The Smoker's Wage Penalty Puzzle —
Evidence from Britain

Lasse F. Brune

ISER Working Paper
2007-31

Acknowledgement:

I am grateful to Stephen Jenkins for helpful comments on earlier versions of this work.

Readers wishing to cite this document are asked to use the following form of words:

> **Brune, Lasse F. (December 2007) 'The Smoker's Wage Penalty Puzzle — Evidence from Britain', ISER Working Paper 2007-31. Colchester: University of Essex.**

The on-line version of this working paper can be found at http://www.iser.essex.ac.uk/pubs/workpaps/

The Institute for Social and Economic Research (ISER) specialises in the production and analysis of longitudinal data. ISER incorporates

- MISOC (the ESRC Research Centre on Micro-social Change), an international centre for research into the lifecourse, and
- ULSC (the ESRC UK Longitudinal Studies Centre), a national resource centre to promote longitudinal surveys and longitudinal research.

The support of both the Economic and Social Research Council (ESRC) and the University of Essex is gratefully acknowledged. The work reported in this paper is part of the scientific programme of the Institute for Social and Economic Research.

Institute for Social and Economic Research, University of Essex, Wivenhoe Park,
Colchester. Essex CO4 3SQ UK
Telephone: +44 (0) 1206 872957 Fax: +44 (0) 1206 873151 E-mail: iser@essex.ac.uk
Website: http://www.iser.essex.ac.uk

ABSTRACT

This work investigates the effect of smoking on wages for male workers using panel data from Britain for the period of 1991–2005. The strong negative correlation of smoking and wages found in a cross-sectional analysis reduces substantially when accounting for unobserved individual heterogeneity using Fixed Effects estimation. I find a statistically significant wage penalty that is causally due to smoking of about –2% for smokers over those who quit. Further analysis indicates, however, that the negative effect might be underestimated when comparing with those who never started smoking or quit a long time ago.

NON-TECHNICAL SUMMARY

Smoking is currently an important topic in public policy in many countries, including Britain. Starting with Scotland in March 2006 and ending with England in July 2007, Britain has only recently introduced far-reaching smoking bans for enclosed public places. This paper investigates indirect individual costs of smoking due to lower wages.

The main challenge of measuring the effects of smoking on wages or earnings is to disentangle causation from mere statistical correlation. A number of channels have been proposed that explain why smoking may causally affect earnings or wages, such as lower productivity due to bad health and higher rates of absenteeism from work. But empirically, a correlation may also arise because smokers are different from non-smokers with respect to characteristics relevant in the labour market. Thus, differences in wages may not be due to smoking itself but due to differences in observed characteristics such as education or in differences in unobserved characteristics such as motivation or self-discipline.

Previous studies have identified puzzlingly high effects of smoking on wages and earnings. Using the extensive information on smoking behaviour in the BHPS, this paper investigates for Britain whether cigarette smoking affects wages. To determine the effect I use in particular the structure of the data set with repeated observations per individual to compare individuals' wages when smoking with wages of the same individuals when not smoking.

The results suggest that there is a small causal, statistically significant effect from smoking on hourly wages for men of about –2% when compared to those who quit smoking. The negative effect of persistent smoking, however, may be underestimated when comparing wages with those who never smoked or quit a long time ago.

1 Introduction

Smoking is currently an important topic in public policy in many countries, including Britain. Starting with Scotland in March 2006 and ending with England in July 2007, Britain has only recently introduced far-reaching smoking bans for enclosed public places. The policy is mainly aimed at the protection of the public from harmful effects on health of passive smoking, which also translate into economic costs for society due to e.g. increased costs for health care. But while smokers impose external costs on society, they also face private economic and health costs. The adverse health effects of smoking in terms of higher probability of death and illness are well-established. Private economic costs for the smoker include direct costs from buying cigarettes. Further, there may be *indirect* private costs through lower wages that are due to smoking. This possibility is investigated in the present paper.

The main challenge of measuring the effects of smoking on wages or earnings is to disentangle causation from mere statistical correlation. A number of channels have been proposed that explain why smoking may causally affect earnings or wages, such as lower productivity due to bad health and higher rates of absenteeism from work. But empirically, a correlation may also arise because smokers are different from non-smokers, i.e. selection into the group of smokers is non-random with respect to characteristics relevant in the labour market. To the extent that smokers are different from non-smokers in observed characteristics (e.g. education, age, etc.) this can be controlled for by including them as covariates in a standard Ordinary Least Squares (OLS) regression. However, if selection takes place on unobservable characteristics of the individual (e.g. motivation or self-discipline), estimation of the causal effects of smoking on wages becomes more complicated.

A few studies have addressed the impact of tobacco use on earnings or wages and a negative effect from smoking is generally found for men, even after controlling unobserved heterogeneity in personal characteristics.[1] The magnitude of estimates varies widely across studies and includes statistically significant negative effects on earnings as high as –20% (Auld, 2005; Lokshin and Beegle, 2006). Auld's results lead him to suggest that this is "a new puzzle" (Auld, 2005, p. 506): The apparently high earnings penalty for smokers cannot be credibly explained by the hypothesised mechanisms by which causal effects arise.

The present paper investigates for Britain whether cigarette smoking affects wages. No previous work has addressed this issue using British data. The British Household Panel Study (BHPS) data used in this work has extensive information on smoking behaviour. While previous studies using panel data (Baum et al., 2006; Heineck and Schwarze, 2003; Levine et al., 1997) are limited to at most three repeated observations per individual in irregular intervals, this work draws on data from yearly re-interviews with individuals for up to 15 years. Using the panel structure of the data, a Fixed Effects (FE) approach is used to control for unobserved individual effects. As an alternative approach, the method of Instrumental Variables (IV) is applied. The IV approach uses variables uncorrelated with the unobserved effects to identify the causal effects of smoking on wages. The results from the preferred FE estimation suggest that there is a small causal, statistically significant effect from smoking on hourly wages for men of about –2% when compared to those who quit smoking. Further analysis suggests, however, that the negative effect of permanent smoking is underestimated when comparing with those who never smoked or quit a long time ago.

[1]For women no statistically significant impact of smoking on wages is usually found.

The paper unfolds as follows. Section 2 describes the theoretical arguments concerning the smoking-wage relationship and discusses the previous literature. Section 3 outlines the empirical strategy and provides an overview of the data. The presentation and discussion of the estimation results follows in Section 4. Finally, section 5 provides a summary and conclusions.

2 Theoretical Background and Previous Work

2.1 Theoretical Background

In principle, there are two reasons why one may observe that smokers earn less than non-smokers. One is that smoking *causes* smokers to earn less. The other is that smokers are different from non-smokers, so that selection into smoking is not random. A number of theoretical arguments have been put forward to explain the causal effect.

First, smokers may be less productive and more costly for employers. Smoking itself takes time away from work, in particular if smokers have to go outside for smoking breaks due to indoors smoking bans at the workplace. Smokers also have higher rates of illness and (therefore) higher rates of absenteeism (Bertera, 1991; Halpern et al., 2001; Leigh, 1995). Lower productivity is expected to result in lower pay on average.

Furthermore, employers may already anticipate future lower productivity from accumulating adverse health effects and take into account additional costs due to substitute or replacement recruitment. Thus, they might be less willing to employ smokers, all else equal. Anecdotal evidence suggests that firms are well aware of these issues and take account of the the applicant's smoker status during the recruitment process.[2]

Smoking at the workplace can also entail increased costs for the employer through additional cleaning and other maintenance work, more or special requirements for ventilation, higher fire insurance premia (Kristein, 1983) or the provision of special facilities such as smoking areas or "smoking rooms". Employers may compensate higher costs that a smoker causes with lower pay.[3]

Another reason why smokers and non-smoker may have different wages is discrimination (Levine et al., 1997). Colleagues or clients might object to smoking, e.g. because of annoyance and adverse health effects from passive smoking. In response, some employers could discriminate against smokers. This point could be less relevant in the UK in the future as ever more firms have no-smoking policies and the recent smoking laws banned smoking in public places, so that passive smoking at the workplace will be less of an issue. (Note, however, that more productive time may be lost due to outside smoke breaks.) Nevertheless, for the time frame of this study, 1991-2005, the discrimination argument may still be valid.

