behavior, which influences itself the size of the transmission effect. We therefore propose in a second step an alternative approach that exploits the individual smoking history to build a retrospective panel and allows to implement an estimation controlling for individual fixed effects. This approach absorbs the individual (time-invariant) transmission effect, which allows us to isolate the informational shock effect.

#### 3.1 First approach: evaluating the effect of the timing of the diagnosis

We estimate the following econometric model on a sample of individuals whose parents where diagnosed with cancer:

$$y_i = \beta_0 + \sum_k \delta_k ParentDiag_{k,i} + X_i \beta_1 + \beta_2 OwnDiag_i + \gamma_t + \gamma_t^2 + u_i$$

where  $ParentDiag_{k,i}$  is an indicator taking value 1 if one of the individual’s parents got diagnosed a health shock when s/he was aged  $k = \{< 15; 15 - 18; 19 - 24; 25 - 29; 30 - 39; 40 - 49\}$ . The  $\delta_k$  coefficients measure the effect of having a parent’s health shock at age  $k$  instead of after 49.<sup>17</sup> We control for several individual attributes, including whether the individual was diagnosed herself with a smoking-related cancer

---

<sup>17</sup>In case of multiple diagnosis, we consider the first one.



at some point before the time of the interview. We also control for the individual’s mental health<sup>18</sup> and her subjective perception of her general health status. Additionally, we include standard socio-demographic variables (gender, age, cohabitation status, place of birth, education, employment situation and socio-professional category). We also control for a time trend in a flexible form to account for the action of changing awareness and cultural norms with respect to smoking along time (due to public health policies for instance).<sup>19</sup>

We consider two different outcomes to describe different aspects of smoking behavior:  $y_i$  may be (i) an indicator taking value 1 in case the individual has ever smoked and 0 otherwise or (ii) the number of cigarettes smoked per day. For these two outcomes we run OLS. Additionally, we look at smoking duration for those who have ever smoked. To account for possible censoring of the duration from the first to the last cigarette (in case the individual is still smoking at the time of the interview), we implement a duration analysis. We estimate non-parametric Cox models to assess the impact of the timing of the diagnosis on the hazard until quit, *i.e.* on the yearly probability of quitting.

Given that we cannot control for the parents’ smoking behavior, the parameters of interest,  $\delta_k$ s, capture both the effect of the parent’s health-related behavior (transmission effect at the young age, which leads then to cumulate an addictive capital stock) and the effect of the parent’s diagnosis (informational shock effect). Our guess is that the intensity of the two contradictory effects increases throughout the children’s life cycle, leading to a permanently ambiguous effect. Concerning the transmission effect, we expect that the longer children are exposed to the smoking behavior of their parents, the more likely they are to start smoking. Moreover, the longer children have been smoking, the more difficult is to quit because of addiction. Therefore, the later the diagnosis, the more likely children are to smoke due to the greater transmission effect and because stronger addiction limits the individual’s ability to react to an informational shock. But the effect of the informational shock coming from a parent’s diagnosis could also increase along the life cycle through an identification effect (*i.e.* personalization effect) since the age of onset for lung cancer and other smoking-related cancers is essentially over 40 (see Figure A.3 in Appendix A). Moreover, the informational content of the health shock may differ depending on the type of cancer the parent is diagnosed. We distinguish between lung cancer and other smoking-related cancers. Assuming that the transmission effect is the same for both types of cancers, differences in the  $\delta_k$  coefficients associated with lung cancer and those associated with other smoking-related cancers are indicative of the importance of the informational shock effect of both types of cancers. In case individuals associate more smoking with lung cancer than with other smoking-related cancers, the informational shock effect should be larger with the diagnosis of lung cancer than with the diagnosis of other smoking-related cancers.

### 3.1.1 Smoking behavior and the timing of the parent’s diagnosis

**Cigarette smoking and smoking intensity.** The impact of the age at which parents were diagnosed on the offspring smoking behavior is summarized in Table 4. Among those who have a parent diagnosed with lung cancer (Panel A), the age of diagnosis is not significantly associated to the probability of having ever smoked. When we look at the intensive margins among smokers, individuals receiving the parental diagnosis

<sup>18</sup>Our mental health indicator is the CES-D scale. See Morin et al. (2011) for further details on this indicator.

