Social Capital and Cigarette Smoking: New Empirics Featuring the Norwegian HUNT Data

M. Kamrul Islam[a, b*]

Sherman Folland[c]

Oddvar M. Kaarbøe[d]

[a] *Department of Economics, Fosswinckelsgate 14, University of Bergen, 5007 Bergen, Norway*

[b] *Uni Research Rokkan Centre, Bergen, Norway*

[c] *Department of Economics, Oakland University, Rochester, MI 48309, USA*

[d] *Department of Health Management and Health Economics, University of Oslo, 0373 Oslo, Norway*

[*]**Corresponding author** at: Department of Economics, Fosswinckelsgate 14, University of Bergen, 5007 Bergen, Norway. E-mail address: kamrul.islam@uib.no (M. K. Islam).

Acknowledgements

We thank the editor and two anonymous reviewers for their constructive comments, which helped us to improve the manuscript. The earlier version of the manuscript has benefited from comments after presentations at iHEA congress (Toronto), Lund University and University of Bergen.

Abstract

Using a rich Norwegian longitudinal data set, this study explores the effects of different social capital variables on the probability of cigarette smoking. There are four social capital variables available in two waves of our data set. Our results based on probit (and OLS) analyses (with municipality fixed-effects) show that the likelihood of smoking participation is negatively and significantly associated with social capital attributes, namely, community trust (–0.017), participation in organizational activities (–0.032), and cohabitation (–0.045). Significant negative associations were also observed in panel data, pooled OLS, and random effects models for community trust (–0.024; –0.010) and cohabitation (–0.040; –0.032). Fixed-effects models also showed significant negative effects for cohabitation (–0.018). Estimates of alternative instrumental variables (IV) based on recursive bivariate probit and IV-GMM models also confirmed negative and significant effects for three of its characteristics: cohabitation (–0.030; –0.046), community trust (–0.065; –0.075), and participation in organizational activities (–0.035; –0.046). The limitations of our conclusions are discussed, and the significance of our study for the field of social capital and health is described, along with suggested avenues for future research.

Keywords: cigarette smoking, social capital, longitudinal data, instrumental variables

JEL Classifications: I12; C26; Z13

1. Introduction

The idea that one's social milieu and social experiences affect their physical circumstances and well-being has an ancient history. In our era, the concept of social capital was brought forward by Glenn Loury (1977), Pierre Bordieu (1985), James Coleman (1988), and most prominently in the recent work of Robert Putnam (2000, 1995, and 1993). Putnam et al. (1993) defined social capital as referring to the "features of social organization, such as trust, norms and networks that can improve the efficiency of society by facilitating coordinated actions" (p.167). In general, four main theoretical ingredients can be identified in the definition of social capital: social trust/reciprocity, collective efficacy, participation in voluntary organizations and social integration for mutual benefit (Lochner et al., 1999).

Social capital is a multidimensional phenomenon. Nevertheless, the notion can be broadly divided into *cognitive* and *structural components*, where the former component includes norms, values, attitudes and beliefs, and the later includes established roles and social networks supplemented by rules, procedures and precedents. The cognitive component of social capital measures people's perceptions of the level of interpersonal trust, sharing, and reciprocity, and the structural component studies the extent and intensity of associational links and activity in society such as measures of informal sociability, density of civic associations, and indicators of civic engagement (Islam et al., 2006a). Social capital can also be separated as *horizontal* and *vertical* social capital. Within horizontal social capital, *bonding social capital* and *bridging social capital* can also be categorized. *Bonding social capital* refers to the relations within homogenous groups (the strong ties that connect family members, neighbors, and close friends, etc.), while *bridging social capital* is heterogeneous by definition (for detailed distinctions with examples, see Islam et al, 2006a, Figure 1). Bonding and cognitive social capital are hypothesized to be vital for establishing and favoring healthy norms and controlling abnormal social behavior. Structural social capital (e.g., social participation) is supposed to complement attributes of cognitive social capital.

Social capital research has described the effects of social ties on the family, friend networks, and relationships with the community as being important contributors to socio-economic outcomes that cannot be explained by a conventional, rational economic model. For example, educational outcomes, crime rates, TV watching by children, and health measures have all proven to be beneficially associated with social capital. The health connection alone has attracted many researchers and resulted in a vast literature; for example, see Kawachi (1999), Kawachi, and Berkman (2001), Poortinga (2006), and Turrell, Kavanagh, and Subramanian (2006). Economists have been prominent in conceptual studies of social capital; for example, see Glaeser et al. (2002), Becker and Murphy (2000), Durlauf, (2002), and Akerlof (1998). But only recently have health economists focused on the social capital and health hypothesis, which has given more attention to econometric issues. Among these more recent studies are Mellor and Milyo (2005), Islam et al. (2006b, 2008), Folland (2006, 2007, 2008), Brown et al. (2006), Laporte and Ferguson (2004), d'Hombres, Rocco, Suhrcke, and McKee (2010), Ljunge (2014), and Fiorillo and Sabatini (2015).

The social capital and health hypothesis simply stated is the proposition that improvements in a person's individual or community social capital (CSC) will cause, *ceteris paribus*, improvements in the person's health. The hypothesis poses econometric difficulties because it requires success in sorting out influences of other variables and addressing questions of endogeneity. However, the hypothesis remains plausible because there are known pathways by which social ties could have the hypothesized effect: 1) friendship and sociability often reduce stress, and stress is known to adversely affect health (Sapolsky, 1998); 2) social contacts provide new information sources about healthy behaviors and medical procedures; 3) following Coleman (1988), social ties enhance one's sense of obligation to loved ones, friends, and by implication, oneself (also Folland, 2006); and 4) by joining in groups, people may be able to better influence the development of better community health services (Mellor and Milyo, 2005; Kawachi, Kennedy, and Lochner, 1997).

Although there have been several studies on the relationship between social capital and smoking, only a few have been developed in an economic framework. First, we consider the literature from outside of economics, namely studies from the medical, epidemiological, and psychological literature.

Sapag and colleagues (2010) found a significant inverse relation between a measure of "trust in neighbors" and cigarette smoking in an urban setting in Santiago, Chile. Afifi, Nakkash, and Khawaja (2010) also studied an urban setting, in this case in Beirut, Lebanon, and in low-income neighborhoods, where they found negative correlations with smoking and "trust of friends and neighbors." Giordano and Lindstrom (2010) found an association between smoking cessation and trust and participation in groups. Lindstrom, Isacsson, and Elmstahl (2003) found that social participation is positively associated with smoking cessation. Lee and Kahende (2007) studied factors that encouraged quitting smoking in the United States and found that, *inter alia*, being married or living with a partner were important (see also Homish and Leonard, 2005). Studies of adolescents (Glaser, Sheton, and van der Brie, 2010; Stewart-Knox et al., 2005; Page et al., 2006) have found that peer smoking behaviors are especially influential, and the social effect may encourage smoking if the peer cohort smokes. These findings may support the social capital and health hypothesis, but they do not address econometric issues such as the role of fixed- or random effects or the endogeneity issue.

Health economists have also recently studied social capital and smoking. Folland (2006) theorized how social capital would alter one's desired rate of trade between risk and reward, supported by data on several health risky behaviors in a cross-section of the United States. Brown et al. (2006) developed a "Petris Social Capital Index" and showed that a portion of this was associated with membership in a religious organization, which seemed to encourage reduced smoking. Folland (2008) developed an economic model based on utility maximization and showed that when smoking damages health and lowers the probability of survival, exogenous increases in social capital can promote reduced smoking. That model was tested using data from the National Longitudinal Survey of Youth 79, where a CSC measure proved to be negatively associated with smoking. Recently, Folland et al. (2014) developed a theoretical model of how one invests in social capital. Social capital is described as a form of leisure choice, which implies that it entails an opportunity cost. In the most general model, the relationship between cigarette smoking and social capital is ambiguous. Specifically, it is shown that since social capital improves the probability of survival, it might also increase smoking because it extends the number of hours one can enjoy their cigarettes.

By recognizing social capital as a complex construct, previous social capital health research has warned of the likelihood of *construct bias* if researchers use a single measure to estimate the effect of social capital on health (Sundquist and Yang, 2007). To address the aforesaid concerns, this paper examines the effects of the following three different forms/dimensions of social capital on smoking behavior: *bonding social capital* (family and friends), *cognitive social capital* (community trust), and *structural social capital* (social participation). In addition, it explores new empirics on an extraordinarily rich longitudinal Norwegian data set not previously applied in this field: The North-Trøndelag Health Study (known as the HUNT data set). The aim of this study was to integrate many of the ideas from the vast interdisciplinary literature on the role of social networks in terms of risky health behaviors. In particular, of the many papers on social capital and smoking, few attempt to address causality. The health economic papers that do measure the causal effects of social capital do not address smoking (e.g., Folland, 2007; d'Hombres et al., 2007; Fiorillo and Sabatini, 2015). The primary contribution of this paper is the use of a longitudinal (panel) Norwegian data set, which permits an examination of fixed/random effects and an instrumental variable (IV) approach to causality.