A number of arguments suggest that unobserved heterogeneity could explain the empirical relationship of smoking and wages — smokers and non-smokers might differ in unobservable characteristics. Regular smoking can be characterised by predominantly short-term benefits that are chosen over the known long-term costs. Smokers may thus have higher rates of time preference, i.e. in comparison with non-smokers, smokers may value the present higher relative to the future. Fuchs (1982) shows a

[2]As an example, one method to illicit this information is to leave applicants into an "artificial" break in which the employer then observes if the applicant uses the time to smoke.

[3]At the relevant margin, however, a worker who smokes may increase those costs only slightly since e.g. extra smoking areas and increased fire insurance premia are largely fixed costs.

relationship between time preference and investment in health as measured in part by smoking habits, though small in size and statistically weak. Since decisions about education are also believed to be influenced by an individual's time preference (see e.g. Evans and Montgomery, 1994), a higher discounting of the future may be related to investment in human capital in general such as job training. Less investment in human capital would result in lower wage growth over the working life and thus, all else equal, in lower wages.

Using US data Munasinghe and Sicherman (2005) also argue that the wage gap between smokers and non-smokers is due to a gap in wage growth. They develop a theory of career choice and devise a test to discriminate between time preference and learning ability as sources of unobserved heterogeneity. From their empirical results they conclude that smoking does proxy for time preference with the consequence that smokers choose flatter wage profiles.

A recent study by Khwaja et al. (2007), however, suggests that smokers and non-smokers may not differ in what has been used traditionally in economic theory as a measure for time preference — discount rates implied by trade-offs between present and future costs and benefits. Discount rates are found to be uncorrelated with the smoker status. The authors produce evidence that suggests that more general measures of time preference and self control such as planning horizons and impulsivity are related to the decision to smoke.

Smoking is generally known to be health damaging. But full anticipation of the consequences or competence to acquire this information may be linked to individual characteristics such as intelligence or as Levine et al. (1997, p.496) suggest, poor judgement. Research in psychology (cf. Gilbert, 1995, chap. 4) indicates that smoking is linked to certain personality traits that are relatively stable over time such as higher impulsivity, lack of conscientiousness or neuroticism.

Hence, smokers and non-smokers may differ in various unobservable characteristics that are relevant in the labour market. It is useful to note that all individual characteristics suggested above have in common that they are usually expected to cause smokers to have lower wages than non-smokers, i.e. smokers are expected to have a higher discounting rate of the future, to be more impulsive, to have less self-control and self-discipline or to exhibit poorer judgment, etc. If true, then this implies a *negative* selection bias if not accounted for — smokers are those that earn less, but smoking is not the cause but a *proxy* for the cause.

2.2 Previous Work

A growing body of literature investigates the relationship between smoking and wages. The first study comes from the medical literature and explores the statistical associations of smoking and being overweight with earnings (Leigh and Berger, 1989). Using US data for the period 1972-1973 no associations for either smoking or being overweight were found to be significant.

The subsequent studies from the economics literature all attempt to indentify causal effects from smoking by controlling for unobserved heterogeneity. Levine et al. (1997) is the first paper from the economics literature on the topic. The authors draw on the National Longitudinal Survey of Youth which allows them to control for an extensive range of personal and family background characteristics (including results from an aptitude test). With observations for the same individuals including information about smoking behaviour in two different years, 1984 and 1991, individual unobserved effects are controlled for by differencing individual observations across time. Additionally, unobserved family background effects are netted out by taking differences across siblings. The authors find a statistically significant reduction in the hourly wage rate for regular smokers in the range of 4%–8% compared to non-smokers depending on the specification. Using the same data set for the years 1992, 1994 and 1998, Baum et al. (2006) jointly estimate the wage effects of smoking jointly with obesity. They find no statistical significant effects of smoking on wages when controlling for unobserved heterogeneity using panel data methods.

Heineck and Schwarze (2003) also use panel data to control for unobserved individual characteristics. Using the years 1998, 1999 and 2001 from the German Socio-Economic Panel Study they find a small negative difference in earnings between smokers and non-smokers with OLS. Controlling for unobserved effects only for male smokers aged 25–35 years compared to their non-smoking counterparts a small *positive*, but only weakly significant effect is found. Using a cross-section from the same data set for the year 2002, Anger and Kvasnicka (2006) employ an IV approach to control for unobserved heterogeneity. Discerning explicitly never-smokers and former smokers, they find a wage penalty of –6%, albeit non-significant, for smokers compared to those who never smoked.

A few studies have investigated the causal effect of smoking on wages jointly with the effect of alcohol use. Lee (1999) uses differences across twins in an Australian data set from 1988/1989. Net of unobserved genetic and family background effects, the effect of smoking on earnings *increases* in magnitude to –5% from around –3% obtained with OLS esimation.[4] In a paper that draws upon data from 2001 in the

[4]The journal article based on the cited working paper (Lee, 2003) does not include results concerning the effects of smoking but focuses on the impact of alcohol use.

Netherlands, Van Ours (2004) also estimates the effects on wages of both smoking and drinking. The results from a simultaneous equation approach with instrumental variables to control for unobserved heterogeneity suggest a negative effect on hourly wages for male smokers of about –10%. The latter is larger in magnitude than the effect obtained in OLS which are estimated to be around –6%. No effect of smoking is found for women.

Auld (2005) conducts another study on the joint impact of smoking and wages using Canadian data with two pooled cross sections of prime age male workers for the years of 1982 and 1992. Like Van Ours (2004), Auld finds that the estimate of the negative effect of smoking on earnings from OLS is less negative than when controlling for unobserved heterogeneity via an IV approach with which estimates increase in magnitude from –8% using OLS to about –24%. Using recent data from Albania, Lokshin and Beegle (2006) reports similar results also for prime age male workers: The OLS estimate of –8% using OLS increase in magnitude to –20% when controlling for unobserved heterogeneity using IV.

3 Empirical Strategy and Data

3.1 Empirical Strategy

I now discuss the empirical strategy used to identify the effect of cigarette smoking on wages. This paper aims at identifying causal effects as opposed to mere statistical correlations. In particular, unobserved individual effects are taken into account. First, OLS is presented as the baseline estimation method and its shortcoming are described. Subsequently, FE and the method of IV are introduced to remedy the problems of OLS.

3.1.1 The Baseline Specification: Pooled Ordinary Least Squares

In order to estimate the effect of smoking on wages, a wage regression with log hourly wages as the dependent variable and a set of human capital variables and other controls on the right-hand side is modified to include variables summarizing smoking behaviour. Given the panel nature of the data the following linear static panel data model for an individual randomly drawn from the population is a natural starting point and serves as the baseline case. Cross-sections are pooled across time to make full use of the data:

$$w_{it} = x_{it}\beta + s_{it}\gamma + \epsilon_{it}, \qquad i = 1, 2, \ldots, N,\ t = 1, 2, \ldots, T_i, \tag{1}$$

with $\max(T_i) = 15$. The dependent variable, log of hourly wages, for individual i at survey year ("wave") t is denoted as w_{it}. Hourly wages are preferred over total earnings for two reasons. First, the theoretical arguments centre on the effects on productivity, and differences in a worker's productivity are expected to be reflected in wages per hour. Second, total wages may overstate a worker's earnings capacity if he works many hours, and the number of hours worked is potentially endogenous as a control variable as it is likely be jointly determined with total wages. The log form accounts for the right-skewed distribution of wages and restricts the model to non-negative values for the dependent variable.

x_{it} is a row vector with control variables that includes the explanatory variables and a constant plus a set of wave dummies to capture year fixed effects that are the same for all individuals in a given survey year. Following the standard human capital model of wage determination (cf. Mincer, 1974; Willis, 1986), x_{it} contains variables for education as well as age and age squared (as an approximation for experience).

Further control variables include marital status, ethnicity and region of residence. The principal interest is in the effect of smoking on wages. Smoking behaviour is represented by the variable s_{it}, which is a dummy variable that takes the value of 1 if the individual is a current smoker and 0 otherwise.