<sup>19</sup>We include a linear trend and a non-linear trend on the date of inclusion as well as a trend capturing the date of birth.

before the age of 50 smoke significantly more cigarettes than individuals receiving the parental diagnosis after 49. These results suggest that the relative sizes of the information and addictive capital stock effects do not differ for individuals receiving the parental diagnosis before and after the age of 50 when considering the probability of ever smoking, but that the informational effect significantly dominates more for individuals receiving the parental diagnosis after 49. These individuals smoke significantly less cigarettes than those receiving the diagnosis when they were younger, even if they do not differ in the probability of having ever smoker. Therefore, the impact of the parental diagnosis differs between the extensive margin (decision to smoke or not) and the intensive margin (number of cigarettes). Note that when we look at the intensive margins, we work on a sample of individuals who keep smoking despite the diagnostic received by their parents and who are therefore less likely to be responsive to informational shocks.

The timing of the diagnosis matters however for other smoking-related cancers (Panel B): individuals whose parents were diagnosed when they were less than 39 years old are more likely to have ever smoked and smoke significantly more cigarettes than people who were above 49 at the age of the parental diagnosis (people who were between 40 and 49 years do not significantly differ in their smoking behavior with respect to the reference category). These estimations suggest that the informational effect associated with a smoking related cancer is significantly stronger for people receiving the parental diagnosis after 40 years old.

The differences between Panels A and B indicate that the two types of diagnosis do not have the same informational power about the health risks of smoking. Lung cancer appears to be more informative than other smoking-related cancers. This may be due to the fact that the general opinion about lung cancer is overall clearly associated with smoking, when the connection between smoking and the other types of cancer may be less understood. This calls for a better information about different health hazards of smoking.

**Table 4:** Influence of the age at which the individual's parents were diagnosed with cancer on the individual's smoking behavior.

	(1)	(2)
	Having ever smoked	Number of cigarettes per day
<b>Panel A: Lung cancer</b>		
Under 15 years old	0.010 (0.048)	1.988* (1.133)
15-18 years old	0.011 (0.045)	1.226 (0.984)
19-24 years old	0.021 (0.032)	2.039*** (0.781)
25-29 years old	0.012 (0.030)	1.712** (0.740)
30-39 years old	0.026 (0.025)	1.688*** (0.625)
40-49 years old	0.022 (0.025)	1.106* (0.652)
Observations	4658	2389
R-squared	0.059	0.101
<b>Panel B: Other smoking-related cancer</b>		
Under 15 years old	0.042** (0.016)	1.204*** (0.408)
15-18 years old	0.031* (0.017)	0.724* (0.381)
19-24 years old	0.041*** (0.013)	0.659** (0.323)
25-29 years old	0.033** (0.013)	0.134 (0.306)
30-39 years old	0.022** (0.011)	0.463* (0.270)
40-49 years old	0.009 (0.011)	0.027 (0.276)
Observations	24579	11481
R-squared	0.045	0.105
Trends (time, time2, occupation)	Yes	Yes
Department FE	Yes	Yes
Demographics	Yes	Yes
Educational attainment	Yes	Yes
Employment-related controls	Yes	Yes
Parents occupation	Yes	Yes
Health-related controls	Yes	Yes
Work-related controls	No	No
Sample	All	All

Source: Constances (2012-2018). Notes: Robust standard errors in parentheses. Statistical significance:

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

**Heterogeneity by gender and by cohort.** We replicate the analysis by gender, as the literature highlights the possibility of gender-specific smoking and transmission patterns (see section B of the Online Appendix). In contrast with Table 4, we find that for males lung cancer and other smoking related cancers do not essentially differ on their informational power. Findings for women are more consistent with results from Table 4, with the major different that when considering lung cancer, there are no significant differences across people receiving the parental diagnosis at different ages whether we consider the extensive or the intensive margin.

We also analyze heterogeneous effects with respect to the birth cohort. The motivation for this analysis is that the individual’s perceptions of the health risk associated with smoking and of the connection between smoking-related disease and smoking are likely to have evolved along time thanks to scientific advances and public health campaigns. Our main specification includes a linear trend on the age of birth, allowing to account for such trends, but it does not allow the transmission and informational shock effects to vary across cohorts. To do so, we augment our main model with interaction terms between the parents’ diagnosis dummies and the indicator of belonging to one of four birth cohort groups.<sup>20</sup> For a given age at the moment of the parents’ diagnosis (lung or other smoking-related cancers), younger generations are more likely to decrease their risky behavior than older generations (see section C of the Online Appendix). This holds particularly for cohorts located in the second and third quartile of the age distribution, since for the first quartile we find that the average decrease (with respect to the fourth quartile) in the probability of ever smoke is eventually compensated by the effects associated to the particular age of the diagnosis. This suggests a relatively stronger informational shock effect among younger generations, which have been actually exposed to more public health campaigns on the consequences of smoking.

