DISCUSSION PAPER SERIES

No. 6780

TOBACCO AND ALCOHOL: COMPLEMENTS OR SUBSTITUTES? A STRUCTURAL MODEL APPROACH

Silja Göhlmann, Till Requate, Christoph M Schmidt and Harald Tauchmann

PUBLIC POLICY

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP6780.asp**

ISSN 0265-8003

TOBACCO AND ALCOHOL: COMPLEMENTS OR SUBSTITUTES? A STRUCTURAL MODEL APPROACH

Harald Tauchmann, RWI Essen
Silja Göhlmann, RWI Essen
Till Requate, University of Kiel
Christoph M Schmidt, RWI Essen, Ruhr-University Bochum and CEPR

Discussion Paper No. 6780
April 2008

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 6780

April 2008

ABSTRACT

Tobacco and Alcohol: Complements or Substitutes? – A Structural Model Approach*

The question of whether two drugs – namely alcohol and tobacco – are used as complements or substitutes is of crucial interest if side-effects of anti-smoking policies are considered. Numerous papers have empirically addressed this issue by estimating demand systems for alcohol and tobacco and subsequently calculating cross-price effects. However, this traditional approach often is seriously hampered by insufficient price-variation observed in survey data. We therefore suggest an alternative instrumental variables approach that statistically mimics an experimental study and does not rely on prices as explanatory variables. This approach is applied to German survey data. Our estimation results suggest that a reduction in tobacco consumption results in a moderate reduction in alcohol consumption. It is shown theoretically that this implies that alcohol and tobacco are complements. Hence, we conclude that successful anti-smoking policies will not result in the unintended side-effect of an increased (ab)use of alcohol.

JEL Classification: C31, D12 and I12
Keywords: complements or substitutes, instrumental variables approach, insufficient price-variation, interdependence in consumption and tobacco and alcohol

Silja Göhlmann
RWI Essen
Hohenzollernstrasse 1-3
D-45128 Essen
Germany
Email: silja.goehlmann@rwi-essen.de

Till Requate
Department of Economics
University of Kiel
University of Kiel
D-24118 Kiel
GERMANY
Email: r requate@bwl.uni-kiel.de

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=165311

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=154872

Christoph M Schmidt
RWI Essen
Hohenzollernstrasse 1-3
D-45128 Essen
Germany
Email: schmidt@rwi-essen.de

Harald Tauchmann
RWI Essen
Hohenzollernstraße 1-3
D-45128 Essen
GERMANY
Email:
harald.tauchmann@rwi-essen.de

* We would like to thank Boris Augurzky, Lars Siemers, and two anonymous referees as well as participants at the Scottish Economic Society Annual Conference 2006, the European Society for Population Economics 20th Annual Conference, and the Verein f¨ur Socialpolitik Annual Conference 2006 for many useful comments. Any errors are our own.

Submitted 19 March 2008

1 Introduction and Literature Review

Tobacco consumption has been identified as a major cause of health problems in industrialized countries. Consequently, smoking has been placed under severe restrictions. Even in Germany, one of last developed countries to consider such restrictions, smoking has recently been banned from public sector buildings and public transport, and a similar ban is discussed for bars and restaurants. But even if such policies successfully manage to reduce tobacco consumption, can they be viewed in isolation? In principle, restricting the access to tobacco may only encourage potential drug users to turn to other substances, notably the socially undisputed licit drug alcohol.

Using German micro data, this paper therefore asks whether tobacco and alcohol are complements or substitutes in consumption. If tobacco and alcohol were substitutes, an isolated policy measure aiming at a reduction of smoking would tend to increase the consumption of alcohol. If the two drugs were complements, though, a smoking ban could have a desirable side effect on the consumption of alcohol. Yet, answering this question is far from straightforward. The standard approach of estimating cross-price effects is typically precluded in micro data due to an insufficient variation of prices across observation units. To solve this problem, we suggest an alternative approach based on a structural model of quantities whose parameters are estimated via instrumental variables. Our results point at a complementary relation of tobacco and alcohol, thus promising positive side effects of smoking bans in the form of reduced alcohol consumption.