The remainder of this paper is organized as follows. The next section describes the data, the social capital variables, and the hypotheses. Section 3 describes the econometric approach used in this study. Section 4 presents the results, and the final section provides a discussion and conclusion and suggests implications and avenues for future research.

2. Data and variables

2.1 Data

The HUNT surveys (Helseundersøkelsen i Nord-Trøndelag) of a region in Norway contain extensive individual data about health and health-risk factors (for more information about HUNT, see Holmen et al. (2003), and the HUNT website (https://www.ntnu.edu/hunt). The HUNT data provide information about self-rated quality of life and health, body weight and height, lifestyle factors (alcohol and tobacco use and physical activities), family-related social capital (co-residence with spouse, partner, parents, etc.), labor market status, and other community-level social capital attributes.

The HUNT data collection was done in three waves. The present paper uses information compiled in HUNT 2 and HUNT 3. HUNT 2 was conducted from 1995 to 1997 and comprised 71.2 percent of the population aged 20 years and older (66,140 persons) as well as 8,984 pupils aged 13–19 years. HUNT 3 was conducted from October 2006 to June 2008. All inhabitants in Nord-Trøndelag 13 years and older were invited to participate in HUNT 3. The overall response rate was about 56 percent HUNT 3 was constructed in the same manner as HUNT 2, but includes several other topics.

In total, 55,629 individuals responded in HUNT 2 and 41, 983 in HUNT 3. Nevertheless, in our data set, education level (important socioeconomic information) was missing from HUNT 3; hence, we used lags of this covariate. Therefore, in analyzing HUNT 3, we used data on the individuals who participated in both HUNT 2 and HUNT 3. After dropping missing information on different covariates, our final analyses consist of 36,308 individuals in HUNT 2 and 36,308 in HUNT 3, representing a balanced panel of a total 76, 612 observations. This methodology also facilitates the use of "the lag of the variables", which is considered an instrument of some endogenous variables, and may protect us from a possible endogeneity bias of the estimates.

2.2 Variables

2.2.1 Dependent variable

'Cigarettes' is a dummy variable that takes the value of 1 if an individual reports smoking cigarettes daily and 0 otherwise.

2.2.2 Independent variables

Social capital variables and hypotheses

We consider three different dimensions/forms of social capital with four indicators to be representative of social capital measures available in two waves of our data set. For the first two types of social capital measures are available in HUNT 2 and HUNT 3, while the information for the third type (structural social capital) is only from HUNT 3:

i. Bonding social capital: family and friends

a. *Cohabit*: The variable cohabit (c*ohabit*) is defined as the cohabiting status of individuals: 1= living with spouse or partner and 0= living alone. We hypothesized that the presence of a spouse or partner would influence the individual to reduce smoking.

b. *Friends*: This variable was constructed on the basis of the question: "Do you have friends that you can speak to confidentially?", with yes=1 and no=0. We believed that having close friends would have an ambiguous influence on the individual regarding smoking.

ii. Cognitive social capital

Community Trust *(co_trust):* This social capital variable was created based on an individual's perception regarding community-level trust. The following statement was rated on a Likert scale from 1=strongly agree to 5=strongly disagree: "People can't trust each other here." We also created a dummy variable where somewhat disagree/agree or strongly disagree/not sure =1 and strongly agree=0. Some people are inherently more trusting than others; however, trust of the community will ultimately be a perception of and response to a specific community at hand. The individual has only a modest influence over this, and they bear no monetary cost for it. We hypothesized that increased trust in the community would influence the individual to reduce smoking.

iii. **Structural social capital**

Participation in Community Organizations (*part_org)* was defined as whether individuals participated in an association, club meeting, or activity. In particular, individuals were asked: "How many times in the last 6 months have you participated in an association or club meeting/activity?" From the alternative answers, we constructed a dichotomous variable where more than one time per week or one time/month or one to three times/month or one to five times/6 months=1 and never =0. This form of social capital is supposed to operate on both the propensity to smoke and budget constraints; hence one may expect that there may be an opportunity cost to participation in community organizations (i.e., this variable is believed to be endogenous). We hypothesized that participation in the community would have an ambiguous influence on individual participation regarding smoking.

2.2.3 Other covariates

In the empirical analysis, we included a large number of control variables. In particular:

Female –a dummy variable where women=1, men=0.

Age –(measured in years) at participation.

SES–socioeconomic status (SES) may influence an individual's decision to smoking or not. We controlled for education and childhood environment, which may be proxy variables for SES:

Education level is categorized into four levels and defined as follows:

educ1= 7 years of primary school or less, considered as an omitted category

educ2= High school, intermediate school, vocational school, 1–2 years of high school

educ3= University qualifying examination, junior college, A-levels

educ4= University or other post-secondary education, less than 4 years of university/college, 4 years or more

Indoor_smoke–As an indicator of individual SES, we included another variable to capture individuals' childhood environments, namely, *indoor_smoke*: "Did any of the adults where you grew up smoke indoors?", with yes=1 and no=0.

QoL: The quality of life (QoL) of an individual may also influence their decision to smoke or not, so we controlled for this attribute in the analysis. For this purpose, we included self-assessed quality of life (*QoL*) in the regression analyses. The following question was asked: "Thinking about your life at the moment, would you say that, by and large, you are satisfied with your life, or are you mostly dissatisfied?" We defined very satisfied/satisfied/somewhat satisfied=1 and neither satisfied nor dissatisfied/somewhat dissatisfied/dissatisfied/very dissatisfied=0.

BMI –We also controlled for individual body mass index (*BMI*), which is also thought to influence an individual's decision to smoke.

Table 1 provides the definitions and descriptive statistics of the variables used in the analyses.

<<<Table 1 about here>>>

3. Econometric approach

The modeling of a dichotomous dependent variable—smoking participation decision, namely whether an individual smokes cigarettes daily—was modeled as a function of social capital, along with their personal characteristics (age, gender, education, experience of indoor smoking, self-assessed QoL, and BMI). Smoking participation for individual *i* can then be expressed as:

$$c_{ijt} = \alpha + s_{ijt}\beta + x_{ijt}\delta + \varepsilon_{ijt} \quad (1)$$

where i= 1, 2…, n, indicates the individual, j=1, 2….M the municipalities, t=1, 2 the time period, c_{ijt} the smoking variable, s_{ijt} the vector of the four social capital measures, x_{ijt} the vector of personal characteristics, ε_{ijt} the error term with a mean of zero and constant variance, and β and δ are the parameters to be estimated.

Smoking participation outcomes within each municipality are likely to be correlated. To correct for this, we further estimated the municipality fixed-effect model, where ϕ_j is the dummy variable of municipality *j*:

$$c_{ijt} = \alpha' + s_{ijt}\beta' + x_{ijt}\delta' + \phi_j + \xi_{ijt} \quad (2)$$

The probability that over the sample period, individual *i* currently smokes cigarettes is given by: $prob(c_{ijt} = 1) = F(s_{ijt}\beta + x_{ijt}\delta)$, where $prob$ denotes probability, and *F* is the cumulative distribution function of the standard normal distribution (assuming that the error term in this equation has a standard normal distribution) gives the probit model. The parameters are estimated by maximum likelihood. For comparison with other models, the parameter vectors have also been estimated using ordinary least squares (OLS), i.e., using linear probability models (LPM).

Standard OLS and probit estimates of the coefficients associated with c_{ijt} yield unbiased results if $E(c_{ijt}\varepsilon_{ijt}) = 0$ holds. However, it was suspected that there are some specific reasons (mainly three) why the orthogonality condition could fail and prevent analysts from interpreting such a relationship as causal. First, it is typically difficult to distinguish the social capital effects from other local effects potentially influencing the smoking decision. Second, social capital and smoking might both be linked to other characteristics of the individual, some of which may be observable and potentially controlled for, but others (such as individual motivation) may not be, depending on specific and unobservable preferences. Hence, theoretically, social capital indicators may yield biased coefficients. Third, there is concern about possible reverse causality, that is, more smoking might lead to the accumulation of more or less social capital. For example, smokers might have more or fewer friends if their social milieu consists of mainly more or fewer smokers.

