FORGONE EARNINGS FROM SMOKING:

EVIDENCE FOR A DEVELOPING COUNTRY

Michael Lokshin and Kathleen Beegle[*]
Development Research Group, World Bank

ABSTRACT

This paper estimates the economic losses related to the negative effect of smoking on wages in a context of a developing country. Using data from 2005 Albania Living Standards Monitoring Survey, we jointly estimate a system of three equations: the smoking decision and two separate wage equations for smokers and non-smokers. The results show that, after controlling for observed characteristics and taking into account unobserved heterogeneity in personal characteristics, smoking has a substantial negative impact on wages. On average smokers' wages are 20 percent lower than the wages of similar non-smokers, providing strong evidence for the potential policy relevance of tobacco control initiatives for developing countries.

September 4, 2006

1.   Introduction

In 2000, about one out of three adults, or 1.1 to 1.2 billion people worldwide, smoked (WHO 1999). By 2010 the number of smokers is expected to grow to around 1.3 billion (FAO 2003). The vast majority of the world's smokers (82 percent) live in developing countries where, in sharp contrast to developed countries, the consumption of tobacco is on the rise. Tobacco control is increasingly seen as a public health priority, motivating campaigns such as the World Health Organization's Tobacco Free Initiative as well as country-specific tobacco control campaigns. These efforts are justified not only on grounds related to well-established adverse effects of smoking on health, but more generally on economic grounds.[1] The gross healthcare cost of smoking for the developing countries is estimated to be as high as 1.1 percent of GDP (Jha and Chaloupka 2000). The economic costs associated with losses in productivity because of smoking could be at least as large and be disproportionately incurred by the poor (The World Bank 1999).

Estimates of the economic costs of smoking tend to reflect costs associated with increased premature death and health-care costs related to smoking-induced chronic diseases (which can affect smokers as well as non-smokers exposed to smoke). In addition to these pathways, there is a third category of costs that has been considered: the earnings and wage reductions associated with smoking. The literature on the relationship between smoking and labor market performance is very limited and based on data from Western Europe and the US. Berger and Leigh (1989) examine the effect of smoking and being overweight on earnings using the 1973 Quality of Employment survey in the US. Their study finds no statistically significant association between smoking and earnings for male and female workers. Levine *et al*. (1997) investigate the effect of smoking on wages using panel data from US. They compare wages of continuous smokers with the wages of workers who quit smoking. Controlling for observable and time-invariant unobservable characteristics, smoking is estimated to lower a worker's wage by about 4 – 8 percent. Drawing on data from Germany, Heineck and Schwarze (2003) examine whether smoking affects earnings of employees 25 to 55 years old. Depending on the

---

[1]   The WHO estimates that tobacco is the second major cause of death worldwide and the fourth most common risk factor for disease worldwide.

model specification, the estimated losses in earnings for smokers range from 2 to 8 percent compared to non-smokers.

Several studies examine the simultaneous effect of alcohol and smoking on earnings. Van Ours (2002) focuses on the simultaneous wage effects of alcohol and tobacco using the sample of respondents 16 years and older in Netherlands. Controlling for unobserved heterogeneity and endogeneity of drinking and smoking, the study estimates that the wages of smokers are 10 percent lower than the wages of otherwise identical non-smokers. Auld (2005) estimates the model of drinking, smoking and wages using the sample of employed Canadian men. After accounting for endogenous nature of the decision to use alcohol and tobacco and their simultaneous effects on wages, he finds that daily smokers have wages about 30 percent lower than non-smokers.

The limited set of studies of the wage or earnings consequences of smoking consistently show significant negative effects. However, these results are restricted to high-income settings; estimates of the economic costs of smoking in developing countries are non-existent to our knowledge. This paper addresses this gap in the literature by estimating the economic losses related to the negative effect of smoking on wages in a developing country, namely Albania, where 60 percent of adult males and 18 percent of females smoked in 2002 and the number of smokers has been increasing rapidly over the last decade (WHO 2002). Anti-smoking policies are high on the country agenda. In January 2006 the Government of Albania submitted a new draft law on tobacco and smoking that outlaw sales of cigarettes to minors younger than 18, and present stricter rules for tobacco industry regarding sale and advertising of tobacco products (ENSP 2006).

Using recent household survey data, we estimate a system of three equations: the smoking decision (never having smoked and current smokers) and two separate wage equations, one for smokers and one for non-smokers. By jointly estimating this system by Full Information Maximum Likelihood we control for the effects of observed and unobserved characteristics that may be correlated with wages and smoking status. Moreover, we introduce new and arguably improved instrumental variables to identify the causal effect of the smoking decision on wages. The results show that, after controlling for observed characteristics and taking into account unobserved heterogeneity

in personal characteristics, smoking has a substantial negative impact on wages. On average smokers' wages are 20 percent lower than the wages of similar nonsmokers, providing strong evidence for the potential policy relevance of tobacco control initiatives for developing countries.

The structure of the paper is as follows: the methodology and empirical approach are discussed in Section 2. The data are discussed in Section 3. Section 4 presents the empirical results and conclusions are drawn in the last section.

2. Methodology and empirical specification

Several theories have been proposed to explain the lower wages associated with smoking. Health effects of smoking are arguably the most frequently noted link; smoking causes various morbidities, which make smokers less healthy and could therefore result in lower labor productivity. A second, partially health-related channel, is the higher costs to employers of hiring a smoker relative to a non-smoker (for example, due to healthcare costs induced by smoking). Even if smokers are equally productive or do not incur more costs in employment, discrimination against smokers in the workplace may result in lower earnings for those who smoke. Finally, smoking could be correlated with lower earnings because of the differences in preferences over present and future consumption between smokers and non-smokers. We discuss each of these theories in turn.

The adverse health implications of smoking could affect a worker's productivity and, therefore, labor market outcomes. Levine *et al.* (1997) report that smoking is associated with lower physical endurance even for the young workers. Young smokers are also more likely to experience respiratory infections (Conway and Cronan 1992; Hoad and Clay 1992). The negative effects of smoking accumulate with age and the productivity consequences of smoking could be larger for older workers. The higher incidence of respiratory diseases among young smokers and more serious illnesses at older ages can result in lower productivity and higher absenteeism (Bertera 1991).

Net of productivity effects, employers face higher costs to hire a smoker. The cost of medical insurance for smokers is higher (Bertera 1991). In turn, larger health insurance claims would reflect negatively on wages of smokers, as employers are unwilling to pay

for all or part of the higher costs of employing smokers. At the same time, workers who are aware of the negative consequences of smoking might select themselves into occupations that provide better health coverage, accepting lower wages in exchange. The cost of the smoking worker for his employer could also rise given the need for additional facilities for smokers that would increase air cooling, heating, and ventilation cost, increased fire hazard and fire insurance premiums, and smoke pollution leading to larger cleaning and maintenance costs (CHPP 2000).

Discrimination against smokers is another potential reason for observed differences in wages between smokers and non-smokers. Employers might discriminate against smoking employees because of the possible effect of second-hand smoking on the other employees and on the customers. Statistical discrimination may also exists if employers assume that an individual smoker is less productive based on the average health effects of smoking[2].