Our contribution adds to a growing literature on the joint consumption of legal drugs. The vast majority of econometric analyses addressing tobacco and alcohol is based on estimating demand functions and calculating cross-price effects from estimated price and income coefficients. Jones (1989), Florkowski and McNamara (1992), Goel and Morey (1995), Dee (1999) and Bask and Melkerson (2004) rely on aggregate data at regional or national level. Several other studies use survey data at the level of individual consumers; e.g. Jimenez and Labeaga (1994), Decker and Schwartz (2000), Cameron and Williams (2001), Zhao and Harris (2004), and Picone et al. (2004).[1] Since prices are generally not consumer-specific such analyses typically have to rely solely on price-variation across periods and/or across regions, and therefore quite regularly exhibit serious difficulties in disentangling genuine price effects from time or regional effects. Irrespective of the level of aggregation and the country considered, most of these studies find negative cross-price effects and therefore conclude that alcohol and tobacco are complements. As the only exception, Goel and Morey (1995) find positive and significant cross-price elasticities.

The remainder of paper is organized as follows. Section 2 discusses the econometric approach, section 3 introduces the data material, section 4 reports the empirical results, and section 5 derives conclusions for economic and health policy.

2 The Econometric Framework

2.1 A Structural Model of Complementarity

Our micro data comprise quantities consumed of tobacco and alcohol, and a range of individual-level background variables. In Germany, the prices of tobacco and alcohol do not display any remarkable variation over time and across regions, and no variation

[1]Chaloupka and Laixuthai (1997), DiNardo and Lemieux (2001), and Williams et al. (2001) address the interdependency of the consumption of alcohol and drugs others than tobacco, for instance, marijuana. Moreover, several related papers do not use prices as explanatory variables and are therefore concerned with correlation of drinking and smoking rather than interdependency, e.g. Su and Yen (2000), Lee and Abdel-Ghany (2004), and Yen (2005).

whatsoever at the level of individual consumers. In order to tackle this problem of insufficient price variation, our empirical analysis avoids relying on prices as explanatory variables. Instead of specifying a conventional demand system, the analysis is based on a structural, interdependent model of the consumption of both commodities. Using subscripts i and t to indicate individuals and survey periods, respectively, we express the demand for alcohol a_{it} as a linear function of the consumption of tobacco c_{it} and common explanatory variables x_{it} as well as of alcohol-specific variables z_{ait}. Correspondingly, there is a demand equation for tobacco which comprises the consumption of alcohol, the common explanatory variables, and some tobacco-specific variables z_{cit} as explanatory variables:

$$a_{it} = \gamma_a c_{it} + \beta_a' x_{it} + \delta_a' z_{ait} + \varepsilon_{ait} \tag{1}$$

$$c_{it} = \gamma_c a_{it} + \beta_c' x_{it} + \delta_c' z_{cit} + \varepsilon_{cit} \tag{2}$$

Here ε_{ait} and ε_{cit} represent random error terms while time and regional effects, including those due to temporal and regional price variation, are accounted for by including sets of dummy variables in the vector x_{it}. Similar structural models have been formulated by Dee (1999) and Bask and Melkerson (2004). However, in contrast to the analysis presented here those analyses still critically rely on price data that serve as instrumental variables and they ultimately aim at estimating cross-price effects.

In structural equation (1) the coefficient γ_a measures what would happen to the consumption of alcohol if the consumption of tobacco were exogenously reduced by one unit.[2] This interpretation analogously applies to γ_c. We use these coefficients as a measure of complementarity in consumption, since they exactly answer the relevant question pertaining to the possible side-effects of drug related regulation: "Imagine the

[2]If feedback-effects are taken into account, one might think of $(1-\gamma_a\gamma_c)^{-1}\gamma_a$ as the more appropriate measure. For model stability, the condition $1-\gamma_a\gamma_c > 0$ needs to be satisfied.

regulator could manage to reduce individual levels of smoking by a certain amount, how would this typically affect the consumption of alcohol?" Appendix A demonstrates that in qualitative terms our proposed measure is equivalent to cross-price derivatives of Hicksian demand functions, the standard measures for complementarity, since it necessarily exhibits the opposite sign.