### 3.1.2 Smoking duration

We now analyze the impact of the timing of parent diagnosis on the survival time from smoking onset to smoking cessation. We restrict the analysis to individuals who have ever smoked and use information on the ages at which they start and eventually quit smoking. As the smoking spell may be censored if the individual has not quit when she enters the Constances cohort, we use duration analysis and estimate non-parametric Cox models. These estimations complement the analysis presented in the previous section, where we focused on the probability of have ever smoked and the number of cigarettes. While the parental diagnosis may not be associated to these two smoking behavior indicators differently depending on the timing of the diagnosis, there could be an association with the duration during which the offspring keeps smoking when receiving the parental diagnosis.

We include demographic and educational controls, department fixed effects, as well as fixed effects for year and age of smoking onset. Results are displayed in Table 5. The first two columns consider lung cancer and the last two columns other smoking-related cancers. From column (1), we conclude that having a parent diagnosed with lung cancer when the individual is between 15 and 18 is significantly associated with higher instantaneous probability of quitting smoking, leading to shorter smoking duration, compared to the situation where the individual is above 49 at diagnosis.<sup>21</sup> This indicates that the informational shock

---

<sup>20</sup>To create the birth cohort groups, we compute the age distribution over the whole sample and define four quartiles of the age distribution.

<sup>21</sup>For people between 40 and 49 years old, the coefficient is significant at 10%.

effect of a parental lung cancer diagnosis is relatively more important than the transmission together with addiction effects for this age category than for the reference category (*i.e.* people receiving the parental diagnosis when they are above 49 years old). This pushes people between 15 and 18 years old to have shorter smoking durations than the reference category and the other age categories of parental diagnosis. In contrast, the age of the parent’s diagnosis for other smoking-related cancers does not promote significantly different smoking quitting behaviors between the reference category and the other age categories of parental diagnosis (see column (4) in Table 5). Smoking duration is not significantly associated with the age of the parental diagnosis when considering other smoking related cancers, suggesting that the informational content of this cancer type is smaller with respect to lung cancer. Columns (2) and (4) focus on people who started smoking before the diagnosis. We find the same pattern of results as for the entire sample of smokers, which also includes people who started smoking after diagnosis.

**Table 5:** Influence of the age at which the individual’s parents were diagnosed with lung cancer or other smoking-related cancer on the individual’s smoking duration: all smokers

Hazard ratios	Lung cancer		Other smoking-related cancer	
	(1)	(2)	(3)	(4)
Under 15 years old	1.046 (0.173)	1.105 (0.804)	0.961 (0.051)	0.802 (0.192)
15-18 years old	1.720*** (0.258)	1.689*** (0.291)	1.029 (0.055)	1.087 (0.071)
19-24 years old	1.023 (0.113)	1.010 (0.113)	1.014 (0.042)	0.999 (0.042)
25-29 years old	1.152 (0.119)	1.143 (0.117)	1.027 (0.040)	1.030 (0.040)
30-39 years old	1.078 (0.089)	1.068 (0.087)	0.984 (0.032)	0.988 (0.032)
40-49 years old	1.152* (0.094)	1.143* (0.092)	1.016 (0.033)	1.019 (0.033)
Smoking starting year FE	Yes	Yes	Yes	Yes
Smoking starting age FE	Yes	Yes	Yes	Yes
Department FE	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes
Educational attainment	Yes	Yes	Yes	Yes
Sample	Smokers	Smokers before diagnosis	Smokers	Smokers before diagnosis
Model	Cox PH	Cox PH	Cox PH	Cox PH
Observations	2,803	2,655	16,405	14,924

Source: Constances (2012-2018). Notes: Robust standard errors in parentheses. These regressions do not include employment-related and health-related controls. Columns (1)-(2) and (4)-(5) include all smokers with at least one parent diagnosed with lung cancer or other smoking-related cancer, respectively, while columns (3) and (6) only include those among them who started smoking before diagnosis. (N)PH stands for (non) proportional hazards. Statistical significance: \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

### 3.2 Second approach: a fixed effects model to relax the identifying assumption about the timing of the diagnosis

The timing of the diagnosis may not be exogenous in case individuals with riskier behaviors or with specific life conditions are diagnosed earlier than others. In that case, there may be confounding factors and individuals whose parents were diagnosed at earlier ages are not comparable to those whose parents were diagnosed later. To address this issue, we use our data to build a retrospective panel and estimate a fixed effects model. Individual fixed effects control for time-invariant individual unobserved heterogeneity, such as time preference, family situation in the childhood and parents' health behaviors in the childhood. Given that we have the age of the first cigarette, in case the individual is a smoker, and the age of the last cigarette, in case the individual quits smoking, we are able to construct the smoking history of the individual. At each age, we know if the individual is a smoker or not and if her parents have already been diagnosed or not. Given that we do not have retrospective information about the number of cigarettes smoked per day each year, we cannot perform this analysis at the intensive margin.