These three theories linking smoking to wage outcomes propose that smokers, all things equal, will realize lower wages. A fourth theoretical consideration would be that smoking is a proxy for some other personal characteristics that could affect wages. For example, individuals with higher valuation of present versus future consumption may be more likely to smoke and less likely to invest in productivity-enhancing human capital. This theory centers on the observation that the costs of smoking are related to the negative effects of smoking on health which often are realized only in the longer-term. On the other hand, the perceived benefits of smoking are predominantly immediately realized and might include stimulation, enhancement of concentration and short-term memory, and, to a certain degree, weigh control (Rohsenow *et al.* 2003). Smokers expect desirable and positive effects of smoking in terms of facilitating of social interactions and relaxation (Brandon and Baker 1991). Mood-altering ability of nicotine tends to alleviate boredom and reduce stress as well as aggressive responses to stressful events (Senagore 2006). So, individuals who smoke are choosing the short-run benefits over the long-run costs, thereby perhaps revealing a higher rate of time preference than non-smokers. Fuchs

---

[2]   Discrimination may be much more relevant in the US context where social norms make smoking less acceptable as reflected in state-level bans against smoking in public spaces compared to Albania where there are no restrictions against smoking. Heineck and Schwarze (2003) make the point that in Germany discrimination against smokers is more moderate than in the US; arguably one could make similar case for Albania where smoking prevalence is even higher.

(1982) suggests that smoking is therefore an indication or proxy of a high rate of time preference resulting in lower human capital accumulation and lower earnings, as evidenced by Evans and Montgomery (1994) and Lahiri and Song (2000). Likewise smokers may have flatter earning profiles if they have a higher rate of time preference. In the "Rational Addiction Theory" (Becker and Murphy 1988) a decision to smoke is based on comparison of the present and future costs and benefits of smoking (Chaloupka 1991, Becker *et al.* 1994).

Given these theories that link smoking to wages, one empirical approach would be to estimate a wage equation based on a model of wage determination that relates wages with the individual human capital characteristics (Mincer 1974) modified to include smoking. In the simplest case the wage function has a form:

$$\ln W_i = X_i \beta + S_i \gamma + \varepsilon_i \qquad (1)$$

where W_i represents monthly wages of an individual i, $\mathbf{X}$ is a vector of exogenous individual characteristics, S_i is dummy variable of the individual smoking status, and ε_i is an i.i.d., mean-zero error term, and β is a vector of parameters. In this specification, parameter γ could be interpreted as an effect of smoking on wages.

While the difference in observable characteristics of smokers and non-smokers could be controlled for by a standard regression approach, the presence of unobservable factors affecting both smoking and earnings would lead to the endogeneity of S_i in (1). As a result of such endogeneity the estimation of the effect of smoking on wages would be biased.[3] For example, individuals who choose to smoke may be different from non-smokers in some unobserved dimensions that are negatively (positively) correlated with wages. As noted above, rate of time preference is one possible trait that is unobserved to the analyst and its omission biases estimates in equation (1). The challenge for the empirical strategy is to estimate the effect of smoking on wages controlling for such unobserved factors. We use two alternative methods to address this: a standard two-stage

---

[3] The reverse causality of the effect of wages on smoking is partly mitigated by the evidence of the addictive nature of smoking that could be especially strong in the developing countries. Lance *et al.* (2004) report a price-elasticity of cigarette consumption for China and Russia to be in the range of 0 to -0.15, estimates which are much lower than the estimates of the elasticity in the US (0.40). The authors note that this low elasticity demonstrates the strong addiction of consumers to cigarettes. Developing countries often lack the defined regulations on production of cigarettes. As a result, cigarettes are likely to have higher nicotine content and to be more addictive on average than in the developed countries.

least square instrumental variables (2SLS) regression and the switching regression approach (e.g., Lee 1978).

The 2SLS model is appealing for at least two reasons. First, the previous studies noted above use 2SLS to estimate the effect of addictive behavior on wages. Second, 2SLS produces an easily-interpreted measure of the effect of smoking: γ. On the other hand, this specification assumes that the effect of personal characteristics on wages is linear and independent of smoking status. This assumption can be relaxed by allowing a complete interaction of a variable for the smoking status and exogenous variables in the wage equation.[4] The simultaneous estimation of the resulting two wage equation, one for smokers and another for non-smokers, and the equation for the smoking status represents the switching regression model formally expressed as:

$$
\begin{aligned}
S_i^* &= Z_i \eta + u_i \\
\ln W_i^1 &= X_i^1 \beta_1 + \varepsilon_{1i} \qquad observed\ if\ S_i^* \geq 0\ (smokes) \\
\ln W_i^2 &= X_i^2 \beta_2 + \varepsilon_{2i} \qquad observed\ if\ S_i^* < 0\ (does\ not\ smoke)
\end{aligned}
\tag{2}
$$

where S_i^* is a latent propensity to smoke, **Z** is a vector of exogenous characteristics, u_i, ε_{1i}, and ε_{2i} are the normally distributed error terms, η is vector of parameters, and superscript 1 identifies smokers and 2 identifies non-smokers.

The identification conditions in (2) are that disturbance terms u and $\varepsilon's$ are independent, or else there is at least one variable in **Z** that is not included in **X**. Assuming that (X, ε_1, ε_2, and u) are i.i.d., and ε_1, ε_2 and u have, conditional on **X**, a joint normal distribution with mean zero and positive definite covariance matrix, the system of equations (2) could be estimated by Full Information Maximum Likelihood method (FIML). The identification restrictions for the 2SLS estimation are similar.

Several studies use instrumental variable approach for identifying the effect of smoking and alcohol consumption and smoking on wages. Lye and Hirschberg (2000) applied the two-step Heckman (1979) correction approach to control for endogeneity of smoking and alcohol consumption. The paper does not explicitly discuss the

---

[4]   We can also consider the intermediate model that is a FIML estimation of the effect of an endogenous binary variable on another endogenous continuous variable, conditional on two sets of independent variables – the so-called "treatment effect" regression model (e.g., Maddala 1983). That model is more restrictive in comparison with the FIML switching regression model and brings no new insights into our analysis. The results of "treatment effect" regression are available from the authors on request.

identification conditions, but seems to rely on identification by the functional form. Such an approach could be problematic due to the "weak instruments bias" (Staiger and Stock 1997). The instrumental variables in Van Ours (2002) include whether the respondent has a partner and whether the respondent indicated the 'early start' of smoking. Auld (2005) includes religion and the price of tobacco as instrumental variables. Heineck and Schwarze (2003) use such variables as 'strong religious conviction' and martial status as instruments in their 2SLS estimation. In our view, it is hard to justify the validity of these instruments for identification of the effect of smoking on wages. Empirical research has consistently shown that married men have substantially higher wages, on average, than otherwise similar unmarried men, the so-called "marriage wage premium" phenomenon (e.g., Cornwell and Rupert 1997, Daniel 1991, Gray 1997). 'Early start' of smoking may be an invalid instrument if length of smoking exposure directly affects wages through health; long-term adverse effects of teenage smoking on health are well-documented in the medical literature (e.g. CDC 1994). Using the price of tobacco as an instrument could also be problematic because of the addictive nature of smoking (see footnote 3). Finally, it is difficult to argue that religious beliefs, while affecting the individual's probability to smoke, have no effect on earnings given the extensive literature that demonstrates the opposite (e.g., Nigel 1984). Indeed, Van Ours (2002) concludes that "… *it is difficult to find good instrumental variables*" and Heineck and Schwarze (2003) state that their "…*identifying instruments … perform poor[ly] in the first stage*" of 2SLS regression.[5]