To handle the first problem, we considered municipality fixed-effects by including municipality dummies in the regression analyses (equation 2). Moreover, to account for the problem with endogeneity caused by unobserved individual level heterogeneity (which generates much of the endogeneity problem), we use the panel data regression approach, namely, individual *fixed-effects* or *random effects* estimators. Equation (1) for the fixed-effects model therefore becomes:

$$c_{ijt} = \alpha'' + s_{ijt}\beta'' + x_{ijt}\delta'' + \lambda_i + v_{ijt} \qquad (3)$$

To control for individual *fixed-effects,* we estimate the models with individual dummies, λ_i, or an equivalent approach, namely, where we "difference out" the individual effects by subtracting the individual means of the outcomes ($\bar{c}_{ijt}$), covariates ($\bar{s}_{ijt}$), and ($\bar{x}_{ijt}$) from c_{ijt}, s_{ijt}, and x_{ijt}, respectively. The *fixed-effects* model is useful when one wants to control for omitted variables that differ between individuals but are constant over time. These estimators, however, use only within-group (different from group mean) variations and ignore between-group variations. Alternatively, if there is reason to believe that some omitted variables may be constant over time but vary between individuals, and others may be fixed between individuals but vary over time;

one can include both types by using the *random effects* estimation approach. In that case, the cross-sectional variation may be more reliable than the within-group time-variation, indicating that fixed-effects estimation may be worse than the *OLS* or *random effects* alternatives[1]. Nevertheless, *fixed-effects* specifications are always a reasonable approach when using panel data (as they produce consistent results), but they may not be the most efficient model to run. Since this paper uses only two waves of data, one may not expect much within-group variation. We report both *fixed-effects* and *random effects* results.

To address the bias issues, we turn to IV estimates for social capital measures (s_{ijt}). To test and control for endogeneity, we first use an IV regression approach based on a recursive bivariate probit model (RBPM) in the following form:

$$c*_{ij} = \alpha^c + s_{ijt}\beta^c + x_{ijt}\delta^c + \phi_j^c + \mu^c{}_{ij}$$

$$s*_{ij} = \alpha^s + z_{ij}\kappa + x_{ijt}\delta^s + \mu^s{}_{ij}$$

$$E[\mu^c{}_{ij}] = E[\mu^s{}_{ij}] = 0$$

$$\mathrm{Var}[\mu^c{}_{ij}] = Var[\mu^s{}_{ij}] = 1$$

$$Cov[\mu^c{}_{ij,}\mu^s{}_{ij}] = \rho$$

where $s*_{ij}$ is an unobserved latent variable such that $s*_{ij}$ ≥0 if s_{ij} =1 and $s*_{ij}$ <0 if s_{ij} =0, β and δ are coefficients, and κ is a vector of IV that is correlated with s but not with μ. In our case, the coefficient of interest is β. The parameter ρ is the correlation between the error terms in the social capital variables and smoking equations. A Wald test of the significance of ρ is a direct test

[1] To decide between fixed or random effects one may run a Hausman test where the null hypothesis is that the preferred model is *random effects* vs. the alternative the *fixed-effects* (see Wooldridge, 2002).

of the endogeneity of *c* and *s* (Wooldridge, 2002, p.478). If $\rho=0$, then it is appropriate to use the univariate probit model as described in equation (2). Alternatively, the significant non-zero ρ value indicates that the social capital indicator and smoking decision are endogenous, the univariate probit results are biased, and the bivariate model should be used.

To check robustness, we also implement IV estimations using the generalized method of moments (GMM).

Notice that, in order for a variable, e.g., *z*, to serve as a valid instrument for *s*, the instrument must be exogenous, i.e., $\text{Cov}(z, \varepsilon) = 0$, and the instrument must be correlated with the social capital variable, *s*, i.e., $\text{Cov}(z,s) \neq 0$. Finding a valid instrument is often a difficult task. The longitudinal (panel) data allow us to consider lagged social capital attributes. Following earlier work, we use long lags of social capital indicators as potential instruments. These are cohabiting status in a previous wave, i.e., in HUNT 2 (*cohabit_96*), friends in a previous wave, (*friend_96*), and community trust perception in HUNT 2 (*co_trust_96*), for instrumenting *cohabit*, *friends*, and *co_trust* variables, respectably.

Distant lags offer a better instrument because a longer lag may reduce correlations between the instrument and the disturbances in the error term of the original OLS regression (Hall, 1988; Yogo, 2004; Murrey, 2006). Moreover, potential reverse causality, that is, more smoking participation leading to the accumulation of lower lagged social capital, may not be an issue here. However, it is also expected that more distant lags are more likely to be weakly correlated with the endogenous variable. Consequently, the validity of distant-lagged variable values as instruments has to be particularly convincing in order for such IV results to be trustworthy (Murrey, 2006).

The other endogenous social capital indicator, participation in organizational activities (*part_org*), is only available in HUNT 3. We use an individual attribute—how regularly an individual was participating in music/singing during the last 6 months (*music*)—as an instrument for *part_org*[2]. We construct the variable *music* as follows: if an individual participated sometime in a week/month=1 and never=0. The theoretical rationale behind choosing this variable as an instrument was that individuals who participate in different organizational activities may also participate in different cultural activities such as music and theatre. However, their affection for these sorts of cultural activities may not affect their odds of being a regular cigarette smoker or not. The plausibility of this assumption may be arguable. Nevertheless, the "relevance condition" (i.e., instruments are correlated with social capital variables) can be tested empirically. For this purpose, we illustrate the first stage regression summary statistics.

Moreover, to handle endogeneity issues for an important SES control variable, i.e., education, we use lagged values of these attributes (i.e., predetermined covariates) in all of our analyses. In some of the reduced form specifications, we also use lagged values of other individual attributes, including social capital, socio-demographic, BMI and QoL variables.

4. Results

4.1 OLS and probit estimates

Table 2 provides the OLS and probit estimates on individual smoking participation decisions while controlling for *municipality fixed-effects*. As seen in Table 2, smoking participation is regressed on HUNT 2 wave characteristics (i.e., lag of all independent variables) in the first two

---

[2] In HUNT 3, the following question was asked*: How many times in the last 6 months have you participated in music/singing/theatre?*

models (Models 1 & 2). Model 1 indicates that the lag of social capital attribute *friend* was not significant, but *cohabitation* (–0.010) and community trust were negatively and significantly associated with smoking participation in both the OLS and probit models. In Model 2, where we drop the social capital covariate *cohabitation* and other control variables, namely, QoL and BMI (those variables are believed to be more endogenous), we find *significant* associations between lagged *community trust* attributes and smoking status. In the other two models (Models 3 & 4), where smoking participation regressed on current covariates (i.e., values in HUNT 3), all social capital variables (except *friend*, which was found to be positively associated) were significantly and negatively associated with probability of smoking participation. These significant negative associations are also remained significant in Model 4 for *community trust (*–0.029) and *participation in organizational activities* (–0.032), where we drop more endogenous control variables, as in Model 2. In Models 3 and 4, *friend* was also found to be positively associated (only at the 10 percent significance level) with an individual's likelihood of cigarette smoking.

<<<Table 2 about here>>>

Other covariates were found to be associated with the smoking participation decision with the expected signs in all alternative models. In particular, the likelihood of cigarette smoking was lower for those who were female or aged and those with a higher level of education (lag variable). By contrast, if an individual grew up in a household where adults smoked indoors, this was positively associated with a higher likelihood of smoking participation. Finally, better self-assessed QoL and higher BMI were significantly associated with lower smoking participation in all specifications (whether or not we controlled for lagged or current covariates) and alternative estimation approaches (OLS and probit), and with similar magnitudes of the coefficients.

4.2 Panel data estimates: pooled OLS and fixed- and random effects estimates

Table 3 illustrates the estimates of alternative panel data regression results based on the *pooled OLS*, *fixed-effects,* and *random effects* models. For each panel data approach, we also estimate two alternative models (with and without more endogenous controls), as specified for OLS and probit approaches. For the pooled OLS model, we find similar results for all social capital variables. In Model 1, other than *friend*[3] (0.013), *community trust* (–0.024) and *cohabitation* (–0.040) were negatively and significantly associated with smoking participation. In Model 2, the *community trust* (–0.035*)* coefficient remained negative and significant with a higher absolute magnitude of the coefficient.

<<<Table 3 about here>>>

The *fixed-effects* estimates show that, other than *cohabitation* (–0.019), no other social capital attributes are significantly associated with smoking participation in Model 1 (Table 3). On the other hand, the *random effects* estimates are rather matched to the *pooled OLS* results-C*ommunity trust* (–0.010)[4] and *cohabitation* (–0.032) are negatively and significantly associated with smoking participation. In Model 2, *community trust* (–0.016) also remained significant at the 1 percent level with a slightly higher magnitude of the coefficient.

In the *fixed-effects* model, QoL is not significantly associated with smoking participation; however, all other control variables are significantly associated with the smoking participation decision with the expected signs, and similar results are seen for the OLS and probit estimates in Models 3 and 4 (Table 2).

---

[3] Significant only at the 10 percent level
[4] Significant only at the 10 percent level

To choose between *fixed-effects* or *random effects* models, we ran a Hausman test where the null hypothesis is that the preferred model is the *random effects* versus the *fixed-effects.* The Hausman test was highly significantly different from zero [$\chi^2(4)=207.29$], suggesting that the *fixed-effects* specifications are more reasonable here than the *random effects*[5].