ϵ_{it} denotes the random disturbance which can assumed to have a mean of zero due to the inclusion of a constant: $E(\epsilon_{it}) = 0$ for all i and t. If individuals are randomly sampled from the population then, in any particular wave, the ϵ_{it} are independent and identically distributed (i.i.d.), so $E(\epsilon_{it}\epsilon_{jt}) = 0$ for $i \neq j$ at any one t. However, independence of disturbances of an individual across time is not presumed as in a panel data set the repeated observations for the same individual over time are not independent. It is also not assumed that ϵ_{is} and ϵ_{it} come from the same distribution for $s \neq t$, their respective variances are allowed to differ across the waves of the panel. While the last two points do not prevent consistent estimation via OLS, the resulting serial correlation and heteroskedasticity respectively would mean that the usual OLS standard errors would not be consistent. Hence, a flexible covariance structure is used to derive correct standard errors for statistical inference.[5]

In order to derive consistent estimates of the coefficients of the above model via Pooled Ordinary Least Squares (POLS), one would have to assume that the regressors are contemporaneously exogenous in the sense that, as a minimum, $E(x_{it}\epsilon_{it}) = 0$ and $E(s_{it}\epsilon_{it}) = 0$ for $t = 1, 2, \ldots, T_i$. This assumption might be violated if variables are omitted from the model's specification that are correlated with one or more of the regressors *and* the outcome variable. The discussion at the end of section 2.1 provided a number of reasons why, in the case of smoking and wages, there may be

[5]The Stata option `cluster` is used to account for dependence of repeated observation of an individual.

unobserved (and unobservable) individual characteristics that make an individual more likely to smoke and to receive a lower wage at the same time.

To make this point more precisely and to present a specification which can address the resulting endogeneity, the unobserved heterogeneity is modelled explicitly. For this purpose, "fixed effects" — individual specific (time invariant) effects — are introduced. Suppose the random disturbance in (1), ϵ_{it}, is composed of an *unobserved* random variable c_i, the "fixed effect", and an idiosyncratic disturbance that is random across individuals *and* across time: $\epsilon_{it} = c_i + \nu_{it}$. If the smoking variable is correlated with the unobserved effects, i.e. $E(s_{it}c_i) \neq 0$, then the exogeneity assumptions of the baseline model in equation (1) can no longer hold: $E(s_{it}\epsilon_{it}) \neq 0$. Since then plim $\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T_i} s_{it}\epsilon_{it} \neq 0$, the POLS estimator is inconsistent.

Two ways of dealing with endogeneity caused by unobserved heterogeneity are presented in the following sections. First, the preferred specification with "fixed effects" is introduced. The appropriate FE estimator explicitly uses the panel structure of the data and eliminates endogeneity resulting from unobserved time invariant individual effects. Second, the method of IV is presented which can solve cases of endogeneity more generally.

3.1.2 The Preferred Specification: Fixed Effects

The assumption of contemporaneous exogeneity of the smoking variable, which is necessary for Pooled OLS to produce consistent estimates, is likely to be violated in the proposed wage equation. Modelling the unobserved individual effect as above, with $\epsilon_{it} = c_i + \nu_{it}$, equation (1) becomes

$$w_{it} = x_{it}\beta + s_{it}\gamma + c_i + \nu_{it}, \qquad i = 1, 2, \ldots, N, \ t = 1, 2, \ldots, T_i, \tag{2}$$

The notation remains essentially unchanged from (1). The random disturbance, ν_{it}, is again assumed to have mean zero. The independent sampling assumption then implies that, in any one wave, the ν_{it} are i.i.d., so $E(\nu_{it}\nu_{jt}) = 0$ for $i \neq j$ at each t. It is assumed that $E(x_{is}\nu_{it}) = 0$ and $E(s_{is}\nu_{it}) = 0$ as well as $E(c_i\nu_{it}) = 0$ for all $s, t = 1, 2, \ldots, T_i$ and all i.

While POLS estimation requires contemporaneous exogeneity of the regressors, for FE to produce consistent estimates it is necessary to assume a form of strict exogeneity: at a minimum, $E(x_{is}\nu_{it}) = 0$ and $E(s_{is}\nu_{it}) = 0$ for all i and $s, t = 1, ..., T_i$.[6] The important difference to POLS is that, while correlation between a regressor and c_i renders POLS inconsistent, for FE estimation any such correlation is allowed. In fact, to obtain the FE estimator, c_i is eliminated by performing the so called "within transformation" and subsequently applying OLS to the transformed model. The within transformation consists of averaging each individual's observation over time and then subtracting the individual's average over time from each of the individual's observation. The estimation equation can then be expressed as:

$$\begin{gathered} w_{it} - \bar{w}_i = (x_{it} - \bar{x}_i)\,\beta + (s_{it} - \bar{s}_i)\,\gamma + \nu_{it} - \bar{\nu}_i, \\ i = 1, 2, \ldots, N,\ t = 1, 2, \ldots, T_i. \end{gathered} \tag{3}$$

The unobserved fixed effects are thus eliminated from the model and hence identification of the coefficients of the remaining (time varying) variables does not depend on the statistical properties of c_i; in particular, correlation with any of the regressors is no longer a problem.

[6] This is the minimal requirement for consistency; standard asymptotic inference assumes independence of disturbances.

By the nature of the FE estimator, identification of the model's parameters works through within-group variation, i.e. changes in the respective explanatory variables for an individual over time. With POLS, in contrast, identification comes mainly from between-group differences, i.e. variation across individuals.[7] The implications of this distinction are discussed in detail in section 4.4.

Despite the FE estimator's ability for consistent estimation in the presence of unobserved heterogeneity, other sources of endogeneity are assumed not to be present. The present model assumes that there is a one way relationship between smoking and wages in the sense that smoking influences wages but not the other way around. If higher wages cause a higher or lower incidence of smoking, then in a single equation model γ is no longer consistently estimated via FE. In early theoretical work, the demand for "good health" was predicted to increase in wages Grossman (1972, pp. 241ff). If quitting smoking or not starting smoking are means to "produce" good health, then higher wages should lead to less smoking. On the other hand, cigarettes may be a normal good, so that the income elasticity of cigarette demand may be positive. But since *price* elasticities are found to be low (cf. Chaloupka and Warner, 2000), and assuming that the substitution effect is at most moderate in magnitude, then the income effect cannot — by standard economic demand theory — be large either. Thus a change in income over time (FE uses within-group variation) is unlikely to induce many individuals to quit or start smoking. In conclusion, two-way causation may exist on theoretical grounds. But the magnitude of the bias can be assumed to be small empirically.

Statistical inference with FE is based on asymptotic standard errors that are robust to arbitrary intertemporal correlation and heteroskedasticity as introduced

[7] The POLS estimator weights within- and between-group variation equally, but since between group variation is much larger in this context, the between-group differences drive the estimates.

by Arellano (1987). So far this is not common practice in applied work, although recent theoretical work has shown its exigency (cf. Kézdi, 2007). The finding of strong evidence for serial correlation (not reported) using the test procedure devised in Wooldridge (2002, chs. 10.5.4, 10.6.3) makes these adjustments inevitable. Correspondingly, standard errors are much higher than under standard asymptotic inference that assumes zero intertemporal correlation of the ν_{it}, conditional on x_{it} and c_i.

3.1.3 An Alternative Approach: Instrumental Variables

Another approach for obtaining consistent estimates is the method of IV. Assuming that the smoking variable s_{it} is the only endogenous regressor, IV estimation applied to the present model, requires finding one or more variables, say q_{it}, that

1. are exogenous in the wage equation of model (1), i.e. $E(q_{it}\epsilon_{it})$=0 (Validity),

2. (controlling for all other covariates in the model) are correlated with the endogenous regressor, $r_{q_{it},s_{it}|x_{it}} \neq 0$ (Reliability), and

3. do *not* already belong in the wage equation as an explanatory variable in their own right, since otherwise they cannot help to identify other coefficients of the model other than their own.

If variables satisfying these conditions can be found, then the coefficients β and, in particular, γ in equation (1) can be consistently estimated via Two-Stage Least Squares (2SLS). Although estimation is not implemented in two stages, it is convenient to think of 2SLS as being carried out that way. In the first stage the smoking variable is regressed on all other exogenous variables including the instrumental variables via OLS. The second stage entails regressing, again by OLS, the original dependent variable on the predicted dependent variable of the first stage plus controls. The predicted values of the first stage regression are a linear projection of exogenous variables. Using the predicted values in the second stage then essentially has the effect of "purging" the smoking variable of correlation with the disturbance, which hence allows consistent estimation of γ.