We therefore estimate the following fixed effects logistic regression with individual fixed effects:

$$y_{it} = \delta PostParentDiag_{it} + \beta_2 PostOwnDiag_{it} + \beta_3 AgeDummies_i + \beta_4 YearDummies_t + \alpha_i + \epsilon_{it}$$

where  $y_{it}$  equals unity if individual  $i$  is a smoker at date  $t$ , zero otherwise. The variable  $PostParentDiag_{it}$  equals zero if none of the individual's parents has ever been diagnosed with cancer at date  $t$  and equals unity if at least one of the individual's parents has been diagnosed with cancer in the past.<sup>22</sup> This variable equals unity from the period following the parent's diagnosis until present. The variable  $PostOwnDiag_{it}$  is defined in the same way but concerns the interviewee's own diagnosis.  $\alpha_i$  are individual fixed effects. We also include age fixed effects ( $AgeDummies_i$ ) and year fixed effects ( $YearDummies_t$ ).

Individual fixed effects control for the time-invariant individual unobserved heterogeneity and thereby capture the transmission effect, so that the only operating mechanism captured by the  $\delta$  coefficient is now the informational shock effect. This approach allows us to compare the informational content of a health shock depending on the type of illness.

#### 3.2.1 Results: Smoking behavior and occurrence of a parent's diagnosis

Our benchmark results of the fixed effects model are reported in Table 6. Column (1) focuses on lung cancer and column (2) on other smoking-related cancers. Contrary to what is done in the first approach, we do not restrict the sample to those whose parents got a diagnosis, but consider here the entire sample, since the inclusion of individual fixed effects allows to control for the transmission effect.

We find that on average, having a parent diagnosed with a cancer significantly reduces the likelihood of being a smoker, but only if the parent is diagnosed with a lung cancer, the diagnosis of another smoking-related cancer having no significant effect. Whatever the cancer we consider, the coefficients associated with parental diagnosis are much smaller than the ones associated with own diagnosis, confirming that individuals require personalized information to modify their risky behavior. Moreover the effect of parental diagnosis is greater for lung cancer than for other smoking-related cancers, confirming that people associate less the latter cancers with smoking behavior.

---

<sup>22</sup>We consider the first diagnosis in case of multiple diagnosis.

Contrary to the literature, we find an average significant impact of the parents' lung cancer diagnosis on the smoking behavior of the children. An explanation for this departure from the literature is that we analyze the long-run effect of diagnosis. If the impact of the informational shock on the smoking behavior is not instantaneous, then one-year-lag or contemporaneous shocks may have no impact, whereas shocks that occurred previously may have a significant impact.

**Table 6:** Average effect of parents' diagnosis on smoking - FE models

	(1)	(2)
	Lung cancer	Other smoking-related cancer
Parents' diagnosis	-0.092*** (0.025)	-0.017 (0.011)
Own diagnosis	-1.373*** (0.117)	-0.152*** (0.025)
Dependent variable	Smoker	Smoker
Age	Yes	Yes
Year fixed effects	Yes	Yes
Sample	All	All
Observations ( $N \times T$ )	3,861,581	3,861,581

Source: Constances (2012-2019). Notes: Fixed-effect logit models. Robust standard errors are in parentheses and clustered at the individual level. Statistical significance: \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

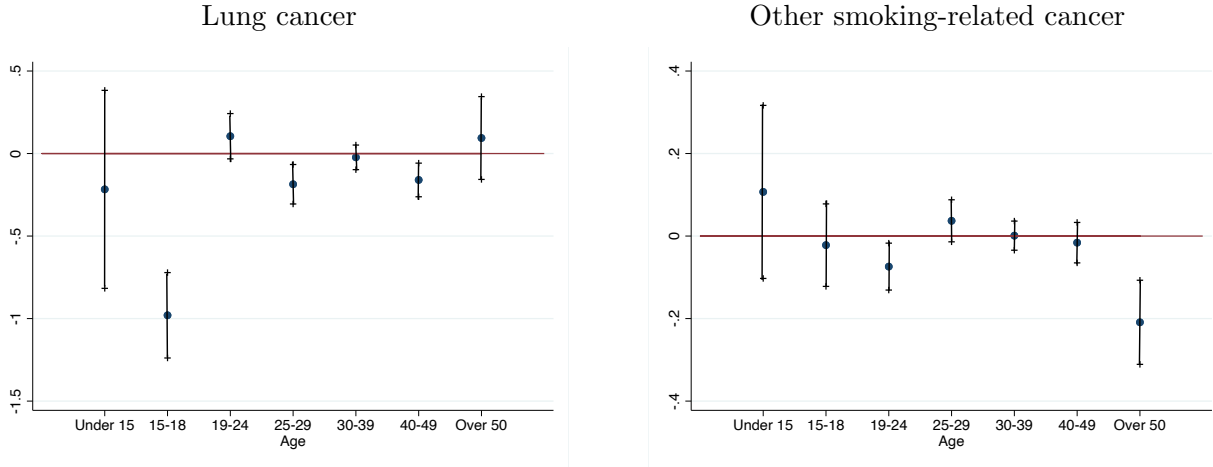
**Heterogeneous effects according to the age at diagnosis.** We replicate the previous analysis, but now allow for heterogeneous effects depending on the age of the individual at the time of parental diagnosis. We report in Figure 1 the estimates for all individuals.