4.3 Instrumental variable (IV) estimates

We used four different social capital variables and seemed that they were closely related. Therefore, it is rather implausible to argue that each of the social capital instruments only operates through the instrumented variable per se. Given that possibility, we run the models where one specific social capital variable is considered an endogenous variable and the other social capital variables are treated as exogenous variables with other control variables. For example, when focusing on *cohabiting status* as an endogenous social capital variable in a specification, the other social capital variables (i.e., *friend, co_trust, part_org)* are also included as covariates, but treated as exogenous variables.

As seen in Table 4, in two alternative RBPMs for the family-related social capital variable, *cohabit*,-show a significant and negative effect on smoking participation. The absolute magnitudes of the average marginal effects of *cohabit* are almost similar: –0.044 and –0.046 for Models 1 and 2, respectively. The magnitudes of the coefficients are similar to those found in the OLS and univariate probit models (see Table 2, Model 3). Notice that even the Wald test cannot reject the null hypothesis that $\rho=0$ for these models. Assuming that the instrument for *cohabit* is valid, the test result may suggest that *cohabit* may need to not be treated as endogenous, and that one may rely on univariate probit estimates for this family social capital variable.

---

[5] We use only two waves of data; therefore, the *fixed-effects* results need to be cautiously interpreted. We did not expect and observe much within-group variation in our data.

<<<Table 4 about here>>>

Moving to the next social capital variable, *friend*, negative but insignificant effects are observed for Models 1 (–0.036) and 2 (–0.035). However, the Wald test cannot reject the null hypothesis that $\rho=0$.

As seen in Table 4, strong negative and significant effects are observed for *community trust* with similar magnitudes of the coefficients (Model 1: –0.066; Model 2: –0.065). The Wald test rejects the null hypothesis that $\rho=0$ at the 1 percent level, which implies that the *community trust* variable is endogenous and that the univariate probit estimates for the *community trust* variable may be biased.

Another social capital variable, *participation in organization* (*part_org*), also negatively and significantly affects smoking likelihood (Model 1: –0.035; Model 2: –0.037). The magnitudes of the coefficients are similar to those found in the OLS (–0.036) and univariate probit (–0.038) models (see Table 2, Model 3). Moreover, the Wald test also shows non-rejection of the null hypothesis for these models. Assuming that the instrument for *part_org* is valid, the Wald test result suggests that *part_org* may not need to be treated as endogenous.

To check the robustness of our results, we re-estimate the models (Models 1 and 2 in Table 4) using the IV-GMM approach. We now consider all four social capital variables together, but one important difference here is that all social capital variables are treated as endogenous variables and are included simultaneously in a model.

As seen in Table 5, IV-GMM estimates show that the significant negative effects for the three social capital indicators with magnitude of the effects are quite similar to those estimated in the RBPMs (Table 4). In particular, the probability of smoking participation was *reduced* by 3 percent (3.1 percent in Model 2) if individuals are cohabiting, by 7.3 percent (7.5 percent in Model 2) if they live in a community with higher *community trust*, and by 4.5 percent (4.6 percent in Model 2) if they *participated in associations*, club meetings, or activities.

<<<Table 5 about here>>>

Questions could be raised about whether our instruments are sufficiently reasonable and valid to allow quantitative estimates of these effects. In response to this concern, **Appendix A1** shows the first stage regression results for four different social capital variables. It is evident that all social capital instruments used in the analyses are highly significantly correlated with their social capital measures. In addition, intuitive arguments in favor of our instruments are provided in the Methods section, and the diagnostic tests (relevance condition) for our instruments are provided in the bottom row of **Appendix A1**. We are confident that our instruments satisfy the rule of thumb criterion suggested by Staiger and Stock (1997). The criterion states that the test rejects the null hypothesis of weak instruments if the first-stage F-statistic is bigger than 10. The adjusted R^2s seem rather low for the instruments used for *friend* (0.06) and *community trust* (0.10), and this may indicate that the lag of these social capital attributes may not be strongly associated with their current social capital status. On the other hand, as seen in **Appendix A1**, the first stage summary statistics for the instrument for other social capital measures—*cohabit* (adjusted R^2=0.25 and partial R^2=0.23) and *participation in various organizations* (adjusted R^2=0.25 and partial R^2=0.23)—are comparatively strong.

The second condition for IVs is that the instruments must not be correlated with the structural error term. If the model is just identified, then we cannot perform a test of overidentifying restrictions. If the model is overidentified, then we can test whether the instruments are correlated with the error term. For such an exercise, we estimate a single model in which we include one extra instrument (perception of the community) with our four social capital variables as the model to be overidentified and then run an overidentification test. The non-rejection of the test [Hansen's J $\chi^2(1)=0.222$ (p=0.638)] of overidentification may suggest that our set of instruments is reasonable and valid. Moreover, the estimated coefficients for our social capital variables were also found to be similar (–0.030, –0.072, and –0.046 for *cohabit*, *co_trust*, and *part_org*, respectively)[6].

In addition, to test the sensitivity of our results, using both IV approaches—RBPM and IV-GMM—we further estimate our social capital variables separately. Irrespective of our chosen specifications, we find that the magnitudes of the effects are generally comparable with alternative IV approaches. As seen in Appendix Tables A2 and A3, both RBPM and IV-GMM estimates show that the probability of smoking participation is reduced by around 4.6 percent (4.5 percent) if individuals are cohabiting, by 6.9 percent (7.0 percent) if they live in a community with higher *community trust*, and by 3.2 percent (2.8 percent) if they *participated in associations*, club meetings, or activities. Nevertheless, the significant negative effect of *part_org* seems to be sensitive regardless of whether we include cohabiting status as a control variable.

5. Discussion and conclusion

[6]We include one extra instrumental variable, namely *co_like*, defined as an attribute on the perception of liking the community: co_like (People like living here; 1=strongly agree to 5=strongly disagree. We define somewhat agree or strongly agree=1 and not sure/somewhat disagree or strongly disagree=0. However, the assumption that co_like only operates through our social capital variables and not through the individual's response to the question seems rather strong and perhaps difficult to defend. Therefore, we do not emphasize these results, and we do not report the full results here.

We used a substantial Norwegian longitudinal data set, the HUNT data set to test the hypothesis that social capital does reduce the likelihood of cigarette smoking. Our results based on OLS and probit models show that the likelihood of smoking participation is negatively associated with three alternative forms of social capital demonstrated by four indicators: *cohabitation*, *friend, community trust,* and *participation in organizations*. Our reduced form findings based on OLS and probit models are consistent with previous studies that reported negative associations between smoking rates and levels of community trust in different settings (e.g., Sapag et al. 2010; Afifi et al., 2010) and positive associations between smoking cessation and social participation (e.g., Lindstrom et al., 2003) and being married or cohabitation (e.g., Lee and Kahende, 2003). Using a two-part demand-for-cigarettes model to determine the association between CSC and smoking status, Brown et al. (2006) found a strong negative association with the number of cigarettes that smokers consume (with an elasticity of –1.04); however, they did not find any association with smoking participation.

Previous research also indicates that teenagers and young 20-somethings tend to be more vulnerable to the attractions of smoking. Within this endeavor, studies on adolescents have found that peer smoking behaviors are especially influential (e.g., Glaser et al., 2010; Stewart-Knox et al., 2005; Page et al., 2006), and greater social capital may encourage smoking if the peer cohort smokes. We have hypothesized that more friends and acquaintances would have an ambiguous influence on the individual regarding smoking. Our findings support such ambiguity on the relation (either weakly positive or no significant association) between friends and the likelihood of smoking participation.

To account for the problem with endogeneity caused by unobserved individual level heterogeneity, we use panel data regression approaches. Panel data estimation results based on *pooled OLS* and *random-effects* also indicate that the likelihood of smoking participation is negatively associated with community trust and cohabitation, but with lower magnitudes of the coefficients in the *random-effects* estimates. It is important to note that our *fixed-effects* estimates support the hypothesis that *cohabitation* status is negatively associated with the probability of smoking participation. However, the *fixed-effect* estimates do not support such a significant negative relation between community trust and the likelihood of smoking participation. To our knowledge, no previous longitudinal studies have used the panel data approach, so we could not compare our results with those from previous studies. Giordano and Lindstrom (2010) use data from the British Household Panel Survey, but they do not apply panel data regression methods. In particular, they identify smoking and non-smoking cohorts at baseline and construct a dependent variable by defining whether individual smoking behavior had changed or not. Measures of social capital (measured by "trust" and "social participation"), income, employment and marital status were tested for associations with changes in smoking behavior over a 2-year period using multivariate logistic regression models. They found that social capital measures are independently and positively associated with smoking cessation and that remaining single is associated with smoking initiation.