Parental smoking behaviour provides the instrumental variables here. It has been used in a similar fashion by Lokshin and Beegle (2006). Using the BHPS, Loureiro et al. (2006) show that adolescents are more likely to smoke if their parents smoke, even after controlling for endogeneity of the parents' decision to smoke.[8] The authors suggest that the parents' function as role models for their kids, an altered perception of the risks of smoking and easier access to tobacco are potential channels of the transmission of smoking habits across generations. Since most adult smokers started smoking as teenagers,[9] youth smoking behaviour is correlated with current smoking for prime age individuals, the population of interest in this study. Hence, parental smoking can be argued to be a potentially reliable instrument.

For parental smoking to be a valid instrument, parental smoking must not have an impact on the offspring's earnings capacity other than through the intergenerational transmission of smoking behaviour. But I have argued before that smoking may be be correlated with unobserved, job market relevant characteristics. After all, that is why the method of IV is applied in the first place. These unobserved characteristics

[8]Loureiro et al. (2006) find stronger same sex relationships than between different sexes which suggests using only the father's smoking habits as an instrument for male workers in this paper. However, in line with Lokshin and Beegle (2006), both parent's smoking behaviour will be used due to sample size restriction.

[9]In the data set used in this work, over 80% of current smokers in 1999 had smoked their first cigarette by the age of 18.

may be subject to intergenerational transmission as well. Suppose self-discipline is one such characteristic. If smoking acts as a proxy for a lack of self-discipline, then it is conceivable that a lack of self-discipline of parents is passed on to their children due to early upbringing and socialisation in the family environment. The validity of the instrument is a serious concern for the instrumental variable estimation technique. Therefore, similarly to Lokshin and Beegle (2006), dummy variables for parents' education are included in the wage equation as additional exogenous regressors. The idea is that parents' education should proxy for unobserved characteristics of parents that are subject to intergenerational transmission, so that controlling for parents' education, parents' smoking habits are assumed to be uncorrelated with the error term in the offspring's wage equation.

The method of IV is used in parallel to POLS as a *Pooled* 2SLS estimator. By eliminating the correlation of s_{it} with the error term ϵ_{it}, the contemporaneous exogeneity assumption holds again, and cross-sections are pooled across time to maximise data use. As before, standard errors have to be adjusted to account for non-independence of observations over time.

An advantage of the FE estimator over the method of IV as applied here is that the former eliminates all endogeneity from unobserved effects. If other variables in the wage equation are correlated with the error term, eliminating only correlation of the smoking variable with the error term does not generally suffice to consistently estimate the smoking coefficient γ. To see this clearly, consider the probability limit of the POLS estimator of the coefficient on smoking, for $N \to \infty$:[10]

$$
\begin{aligned}
\operatorname{plim} \hat{\gamma}_{POLS} = \gamma + & \left\{ \operatorname{plim} \frac{1}{N} \left(O_{ss} - O_{sx} O_{xx}^{-1} O_{xs} \right) \right\}^{-1} \\
& \times \left\{ \operatorname{plim} \frac{1}{N} O_{s\epsilon} - \left[\operatorname{plim} \frac{1}{N} \left(O_{sx} O_{xx}^{-1} O_{xs} \right) \cdot \operatorname{plim} \frac{1}{N} O_{x\epsilon} \right] \right\}
\end{aligned} \tag{4}
$$

where $O_{ab} = \sum_{i=1}^{N} \sum_{t=1}^{T_i} a'_{it} b_{it}$. Even if s_{it} were *not* correlated with the unobserved effects, e.g. by instrumenting for it by using 2SLS, equation (4) reveals that endogeneity of any other regressor in the wage equation (e.g. education) would result in $\hat{\gamma}$ being inconsistent. The second summand in (4) does not generally vanish if $\operatorname{plim} \frac{1}{N} O_{x\epsilon} \neq 0$ even if $\operatorname{plim} \frac{1}{N} O_{s\epsilon} = 0$ — unless, of course, the endogenous regessors are orthogonal to s_{it}, i.e. $\operatorname{plim} \frac{1}{N} O_{xs} = 0$.

It is often argued that, in a wage regression, education is likely to be correlated with the error term through unobserved effects. If education is correlated with smoking (e.g. via changes in perception of health risks from smoking or peer group effects) then the endogeneity of the education variables will impede consistent estimation of the coefficients on smoking even in an IV approach that eliminates the correlation of the smoking variable with unobserved individual effects. Since the FE estimator, on the other hand, is consistent in the presence of any endogeneity due to time invariant unobserved effects that are correlated with *any* regressor, estimation with FE can be seen as more reliable than IV as applied here.[11,12]

[10] See appendix for derivation.

[11] IV could generally also applied in a FE context, but here IV is referred to as applied in a cross-sectional context via e.g. Pooled 2SLS (the instruments used here are time invariant).

[12] A disadvantage of FE over IV is that the former aggravates the attenuation bias arising from measurement error in explanatory variables. I believe that the smoker status variable is not especially prone to measurement error, so that this disadvantage is not essential here.

3.2 Data

I use data from the British Household Panel Study (BHPS) drawing on all 15 currently available waves that cover the years between 1991 and 2005.[13] The BHPS is a nationally representative survey of British individuals and households. In the first wave in 1991 some 5 500 households and about 10 300 individuals were randomly selected and have since been re-interviewed annually. Additional samples were added in later years, but this study focuses on the original "Essex" sample of households that was selected in the first wave. Respondents are made up of all individuals of the sampled households of age 16 and older. Children become part of the panel once they turn 16. Original interviewees and their offspring are followed and reinterviewed when they move out of the original interview household.

The population of interest for this study consists of male workers and is restricted to those full time employed, between 25 and 55 years of age and with a minimum monthly wage or salary of £400 (in prices of 2005). The restrictions focus the analysis on the more homogenous group of prime age male workers that have for most parts completed their education. This is done to avoid complications from issues of life-cycle labour supply such as full time education, early retirement and —especially for women— child rearing.[14] In addition, the restrictions ensure comparability with other studies that generally restrict their sample to prime age individuals, with the majority based on samples of men.

[13]The description of the BHPS data is based on Taylor et al. (2006). An overview of all derived variables is provided in the appendix.

[14]Data on labour market experience is be less reliable for women, and age does not proxy well for experience due to non-continuous labour market spells. In the absence of a variable for actual labour market experience this is a problem. Preliminary results for women showed small or no wage effects from smoking. This is in line with previous work, but since results are less reliable and to economise on space results are not reported.

For the dependent variable, log of hourly wages, monthly wages and weekly working hours are combined to compute wages per hour. Earnings, indexed to prices of 2005,[15] are measured using the usual monthly wage or salary before tax based on the employee's last payment for the current main job. Usual hours per week, including paid overtime, are then used to compute the wage per hour.[16]

The data set contains detailed information about current cigarette smoking behaviour of a respondent. In all of its 15 waves, except in wave 9 in which questions were more detailed, respondents were asked if they smoked cigarettes at all and, if so, a follow up questions asked for the usual number of cigarettes consumed per day. In line with the literature, I classifiy an individual as a current regular smoker in wave t, if he/she answered affirmatively to the first smoking question and indicated that he/she was smoking at least one cigarette per day on average. Additionally, a categorical variable "intensity of smoking" is constructed from the daily number of cigarettes that sorts individuals into 5 groups: non-/non-regular smokers, and those who smoke 1-10, 11-20, 20-30 or 30 and more cigarettes per day respectively.

In wave 9, a series of more detailed questions about current and past smoking were asked, including how long ago a respondent stopped smoking if he/she was not currently smoking regularly but had done so in the past. The information about past smoking behaviour is used to further differentiate non-smokers into groups of never-smokers and former smokers where former smokers are split up into finer groups indicating roughly how long ago they stopped smoking. This information is used in section 4.4 where the estimation results are discussed.

[15]Based on Consumer Price Index provided by the Office for National Statistics (www.statistics.gov.uk).

[16]See appendix, table 7 for an exact definition and an overview of BHPS variables used.