Our identification strategy relies on two variables about smoking behavior of the respondent's parents. We had an opportunity to introduce these variables in the ALSMS questionnaire specifically for the purpose of using them as instruments in the current analysis. We argue that the parental smoking history is correlated with the probability to smoke and has no direct impact on individual's earnings[6]. The effect of parental smoking on offspring tobacco use is well established in the literature (e.g., Andrews *et al.* 1993; Pedersen and Lavik 1991). Smoking parents may present a model of smoking behavior to

---

[5]   See Van Hours (2002), page 11; Heineck and Schwarze (2003), page 8.
[6]   MacDonald and Shields (2001) use similar instruments in the analysis of the impact of alcohol use on occupational attainments in England. Here we discuss these instruments in the context of probability of being a current smoker with the caveat that these instruments are arguably stronger for identification of non-smokers who never having smoked as opposed to non-smokers who had previously smoked and quit.

their children, who are therefore at increased risk of tobacco use when adults (Fagan *et al.* 2005).

Perhaps the strongest theoretical argument against the validity of our instruments is the adverse health effect of secondhand smoking that could affect the future productivity of children of smoking parents. While there exists an extensive literature on the short-term effects of secondhand smoking on health we are unaware of any study that demonstrates the direct negative impact of the secondhand smoking when a child on the child's labor productivity in adulthood.

One could also argue that smoking is an indication of certain personal traits of a parent, potentially unrelated to smoking, that could affect outcomes of her children (e.g., schooling attainment, motivation, etc.), which, in turn, affect the children's future earnings[7]. For example, a rate of time preference might be transferred from parents to their children genetically or acquired in early childhood. Then the observed low wages of the smokers relative to non-smokers could be caused not by the smoking directly but by the higher rate of personal time preferences of smokers that was somehow acquired from the parents. Omitting these measures in the estimation could potentially invalidate our instruments. We partially address this by including education of parents in our empirical specification – the only additional parental information that is available in the data. The rates of time preferences of the parents would be reflected in their educational achievements, thus clearing our instruments from the component potentially correlated with respondents' wages. The educational achievements of the parents are quite similar between the groups of smokers and those who never smoked, indirectly lending support to our chosen instrument. Nevertheless, given the limitations of our data, we can never be sure that our instruments capture no unobserved characteristics of the parents that could be correlated with the respondents' labor market outcomes. Thus, our estimates could reflect not only the direct effect of smoking on wages but also the effects of these unobservables.

The effect of smoking on wages is estimated differently in our two empirical models. The coefficient on the endogenous dummy variable S_i in the 2SLS regression

---

[7] We know of no study confirming that multi-level effect of parental smoking on children's future earnings. In addition, we would think that the influence of these factors on the future productivity would be less pronounced in Albania, where more than 60 percent of respondents grew up with smoking parents.

estimates the effect directly. In the FIML model the effect could be derived by comparing the predicted expected wages of smokers with the expected wages they would receive had they been non-smokers or quitters (the counterfactual wage). The expected wage of a smoker could be expressed as:

$$E(\ln w_1 \mid X_i^1, S = 1) = \beta_1 X_i^1 + \rho_1 \sigma_1 \frac{\varphi(\eta Z_i^1)}{\Phi(\eta Z_i^1)} = \beta_1 X_i^1 + \beta_{\lambda 1} \lambda(Z_i^1 \mid S = 1) \qquad (3.1)$$

Similarly, the expected wage of non-smoker is expressed as:

$$E(\ln w_2 \mid X_i^2, S = 0) = \beta_2 X_i^2 + \rho_2 \sigma_2 \frac{-\varphi(\eta Z_i^2)}{1 - \Phi(\eta Z_i^2)} = \beta_2 X_i^2 + \beta_{\lambda 2} \lambda(Z_i^2 \mid S = 0) \qquad (3.2)$$

where and σ_1^2 and σ_2^2 are variances of the error terms in the wage equations, and ρ_1 and ρ_2 are the coefficients of correlation between ε_1, ε_2 and u, and λ's are the so-called Mills ratios. From (3.1-2) the expected counterfactual wage for the sub-sample of smokers is:

$$E(\ln w_1 \mid X_i^1, S = 0) = \beta_2 X_i^1 + \rho_2 \sigma_2 \lambda(Z_i^1 \mid S = 1) \qquad (4)$$

The interpretation of our results depends on how we model our counterfactual. There is no consensus on the choice of the comparison group in the literature on smoking and wages. The majority of the literature compares wages of current smokers with everybody else (Lye and Hirschberg 2004, Van Ours 2002, Heineck and Schwarze 2003). Levine *et al.* (1997) compares wages of smokers only with those who quit smoking. The main goal of our study is to evaluate the component of the economic cost of smoking related to the effect of smoking on wages. We define the economic cost of smoking as measured by the higher wages of smokers had they never started smoking in which case the counterfactual of smokers is the group of those who have never smoked.

There are, of course, alternative counterfactuals that could be framed. The most obvious alternative would be to evaluate the economic cost of continued smoking as measured by the increased wages for smokers if they quit. In this case the counterfactual of smokers is the group of quitters. However, we failed to find in our data convincing instruments to identify the decision to quit or continue smoking. One could also assess the changes in wages associated with never-smokers who begin to smoke. These alternative counterfactuals reflect future benefits of successful anti-smoking campaigns,

whereas our chosen counterfactuals measure the costs of smoking in terms of current losses in productivity.

3.   Data and descriptive statistics

The data for our analysis comes from the fourth round of Albania Living Standard Monitoring Survey (ALSMS) conducted in April-July 2005 by the Albanian Institute of Statistics with the technical support of the World Bank. The ALSMS uses a two-stage stratified clustered sample of 3640 households, and is representative at the national, urban, rural, and regional levels.[8] The ALSMS collects information on the demographic characteristics of the household members, their labor market activities, education, heath characteristics, and the household's access to social services. The questionnaire also gathers information on individual and household income and consumption expenditures as well as on ownership of assets. A special section of ALSMS collects information about individual smoking behavior including: whether the respondent smokes on a daily basis or is an ex-smoker, age at which the respondent started smoking, number of cigarettes smoked per day, and whether the respondent's farther and mother ever smoked.

For the purpose of the current analysis we restrict our sample to employed prime-age adults 25 to 60 years of age. By the age of 25 most of the adults in Albania have completed their education and are actively involved in the labor market. An individual is considered to be employed if he had worked in a permanent job in the week preceding the survey or was temporary absent from his permanent job due to vacation or illness. Our sample excludes employers, self-employed and individuals working in a household farm because the wage data is not reported for this group of respondents. The individual wages were measured as net monthly payments or earnings and bonuses.