To interpret whether smoking likelihood and social capital relationships are causal, we further tested and tried to tackle the endogeneity problem according to different social capital indicators using two alternative IV techniques: the recursive bivariate probit approach and the IV-GMM approach. Our alternative IV estimates also show negative and significant effects for family-related social capital variables and *cohabiting status*, which supports our hypothesis and is in line with previous studies. For example, in a longitudinal study, Burt (2010) investigated the behavior of young men prior to and after marriage. Among unmarried men, those who would become married showed less antisocial behavior. Once married, their antisocial behaviors again declined. This negative effect of cohabitation on smoking participation seems consistent, controlling for with or without more covariates (believed to be more endogenous, such as QoL and BMI). Two other forms of social capital, community trust and participation in organizational activities, also negatively affect the likelihood of smoking, and thus support our hypothesis. Although the negative effect of greater community trust on smoking participation seems unaffected with or without more covariates in the models, the significant effect of social participation on smoking participation appears sensitive with or without the inclusion of the cohabiting variable in the models. It is important to mention that in previous social capital and smoking research, along with age, gender, and self-rated health/psychological health/QoL, cohabitation/marital status was also included as a covariate and as variables considered to be potential confounders to the relation (e.g., Giordano and Lindstrom, 2010; Brown et al., 2006).

Our overall conclusion is that for the Norwegian population studied, an individual's smoking decision is modified beneficially through most if not by all attributes of social capital. This finding suggests ideas for future research. One may think that the expected effect of social capital variables to some extent depends on the context. In particular, for young people, peer pressure may play an important role in determining smoking participation. Due to peer pressure, greater social capital could well increase the likelihood of smoking participation within peer groups, specific forms of social capital (e.g., friend, spouse, or cohabitant) may operate through peer effects, and methods of studying peer effects may be applied. Future research should focus on others' risky health behaviors such as obesity or the use of heroin, crack cocaine, ecstasy, and other illicit narcotics.

References

Afifi, R.A., Naakkash, R.T., Khawaja, M. 2010. Social capital, women's autonomy and smoking among married in low-income urban neighborhoods of Beirut, Lebanon, *Women's Health Issues*, 20, 156-67.

Akerlof, G.A. 1998. Men without children, *The Economic Journal*, 115, 715-753.

Becker, G.S., Murphy, K.M. 2000. Social Economics. Market Behavior in a Social Environment. The Belknap Press of Harvard University Press, Cambridge, MA.

Bourdieu, P.1985.The forms of capital, in Richardson, J. ed, Handbook of Theory and Research in the Sociology of Education. New York.

Brown TT**,** Scheffler RM, Seo S., Reed M. 2006. The Empirical Relationship between Community Social Capital and the Demand for Cigarettes. *Health Economics,* 15, 1159-72.

Burt, S.A. 2010, Does marriage inhibit antisocial behavior? An examination of selection vs causation via a longitudinal twin design. Archives of General Psychiatry, 67, 1309-1315.

Coleman, J. 1988, Social capital in the creation of human capital. *The American Journal of Sociology: Supplement*, 94, S95-S120.

d'Hombres, B., Rocco, L., Suhrcke, M., and McKee, M. 2010. Does social capital determine health? evidence from eight transition countries. *Health Economics*. 19, 56-74.

Durlauf, S. N. 2002, On the empirics of social capital, *The Economic Jou*rnal. 112, F-439-479.

Fiorillo, D, Sabatini, K., 2015, Structural social capital and health in Italy*, Economics & Human Biology*, 17, 129-142.

Folland, S. 2006, Value of life and behavior toward health risks: An interpretation of social capital, *Health Economics*, 15, 159–171.

Folland, S. 2007. Does community social capital contribute to population health? *Social Science & Medicine*, 64, 2342-2354.

Folland, S. 2008. An economic model of social capital and health. *Health Ecoonomics, Policy and Law*. 3, 1-15.

Folland, S. Kaarbøe, O., Islam, MK. 2014. How Do We Invest in Social Capital? An Exploration of an Economic Model of Social Capital and Health, In Sherman Folland & Lorenzo Rocco (ed.), The Economics of Social Capital and Health. World Scientific. ISBN 978-981429339-6. Chapter 3. s 17 – 26.

Giordano, G.N., Lindstrom, M. 2010. The impact of social capital on changes in smoking behavior: a longitudinal cohort study. European Journal of Public Health, 21, 347-354.

Glaeser, E.L., Laibson, D., Sacerdote, B. 2002. An economic approach to social capital. *The Economic Journal*, 112, F437–F458.

Glaser, B., Shelton, K.H., van den Bree, M.B., 2010. The moderating role of close friends in the relationship between conduct problems and adolescent substance abuse. *Journal of Adolescent Health*, 47, 35-42.

Gruber, J. and J. Zinman, 2002. Youth smoking in the U.S.: Evidence and implications, National Bureau of Economic Research, Working Paper 8872.

Hall, Robert E. 1988. Intertemporal substitution in consumption. *Journal of Political Economy*. 96, 339–57.

Holmen, J. Midthjell, K. Krüger, Ø et al. 2003. The Nord-Trøndelag Health Study 1995-97 (HUNT 2): Objectives, contents, methods and participation. *Norsk Epidemiologi* (The Norwegian Journal of Epidemiology), 13, 19-32.

Homish, G. G., Leonard, K. E. 2005. Marital Quality and Congruent Drinking. *Journal of Studies on Alcohol*, *66*, 488–496.

Islam, M.K., Merlo, J., Kawachi, I., Lindström, M.,Gerdtham, U-G. 2006a. Social capital and health: does egalitarianism matter? A literature review. *International Journal for Equity in Health* 5, 1-28.

Islam, M. K ., Merlo, J., Kawachi, I., Lindström, M., Burström, K, Gerdtham, U-G, 2006b. Does it really matter where you live? A panel data multilevel analysis of Swedish municipality level social capital on individual health-related quality of life. *Health Economics, Policy and Law*, 1, 209-235.

Islam, M. K., Gerdtham, U-G. , Gullberg, B., Lindström, M., Merlo, J. 2008. Social capital externalities and mortality in Sweden. *Economics and Human Biology*, 6, 19–42.

Kawachi, I. 1999, Social capital and community effects in population and individual health. *Annals of the New York Academy of Sciences*, 896, 120–130.

Kawachi, I. and Berkman L.F, 2001, Social ties and mental health, *Journal of Urban Health*, 78: 458–467.

Kawachi, I., Kennedy, B.P., Lochner, K., Prothrow, S.D. 1997. Social capital income inequality, and mortality. *American Journal of Public Health*, 87(9):1491-8.

LaPorte, Audrey and Brian Ferguson, 2004.Testing the impact of social capital on population health using a Grossman framework, mimeo.

Lee, C.W., Kahende, J., 2007. Factors associated with successful smoking cessation in the United States, 2000, *American Journal of Public Health*, 97, 1503-9.

Lindstrom, M., Isacsson, S-O, Elmstahl, S. 2003. Impact of different aspects of social participation and social capital on smoking cessation among daily smokers: a longitudinal study, Tobacco Control 12, 274-81.

Ljunge, M. 2014. Social capital and health: Evidence that ancestral trust promotes health among children of immigrants. *Economics & Human Biology*, 15, 165-186.

Lochner K, Kawachi I, Kennedy BP. 1999. Social capital: a guide to its measurement. Health & Place, 5, 259-270.

Loury, G., 1977. A dynamic theory of racial income differences. in Wallace, P.A. and LaMond, A., eds, Women, minorities and employment discrimination. Lexington, MA: Lexington Books.

Mellor, J.M., Milyo, J. 2005. State social capital and individual health status. *Journal of Health Politics Policy and Law* 30, 1101–1130.

Murrey,. M. P. 2006. Avoiding invalid instruments and coping with weak instruments. *Journal of Economic Perspectives.* 20,111–132.

Page, R.M., et al., 2006. Cigarette smoking, friendship factors, and social norm perceptions among central and Eastern European high school students. *Journal of Drug Education*, 36, 213-31.

Poortinga, W. 2006. Social capital: An individual or collective resource for health? *Social Science & Medicine*, 62, 292–302.

Putnam, R.D., 1995. Bowling alone: America's declining social capital. *Journal of Democracy* 6, 65-78.

Putnam, R.D., 2000, Bowling alone, New York: Simon & Schuster.

Putnam, R.D., Leonardi, R., Nanetti, R., 1993. Making democracy work: Civic tradition in modern Italy. Princeton: Princeton University Press.

Sapag J.C. Poblete, F.C., Eicher C., Aracena M., Caneo C., Vera G., Martínez M., Hoyos R., Villarroel L., Bradford E. 2010. Tobacco smoking in urban neighborhoods: exploring social capital as a protective factor in Santiago, Chile. *Nicotine and Tobacco Research*, 12, 927-36.

Sapolsky, Robert M., 1998. Why zebras don't get ulcers, An updated guide to stress, stress related diseases and coping. New York: W. H. Freeman and Company.

Staiger, D. and Stock, J.H. 1997. Instrumental variables regression with weak instruments". *Econometrica*. 65, 557-586.

Stewart-Knox, B.J., et al., 2005. Smoking and peer groups: results from a longitudinal qualitative study of young people in Northern Ireland, *The British Journal of Social Psychology*, 44, 397-414.

Sundquist K, Yang M. 2007. Linking social capital and self-rated health: a multilevel analysis of 11,175 men and women in Sweden. Health Place, 13, 324-334.