Definitions of the remaining control variables that summarize age, education, marital status, ethnicity and region of residence are defined in table 7 (see appendix). Subject to the sample selection criteria, the variables for estimation of the POLS and FE models are available for 3,707 individuals with a total of 23,381 individual-year observations. Averaged over all waves the prevalence of regular smoking is 26.2%. While this is less than the maximum in the first wave in 1991, there is no significant downward time trend for the estimation subsample. Table 1 provides descriptive statistics for the main variables for an overview of the data separated by smoker status. On average, hourly wages for smokers are 22.1% lower than for non-smokers.[17] Smokers are also younger and less educated on average. In comparison with non-smokers, a smaller fraction of smokers is married, more are divorced, widowed or live separated and relatively more are singles. The differences in region and ethnicity are not statistically significant.

To create the instrument variables for the IV approach an intergenerational sample was constructed. For respondents that lived in one household with one or both of their parents at any one interview, it is possible to match children with their parents. The instrumental variables draw on parents' smoking behaviour: A set of dummy variables is created from a categorical variable that is the maximum value over both parents' smoker types over the time that the parents are observed in the sample.[18] The intergenerational subsample also allows the construction of a variable for parents' education which is coded, analogously to the education variable for all other respondents, as a categorical variable indicating the highest education achieved by

[17] $22.1 = exp(2.07 - 2.32) = 0.221$.

[18] Due to the sample size restrictions distinguishing between mother and father was not feasible. The current coding is was chosen because it provided the most reliable and most valid instruments in the specifications tests. Results, however, were not completely robust to alternative codings of the instruments (e.g. modulus instead of maximum over time); see Results section.

Table 1: Means of main variables by smoker status

Variable	Smoker		Non-smoker
Log hourly wage (log £)	2.07	***	2.32
Age (years)	37.47	***	39.03
Education (%)			
Degree	9.5	***	22.2
Further education	29.9	**	33.7
A-levels	13.2		12.6
O-levels	20.8	***	16.0
Other qualification	9.3	***	6.2
None	17.3	***	9.3
Marital Status (%)			
Married/cohabiting	76.9	***	82.3
Divorced/widowed/sep.	7.6	***	4.8
Single	15.4	**	12.9
Non-white (%)	3.5		3.0
Region (%)			
London	10.0		8.9
South East	20.8		19.2
Rest	69.3		71.9
Total observations	6 138		17 243

Notes: *** Means different at 1% level of significance based on standard errors clustered by individual, ** at 5% level. BHPS 1991-2005. See text for a description of the underlying sample.

the parents (if both parents are available the higher level of education is chosen). The additional sample selection criteria (matching with parents and availability of smoking parents' behaviour and education) substantially reduces the sample size so that the IV subsample consists of 635 individuals with in total 3 317 individual-year observations.

4 Estimation Results

4.1 Pooled Ordinary Least Squares

POLS estimates are presented in Table 2, column (1). All included variables are highly statistically significant. The adjusted R^2 indicates a reasonable fit of the model. The estimated coefficients for the education dummies show a gradient in the wage differential by education — the higher the educational achievement, the higher the wage premium. The markup is greatest for those who have a first or higher degree, who earn 114% more per hour on average compared to those with no education.[19] Those who are married or live with their partner have 15% higher hourly wages than singles, all else equal. The widowed, divorced or separated also earn about 7% more per hour compared to singles. Residents of London and of the South East region make ca. 30% and 18% more, respectively, than those who live neither in London nor in the South East. Controlling for other covariates in the model, older age is associated with higher wages for those under 45 years old with decreasing marginal returns to age.[20]

Current regular smoking is associated with a wage penalty. The coefficient on the dummy for current regular smoking indicates that, all else equal, smokers have

[19] $114\% = exp(0.761) - 1 = 1.140$ and analogously for the remaining dummies.
[20] $45 \approx \text{argmax}(0.076 \cdot age - 0.085 \cdot age^2/100 = 44.71)$

Table 2: Pooled Ordinary Least Squares and Fixed Effects

	(1) POLS		(2) FE	
Dependent variable:	Log hourly wage		Log hourly wage	
Current regular smoker	–0.136	(9.35)***	–0.020	(1.72)*
Degree	0.761	(28.56)***	0.097	(1.53)
Further education	0.417	(17.75)***	–0.021	(0.62)
A-level	0.376	(13.40)***	0.008	(0.20)
O-level	0.234	(9.73)***	0.024	(0.65)
Low education	0.155	(5.44)***	0.005	(0.11)
Age	0.076	(12.74)***	0.067	(5.82)***
Age squared/100	–0.085	(11.13)***	–0.098	(14.67)***
Married/couple	0.143	(7.88)***	0.027	(1.62)
Widowed/divorced/separated	0.069	(2.39)**	–0.024	(1.00)
Non-white	–0.123	(2.85)***	—	—
London	0.259	(9.78)***	0.040	(0.82)
South East	0.167	(8.95)***	0.091	(2.38)**
N individuals	3 707		3 707	
Total observations	23 381		23 381	
Adjusted R^2	0.340		0.256	

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Absolute t statistics in parentheses. p-values of F-test for joint significance in FE specification: 0.076 for educational variables, 0.007 for marital status variables, 0.056 for region of residence. t and F statistics are based on variances robust to heteroskedaticity and arbitrary within-individual correlation over time (see section 3.1). BHPS 1991-2005. The dependent variable is indexed to prices of 2005. Estimation also included a constant and set of year dummies. Reference categories are: "Current non-/non-regular smoker" for "Current regular smoker" ; "No education" for the educational dummies; "never married/single" for marital status; "White" for "non-white" and "Neither London nor South East" for "London" and "South East".

a 12.7% lower wage per hour compared to non-smokers. The estimates imply that a white, married, 38 years old, male worker in 2005 with further education, residing neither in London nor in the South East, earned on average £10.5 per hour if currently regularly smoking, which is £1.5 less than a current non-smoker with otherwise identical characteristics. The raw sample difference in group means for current smokers vs. current non-smokers, as described in the Data section, is 22.1%. Hence, holding education, age, marital status, race and region of residence constant, the difference in mean hourly wage reduces substantially.

4.2 Fixed Effects

Despite the inclusion of controls, there remains a considerable wage gap between smokers and non-smokers according to the POLS estimates. To isolate the variation that is causally due to smoking, unobserved heterogeneity needs to be taken into account. Table 2, column (2), shows the results from the fixed effects model.

Controlling for unobserved individual effects, the education dummies become individually statistically insignificant at all conventional levels, though still jointly significant at the 10% level. But identification in the fixed effects model comes from within-group variation. By the age of 25, which is the minimum age for inclusion in the estimation sample, the vast majority of individuals had completed their formal education. Understandably, the individual coefficients are not precisely estimated and are not statistically significantly different from zero. FE estimation based on within-group variation also lowers the fit of the model as indicated by the lower adjusted R^2 in comparison with POLS, which relies mainly on between-group variation.

The coefficient on the variables for marital status are jointly statistically significant at the 1% level, but not individually. The estimates imply a small positive effect of marriage and a small negative effect from being widowed, divorced or separated compared with singles. Due to a lack of variation over time, estimation of the difference between whites and non-white cannot be estimated with FE. Living in the regions of London or South East has positive effects, though lower in magnitude than with POLS. While not individually significant, the regional dummies are jointly significant at the 10% level. The FE estimates of age and age squared are highly statistically significant and qualitatively similar to those obtained with POLS, the main difference being that the gradient in marginal returns is flatter. The implied maximum return to age is at around 34 years.

Now the negative effect of smoking is statistically significant only at the 10% level and the magnitude drops by a lot compared to POLS. Holding unobserved effects and other observed characteristics constant, a worker that is currently smoking has on average a 2% lower hourly wage than another otherwise identical who is not smoking.[21,22] The drop in magnitude of the estimate from 12.7% obtained from POLS to just under 2% with FE seems to suggest that unobserved heterogeneity plays a large role and not taking it account leads to sizeable bias in the estimated effects of smoking on wages.

[21]The econometric model implicitly assumes that the effect of smoking is the same in every year. An F-test in an estimation with smoker $\times$ year interactions did not reject the Null of equal coefficients (p-value 0.622).

[22]The theoretical arguments about the causal effects of smoking on wages make it plausible that smoking more cigarettes may increase the wage penalty. Therefore, I extended the FE specification by substituting the simple dummy indicating the current smoker status for a set of dummy variables that indicate, in five groups, the number of cigarette smoked per day. The estimates, reported in table 6 in the appendix, suggest that negative effect from smoking increases with the intensity of smoking.