Forty percent of males and about 5 percent of females age 25 to 60 reported to have ever smoked; 32 percent of males and less than 4 percent of females in our sample have categorized themselves as current smokers.[9] These numbers are comparable with

---

[8]   For further information on survey design and sample frame see: http://www.instat.gov.al/
[9]   That low proportion of smoking women may not reflect the actual prevalence of female smoking, as in Albania women are reluctant to admit they smoke because smoking is not seen as being socially desirable or acceptable for females (IMF 2006).

other studies of smoking in Albania. The results of the Population-based Survey show that 38 percent of adult males and 18 percent of females were current smokers in the capital Tirana in 2002 (Shapo *et al.* 2003). IMF (2006) reports that, according to 2002 LSMS survey, 31 percent of Albanian men smoked. Because of a small proportion of female smokers in the sample our further analysis will focus only on working males 25-60 years of age.

The incidence of smoking is similar among urban and rural male population in Albania. The prevalence of current smokers is close to 50 percent among the least-educated males and declines for better-educated men. Only one quarter of men with a university degree smoke. The propensity to smoke varies for men by religion. The highest proportion of smokers (36 percent) is among Albanian Muslims, including Bektashi. Albanian Christians smoke at a rate of about 30 percent.

The left panel of Figure 1 graphs the results of non-parametric regression (running-line least-squares) for the proportion of smokers by age. The incidence of smoking is increasing with age of the respondents. Less than 20 percent of 25 years-old men classified themselves as current smokers. The proportion of smokers increases sharply for older age groups peaking at about 35 percent for men 38 years old and older.

The right panel of Figure 1 shows the relation between monthly log-wages and age for smokers and non-smokers. On average, non-smokers earned about 26,300 Lek (3.27 in logs) per months relative to 24,290 Lek per month (3.19 in logs) for smokers (8.3 percent difference). Both groups have an inverted U-shaped age-wage profile. Wages of younger males are increasing and reach maximum at age of 35 and then decline for the older workers. Wages of non-smokers are higher than wages of smokers for all age groups with the largest gap in wages observed for workers 35 to 50 years of age.

4.   Results

In order to estimate the effect of smoking and the counterfactual as described in Section 3, the dependent variables in the regression estimates are the logs of monthly wages for smokers and those who never smoked, and the binary indicator capturing whether a respondent is a current smoker or has never smoked. Table 1 shows the descriptive

statistics for the dependent and explanatory variables separately for smokers, quitters, and non-smokers. The explanatory variables in our empirical specification are classified into several groups. The first group includes human capital characteristics: age (squared and cubed), educational level, foreign languages spoken, and religious affiliation. The second group of variables consists of socio-demographic characteristics of the household. These include the household size and household composition variables. Conditions on the local labor markets are captured by the set of 36 district dummies and the dummy for the type of locality (urban or rural). We include parental education as an indicator of familial background that could affect labor market motivation or productivity which could otherwise confound interpretation of smoking status in the 2SLS and FIML estimations. Finally, we include total hours worked last month.

Table 2 presents the results of the estimation of the OLS and 2SLS regression models of the effect of smoking on wages. According to the OLS regression, smoking decreases male wages by 4.7 percent (standard error of 0.025); this result is comparable to what was shown in Figure 1. Given concerns that smoking does not randomly occur among men in Albania, we now turn to the results on the probability of smoking.

The probability of smoking is positively and significantly related to whether the father of the respondent smoked; the coefficients on mother's and father's smoking are jointly significant. Older individual are more likely to smoke relative to the younger respondents. Prevalence of smoking is significantly lower among Catholic and significantly higher among atheists relative to men of other religions. Men with the less than four years of primary education are more likely to smoke compared to better-educated men. Estimation also reveals some regional differences of the incidence of smoking among working men.

The key result from the first-stage of the 2SLS regression is the joint significance of the instrumental variables for smoking. The Sargan's (1958) test of over-identifying restrictions fails to reject the null hypothesis that our excluded instruments, variables on the mother's and father's smoking status, are uncorrelated with the error term and are correctly excluded from the wage equation.

Turning to the results of the second stage of the 2SLS estimation, the wage reduction associated with smoking is much larger when smoking is treated as an

endogenous variable. We now find that the size of the wage loss for smokers is equal to a 23.4 percent wage reduction (bootstrapped standard error of 9.8) compared to wages of non-smokers.[10] This large estimated loss in wages is comparable with the estimates by Auld (2005) who finds a 30 percent earning loss for smokers in Netherlands and is higher than the effects found in studies of smoking in the US and Canada. The results show a positive selection bias into smoking as found in other studies (such as Auld, 2005, and Van Ours 2002)[11]; OLS results underestimate the negative wage effect of smoking.

The results of the FIML estimation of the switching regression model (2) are shown on Table 3.[12] The Log-likelihood test of independence is rejected in favor of joint dependence of the error terms in system (2). Hausman over-identification test (Hausman 1983) confirms the validity of our instruments with 5 percent significance level.

Turning to the estimations of the earning functions, controlling for other personal characteristics respondent's age has a negative and significant effect on individual earning. Married non-smoking males earn significantly more than single nonsmokers and whereas marital status of smokers has no effect on their earnings. The relationship between wages and education of non-smokers is stronger (i.e. wage-education profile is steeper) than for smokers. An ability to speak English is associated with a significant wage premium for both groups. Knowledge of Italian and Greek, on the other hand, is beneficial in term of higher wages only for non-smokers. Overall we find a stronger dependence between wages and productive human capital characteristics for non-smokers than for smokers. That could support the hypothesis about the differences in the discount rates of the future between smoking and non-smoking individuals (Evans and Montgomery 1994).

We evaluate the effect of smoking on wages by simulations based on estimates in Table 3 of the parameters of equations (4) and (5). For every smoking respondent we

---

[10] We use simulated effects instead of approximate effects given by the coefficient on the smoker dummy in Table 2.

[11] Levine *et al.* (1997) show mixed results in regards to evidence of the positive selection but find no evidence of negative selection.

[12] The estimation of the switching regression model and simulation are based on the FIML algorithm implemented as a Stata program *movestay* (Lokshin and Sajaia 2004). We also tried to estimate the simultaneous model with three wage equations for smokers, quitters, and those who never smoked by FIML. Aside from the issue of finding valid instruments for identification, the model failed to converge, most likely due to the small number of quitter in our sample.

predict his expected wage conditional on him being a smoker and compare it with a predicted wage conditional on not smoking. According to the results of the switching regression FIML, an average smoker experiences a 22.5 percent reduction in his wages (bootstrap standard error of 9.2) due to smoking. The results are not statistically different than the wage reduction associated with smoking produced from the 2SLS estimates. Moreover, the effect of smoking on individual wages is similar across different age groups and for respondents residing in rural and urban locations. Yet, we do find evidence that the smoking penalty is lower for men with higher education: wages of smokers with a university diploma are 17 percent than their non-smoking counterparts.

Several alternative econometric specifications for model (2) can be tested. In the specification for Table 3 the non-smokers include only men who never smoked. We re-estimate our model on the sample that includes both the respondents who never smoked and those who quit smoking (results shown in Tables A1 and A2 in Appendix). The effect of smoking on wages derived from this unrestricted sample estimation is not different than the results in Table 3; the wage loss from smoking is 27.4 percent (standard error of 9.4).