2.2 Identification using Instrumental Variables

Our approach to estimating the parameters of demand equations (1) and (2) is based on the idea that if – as in a controlled experiment – the consumption-level of one drug could be varied exogenously, the effect of this variation on the consumption of the other drug could be measured directly. However, such experimental data is not available to us. For our empirical application we have to use survey data instead. Therefore, both a_{it} and c_{it} are themselves choice variables and estimates for γ_a and γ_c obtained from naively estimating (1) and (2) via OLS are severely biased. Nonetheless, the coefficients θ of the corresponding reduced-form representation

$$a_{it} = \theta'_{a1} x_{it} + \theta'_{a2} z_{ait} + \theta'_{a3} z_{cit} + \upsilon_{ait} \tag{3}$$

$$c_{it} = \theta'_{c1} x_{it} + \theta'_{c3} z_{ait} + \theta'_{c2} z_{cit} + \upsilon_{cit} \tag{4}$$

can be estimated consistently by OLS. The structural-form coefficients directly translate into reduced-form coefficients as follows:

$$\theta_{a1} \equiv \frac{\gamma_a \beta_c + \beta_a}{1-\gamma_a\gamma_c}, \qquad \theta_{a2} \equiv \frac{\delta_a}{1-\gamma_a\gamma_c}, \qquad \theta_{a3} \equiv \frac{\gamma_a \delta_c}{1-\gamma_a\gamma_c}, \qquad \upsilon_{ait} \equiv \frac{\gamma_a \varepsilon_{cit} + \varepsilon_{ait}}{1-\gamma_a\gamma_c}.$$

The terms for θ_{c1}, θ_{c2}, θ_{c3}, and υ_{cit} are defined analogously.

If z_{ait} and z_{cit} were empty, that is, if we had no instruments for alcohol and tobacco consumption respectively, estimates for θ would be of no value to our principal research question. However, with valid instruments z_{ait} and z_{cit} in hand one can calculate any structural coefficients including γ from estimates for θ, since $\gamma_a = \frac{\theta_{a3k}}{\theta_{c2k}}$ and $\gamma_c = \frac{\theta_{c3k}}{\theta_{a2k}}$ hold.[3] As a more efficient alternative, one can employ the classical two-stage least squares estimator. Evidently, this two-step approach still relies on valid instruments. That is, to estimate the coefficients of the demand equation for alcohol (1) consistently, we need to find variables which affect the consumption of tobacco, but do not affect the consumption of alcohol through any other channel than through tobacco consumption. Similarly, to estimate equation (2) consistently, we need to search for variables affecting alcohol consumption directly, and yet tobacco consumption only indirectly via the consumption of alcohol. In the quest for such instrumental variables we might succeed for one equation and fail for the other.

Indeed careful reasoning suggests that our data comprises variables which can be regarded as valid instruments both for our principal equation of interest, the demand for alcohol. Our reasoning exploits the close link between parental drinking and children's later consumption patterns.[4] For instance, Bantle and Haisken-DeNew (2002) find significant correlations between parental smoking behavior and children's tobacco consumption for Germany. In order to use parental consumption habits as instruments, we argue that the link is only direct for the same substance. Specifically, we presume that parents' smoking habits do influence children's later tobacco consumption, but conditional on children's later smoking behavior (and other observables), they will not have any effect on their drinking habits. Even though parents' tobacco consumption and children's later alcohol use might be correlated, the correlation purely operates through children's own smoking habits (and other observables).

[3]The subscript k indicates the kth element of the corresponding vector. I.e. if the vectors z_a and z_c consist of more than one element, several different estimates for γ_a and γ_c can be calculated.

[4]In principle, the reasoning for the second structural equation (2) is symmetric, although one cannot expect that the argument necessarily holds equally successfully. However, failure to identify equation (2) does not invalidate the approach to equation (1).

When estimating our coefficient of primary interest γ_a via instrumental variables, we have to acknowledge that this coefficient is not necessarily the same for all individuals. Rather, what we can identify if γ_a is heterogenous – given the validity of our exclusion restrictions – are (local) average treatment effects (Imbens and Angrist, 1994). That is, we estimate average patterns of complementarity for those respondents who would exhibit substantially different smoking and drinking habits if their parents had shown different behavior as well (Angrist and Krueger, 2001). One restriction allowing to generalize the estimated pattern of complementarity to the entire population is to assume homogenous effects. In the remainder of this paper, we focus on the interpretation of γ_a in terms of such a structural model parameter, keeping the less restrictive interpretation as a local average effect in mind.