4.3 Instrumental Variables

Table 3 shows the IV estimation results. A set of dummy variables for parental education is introduced as additional control variables. To be able to compare the Pooled 2SLS results with the baseline model, POLS estimation is carried out again on the smaller IV subsample. The estimates, column (1), are comparable with the POLS estimates on the sample underlying Table 2. In particular, the statistically significant wage penalty for smokers over non-smokers is of similar magnitude. The newly-introduced dummmies suggest that parents' education has a positive effect on the offspring's mean wages compared to no education of parents. However, the coefficients are imprecisely estimated (p-value of F-test of joint significance: 0.149).

In column (2) and (3), the estimation results from Pooled 2SLS are shown. In the first stage the only coefficients of interest are those on the instruments. Parents' smoking is positively correlated with the respondent's smoking habits for three of four categories. The four instruments are jointly significant at the 5% level. According to Hansen's J statistic, the null hypothesis of validity of the instruments is not rejected.[23] However, the instruments appear to be weak as the F statistic testing the joint significance of the instruments in the first stage does not exceed 10, which is the threshold level proposed by Staiger and Stock (1997). Weak instruments generally lead to a finite sample bias in the direction of the OLS bias (Bound et al., 1995).

[23]Due to the use of four instrument for one endogenous variable, the model is overidentified: there are more variables than coefficients to estimate. This allows one to test the validity of the instruments, i.e. whether the correlation of the instruments with the disturbances is zero. Hansen's J test statistic is the robust (to non-spherical disturbances) equivalent of the usual Sargan (1958) test of overidentifying restrictions (Baum et al., 2003, p. 18).

In the second stage, the control variables that are statistically significant are comparable to those from POLS in column (1). The Pooled 2SLS estimate of the coefficient on smoking is positive, but completetly insignificant statistically: the associated 95% confidence interval ranges from –0.345 to +0.477. The IV approach is not overly robust: Relatively small changes in the definition of the instruments or controls causes the instruments' validity and reliability as well as the estimate on smoking to vary considerably. The estimates would suggest that there is no statistically significant effect from smoking when controlling for endogeneity of the smoking decision. But due to the unreliability of the IV estimates, the particular results are given little weight in the subsequent discussions.

4.4 Discussion of Results

4.4.1 Different Counterfactuals

In order to correctly interpret the estimates of the coefficient on the smoker status dummy, it is important to consider carefully the implied counterfactuals that arise from different estimation techniques. In the above specifications, a single dummy is used to estimate the influence that smoking has on wages, which takes the value of 1 if a worker is currently a regular smoker and 0 otherwise. But the group of current non-smokers is generally composed of former smokers and never-smokers. Former smokers may still be affected by persistent effects from smoking. They may also differ in unobserved characteristics such as motivation or drive from those who never smoked and from current smokers. Hence, the difference in wages of smokers and non-smokers using a single dummy might vary depending on the comparison group, and the interpretation of a dummy for current smoking depends on who current smokers are actually compared to.

Table 3: Pooled Ordinary Least Squares and Instrumental Variables

	(1) POLS		(2) IV 1st stage		(3) IV 2nd stage	
Dependent variable:	Log hourly wage		Current smoker		Log hourly wage	
Current regular smoker	–0.124	(4.68)***			0.066	(0.31)
Degree	0.653	(7.47)***	–0.263	(2.85)***	0.707	(5.66)***
Further education	0.403	(6.40)***	–0.200	(2.29)**	0.443	(4.94)***
A-level	0.347	(5.22)***	–0.162	(1.55)	0.381	(4.47)***
O-level	0.243	(4.21)***	–0.102	(1.10)	0.264	(3.76)***
Low education	0.189	(2.99)***	–0.099	(0.93)	0.208	(2.95)***
Age	0.047	(2.56)**	–0.031	(1.66)*	0.052	(2.74)***
Age squared/100	–0.044	(1.65)*	0.031	(1.15)	–0.049	(1.84)*
Married/couple	0.134	(4.51)***	–0.004	(0.12)	0.134	(4.45)***
Widowed/divorced/sep.	0.121	(1.77)*	0.077	(1.06)	0.107	(1.52)
Nonwhite	0.030	(0.31)	0.116	(1.31)	0.010	(0.10)
London	0.332	(6.21)***	0.097	(1.67)*	0.314	(5.05)***
South East	0.135	(3.39)***	0.091	(1.86)*	0.118	(2.54)**
Parents: Degree	0.118	(0.86)	–0.128	(1.45)	0.143	(0.97)
Parents: Further educ.	0.072	(1.44)	0.011	(0.16)	0.069	(1.30)
Parents: A-level	0.017	(0.33)	0.058	(0.70)	0.002	(0.03)
Parents: O-level	0.102	(2.19)**	–0.070	(1.42)	0.113	(2.31)**
Parents: Low education	0.070	(1.32)	–0.040	(0.69)	0.080	(1.51)
Parents: 1–10 cig's/day			0.077	(1.08)		
Parents: 11–20 cig's/day			–0.049	(0.94)		
Parents: 21–30 cig's/day			0.121	(1.56)		
Parents: 30+ cig's/day			0.270	(2.35)**		
N individuals	635		635		635	
Total observations	3317		3317		3317	
Adjusted R^2	0.386		0.066		0.355	
1st stage $F[3][634]$			2.77(0.027)			
Hansen's J, $\chi^2[3]$			4.56(0.207)			

Notes: See table (2). Reference category for parents' education is "no education". Degrees of freedom for specification test statistics in square brackets, p-values in round parenthesis.

In an OLS regression using only wave 9 (year 1999) of the BHPS, the coefficient on current smoking does virtually not change when changing the reference group from never- and ex-smokers to only never-smokers by controlling for past smoking (table 4). Furhter, the estimates on the dummies for quitting smoking more than 2 years ago are non-negative and insignificant. Thus, with POLS, which mostly relies on variation across different individuals, the coefficient in an estimation with only one dummy for current smoking might be interpreted as the wage differential between current smokers and those who never smoked or quit a long time ago.

With FE, however, the case is different, since identification of the model's parameters works through within-group variation. Hence, identification of a dummy indicating the current smoker status comes from those individuals who change said status at least once during their observation in the sample. Very few individuals take up the habit of regular smoking for the first time after the age of 25, the minimum age for inclusion in the sample. Drawing again on the more detailed information of wave 9, more than 95% of those who were current regular smokers in 1999 had smoked cigarrettes before the age of 25. Further, only 9 of 829 or about 1.1% of those who have never smoked regularly before as of wave 9 have started smoking during observation in the sample after wave 9. On the other hand, 68% of current non-smokers have smoked regularly before. 10% of current regular smokers in wave 9 in 1999 were not classified as such in 2000 and about 32% reported to be non-smokers at least once in the years following 1999. Hence, estimation via FE primarily relies on those who quit smoking as opposed to those who start and does thus not identify the difference in wages between smokers and never-smokers but between smokers and former smokers.

Table 4: Ordinary Least Squares: Current and past past smoking (wave 9)

	(1)		(2)		(3)	
Dependent variable:	Log hourly wage		Log hourly wage		Log hourly wage	
Current regular smoker	–0.139	(5.96)***	–0.146	(5.84)***	–0.144	(5.82)***
Former smoker			–0.020	(0.76)		
Quit $\leq$ 1 year					–0.172	(2.80)***
Quit $\leq$ 2 years					–0.053	(0.61)
Quit $\leq$ 5 years					0.071	(1.04)
Quit > 5 years					0.007	(0.22)
N	1669		1669		1669	
Adjusted R^2	0.329		0.329		0.331	

Notes: See table (2). BHPS 1999. Reference category for the smoking dummies is "never (regularly) smoked". Not reported are control variables which are those of table (2).

4.4.2 Persistent Effects

With FE, a dummy for current smoking is estimated as the average wage effect of being a smoker vs. being a former smoker, as measured by current smoking for those individuals who change smoking status over time. If the causal wage effects from smoking are at least somewhat persistent over time, e.g. due to persistent health effects, then the coefficient on the current smoker dummy understates the cumulative effect from continuous smoking under the FE estimator (and would overstate the effect if interpreted as the effect from current smoking). Consider the case that a regular smoker temporarily quits in one year and takes up smoking again in the following year. If the effects are persistent, the full benefit from quitting, or accordingly the full penalty from continuous smoking, is not reflected in the wage differences between the year of non-smoking and the years of smoking. To capture more fully the effects from a long-standing smoking habit vs. not smoking at all, I extended the FE specification to include lags of the smoker dummy (table 5).