We also estimate the switching regression model with the specification that includes the occupational dummies (not presented). These variables could be endogenous in the wage equation and so the results of this estimation should be interpreted with caution. Again, the results are largely unchanged; after controlling on occupational choice, an average smoker earns 26.3 percent less than a non-smoker[13]. In sum, these different specifications produce comparable (and statistically similar) estimates of the economic cost of smoking.

Trying to test various hypotheses about the effect of smoking on wages we estimate three wage regressions for smokers and quitters that include the intensity of smoking, expressed in cigarettes per day (only for the current smokers), and smoking experience as total years of smoking (Table 4). We find that within the group of current smokers and the respondents who quitted smoking the effect of intensity of smoking and smoking experience is small and statistically insignificant. If indeed employers discriminate against smokers we would expect to find a stronger negative effect of the

---

[13] The results of these estimations are available from the authors on request.

intensity of smoking on smoker's wages. To account for a potential endogeneity of a smoking experience with respect to wages, we instrument this variable with the instruments we use in our FIML estimations. The rational for the instruments is similar: we argue that parental smoking might affect the intensity of smoking and age at which an individual starts smoking but has no direct effect on individual wages. The results of this IV regression (available from the authors on the request) are similar to the results of OLS estimation, i.e., the effect of smoking experience on wages of current smokers and quitters is insignificant.

The economic cost of smoking is large for an individual smoker and the implied societal cost of smoking could also be substantial. Taking a conservative estimate of the negative effect of smoking on wages for the prime-age males to be about 20 percent and calculating an aggregated loss in wages of 30 percent (about 400,000 workers) of current male smokers, every month the Albanian economy looses approximately 2.6 percent of its GDP because of negative consequences of smoking among the prime age males[14].

5.   Conclusions

For working prime-age men in Albania, our analysis shows that wages of smokers are significantly lower than wages of individuals who never smoked. The simple comparison indicates an 8 percent wage advantage of non-smokers relative to smokers. However, smokers differ from non-smokers in both observable and unobservable characteristics. Consistent with other studies for developed countries, the wage penalty differs depending on how observed and unobserved traits of smokers and non-smokers are addressed. Controlling for the differences in observable characteristics in an OLS regression, the negative wage effect of smoking persists but is reduced to 4 percent. Taking into account both the differences in observable and unobservable characteristics of smokers and non-smokers, the wage penalty for smoking is statistically and economically large: smokers experience wage reductions of 21-28 percent. The large wage reduction caused by smoking is robust to several specifications, including 2SLS and FIML estimations and

---

[14] This simulation is based on figures from CIA (2006). According the official exchange rate the 2005 GDP in Albania was $8.52 B, the total population was 3,500,000 people, and the average monthly wage of non-smokers is about $230.

controls for occupation and intensity of smoking. While better-educated smokers have a slightly lower wage penalty, there are otherwise no differences by other characteristics (age and location). Due to data limitations, we can not empirically assess to what extent each of the competing hypotheses contributes to the large wage reduction from smoking, although, the lack of an age-gradient in the wage penalty suggests that time preferences can not fully explain the wage reduction.

These results suggest that there are large economic costs of smoking in low-income settings. Some caveats to this analysis should be mentioned. First, we restrict our sample to employed males. Employment is likely to be endogenous to smoking: if smokers are more likely than non-smokers to exit the labor force, for example due to poor health caused by smoking, then our estimates of the wage differential between smokers and non-smokers may be biased. We speculate that our results underestimate the actual effect of smoking on earnings as a result of this bias. Finally, excluding women may result in an incomplete computation of the societal cost of smoking if wage-effects of smoking differ between men and women.

References

Andrews, J. A., Hops, H., Ary, D. V., Tildesley, E., and Harris, J. (1993) "Parental Influence on Early Adolescent Substance Use: Specific and Nonspecific Effects." *Journal of Early Adolescence*, Vol. 13(3): 285–310.

Auld, M. (2005) "Smoking, Drinking, and Income." *Journal of Human Resources*, Vol. 40(2): 505-518

Becker, G., and K. Murphy (1988) "A Theory of Rational Addiction." *Journal of Political Economy*, Vol. 96: 675-700

Becker, G., Grossman, M., and K. Murphy (1994) "An Empirical Analysis of Cigarette Addiction." *American Economic Review*, Vol. 84(3): 396-418

Berger, M. and J. Leigh (1989) "The Effects of Smoking and Being Overweight on Current Earnings." *American Journal of Preventive Medicine*, Vol. 5(1): 8-14

Bertera, R. (1991) "The Effects of Behavioral Risks on Absenteeism and Health-Care Costs in the workplace." *Journal of Occupational Medicine* Vol. 33(11): 1119-23

Brandon, T. H., and T. Baker (1991) "The Smoking Consequences Questionnaire: The Subjective Expected Utility of Smoking in College Sstudents." *Psychological Assessment*, Vol. 3: 484–491.

Chaloupka, F. (1991) "Rational Addictive Behavior and Cigarette Smoking." *Journal of Political Economy*, Vol. 99: 722-734

Center for Disease Control (1994) *Preventing Tobacco Use among Young People – A Report of the Surgeon General*. US

Center for Health Promotion and Publications (2000) *The Dollar (and sense) Benefits of Having a Smoke-Free Workplace*. Michigan Tobacco Control Program, Lansing, Michigan

Center Intelligence Agency (2006) *The World Factbook,* online addition http://www.cia.gov/cia/publications/factbook/index.html, Washington, DC

Cornwell, C. and P. Rupert (1997) "Unobserved Individual Effects, Marriage and the Earnings of Young Men," *Economic Inquiry*, Vol. 35(2): 285-94

Daniel, K. (1991) "Does Marriage Make Men More Productive?" mimeo, University of Chicago

European Network for Smoking Prevention (2006) *European News Bulletin No 14.9.* ENSP, www.ensp.org

Evans, W., and E. Montgomery (1994) "Education and Health: Where There's Smoke There's an Instrument." NBER working paper #4949

FAO (2003) *Projections of Tobacco Production, Consumption and Trade to the Year 2010.* Food and Agriculture Organization of the United Nations, Italy, Rome

Fagan, P., Brook, J., Rubenstone, E., and Z. Chenshu (2005) "Parental Occupation, Education, and Smoking as Predictors of Offspring Tobacco Use in Adulthood: A Longitudinal Study." *Addictive Behavior*, Vol. 30: 517-529

Fuchs, V. (1982) "Time Preference and Health: An Exploratory Study." In V.R. Fuchs (Ed.) *Economic Aspects of Health*, Chicago, IL: University of Chicago Press, 93-120

Gray, J. (1997) "The Fall in Men's Return to Marriage: Declining Productivity Effects or Changing Selection?" *Journal of Human Resources*, Vol. 32(3): 481-504.

Hausman, J. (1983) "Specification and Estimation of Simultaneous Equations Models." in Griliches, Z., and M., Intriligator, eds. *Handbook of Econometrics*, Amsterdam: North-Holland

Heckman, J. (1979) "Sample Selection Bias as a Specification Error." *Econometrica,* Vol. 47(1): 153–161.