When considering lung cancer (left-hand side panel in Figure 1), we find that the likelihood to be a smoker is significantly smaller when the diagnosis of their parents is received when the offspring is between 15 and 18 years old (when the decision to start smoking is about to be made). Receiving a parental lung cancer diagnosis between 25 and 49 years old also tends to lower smoking but the effect is overall insignificant or with a much smaller magnitude at these later ages. When considering other smoking related cancers (right hand side panel in Figure 1), we find that a parental diagnosis received between 19 and 24 years old significantly reduces smoking, though the effect is not large in magnitude. The effect of the diagnosis is larger for people who were above 50 at the moment of the parental diagnosis (though the two effects are not significantly different).

All in all, comparing both panels, we easily observe that the largest coefficient (in negative terms) corresponds to a parental diagnosis of lung cancer when the individual is 15-19 years old. Again, the decision to start smoking is about to be made at these ages, and the informational shock induced by the parental health shock seems clearly efficient to modify (*i.e.* reduce) the long-run smoking behavior of the offspring. This finding is also consistent with estimations in Table 5 according to which individuals receiving a lung

cancer parental diagnosis between 15 and 18 years old had shorter smoking spells because they had a higher probability to quit smoking. This stronger effect of the parental diagnosis when it is made around the usual ages of smoking initiation may be due to the fact that the youth have a lower addictive capital stock and may be more able to change behavior. It could also be due to the fact that people aged between 15 and 18 are likely to live with their parents at the time of the diagnosis, which may makes them better realize the reality of the hazards of smoking.<sup>23</sup>

**Figure 1:** Average effect of parents' cancer diagnosis on smoking depending on the age at diagnosis - FE models



Source: Constances (2012-2019). Notes: Estimated coefficient and 95%-confidence intervals from fixed-effect logit models. Robust standard errors clustered at the individual level. Estimations include controls for age, own diagnosis and year fixed effects.

## 4 Conclusion

In this paper, we analyze to what extent a parental smoking-related health shock can be associated with an informational shock in their offspring and reduce their smoking behavior. In contrast with the existing literature, we consider that the effect of this informational shock may vary along the life cycle. We investigate this variability of the informational shock effect by assessing the impact of parental diagnosis of smoking-related cancer (we distinguish between lung cancer and other smoking related cancer) on the offspring's smoking behavior, depending on offspring's age at diagnosis.

Duration models reveal that individuals receiving a parental diagnosis of lung cancer at the age of 15-18 years old have shorter duration spells. This finding is also confirmed when controlling for the transmission effect through a retrospective panel approach including individual fixed effects. We find that offsprings receiving a parental lung cancer diagnosis at the age of 15-18 years old are less likely to smoke than individuals that have not received a parental diagnosis. Finally, the informational effect associated with lung cancer

<sup>23</sup>We do not have information about how close people are to their parents, whether they live together or meet often, so we cannot test further the relevance of this explanation.



systematically arises as stronger than the informational effect associated to other smoking related cancers. All in all, given that the informational shock is effective when it arrives at the age when the individual decides to start smoking (15-18 years old), public health policies should target this young population subgroup with a strong information campaign and to frame the information in a way that it makes them respond to it. Moreover, the smaller changes in smoking behavior following the diagnosis of other smoking-related cancers calls for more comprehensive public health communication on the health hazards associated with smoking.