The inclusion of lags substantially reduces the sample size, since not for all individuals two lags or more of all variables are observed. In column (1) the original

Table 5: Fixed Effects: Current smoker and lags

	(1)		(2)		(3)	
Dependent variable:	Log hourly wage		Log hourly wage		Log hourly wage	
Current regular smoker	–0.020	(1.30)	–0.015	(1.08)	–0.015	(1.05)
Lag (1 year) of smoker			–0.016	(1.07)	–0.014	(0.99)
Lag (2 years) of smoker					–0.010	(0.77)
N individuals	2 264		2 264		2 264	
Total observations	13 018		13 018		13 018	
Adjusted R^2	0.210		0.210		0.210	

Notes: See table (2). Not reported are the estimates of control variables which are those of table (2).

specification without lags is presented for comparison. The coefficient is practically identical to previous estimates with the full sample, but the standard errors are much higher (p-value of t-test 0.195). Correspondingly, the following interpretation of the specifications including lags are merely suggestive.

Including a one year lag (column 2), the effect on the hourly wage from contemporaneous smoking is estimated to be –1.5% on average and the effect from smoking one year ago is estimated to be –1.6% compared to not smoking a year ago. For a current smoker who also smoked in the previous year the effects add up, so that he would have a 3.1% lower wage than when smoking neither now nor last year. The estimates of wage effects from contemporaneous and previous year smoking decline slightly when a one- and a two-year lag are included. The estimate on the coefficient for smoking two years ago is negative but of smaller magnitude than the effects of current and previous year smoking. Adding up the three estimates means a wage reduction of –3.7%, all else equal, for someone who smokes currently and has done so during the last two years in comparison with someone who did not smoke at all. In conclusion, the results, while only tentative, are supportive of the idea that smoking effects are persistent. In combination with the results from the analysis with past smoking data from 1999, however, the estimates from the lag specification do also suggest that the persistence is limited. The coefficient on the two-year lag in the extended FE specification is smaller than for the one-year lag. And, while not controlling for unobserved effects, OLS on the data from wave 9 seems to suggest that quitting smoking more than two years ago has little or no effect on current wages.

4.4.3 Time-varying Unobserved Effects

A shortcoming of FE estimation is that a specification with unobserved individual effects explicitly assumes that these effects are time invariant — c_i has the same impact for individual i in every year. If the job market relevant unobserved characteristics that are correlated with the regressors were also to vary over time, then the FE estimator would not be consistent. The time varying parts would then be part of the idiosyncratic random disturbance (ν_{it}) and the exogeneity assumption would be violated. If e.g. a boost in motivation and self-discipline has positive effects on wages, but is not caused by the decision to quit smoking then the FE estimate of the effect of quitting would be upward biased. The method of IV does not rely on unobserved effect to be time-invariant to consistenly estimate γ. Unfortunately, the method's application is limited here by the small sample size, so that the resulting unstable estimates are not comparable with the FE estimates to assess the degree of the potential bias.

5 Summary and Conclusions

With reference to the preferred FE estimation, the analysis of the present paper suggests that there are small negative effects from smoking on wages. However, there appears to be no "puzzle". Despite the large sample size, the coefficients on smoking are only weakly significant and at the same time economically plausible. The observed large correlation between smoking and wages are mainly due to non-random selection into smoking. Controlling for unobserved effects, the wage penalty due to current smoking is estimated to be –2% over someone who has quit smoking.

Since identification with FE is based primarily on within-group variation of smokers quitting or former smokers starting again, the large difference between POLS and FE estimates are not only due to unobserved effects but partially because the respective techniques imply different counterfactuals. Additionally, further analysis suggestes that the effect of long-term smoking compared to not smoking at all is somewhat underestimated by the FE approach as applied here.

The analysis in this paper is conditional on the respondent being in the workforce and being in the data sample. The estimated effect does not take into account the more severe cases of deteriorated health due to smoking that leave workers incapacitated for work. Further, smokers may be more likely to drop out of the sample than non-smokers as they differ in personal unobserved characteristics. Both issues would result in a tendency of the analysis to underestimate the negative effect of smoking on wages. Correcting for these selection issues, however, requires more elaborate models than the specifications applied here (e.g. Heckman type sample selection models).

In this work, I have investigated *if* smoking causally affects wages. In order to test hypotheses about the "*why*", i.e. the channels of causation, more detailed data is necessary. Data on workers' rates absenteeism or time spent in smoking breaks, for example, was not available. The particular causal mechanisms of the negative effect of smoking on wages are an interesting topic for future research.

A Appendix

A.1 Derivation of equation (4), page 17

For the ease of exposition I assume that the data is a balanced panel. The results are qualitatively identical to the unbalanced case but notationally less tedious to derive.

Let X be a matrix with the K_1-dimensional row vectors of control variables x_{it} stacked by individuals and time, and define analogously S, w and ϵ for the smoking variable(s), the dependent variable and the disturbances respectively:

$$\underbrace{X}_{(NT\times K_1)} = \begin{bmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1T} \\ \vdots \\ x_{N1} \\ \vdots \\ x_{NT} \end{bmatrix}, \quad \underbrace{S}_{(NT\times K_2)} = \begin{bmatrix} s_{11} \\ s_{12} \\ \vdots \\ s_{1T} \\ \vdots \\ s_{N1} \\ \vdots \\ s_{NT} \end{bmatrix}, \quad \underbrace{w}_{(NT\times 1)} = \begin{bmatrix} w_{11} \\ w_{12} \\ \vdots \\ w_{1T} \\ \vdots \\ w_{N1} \\ \vdots \\ w_{NT} \end{bmatrix}, \quad \underbrace{\epsilon}_{(NT\times 1)} = \begin{bmatrix} \epsilon_{11} \\ \epsilon_{12} \\ \vdots \\ \epsilon_{1T} \\ \vdots \\ \epsilon_{N1} \\ \vdots \\ \epsilon_{NT} \end{bmatrix}.$$

Define $\tilde{X} = (X \ S)$ and $\delta = (\beta' \ \gamma')'$. Then equation (1) on page 9 can be expressed as

$$w = X\beta + S\gamma + \epsilon = \tilde{X}\delta + \epsilon.$$

The POLS estimator is $\hat{\delta} = (\tilde{X}'\tilde{X})^{-1}\tilde{X}'w$, which after substitution for w becomes

$$\begin{bmatrix} \hat{\beta} \\ \hat{\gamma} \end{bmatrix} = \delta + (\tilde{X}'\tilde{X})^{-1}\tilde{X}'\epsilon = \begin{bmatrix} \beta \\ \gamma \end{bmatrix} + \begin{bmatrix} X'X & X'S \\ S'X & S'S \end{bmatrix}^{-1} \begin{bmatrix} X'\epsilon \\ S'\epsilon \end{bmatrix}.$$

By use of the partitioned inverse (Greene, 2003, A.5, p. 824), the estimator for γ becomes

$$\hat{\gamma} = \gamma + \left\{ S'S - (S'X)(X'X)^{-1}(X'S) \right\}^{-1} \times \left\{ S'\epsilon - \left[(S'X)(X'X)^{-1}(X'\epsilon) \right] \right\},$$

which is equation (4) when written in double (over individuals and time) sum form and taking the probability limit.

A.2 Tables

Table 6: Fixed Effects: current smoker and number of cig.

	(1)		(2)	
Dependent variable:	Log hourly wage		Log hourly wage	
Current regular smoker	–0.020	(1.72)*		
Number of cig./day				
1–10			–0.014	(1.08)
11–20			–0.024	(1.72)*
21–30			–0.035	(1.75)*
30+			–0.036	(0.98)
N individuals	3707		3707	
Total observations	23381		23381	
Adjusted R^2	0.256		0.256	

Notes: See table (2). Not reported are the estimates of control variables which are those of table (2). Reference category for smoker type dummies is Non-/Non-regular smoking.