2.3 Testing for Over-Identifying Restrictions

The validity of our exclusion restrictions is decisive for our empirical analysis. As a minimum we need to justify them, equation by equation, by a priori reasoning. It does not seem implausible that parental smoking behavior might arguably be irrelevant for own drinking habits, given own smoking behavior and parental drinking habits.

Fortunately, with respect to our identifying assumptions we do not have to rely on intuition alone but we have the opportunity of testing them since the vectors z_{ait} and z_{cit} each consist of more than one element, namely the consumption habits of both mothers and fathers.[5] Hence, the structural coefficients γ_a and γ_c are over-identified and one can apply tests for over-identifying restrictions. We apply three different test procedures.

The first approach represents an intuitive quasi-test for the validity of over-identifying restrictions. As pointed out, estimating the reduced form model allows for calculating

[5] In addition, consumption habits of mothers and fathers (expressed in different consumption levels) are parameterized as sets of dummy variables not as two single variables.

as many estimates for γ_a and γ_c as instruments are available. If all instruments are valid, one should expect that these different estimates only differ because of sampling error. Secondly, we employ a regression-based, heteroscedasticity-robust variant of the Hausman test, cf. WOOLDRIDGE (2002). The test procedure is based on the idea that a regression of two-stage least squares residuals on all exogenous variables should not exhibit any explanatory power, given that the over-identifying restrictions are valid.[6] Finally, we employ a third test for over-identifying restrictions (NEWEY AND MCFADDEN, 1994) that rests on the idea that under the null-hypothesis of valid over-identifying restrictions the structural and the reduced form are fully equivalent representations of the same model. Therefore, predicted values obtained from either estimating the reduced or the structural form should deviate only because of sampling error. Conventional χ^2-tests allow for testing the underlying null-hypotheses.

2.4 The Econometric Specification

If only strictly positive values for the consumption of alcohol and tobacco were observed, estimation by standard linear two-stage least squares would be straightforward. Yet, many individuals do not drink or smoke at all and the consumption patterns of both alcohol and tobacco are therefore characterized in our data by large shares of corner solutions. To account for this in the econometric analysis, we suggest two different approaches.

First, we reformulate the equations (1) through (4) in terms of latent demand a_{it}^* and c_{it}^* instead of actual consumption a_{it} and c_{it}. One may interpret latent demand as the inclination to consume. It might well fall below zero if an individual dislikes tobacco or alcohol. Since negative latent demand is reflected by zero actual consumption, the dependent variables are censored at zero. Under the assumption of normally

[6]Since this procedure requires calculating regression residuals, it is not applicable to the Tobit specification of the model, see section 2.4.

distributed errors this leads to an interdependent Tobit model. Several estimators have been proposed for interdependent Tobit models and – more generally – for Tobit models with endogenous regressors. For the former Maddala (1983) and Nelson and Olsen (1978) discuss several variants. For the latter Wooldridge (2002) proposes an efficient full-information maximum-likelihood (FIML) approach, while Newey (1986) and Smith and Blundell (1986) suggest two-step approaches.[7]

The results reported in this paper are based on a particularly simple instrumental-variables procedure proposed by Greene (2002) that directly mirrors two-step least squares in the linear case, i.e. the reduced form equations (3) and (4) are individually estimated by standard Tobit procedures. From the estimates for θ fitted values $\widehat{a}_{it}^*$ and $\widehat{c}_{it}^*$ are calculated. These serve as regressors in the structural equations, which once again are estimated by conventional Tobit procedures. It is important to note that this simple approach does not allow for identifying $\text{var}(\varepsilon_a)$ and $\text{var}(\varepsilon_c)$.[8] Thus, marginal effects on expected actual consumption are not identified either. Yet, since we are primarily interested in the signs of the parameters γ_a and γ_c this limitation is of marginal relevance.

The Tobit approach to the model is quite restrictive, though. Besides normality, it assumes that the discrete outcome whether an individual drinks or does not drink at all is determined by exactly the same mechanism that determines the amount of alcohol consumed conditional on drinking. Analogous restrictions are imposed for the smoking equation. Moreover, the simple two-step Tobit estimator relies on the assumption that latent demand a_{it}^* and c_{it}^* rather than actual consumption a_{it} and c_{it} enters the right-hand side of the structural equations, although one might argue that actual consumption better corresponds with our experiment-like strategy for identification.