## References

- Adda, J. (2007). Behavior towards health risks: An empirical study using the “mad cow” crisis as an experiment. *Journal of Risk and Uncertainty*, 35(3):285–305.
- Adda, J., Bjorklund, A., and Holmlund, H. (2011). The Role of Mothers and Fathers in Providing Skills: Evidence from Parental Deaths. IZA Discussion Papers 5425, Institute of Labor Economics (IZA).
- Alam, S. A. (2015). Parental health shocks, child labor and educational outcomes: Evidence from Tanzania. *Journal of Health Economics*, 44(C):161–175.
- Arcidiacono, P., Sieg, H., and Sloan, F. (2007). Living Rationally under the Volcano? An Empirical Analysis of Heavy Drinking and Smoking. *International Economic Review*, 48(1):37–65.
- Arora, K. (2016). How does dementia onset in parents influence unmarried adult children’s wealth. *Social Science & Medicine*, 152(C):156–165.
- Bünnings, C. (2017). Does new health information affect health behaviour? The effect of health events on smoking cessation. *Applied Economics*, 49(10):987–1000.
- Byrne, S., Katz, S. J., Mathios, A., and Niederdeppe, J. (2015). Do the Ends Justify the Means? A Test of Alternatives to the FDA Proposed Cigarette Warning Labels. *Health Communication*, 30(7):680–693. PMID: 25119144.
- Cas, A., Frankenberg, E., Suriastini, W., and Thomas, D. (2014). The Impact of Parental Death on Child Well-being: Evidence From the Indian Ocean Tsunami. *Demography*, 51(2):437–457.
- Case, A. and Ardington, C. (2006). The impact of parental death on school outcomes: Longitudinal evidence from South Africa. *Demography*, 43(3):401–420.
- Chen, S. H., Chen, Y.-C., and Liu, J.-T. (2009). The Impact of Unexpected Maternal Death on Education: First Evidence from Three National Administrative Data Links. *American Economic Review*, 99(2):149–153.
- Cutler, D. M. and Glaeser, E. (2005). What explains differences in smoking, drinking, and other health-related behaviors? *American Economic Review*, 95(2):238–242.
- Darden, M. (2017). Smoking, Expectations, and Health: A Dynamic Stochastic Model of Lifetime Smoking Behavior. *Journal of Political Economy*, 125(5):1465–1522.
- Darden, M. and Gilleskie, D. (2016). The Effects of Parental Health Shocks on Adult Offspring Smoking Behavior and Self-Assessed Health. *Health Economics*, 25(8):939–954.
- de Vries, H., Engels, R., Kremers, S., Wetzels, J., and Mudde, A. (2003). Parents’ and friends’ smoking status as predictors of smoking onset: findings from six european countries. *Health Education Research*, 18:627–636.

- Gerking, S. and Khaddaria, R. (2012). Perceptions of health risk and smoking decisions of young people. *Health Economics*, 21(7):865–877.
- Göhlmann, S., Schmidt, C. M., and Tauchmann, H. (2010). Smoking initiation in Germany: the role of intergenerational transmission. *Health Economics*, 19(2):227–242.
- Hillebrandt, M.-A. (2022). Impact of changes in relationship status on smoking behavior and body weight. *Economics & Human Biology*, 44:101077.
- Holford, T. R., Levy, D. T., McKay, L. A., Clarke, L., Racine, B., Meza, R., Land, S., Jeon, J., and Feuer, E. J. (2014). Patterns of Birth Cohort-Specific Smoking Histories, 1965–2009. *American Journal of Preventive Medicine*, 46(2):e31–e37.
- Jackson, C. and Henriksen, L. (1997). Do as i say: parent smoking, antismoking socialization, and smoking onset among children. *Addictive Behaviors*, 22:107–114.
- Jin, L., Kenkel, D. S., Lovenheim, M. F., Mathios, A. D., and Wang, H. (2022). Misinformation, consumer risk perceptions, and markets: The impact of an information shock on vaping and smoking cessation. Working Paper 30255, National Bureau of Economic Research.
- Johnston, D. W., Schurer, S., and Shields, M. A. (2013). Exploring the intergenerational persistence of mental health: Evidence from three generations. *Journal of Health Economics*, 32(6):1077–1089.
- Khwaja, A., Silverman, D., Sloan, F., and Wang, Y. (2009). Are mature smokers misinformed? *Journal of Health Economics*, 28(2):385–397.
- Khwaja, A., Sloan, F., and Chung, S. (2006). Learning about individual risk and the decision to smoke. *International Journal of Industrial Organization*, 24(4):683–699.
- Krosnick, J., Malhotra, N., Mo, C., Bruera, E., Chang, L., Pasek, J., and et al. (2017). Perceptions of health risks of cigarette smoking: A new measure reveals widespread misunderstanding. *PLoS ONE*, 12(8).
- Le, H. T. and Nguyen, H. T. (2017). Parental health and children’s cognitive and noncognitive development: New evidence from the longitudinal survey of Australian children. *Health Economics*, 26(12):1767–1788.
- Le, H. T. and Nguyen, H. T. (2018). The Impact of Maternal Mental Health Shocks on Child Health: Estimates from Fixed-Effects Instrumental Variables Models for Two Cohorts of Australian Children. *American Journal of Health Economics*, 4(2):185–225.
- Li, C. and Gilleskie, D. B. (2021). The influence of endogenous behaviors among social pairs: Social interaction effects of smoking. *Journal of Health Economics*, 80:102547.
- Lillard, D. R., Molloy, E., and Sfekas, A. (2013). Smoking initiation and the iron law of demand. *Journal of Health Economics*, 32(1):114–127.
- Lin, W. and Sloan, F. (2015). Risk perceptions and smoking decisions of adult chinese men. *Journal of Health Economics*, 39:60–73.