Turrell G, Kavanagh A, Subramanian S V. 2006. Mortality in Tasmania (Australia): the contributions of socioeconomic disadvantage, social capital and geographic remoteness. *Health and Place*, 12, 291-305.

Wooldridge, Jeffery M. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

Yogo, Motohiro. 2004. Estimating the Elasticity of Intertemporal Substitution When Instruments are Weak. *Review of Economics and Statistics*, 86, 797–810.

Table 1: Definitions and descriptive statistics of the variables used in the analyses

Name of the variable	**Definition of the Variable**	HUNT 2 (N=36,308)		HUNT 3 (N=36,308)	
		Mean /pro	Std. Dev	Mean /pro	Std. Dev
cigarettes	Smokes cigarettes daily: yes = 1. else =0	0.260	0.439	0.185	0.388
friend	Do you have friends that you can speak to confidentially? Yes =1; No = 0	0.973	0.163	0.924	0.264
co_trust	People can't trust each other here; This is a Likert scale with 1=strongly agree to 5=strongly disagree; Strongly disagree=1; Strongly agree/Somewhat agree/disagree or not sure =0	0.300	0.458	0.351	0.477
part_org§	"How many times in the last 6 months have you participated in an association or club meeting/activity?" More than 1 time per week or 1 time/week or 1-3 times/ month or 1-5 times/6 month =1; Never =0	NA		0.591	0.492
cohabit	1 = living with spouse or partner 0= living alone.	0.730	0.444	0.682	0.466
age	Age at participation	46.78	13.38	58.12	13.34
sex	female= 1; male = 0	0.448	0.497	0.448	0.497
Educational level±					
educ1	7 years primary school or less (Omitted category)	0.302	0.459	NA	
educ2	High school, intermediate school, vocational school, 1-2 years high school	0.365	0.481	NA	
educ3	University qualifying examination, junior college, A levels	0.090	0.287	NA	
educ4	University or other post-secondary education. less than 4 years or University/college. 4 years or more	0.220	0.414	NA	
indoor_smoke	"Did any of the adults where you grew up smoke indoors?" Yes =1; No = 0	0.652	0.476	0.660	0.474
QoL	Thinking about your life at the moment, would you say that you by and large are satisfied with life, or are you mostly dissatisfied? Very satisfied/ Satisfied/ Somewhat satisfied=1; Neither satisfied nor dissatisfied/ somewhat dissatisfied/Dissatisfied/ Very dissatisfied=0	0.836	0.371	0.860	0.347
bmi	Body Mass Index	26.17	3.861	27.40	4.290

Note: ± the information is available only in HUNT 2. §only asked in HUNT 3.

Table 2: Smoking participation and social capital: OLS (LPM) and Probit marginal estimates with municipality fixed–effects

Variables	Smoking regressed on lagged independent variables (HUNT 2 wave characteristics)				Smoking regressed on current independent variables (HUNT 3 wave characteristics)			
	Model 1		Model 2		Model 3		Model 4	
	OLS	Probit	OLS	Probit	OLS	Probit	OLS	Probit
friend	**0.015**	**0.012**	**0.007**	**0.005**	**0.015****	**0.015****	**0.014***	**0.014***
	(0.012)	**(0.012)**	**(0.012)**	**(0.012)**	**(0.008)**	**(0.007)**	**(0.008)**	**(0.007)**
co_trust	**–0.015*****	**–0.015*****	**–0.020*****	**–0.020*****	**–0.016*****	**–0.017*****	**–0.028*****	**–0.029*****
	(0.004)	**(0.005)**	**(0.004)**	**(0.005)**	**(0.004)**	**(0.003)**	**(0.004)**	**(0.004)**
part_org	____	____	____	____	**–0.036*****	**–0.038*****	**–0.032*****	**–0.032*****
					(0.004)	**(0.004)**	**(0.004)**	**(0.004)**
cohabit	**–0.010****	**–0.010****	____	____	**–0.044*****	**–0.045****	____	____
	(0.006)	**(0.005)**			**(0.005)**	**(0.005)**		
age	–0.001***	–0.001***	–0.002***	–0.002***	–0.002***	–0.002***	–0.002***	–0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
sex	–0.030***	–0.029***	–0.032***	–0.033***	–0.027***	–0.027***	–0.033***	–0.032***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
educ2	–0.037***	–0.033***	–0.037***	–0.033***	–0.034***	–0.030***	–0.034***	–0.031***
	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.004)	(0.005)
educ3	–0.107***	–0.086***	–0.106***	–0.086***	–0.105***	–0.084***	–0.101***	–0.082***
	(0.008)	(0.006)	(0.008)	(0.006)	(0.008)	(0.006)	(0.006)	(0.006)
educ4	–0.148***	–0.133***	–0.145***	–0.132***	–0.143***	–0.129***	–0.138***	–0.127***
	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)
indoor_smoke	0.070***	0.070***	0.070***	0.070***	0.080***	0.051***	0.076***	0.076***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
QoL	–0.062***	–0.061***	____	____	–0.054***	–0.046***	____	____
	(0.006)	(0.006)			(0.006)	(0.006)		
bmi	–0.006***	–0.005***	____	____	–0.008***	–0.008***	____	____
	(0.001)	(0.001)			(0.000)	(0.000)		
R–squared	0.041	0.045	0.035	0.039	0.051	0.062	0.038	0.042
municipality fixed–effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	36 308	36 308	36 308	36 308	36 308	36 308	36 308	36 308

Note: Robust standard errors are in the parentheses. '***', '**' and '*' indicates significance level at the 1%, 5% and 10% respectively.

Table 3: Smoking participation and social capital: Panel data Pooled OLS, Fixed–effects and Random– effects estimates

Variables	Pooled OLS Models		Fixed–effects Models		Random–effects Models	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
friend	**0.013***	**0.010**	**–0.003**	**0.001**	**0.002**	**0.004**
	(0.007)	**(0.007)**	**(0.005)**	**(0.006)**	**(0.005)**	**(0.005)**
co_trust	**–0.024*****	**–0.035*****	**0.007**	**0.003**	**–0.010***	**–0.016*****
	(0.006)	**(0.007)**	**(0.004)**	**(0.004)**	**(0.005)**	**(0.005)**
cohabit	**–0.040*****	___	**–0.018*****	___	**–0.032*****	___
	(0.003)		**(0.003)**		**(0.003)**	
age	–0.001***	–0.001***	0.003	0.001	–0.001***	–0.001***
	(0.000)	(0.000)	(0.003)	(0.003)	0.000)	(0.000)
sex	–0.035***	–0.039***	___	___	–0.034***	–0.039***
	(0.004)	(0.004)			(0.004)	(0.004)
indoor_smoke	0.101***	0.098***	0.026**	0.030***	0.084***	0.082***
	(0.005)	(0.005)	(0.010)	(0.011)	(0.006)	(0.005)
QoL	–0.070***	___	0.007	___	–0.034***	___
	(0.010)		(0.007)		(0.007)	
bmi	–0.009***	___	–0.023***	___	–0.011***	___
	(0.001)		(0.001)		(0.000)	
year08	–0.057***	–0.063***	–0.083**	–0.086***	–0.056***	–0.064***
	(0.003)	(0.003)	(0.033)	(0.030)	(0.003)	(0.003)
R–squared	0.040	0.029	0.012	0.012	0.037	0.026
Number of observations	72 616	72 616	72 616	72 616	72 616	72 616

Note: Robust standard errors (are adjusted for clusters in municipalities) are in the parentheses.
'***', '**' and '*' indicates significance level at the 1%, 5% and 10% respectively.

Table 4: Smoking participation and social capital: Recursive bivariate probit estimates for social capital variables together: Average marginal effects