Heineck, G., and J. Schwarze (2003) "Substance Use and Earnings: The Case of Smokers in Germany." Discussion Paper # 743, IZA, Germany

Hoad, N., and D. Clay (1992) "Smoking Impairs the Response to a Physical Training Regime: a Study of Officer Cadets." *Journal of the Royal Army Medical Corps,* Vol. 138:115-117

International Monetary Fund (2006) "Albania: Poverty Reduction Strategy Paper—Annual Progress Report." IMF Country Report No. 06/23, IMF, Washington, DC

Jha, P., and F. Chaloupka (2000) *Tobacco control in developing countries*. Oxford University Press Inc., New York

Lance, P., Akin, J., Dow, W., and C. Loh (2004) "Is cigarette smoking in poorer nations highly sensitive to price? Evidence from Russia and China." *Journal of Health Economics*, Vol. 23: 173-189

Lahiri, K., and J. Song (2000) "The Effect of Smoking on Health Using a Sequential Self-Selection Model," *Health Economics*, Vol. 9: 491-511

Lee, L. (1978) "Unionism and Wage Rates: A Simultaneous Equations Model with Qualitative and Limited Dependent Variables." *International Economic Review*, Vol. 19: 415-433

Levine, P., Gustafson, T., and A. Valenchik (1997) "More Bad News for Smokers? The Effect of Cigarette Smoking on Wages," *Industrial and Labor Relations Review*, Vol. 50: 493-509

Lye, J., and J. Hirschberg (2004) "Alcohol Consumption, Smoking and Wages," *Applied Economics,* Vol. 36: 1807-1817

Lokshin, M., and Z. Sajaia (2004) "Maximum-likelihood Estimation of Endogenous Switching Regression Models," *STATA Journal*, Vol. 4(3): 227-234

MacDonald, Z., and M. Shields (2001) "The Impact of Alcohol Use on Occupational Attainment in England." *Economica*, Vol. 68: 427:54

Maddala, G. (1983) *Limited-dependent and Qualitative Variables in Econometrics,* Cambridge: Cambridge University Press

Mincer, J. (1974) *Schooling, experience, and earnings,* Brookfield, VT: Ashgate Publishing Company.

Nigel, T. (1984) "The Effects of Religion and Denomination on Earnings and the Returns to Human Capital," *Journal of Human Resources* Vol.19 (4): 472-88.

Pedersen, W., and N. Lavik (1991) "Role modeling cigarette smoking: Vulnerable working class girls? A longitudinal study," *Scandinavian Journal of Social Medicine*, Vol. 19(2): 110–115

Rohsenow, D., Abrams D., Monti, P., Colby S., Martin R. and R. Niaura (2003) "The smoking effects questionnaire for adult populations. Development and psychometric properties." *Addictive Behaviors*, Vol. 28: 1257-1270

Sargan, J. (1958) "The Estimation of Economic Relationships Using Instrumental Variables." *Econometrica* Vol. 26: 393-415

Senagore, A. (2004) "Smoking Cessation." In Gale *Encyclopedia of Surgery: A Guide for Patients and Caregivers.* Ed. A. Senagore, Gale, Farmington Hills, MI, 2004.

Shapo, L., Gilmore, A., Coker, R., McKee M., and E. Shapo (2003) "Prevalence and Determinants of Smoking in Tirana City, Albania: A Population-based Survey." *Public Health*, Vol. 117(4): 228-36

Staiger, D., and J. Stock (1997) "Instrumental Variables Regression with Weak Instruments." *Econometrica*, Vol. 65(3): 557-586

Van Ours, J. (2002) "A Pint a Day Raises a Man's Pay; but Smoking Blows That Gain Away." Discussion Paper #3308, Centre for Economic Policy Research, Tilburg University

The World Bank (1999) *Curbing the Epidemic: Governments and the Economics of Tobacco Control*. The World Bank, Washington, DC

World Health Organization (2002) *European Country Profiles on Tobacco Control, 2001*. Copenhagen: WHO Regional Office for Europe

_____ (1999) *Making a Difference, World Health Report 1999.* Geneva, Switzerland

Table 1: Summary statistics of the main variables used in the analysis

	Smokers		Never smoked	
	Mean	Std. Error	Mean	Std. Error
Log monthly wage	3.186	0.511	3.273	0.543
Mother ever smoked	0.065	0.247	0.040	0.197
Father ever smoked	0.701	0.458	0.506	0.500
Age in years	40.506	10.147	39.609	10.851
Marital status				
Single	0.123	0.329	0.190	0.393
Married	0.856	0.352	0.808	0.394
Divorced	0.005	0.072	0.001	0.031
Widowed	0.011	0.102	0.000	0.000
Religious believes				
Muslim	0.821	0.384	0.791	0.407
Christian Orthodox	0.077	0.267	0.094	0.292
Christian Catholic	0.049	0.216	0.074	0.261
Bektashian	0.037	0.189	0.033	0.180
Other religions	0.016	0.125	0.008	0.088
Education				
Primary	0.425	0.495	0.350	0.477
Secondary general education	0.193	0.395	0.228	0.420
Vocational education 2-3 years	0.026	0.160	0.029	0.169
Vocational education 4-5 years	0.230	0.421	0.212	0.409
University diploma	0.125	0.331	0.181	0.385
Speaks English	0.028	0.165	0.064	0.244
Speaks Italian	0.058	0.234	0.083	0.277
Speaks Greek	0.079	0.270	0.090	0.287
Hours worked per week	46.525	9.959	46.558	9.786
Household characteristics				
Household size	4.917	1.803	4.864	1.600
Share of children 0 to 6 years	0.079	0.126	0.074	0.126
Share of children 7 to 14 years	0.162	0.177	0.148	0.175
Share of elderly	0.064	0.119	0.072	0.124
Urban household	0.666	0.472	0.677	0.468
Education of the father				
Primary	0.489	0.500	0.462	0.499
Secondary general education	0.244	0.430	0.211	0.408
Vocational education 2-3 years	0.127	0.333	0.114	0.318
Vocational education 4-5 years	0.046	0.209	0.068	0.251
University diploma	0.095	0.293	0.145	0.353
Education of the mother				
Primary	0.548	0.498	0.546	0.498
Secondary general education	0.214	0.411	0.201	0.401
Vocational education 2-3 years	0.116	0.320	0.091	0.288
Vocational education 4-5 years	0.026	0.160	0.036	0.187
University diploma	0.095	0.293	0.126	0.332
Number of observations	**569**		**1009**	

Table 2: 2SLS estimation of the effect of smoking on wages.