- Loureiro, M. L., Sanz-de Galdeano, A., and Vuri, D. (2010). Smoking habits: like father, like son, like mother, like daughter? *Oxford Bulletin of Economics and Statistics*, 72(6):717–743.
- Melchior, M., Chastang, J.-F., Mackinnon, D., Galera, C., and Fombonne, E. (2010). The intergenerational transmission of tobacco smoking – the role of parents long-term smoking trajectories. *Drug and Alcohol Dependence*, 107:257–260.
- Morin, A. J. S., Moullec, G., Maïano, C., Layet, L., Just, J.-L., and Ninot, G. (2011). Psychometric properties of the Center for Epidemiologic Studies Depression Scale (CES-D) in French clinical and nonclinical adults. *Rev Epidemiol Sante Publique*, 59(5):327–340.
- Palali, A. and van Ours, J. C. (2019). The impact of tobacco control policies on smoking initiation in eleven European countries. *The European Journal of Health Economics*, 20(9):1287–1301.
- Pan, J. and Han, W. (2017). Exploring the intergenerational persistence of health behaviour: an empirical study of smoking from China. *BMC Public Health*, 17(1):1–9.
- Senne, J.-N. (2014). Death and schooling decisions over the short and long run in rural Madagascar. *Journal of Population Economics*, 27(2):497–528.
- Skurka, C., Kemp, D., Davydova, J., Thrasher, J. F., Byrne, S., Safi, A. G., Avery, R. J., Dorf, M. C., Mathios, A. D., Scolere, L., and Niederdeppe, J. (2017). Effects of 30% and 50% Cigarette Pack Graphic Warning Labels on Visual Attention, Negative Affect, Quit Intentions, and Smoking Susceptibility among Disadvantaged Populations in the United States. *Nicotine & Tobacco Research*, 20(7):859–866.
- Sloan, F. A., Smith, V. K., and Taylor, D. H. (2003). *The smoking puzzle: Information, risk perception, and choice*. Harvard University Press.
- Slovic, P. E. (2001). *Smoking: Risk, Perception, & Policy*. SAGE Publications.
- Smith, V. K., Taylor, D. H., Sloan, F. A., Johnson, F. R., and Desvousges, W. H. (2001). Do Smokers Respond To Health Shocks? *The Review of Economics and Statistics*, 83(4):675–687.
- Sundmacher, L. (2012). The effect of health shocks on smoking and obesity. *The European Journal of Health Economics*, 13(4):451–460.
- Viscusi, W. K. (1990). Do Smokers Underestimate Risks? *Journal of Political Economy*, 98(6):1253–1269.
- Zins, M., Goldberg, M., and CONSTANCES team (2015). The french constances population-based cohort: design, inclusion and follow-up. *European Journal of Epidemiology*, 30(12):1317–1328.