Variables	Social Capital variable ***cohabit*** treated as endogenous≠		Social Capital variable ***friend*** treated as endogenous§		Social Capital variable ***"co_trust"*** treated as endogenous €		Social Capital variable ***"part_org"*** treated as endogenous±	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
cohabit	**–0.044*** **	**–0.046*** **	–0.046***	–0.050***	–0.032***	–0.037***	–0.043***	–0.047***
	(0.011)	**(0.011)**	(0.005)	(0.005)	(0.006)	(0.006)	(0.005)	(0.005)
friend	0.016*	0.009	**–0.036**	**–0.035**	0.023***	0.016*	0.015*	0.007
	(0.008)	(0.008)	**(0.036)**	**(0.038)**	(0.008)	(0.008)	(0.008)	(0.008)
co_trust	–0.017***	–0.019***	–0.015***	–0.016***	**–0.066*** **	**–0.065*** **	–0.017***	–0.019***
	(0.005)	(0.005)	(0.005)	(0.005)	**(0.018)**	**(0.018)**	(0.004)	(0.004)
part_org	–0.037***	–0.041***	–0.035***	–0.039***	–0.039***	–0.042***	**–0.035**	**–0.037*** **
	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	**(0.010)**	**(0.010)**
age	–0.002***	–0.002***	–0.002***	–0.002***	–0.001***	–0.002***	–0.002***	–0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
sex	–0.027***	–0.032***	–0.029***	–0.034***	–0.027***	–0.032***	–0.027***	–0.032***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
educ2	–0.031***	–0.028***	–0.031***	–0.028***	–0.028***	–0.025***	–0.031***	–0.028***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
educ3	–0.099	–0.094***	–0.099***	–0.094***	–0.095***	–0.090***	–0.099***	–0.094***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
educ4	–0.153***	–0.146***	–0.154***	–0.146***	–0.147***	–0.139***	–0.154***	–0.146***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
indoor_smoke	0.082***	0.081***	0.082***	0.081***	0.082***	0.081***	0.082***	0.081***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
QoL	–0.047***	___	–0.043***	___	–0.044***	___	–0.047***	___
	(0.006)		(0.006)		(0.006)		(0.006)	
bmi	–0.008***	___	–0.008***	___	–0.008***	___	–0.008***	___
	(0.000)		(0.000)		(0.000)		(0.000)	
Municipality fixed–effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ρ	–0.003	–0.007	0.106	0.087	0.122	0.117	–0.008	–0.013
Wald test ρ=0	χ2(1)= 0.000	χ2(1)= 0.059	χ2(1)=2.26	χ2(1)= 1.387	χ2(1)= 7.47	χ2(1)= 6.86	χ2(1)=0.098	χ2(1)= 0.279
[p–values]	[0.991]	[0.807]	[0.132]	[0.239]	**[0.006]**	**[0.009]**	[0.755]	[0.597]

Note: Delta–method standard errors are in the parentheses. '***', '**' and '*' indicates significance level at the 1%, 5% and 10% respectively.

≠ Instruments: cohabiting status in 1996; § Instruments: friends in 1996; € Instrument used is co_trust in 1996; ± Instruments used is how frequent individual has participated in music in HUNT 3.

Table 5: Smoking participation and social capital: Instrumental Variables (IV)–GMM estimates for all social capital variables together

Variables	Model 1	Model 2
cohabit	**–0.030****	**–0.031****
	(0.013)	**(0.013)**
friend	**0.078**	**0.073**
	(0.065)	**(0.063)**
co_trust	**–0.073*****	**–0.075*****
	(0.020)	**(0.020)**
part_org	**–0.045*****	**–0.046*****
	(0.015)	**(0.015)**
age	–0.002***	–0.002***
	(0.000)	(0.000)
sex	–0.025***	–0.030***
	(0.005)	(0.005)
educ2	–0.030***	–0.028***
	(0.006)	(0.006)
educ3	–0.099***	–0.093***
	(0.009)	(0.009)
educ4	–0.135***	–0.126***
	(0.008)	(0.008)
indoor_smoke	0.079***	0.077***
	(0.004)	(0.004)
QoL	–0.055	___
	(0.008)	
bmi	–0.008***	___
	(0.000)	
Municipality fixed–effects	Yes	Yes

Note: Robust standard errors are in the parentheses. '***', '**' and '*' indicates significance level at the 1%, 5% and 10% respectively.

≠ Instruments: cohabiting status in 1996; §Instruments: friends in 1996; € Instrument used is co_trust in 1996; ± Instruments used is how frequent individual has participated in music in HUNT 3.

Appendix

Table A1: First–stage regressions results for all social capital variables together: models estimated in table 5

variables	**cohabit**	**friend**	**co–trust**	**part_org**
cohabit_96[≠]	**0.3716*****	–0.0094***	0.0301***	0.0024
	(0.0054)	(0.0031)	(0.0054)	(0.0051)
friend_96[§]	0.0514***	**0.2050*****	0.0720***	0.0654***
	(0.0127)	**(0.0141)**	(0.0134)	(0.0137)
co_trust_96 [€]	0.0001	0.0279***	**0.2437*****	0.0257***
	(0.0046)	(0.0028)	**(0.0057)**	(0.0052)
part_music[±]	–0.2849***	0.0580***	–0.1262***	**0.4747*****
	(0.0047)	(0.0026)	(0.0048)	**(0.0043)**
age	–0.0024***	–0.0007***	0.0014***	0.0006***
	(0.0002)	(0.0001)	(0.0002)	(0.0002)
sex	0.0495***	–0.0475***	–0.0068	–0.0130***
	(0.0043)	(0.0028)	(0.0049)	(0.0046)
educ2	0.0342***	–0.0036	0.0502***	0.0716***
	(0.0055)	(0.0036)	(0.0062)	(0.0057)
educ3	0.0326***	–0.0009	0.0757***	0.1170***
	(0.0088)	(0.0053)	(0.0098)	(0.0095)
educ4	0.0683***	–0.0103**	0.1069***	0.1593***
	(0.0063)	(0.0041)	(0.0072)	(0.0067)
indoor_smoke	0.0068	–0.0020	–0.0005	–0.0179***
	(0.0046)	(0.0029)	(0.0051)	(0.0048)
QoL	0.1039***	0.0887***	0.0944***	0.0351***
	(0.0067)	(0.0052)	(0.0065)	(0.0065)
bmi	0.0008	–0.0003	–0.0005	0.0036***
	(0.0005)	(0.0003)	(0.0006)	(0.0005)
municipality fixed–effects	Yes	Yes	Yes	Yes
First–stage regression: summary statistics				
Adjusted R^2	0.2525	0.0568	0.0984	0.2461
Partial R^2	0.2330	0.0315	0.0783	0.2292
Robust F (4, 36272)	3083.23	197.412	771.179	3044.24

Note: Robust standard errors are in the parentheses. '***' and '**' indicates significance level at the 1% and 5% respectively. [≠] Instruments: cohabiting status in 1996; [§]Instruments: friends in 1996; [€] Instrument used is co_trust in 1996; [±] Instruments used is how frequent individual has participated in music in HUNT 3.

Table A2: Smoking participation and social capital variables estimates separately: Recursive bivariate probit estimates Average marginal effects

Variables	Social Capital variable *cohabit*≠		Social Capital variable *friend*§			Social Capital variable *co–trust*€			Social Capital variable *part_org*±		
	Model 1	Model 2	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
friend	——	——	**–0.023 (0.043)**	**–0.073* (0.044)**	**–0.073* (0.044)**	——	——	——	——	——	——
co_trust	——	——	——	——	——	**–0.079*** (0.017)**	**–0.081*** (0.017)**	**–0.069*** (0.018)**	——	——	——
part_org	——	——	——	——	——	——	——	——	**–0.000 (0.008)**	**–0.013 (0.009)**	**–0.032*** (0.009)**
cohabit	**–0.039*** (0.011)**	**–0.045*** (0.011)**	——	——	–0.046*** (0.004)	——	——	–0.027*** (0.006)	——	——	–0.051*** (0.005)
age	–0.002*** (0.000)	–0.002*** (0.000)	–0.002*** (0.000)	–0.002*** (0.000)	–0.002*** (0.000)	–0.001*** (0.000)	–0.002*** 0.000	–0.002*** (0.000)	–0.002*** (0.000)	–0.002*** (0.000)	–0.002*** (0.000)
sex	–0.027*** (0.004)	–0.032*** (0.004)	–0.029*** (0.005)	–0.035*** (0.005)	–0.035*** (0.005)	–0.029*** (0.004)	–0.033*** 0.004	–0.033*** (0.004)	–0.028*** (0.004)	–0.033*** (0.004)	–0.031*** (0.004)
educ2	–0.035*** (0.005)	–0.032*** (0.005)	–0.037*** (0.006)	–0.032*** (0.005)	–0.032*** (0.005)	–0.031*** (0.006)	–0.029*** (0.005)	–0.029*** (0.005)	–0.037*** (0.005)	–0.034*** (0.005)	–0.030*** (0.005)
educ3	–0.105*** (0.008)	–0.099*** (0.008)	–0.106*** (0.008)	–0.100*** (0.008)	–0.100*** (0.008)	–0.100*** (0.008)	–0.094*** (0.008)	–0.095*** (0.008)	–0.106*** (0.008)	–0.100*** (0.008)	–0.960*** (0.008)
educ4	–0.162*** (0.007)	–0.155*** (0.006)	–0.167*** (0.006)	–0.155*** (0.006)	–0.155*** (0.006)	–0.154*** (0.006)	–0.147*** (0.007)	–0.147*** (0.007)	–0.166*** (0.006)	–0.157*** (0.006)	–0.149*** (0.006)
indoor_smoke	0.083*** (0.004)	0.086*** (0.004)	0.081*** (0.004)	0.081*** (0.004)	0.081*** (0.004)	0.082*** (0.004)	0.080*** (0.004)	0.081*** (0.004)	0.082*** 0.004	0.080*** (0.004)	0.081*** (0.004)
QoL	–0.050*** (0.006)	——	–0.052*** (0.007)	——	——	–0.045*** (0.006)	——	——	–0.055*** (0.005)	——	——
BMI	–0.008*** (0.000)	——	–0.008*** (0.000)	——	——	–0.008*** (0.000)	——	——	–0.008*** (0.000)	——	——
Municipality fixed–effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ρ	0.0002	0.003	0.074	0.140	0.140	0.142	0.139	0.130	–0.0791	–0.049	–0.022
Wald test ρ=0	χ2(1)= 0.00	χ2(1)= 0.01	χ2(1)= 0.80	χ2(1)= 2.85	χ2(1)= 2.85	χ2(1)= 11.7	χ2(1)= 11.5	χ2(1)= 9.00	χ2(1)= 12.3	χ2(1)= 4.02	χ2(1)= 0.80
[p–values]	[0.993]	[0.919]	[0.373]	**[0.091]**	**[0.091]**	**[0.001]**	**[0.003]**	**[0.001]**	**[0.000]**	**[0.045]**	[0.370]

Note: Delta–method standard errors are in the parentheses. '***', '**' and '*' indicates significance level at the 1%, 5% and 10% respectively.
≠ Instruments: cohabiting status in 1996; §Instruments: friends in 1996; € Instrument used is co_trust in 1996; ± Instruments used is how frequent individual has participated in music in HUNT 3.