| | OLS | | 2SLS | | | |
| | | | First Stage | | Second Stage | |
	Coeff.	Std. Err	Coeff.	Std. Err	Coeff.	Std.Err
Instrument: *Father smoked*			0.103*	0.056♣		
Instrument: *Mother smoked*			0.153***	0.025♣		
Current smoker	-0.048*	0.026			-0.256	0.154
Respondents who never smoked			*Reference category*			
Age	0.098*	0.048♣	0.109*	0.047♣	0.120*	0.051♣
Age^2/100	-0.265*	0.125♣	-0.279*	0.123♣	-0.321*	0.131♣
Age^3/10000	0.216*	0.105♣	0.223*	0.103♣	0.260*	0.109♣
Marital status						
Single			*Reference category*			
Married	0.098*	0.056	0.046	0.056	0.111*	0.057
Divorced/Widowed	-0.324	0.257	0.405	0.254	-0.228	0.266
Religious believes						
Muslim			*Reference category*			
Christian Orthodox	0.036	0.048	-0.050	0.047	0.027	0.048
Christian Catholic	-0.003	0.059	-0.122*	0.058	-0.032	0.062
Bektashian	-0.037	0.073	0.027	0.072	-0.037	0.073
Other	-0.035	0.118	0.236*	0.117	0.011	0.123
Education						
Primary			*Reference category*			
Secondary general education	-0.007	0.034	-0.080*	0.033	-0.026	0.036
Vocational education 2-3 years	0.058	0.076	-0.035	0.075	0.045	0.076
Vocational education 4-5 years	0.101**	0.034	-0.043	0.034	0.091*	0.035
University diploma	0.340***	0.042	-0.106*	0.041	0.314***	0.046
Speaks English	0.288***	0.066	-0.082	0.065	0.273***	0.067
Speaks Italian	0.083	0.051	0.017	0.051	0.085	0.051
Speaks Greek	0.186***	0.046	-0.019	0.045	0.181***	0.046
Log hours worked per week	0.404***	0.053	0.012	0.053	0.404***	0.053
Household characteristics						
Log household size	-0.140	0.178	-0.030	0.177	-0.159	0.179
Log household size^2	0.048	0.058	0.018	0.058	0.056	0.059
Urban household	0.024	0.030	-0.014	0.030	0.021	0.030
Share of children 0 to 6 years	-0.231*	0.127	-0.140	0.125	-0.258*	0.128
Share of children 7 to 14 years	0.053	0.089	-0.019	0.087	0.048	0.089
Share of elderly	-0.178	0.115	0.091	0.114	-0.173	0.115
Constant	0.397	0.645	-1.092*	0.638	0.207	0.659
F-statistic for excluded instruments (P>F)			16.91	(0.000)		
R^2	0.222		0.102		0.152	
Sargan overid test (χ^2)			0.001	(0.981)		
Number of observations	1588		1588			

Note: * is significant at 10% level; ** at 5% level; *** at 1% level. Regressions also include complete set of geographic dummy variables for the 38 districts and a set of dummy variables reflecting educational achievements of respondent's parents. ♣ indicates joint significance of coefficients at 10% level.

Table 3: FIML estimation of the switching regression model.

	Probability to smoke		Wage of smokers		Wage of never smokers	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
Instrument: *Father smoked*	**0.297***	**0.144♣**				
Instrument: *Mother smoked*	**0.450*****	**0.079♣**				
Age	0.324*	0.139	0.080	0.094	0.136*	0.061
$Age^2/100$	-0.805*	0.363	-0.197	0.241	-0.375*	0.160
$Age^3/10000$	0.635*	0.303	0.145	0.197	0.314*	0.135
Marital status						
Single			Reference Category			
Married	-0.004	0.153	0.049	0.080	0.162*	0.074
Divorced/Widowed	0.874	0.810	-0.506	0.332	0.457	0.520
Religious believes						
Muslim			Reference Category			
Christian Orthodox	-0.160	0.141	0.103	0.085	-0.017	0.059
Christian Catholic	-0.378*	0.174	-0.147	0.116	-0.013	0.073
Bektashian	0.168	0.204	0.058	0.113	-0.077	0.096
Others	0.644*	0.325	0.073	0.177	-0.007	0.177
Education						
Primary			Reference Category			
Secondary general education	-0.237*	0.096	-0.001	0.061	-0.031	0.044
Vocational education 2-3 years	-0.096	0.215	0.031	0.126	0.054	0.094
Vocational education 4-5 years	-0.148	0.096	0.055	0.054	0.108*	0.046
University diploma	-0.305*	0.119	0.214*	0.081	0.335***	0.055
Speaks English	-0.303	0.202	0.792***	0.143	0.145*	0.077
Speaks Italian	0.021	0.149	-0.022	0.087	0.111*	0.064
Speaks Greek	-0.044	0.131	0.044	0.079	0.243***	0.057
Log hours worked per week	0.007	0.152	0.358***	0.084	0.424***	0.067
Household characteristics						
Log household size	-0.206	0.501	-0.449	0.284	-0.003	0.232
Log household size^2	0.079	0.164	0.135	0.091	0.011	0.077
Urban household	-0.039	0.085	0.064	0.050	0.003	0.038
Share of children 0 to 6 years	-0.246	0.356	0.096	0.204	-0.460**	0.161
Share of children 7 to 14 years	0.014	0.248	0.196	0.144	-0.006	0.112
Share of elderly	0.353	0.327	-0.113	0.195	-0.150	0.142
Constant	-4.475*	1.888	0.545	1.321	-0.072	0.803
$\sigma_{1,2}$			0.485	0.080	0.479	0.015
$\rho_{1,2}$			0.703*	0.350	0.242**	0.109
LR test of independent equations						
$\chi2$ (Prob> $\chi2$)	3.995**	0.012				
Number of observations	1588					

Note: * is significant at 10% level; ** at 5% level; *** at 1% level. Regressions also include complete set of geographic dummy variables for the 38 districts and a set of dummy variables reflecting educational achievements of mother and father of the respondents. ♣ indicates joint significance of coefficients at 10% level.

Table 4: OLS regressions of individual wages for smokers and quitters

	Smokers Specification 1		Smokers Specification 2		Quitters	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
Years of smoking	-0.008	0.005	-0.008	0.005	0.009	0.010
Cigarettes per day			0.002	0.002		
Age	0.034	0.082	0.031	0.082	-0.291	0.305
Age^2/100	-0.070	0.212	-0.061	0.212	0.723	0.747
Age^3/10000	0.051	0.175	0.044	0.175	-0.605	0.591
Marital status						
Single			*Reference Category*			
Married	0.018	0.087	0.017	0.087	0.260	0.450
Divorced/Widowed	-0.272	0.195	-0.280	0.195		
Religious believes						
Muslim			*Reference Category*			
Christian Orthodox	0.119	0.085	0.121	0.085	0.620*	0.286
Christian Catholic	-0.076	0.108	-0.085	0.109	0.453	0.333
Bektashian	0.099	0.112	0.103	0.112	-0.143	0.293
Others	-0.053	0.158	-0.044	0.159	-0.013	0.431
Education						
Primary			*Reference Category*			
Secondary general education	0.040	0.057	0.044	0.057	-0.081	0.230
Vocational education 2-3 years	0.086	0.126	0.077	0.127	-0.170	0.346
Vocational education 4-5 years	0.074	0.053	0.075	0.053	0.129	0.170
University diploma	0.279***	0.072	0.282***	0.072	0.238	0.219
Speaks English	0.813***	0.136	0.805***	0.137	0.208	0.530
Speaks Italian	-0.020	0.087	-0.015	0.087	0.803*	0.340
Speaks Greek	0.071	0.079	0.073	0.079	0.082	0.427
Log hours worked per week	0.407***	0.086	0.402***	0.087	0.519	0.310
Household characteristics						
Log household size	-0.313	0.287	-0.281	0.289	1.375	1.691
Log household size^2	0.087	0.091	0.080	0.091	-0.392	0.511
Urban household	0.074	0.050	0.072	0.050	-0.146	0.163
Share of children 0 to 6 years	0.194	0.208	0.201	0.208	0.227	0.844
Share of children 7 to 14 years	0.246*	0.147	0.228	0.148	0.337	0.591
Share of elderly	-0.135	0.192	-0.141	0.192	0.250	0.663
Constant	1.043	1.127	1.020	1.127	2.681	3.973
Adjusted R^2	0.349		0.350		0.617	
Number of observations	569		569		95	