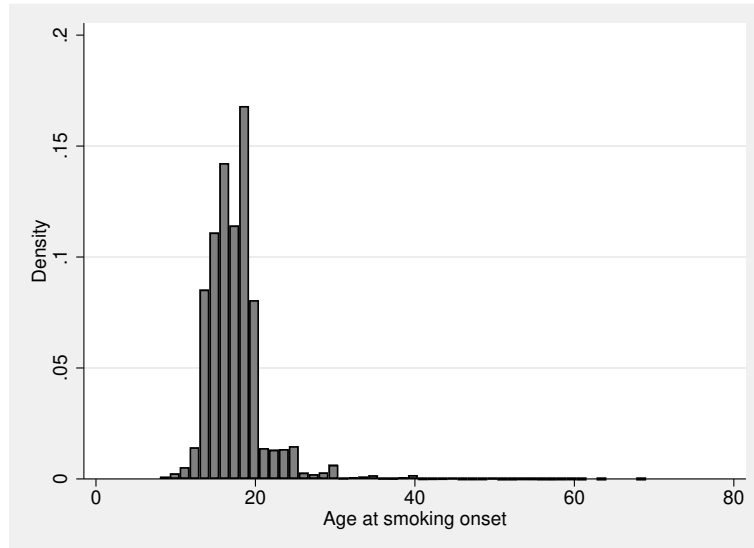
## A Appendix: Descriptive statistics

### A.1 Sample composition

**Table A.1:** Sample composition by gender- Smoking Behaviors

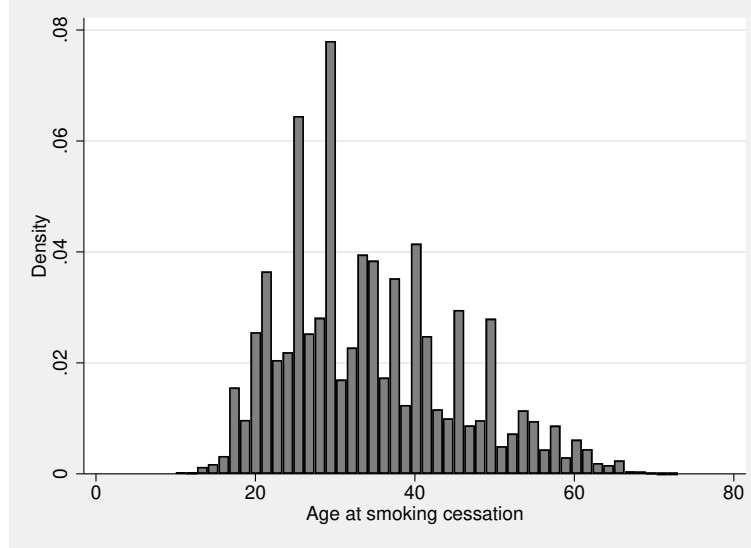
	Males	Obs	Females	Obs
<b>Panel A: Ever smoked and smoking at first wave</b>				
Ever smoked	0.575	48,908	0.478	47,106
Smoker	0.196	16,667	0.175	17,256
<b>Panel B: Among ever smoked</b>				
Smoker	0.341	16,667	0.366	17,256
Age at smoking onset	17.445	48,908	17.430	47106
Number of cigarettes per day	13.324	48,908	10.569	47106
<b>Panel C: Among no longer smokers</b>				
Age at smoking onset	17.318	32,241	17.383	29,850
Number of cigarettes per day	14.316	32,241	11.194	29850
Age at smoking cessation	35.332	32,241	33.388	29,850

**Figure A.1:** Distribution of the age at which people start smoking



Source: Constances (2012-2018). Authors' computations.

**Figure A.2:** Distribution of the age at which people quit smoking

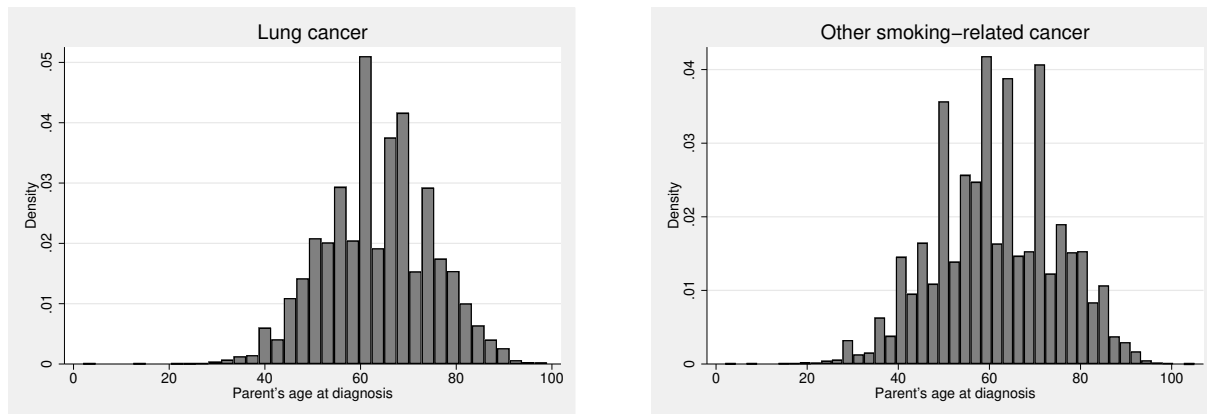


Source: Constances (2012-2018). Authors' computations.

**Table A.2:** Proportion of the population with parents having smoking-related cancer by gender

	Males	Obs	Females	Obs
Parents with lung cancer	0.041	3,439	0.040	3,972
Parents with other smoking-related cancer	0.204	17,305	0.217	21,329

**Figure A.3:** Distribution of the age at which parents were diagnosed with lung cancer or other smoking-related cancer



Source: Constances (2012-2018). Authors' computations.