Table A3: Smoking participation and social capital: Instrumental Variables (IV) GMM estimates for each SC variable separately

Variables	IV models for SC variable *cohabit*≠		IV models for SC variable *friend*§			IV models for SC variable *co–trust*€			IV models for SC variable *part_org*±		
	Model 1	Model 2	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
friend	——	——	**0.035 (0.057)**	**0.025 (0.056)**	**0.024 (0.056)**	——	——		——	——	——
co_trust	——	——	——	——		**–0.077*** (0.017)**	**–0.079*** (0.017)**	**–0.070*** (0.017)**	——	——	——
part_org	——	——	——	——		——	——		**0.005 (0.009)**	**0.005 (0.009)**	**–0.028*** (0.017)**
cohabit	**–0.045*** (0.012)**	**–0.046*** (0.012)**	——	——	–0.043*** (0.005)	——	——	–0.027*** (0.006)	——	——	–0.050*** (0.005)
age	–0.003*** (0.000)	–0.002*** (0.000)	–0.002*** (0.000)	–0.002*** (0.000)	–0.002*** (0.000)	–0.002*** (0.000)	–0.002*** (0.000)	–0.002*** (0.000)	–0.002*** (0.000)	–0.002*** (0.000)	–0.002*** (0.000)
sex	–0.028*** (0.004)	–0.031*** (0.004)	–0.027*** (0.005)	–0.032*** (0.005)	–0.030*** (0.005)	–0.030*** (0.004)	–0.034*** (0.004)	–0.033*** (0.004)	–0.030*** (0.004)	–0.033*** (0.005)	–0.032*** (0.004)
educ2	–0.038*** (0.007)	–0.035*** (0.007)	–0.041*** (0.006)	–0.038*** (0.006)	–0.035*** (0.006)	–0.035*** (0.006)	–0.033*** (0.006)	–0.031*** (0.006)	–0.041*** (0.006)	–0.038*** (0.006)	–0.033*** (0.006)
educ3	–0.111*** (0.008)	–0.105*** (0.008)	–0.113*** (0.008)	–0.107*** (0.008)	–0.105*** (0.008)	–0.106*** (0.008)	–0.100*** (0.008)	–0.100*** (0.008)	–0.113*** (0.008)	–0.108*** (0.008)	–0.101*** (0.008)
educ4	–0.152*** (0.006)	–0.144*** (0.006)	–0.156*** (0.006)	–0.148*** (0.006)	–0.143*** (0.006)	–0.146*** (0.006)	–0.137*** (0.006)	–0.135*** (0.006)	–0.157*** (0.006)	–0.149*** (0.006)	–0.138*** (0.006)
indoor_smoke	0.081*** (0.004)	0.078*** (0.004)	0.079*** (0.004)	0.077*** (0.004)	0.078*** (0.004)	0.080*** (0.004)	0.077*** (0.004)	0.078*** (0.004)	0.079*** (0.004)	0.077*** (0.004)	0.077*** (0.004)
QoL	–0.056*** (0.006)	——	–0.065*** (0.009)	——	——	–0.053*** (0.006)	——	——	–0.065*** (0.009)	——	——
BMI	–0.008*** (0.000)	——	–0.008*** (0.000)	——	——	–0.008*** (0.000)	——	——	–0.008*** (0.000)	——	——
municipality fixed–effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Robust standard errors are in the parentheses. '***', '**' and '*' indicates significance level at the 1%, 5% and 10% respectively.
≠ Instruments: cohabiting status in 1996; §Instruments: friends in 1996; € Instrument used is co_trust in 1996; ± Instruments used is how frequent individual has participated in music in HUNT 3.

Table A4: First–stage regressions results for alternative social capital variables and models estimated in table A3

Variables	SC variable *cohabit*		SC variable *friend*§			SC variable *co–trust*€			SC variable *part_org*±		
	Model 1	Model 2	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
cohabit_96≠	**0.395*** (0.005)**	**0.398*** (0.005)**	——	——	——	——	——	——	——	——	——
friend_96§	——	——	**0.211*** (0.014)**	**0.220*** (0.014)**	**0.219*** (0.014)**	——	——	——	——	——	——
co_trust_96 €	——	——	——	——	——	**0.254*** (0.006)**	**0.259*** (0.006)**	**0.240*** (0.006)**	——	——	——
part_music±	——	——	——	——	——	——	——	——	**0.474*** (0.021)**	**0.474*** (0.004)**	**0.455*** (0.004)**
cohabit	——	——	——	——	–0.040*** (0.003)	——	——	0.219*** (0.005)	——	——	–0.062*** (0.005)
age	–0.003*** (0.000)	–0.003*** (0.000)	–0.001*** (0.000)	–0.001*** (0.000)	–0.001*** (0.000)	0.001*** (0.000)	–0.001*** (0.000)	–0.002*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)
sex	–0.049*** (0.004)	–0.051*** (0.005)	–0.047*** (0.005)	–0.046*** (0.003)	–0.045*** (0.003)	–0.009* (0.004)	–0.008* (0.004)	–0.016*** (0.005)	–0.014*** (0.005)	–0.012*** (0.005)	–0.010** (0.005)
educ2	–0.036*** (0.007)	–0.036*** (0.006)	–0.003 (0.006)	–0.002 (0.006)	0.000 (0.004)	0.053*** (0.006)	0.054*** (0.006)	0.041*** (0.006)	0.073*** (0.006)	0.072*** (0.006)	0.076*** (0.006)
educ3	–0.025*** (0.008)	–0.026*** (0.009)	–0.002 (0.005)	0.003 (0.005)	0.006 (0.005)	0.073*** (0.010)	0.075*** (0.010)	0.067*** (0.010)	0.119*** (0.009)	0.117*** (0.009)	0.120*** (0.009)
educ4	–0.056*** (0.007)	–0.056*** (0.006)	–0.005 (0.004)	–0.003 (0.004)	0.000 (0.004)	0.104*** (0.007)	0.107*** (0.007)	0.087*** (0.007)	0.163*** (0.007)	0.160*** (0.007)	0.167*** (0.007)
indoor_smoke	0.014*** (0.005)	0.016*** (0.005)	–0.003 (0.003)	0.003 (0.004)	0.001 (0.003)	0.004 (0.005)	0.006 (0.005)	0.000 (0.005)	–0.018*** (0.005)	–0.016*** (0.005)	–0.015*** (0.005)
QoL	0.104*** (0.007)	——	0.090*** (0.009)	——	——	0.097*** (0.007)	——	——	0.039*** (0.006)	——	——
BMI	0.001 (0.001)	——	–0.000 (0.000)	——	——	–0.000 (0.000)	——	——	0.004*** (0.001)	——	——
municipality fixed–effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First–stage regression summary statistics	Adjusted (partial) R^2=0.164 (0.141)	Adjusted (partial) R^2= 0.158 (0.143)	Adjusted (partial) R^2=0.044 (0.017)	Adjusted (partial) R^2= 0.030 (0.019)	Adjusted (partial) R^2=0.036 (0.019)	Adjusted (partial) R^2= 0.081 (0.060)	Adjusted (partial) R^2=0.076 (0.062)	Adjusted (partial) R^2=0.120 (0.056)	Adjusted (partial) R^2=0.246 (0.228)	Adjusted (partial) R^2=0.244 (0.228)	Adjusted (partial) R^2= 0.247 (0.196)
	F =5198	F =5277	F =218	F =231	F =232	F =2066	F =2155	F = 1925	F =12090	F =11927	F =8908

Note: Robust standard errors are in the parentheses. '***', '**' and '*' indicates significance level at the 1%, 5% and 10% respectively.
≠ Instruments: cohabiting status in 1996; §Instruments: friends in 1996; € Instrument used is co_trust in 1996; ± Instruments used is how frequent individual has participated in music in HUNT 3.