Note: * is significant at 10% level; ** at 5% level; *** at 1% level. Regressions also include a geographic dummy variable for living in Tirana and a set of dummy variables reflecting educational achievements of mother and father of the respondents. ♣ indicates joint significance of coefficients at 10% level.
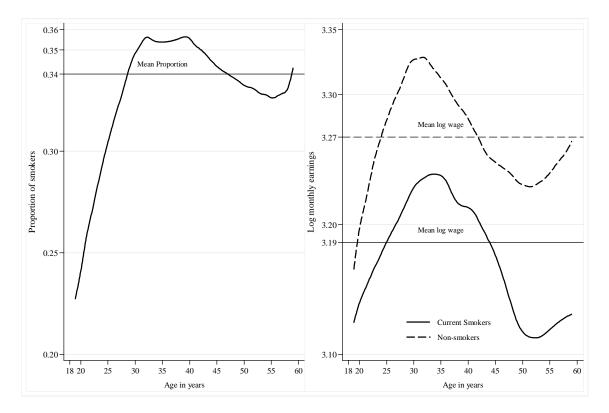
Figure 1: Proportion of smokers and average log monthly earnings for smokers and non-smokers by age. Working males 25 – 60 years of age.

Appendix

Table A1: Summary statistics for the sample of quitters

	Quitters	
	Mean	Std. Error
Log monthly wage	3.200	0.531
Mother ever smoked	0.137	0.346
Father ever smoked	0.842	0.367
Age in years	44.695	8.838
Marital status		
Single	0.084	0.279
Married	0.916	0.279
Divorced	0.000	0.000
Widowed	0.000	0.000
Religious believes		
Muslim	0.768	0.424
Christian Orthodox	0.095	0.294
Christian Catholic	0.053	0.224
Bektashian	0.063	0.245
Other religions	0.021	0.144
Education		
Primary	0.263	0.443
Secondary general education	0.147	0.356
Vocational education 2-3 years	0.042	0.202
Vocational education 4-5 years	0.316	0.467
University diploma	0.232	0.424
Speaks English	0.021	0.144
Speaks Italian	0.042	0.202
Speaks Greek	0.032	0.176
Hours worked per week	46.337	9.718
Household characteristics		
Household size	4.916	1.616
Share of children 0 to 6 years	0.058	0.103
Share of children 7 to 14 years	0.173	0.167
Share of elderly	0.048	0.111
Urban household	0.726	0.448
Education of the father		
Primary	0.516	0.502
Secondary general education	0.137	0.346
Vocational education 2-3 years	0.179	0.385
Vocational education 4-5 years	0.042	0.202
University diploma	0.126	0.334
Education of the mother		
Primary	0.589	0.495
Secondary general education	0.158	0.367
Vocational education 2-3 years	0.105	0.309
Vocational education 4-5 years	0.021	0.144
University diploma	0.126	0.334
Number of observations	**95**	

Table A2: FIML estimation of the switching regression model, *including quitters*

	Probability to smoke		Wage of smokers		Wage of non-smokers	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
Instrument: *Father smoked*	0.180	*0.127*♣				
Instrument: *Mother smoked*	0.397***	*0.077*♣				
Age	0.357*	*0.136*♣	0.105	*0.098*♣	0.123*	*0.059*♣
Age^2/100	-0.898*	*0.354*♣	-0.261	*0.253*♣	-0.332*	*0.155*♣
Age^3/10000	0.714*	*0.295*♣	0.198	*0.207*♣	0.273*	*0.129*♣
Marital status						
Single			*Reference Category*			
Married	-0.006	*0.151*	0.052	*0.083*	0.150*	*0.072*
Divorced/Widowed	0.903	*0.791*	-0.447	*0.348*	0.493	*0.517*
Religious believes						
Muslim			*Reference Category*			
Christian Orthodox	-0.134	*0.137*	0.104	*0.086*	0.005	*0.057*
Christian Catholic	-0.402*	*0.171*	-0.169	*0.118*	0.003	*0.071*
Bektashian	0.094	*0.194*	0.038	*0.117*	-0.105	*0.088*
Other	0.627*	*0.307*	0.093	*0.177*	-0.032	*0.159*
Education						
Primary			*Reference Category*			
Secondary general education	-0.244*	*0.094*	-0.009	*0.062*	-0.045	*0.043*
Vocational education 2-3 years	-0.145	*0.208*	0.015	*0.130*	0.027	*0.089*
Vocational education 4-5 years	-0.180*	*0.093*	0.045	*0.056*	0.103*	*0.043*
University diploma	-0.360**	*0.116*	0.187*	*0.085*	0.316***	*0.052*
Speaks English	-0.251	*0.199*	0.787***	*0.144*	0.149*	*0.076*
Speaks Italian	0.045	*0.145*	-0.014	*0.089*	0.125*	*0.062*
Speaks Greek	-0.017	*0.129*	0.047	*0.080*	0.244***	*0.056*
Log hours worked per week	-0.010	*0.146*	0.352***	*0.086*	0.432***	*0.065*
Household characteristics						
Log household size	-0.269	*0.489*	-0.478	*0.291*	0.011	*0.226*
Log household size^2	0.095	*0.160*	0.143	*0.093*	0.005	*0.075*
Urban household	-0.047	*0.082*	0.058	*0.052*	-0.023	*0.036*
Share of children 0 to 6 years	-0.242	*0.349*	0.081	*0.210*	-0.383*	*0.156*
Share of children 7 to 14 years	0.006	*0.243*	0.191	*0.148*	-0.004	*0.108*
Share of elderly	0.400	*0.319*	-0.096	*0.199*	-0.107	*0.136*
Constant	-4.771*	*1.839*	0.222	*1.364*	-0.044	*0.776*
$\Sigma_{1,2}$			0.513	*0.077*	0.485	*0.017*
$P_{1,2}$			0.843*	*0.432*	0.349*	*0.176*
LR test of independence equations X2 (Prob> χ2)	5.943**	*0.014*				
Number of observations	1683					

Note: * is significant at 10% level; ** at 5% level; *** at 1% level. Regressions also include complete set of geographic dummy variables for the 38 districts and a set of dummy variables reflecting educational achievements of mother and father of the respondents. ♣ indicates joint significance of coefficients at 10% level.