Table 7: Definition of Derived Variables

Variable	Definition	Based on BHPS Variable(s)
Log hourly wage	Log of gross hourly wage deflated to prices of 2005, assuming overtime rate is 1.5. Hourly wage = (usual gross wage or salary per month * 12/52) / [(normal standard weekly hours)+ 1.5*(normal weekly paid overtime hours)].	PAYGU, JBHRS, JBOT and JBPOT
Current regular smoker	Dummy variable =1 if respondent smokes at least one cigarette per day in t, 0 otherwise.	SMOKER, NCIGS, SMEVER (wave 9), SMNOW (wave 9)
Intensity of smoking	Dummy variables =1 if respective category applies and =0 otherwise. Categories are None or <1 cig./day and 1-10, 11-20, 21-30 and 30 or more cigarettes daily.	SMOKER, NCIGS, SMEVER (wave 9), SMNOW (wave 9)
Time since quitting	Dummy variables =1 if respective time frame applies and =0 otherwise (only wave 9): $\leq$ 1 year, $\leq$ 2 years, $\leq$ 5 years, > 5 years.	SMSTOP (wave 9)
Education	Highest educational qualification: Dummy variables =1 if respondent has qualification, =0 if not.	QFEDHI
Degree	First or higher degree	
Further education	Teaching, nursing or other higher qualification	
A-levels	A-levels or equivalent	
O-levels	O-levels or equivalent	
Other qualification	Commercial qualification, apprenticeship, other qualification	
None	None of the above	
Age	Age in years at date of interview.	AGE
Marital Status	Dummy variables =1 if respective status applies and =0 otherwise. Categories are "married/cohabiting", "divorced/widowed/separated" and "single/never married".	MASTAT
Non-white	Dummy variable =1 if non-white, =0 if white.	RACE
Region	Dummy variables =1 if respondent lives in region, and =0 otherwise. Regions are London (inner and outer London), South East and neither London nor South East.	REGION

Notes: BHPS 1991-2005. Creating the instruments on the basis of "intensity of smoking" of parents also required variables that link children with their parents. Coding of wage and education variables follows Bryan and Sevilla Sanz (2007).

References

Anger, Silke and Michael Kvasnicka (2006). "Biases in Estimates of the Smoking Wage Penalty". Discussion Papers of DIW Berlin 654, DIW Berlin, German Institute for Economic Research.

Arellano, Manuel (1987). "Computing Robust Standard Errors for Within-Group Estimators". *Oxford Bulletin of Economics and Statistics*, 49:431–434.

Auld, Christopher M. (2005). "Smoking, Drinking, and Income". *Journal of Human Resources*, 40(5):505–518.

Baum, Charles L., William F. Ford, and Jeffrey D. Hopper (2006). "The Obese Smoker's Wage Penalty". *Social Science Quarterly*, 87:863–881.

Baum, Christopher F., Mark E. Schaffer, and Steven Stillman (2003). "Instrumental Variables and GMM: Estimation". *The Stata Journal*, 3(1):1–31.

Bertera, Robert L. (1991). "The Effects of Behavioral Risks on Absenteeism and Health-Care Costs in the Workplace". *Journal of Occupational Medicine*, 33(11):1119–1124.

Bound, J., D. A. Jaeger, and R. M. Baker (1995). "Problems with Instrumental Variables Estimation when the Correlation between the Instruments and the Endogenous Explanatory Variable is Weak". *Journal of the American Statistical Association*, 90(430):443–450.

Bryan, Mark L. and Almudena Sevilla Sanz (2007). "Does Housework Lower Wages and Why? Evidence for Britain". British household panel survey research conference, ISER.

Chaloupka, Frank J. and Kenneth E. Warner (2000). "The Economics of Smoking". In Culyer, A. J. and J. P. Newhouse (eds.) "Handbook of Health Economics", vol. 1 of *Handbook of Health Economics*, chap. 29, 1539–1627. Elsevier.

Evans, William N. and Edward Montgomery (1994). "Education and Health: Where There's Smoke There's an Instrument". NBER Working Papers 4949, National Bureau of Economic Research.

Fuchs, Victor R. (1982). "Time Preference and Health: An Exploratory Study". In Fuchs, Victor R. (ed.) "Economic Aspects of Health", 93–120. University of Chicago Press.

Gilbert, David G. (1995). *Smoking: Individual Differences, Psychopathology, and Emotion*. The Series in Health Psychology and Behavioral Medicine. Taylor & Francis, Washington, DC.

Greene, William H. (2003). *Econometric Analysis*. Prentice Hall, Englewood Cliffs, 5th ed.

Grossman, Michael (1972). "On the Concept of Health Capital and the Demand for Health". *The Journal of Political Economy*, 80(2):223–255.

Halpern, Michael T., Richard Shikiar, Anne M. Rentz, and Zeba M. Khan (2001). "Impact of Smoking Status on Workplace Absenteeism and Productivity". *Tobacco Control*, 10(3):233–238.

Heineck, Guido and Johannes Schwarze (2003). "Substance Use and Earnings: The Case of Smokers in Germany". IZA Discussion Papers 743, Institute for the Study of Labor (IZA).

Kézdi, G. (2007). "Robust Standard Error Estimation in Fixed-Effects Panel Models". *Hungarian Statistical Review Special*, 9:96–116.

Khwaja, Ahmed, Dan Silverman, and Frank Sloan (2007). "Time preference, Time Discounting, and Smoking Decisions". *Journal of Health Economics*, 26(5):927–949.

Kristein, Marvin M. (1983). "How Much can Business Expect to Profit from Smoking Cessation?" *Preventive medicine*, 12:358–381.

Lee, Yew Liang (1999). "Wage Effects of Drinking and Smoking: An Analysis Using Australian Twins Data". Discussion Paper 22, University of Western Australia.

——— (2003). "Wage Effects of Drinking in Australia". *The Australian Economic Review*, 36(3):265–282.

Leigh, J. Paul (1995). "Smoking, Self-Selection and Absenteeism". *The Quarterly Review of Economics and Finance*, 35(4):365–386.

Leigh, Paul J. and Mark Berger (1989). "Effects of Smoking and Being Overweight on Current Earnings". *American Journal of Preventative Medicine*, 5:8–14.

Levine, Phillip B., Tara A. Gustafson, and Ann D. Velenchik (1997). "More Bad News for Smokers? The Effects of Cigarette Smoking on Wages". *Industrial and Labor Relations Review*, 50(3):493–509.

Lokshin, Michael and Kathleen Beegle (2006). “Forgone Earnings from Smoking: Evidence for a Developing Country”. Policy Research Working Paper Series, The World Bank, Washington, D.C.

Loureiro, Maria L., Anna Sanz de Galdeano, and Daniela Vuri (2006). “Smoking Habits: Like Father, Like Son, Like Mother, Like daughter”. IZA Discussion Papers 2279, Institute for the Study of Labor (IZA).

Mincer, J. (1974). *Schooling, Experience and Earnings.* National Bureau of Economic Research, New York.

Munasinghe, Lalith and Nachum Sicherman (2005). “Wage Dynamics and Unobserved Heterogeneity: Time Preference of Learning Ability?” NBER Working Papers 11031, National Bureau of Economic Research.

Sargan, John D. (1958). “The Estimation of Economic Relationships Using Instrumental Variables”. *Econometrica*, 26(3):393–415.

Staiger, D. and J. H. Stock (1997). “Instrumental Variables Regression with Weak Instruments”. *Econometrica*, 65(3):557–586.

Taylor, Marcia Freed, John Brice, Nick Buck, and Elaine Prentice-Lane (2006). *British Household Panel Survey User Manual Volume B: The Codebook.* The Univserity of Essex, Colchester. Url: www.iser.essex.ac.uk/ulsc/bhps/doc.

Van Ours, J. (2004). “A Pint a Day Raises a Man’s Pay; But Smoking Blows That Gain Away”. *Journal of Health Economics*, 23(5):863–886.

Willis, Robert J. (1986). “Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions”. In Ashenfelter, O. and R. Layard (eds.) “Handbook of Labor Economics”, vol. 1 of *Handbook of Labor Economics*, chap. 10, 525–602. Elsevier.

Wooldridge, Jeffrey M. (2002). *Econometric Analysis of Cross Section and Panel Data.* The MIT Press, Cambridge, Massachusetts.