[7]Hard-coded procedures for the FIML as well as the two-step approach are currently available in econometric software packages like Stata; see e.g. Winter-Ebmer (2006) for a recent application.

[8]Cf. Rivers and Vuong (1988) for an equivalent problem in the case of the Probit model.

In order to relax these restrictive assumptions, we alternatively estimate the equations (1) and (2) as conditional on $a_{it} > 0$ and $c_{it} > 0$.[9] That is, the model explains the interdependence in consumption for those individuals that do both smoke and drink. Yet, whether an individual actually is a smoker or a drinker is determined by a separate upstream model, which may be specified as e.g. a bivariate Probit model. Under the assumption of joint normality the resulting joint model represents the bivariate generalization of Poirier (1980) to the Heckman (1976) sample selection (Heckit) model.

In the Heckit model identification critically relies on valid exclusion restrictions. If however the same set of explanatory variables enters the discrete and the continuous model, identification solely rests on non-linearity. In this case, a simple two-part model that ignores error-correlation across both model components and estimates them separately may clearly outperform the Heckit model in terms of the mean squared error (Dow and Norton, 2003). In our data we cannot identify any variable that may legitimately be excluded only from the continuous model. For this reason we prefer the two-part specification to the Heckit model,[10] though it is hard to defend consistency of this estimator unless one assumes uncorrelated errors (Jones, 2000). In essence, we apply the standard linear two-stage least squares approach simply excluding non-smokers and non-drinkers from the sample. In the following we refer to this specification as the conditional linear model.

All previously suggested econometric specifications use two-step procedures for estimating the structural model equations. This requires some caution in calculating valid standard errors. Either an appropriate correction procedure, cf. Murphy and

[9] One might think of conditioning equation (1) only on $a_{it} > 0$ and equation (2) only on $c_{it} > 0$. However, if we chose different conditions for both structural equations, for determining the reduced-form representation a differentiation of cases were required. Therefore, equation-wise conditioning on either $a_{it} > 0$ or $c_{it} > 0$ is not consistent with the basic structure of the model.

[10] Nonetheless, we also estimated the generalized Heckit model. Yet, the Dow and Norton (2003) MSE-criterion argues in favor of the two-part specification.

Topel (1985), is required or bootstrapping, which encompasses both stages of the estimation procedure. We choose the latter strategy and report bootstrapped standard errors for the structural model parameters.

3 The Data

3.1 Data Sources

This analysis uses data from the "Population Survey on the Consumption of Psychoactive Substances in Germany"[11] collected by IFT[12] Munich; see Kraus and Augustin (2001) for a detailed description. The data originally comprises eight separate cross sections at the level of individual consumers, collected by mail at irregular intervals in the years 1980, 1986, 1990, 1992, 1995, 1997, 2000, and 2003. The sample size varies significantly from 4,455 in 1992 to 21,632 in 1990. While the first two surveys concentrate solely on West Germany, the 1992 survey exclusively deals with the former East German GDR. All other waves cover Germany as a whole. Until 1992 only German citizens were interviewed, immigrants not holding the German citizenship were disregarded. Later on, the complete German speaking population was included in the survey, irrespective of citizenship. The data provides comprehensive information with respect to various legal as well as illicit drugs regarding prevalence, frequency and intensity of consumption, consumption habits and age at first use. Additionally, detailed information on socioeconomic characteristics is provided along with information on attitudes towards several drug-related issues.

Unfortunately, both the questionnaire and the study's target population have changed over time. The first wave focuses on teens and young adults aged 12 to 24. In

[11]Bundesstudie "Repräsentativerhebung zum Gebrauch psychoaktiver Substanzen in Deutschland"

[12]Institute for Therapy Research (Institut für Therapieforschung)

subsequent waves the upper age limit was successively raised up to 39 in 1990. Since 1995 the target population solely consists of adults aged 18 to 59. As a consequence, consumers' family background increasingly became a minor issue and therefore smoking as well as drinking habits at the parental home are not reported in waves after 1992. The recent waves therefore lack those instrumental variables that are decisive for our econometric model and, consequently, our analysis has to rely on data collected in 1980, 1986, 1990, and 1992. We also do not consider individuals younger than 16 years for estimating the model. Though numerous people from this age group do report having consumed alcohol or tobacco, this often may reflect experimenting rather than already settled consumption patterns. After excluding observations with missing data the sample consists of 26,516 individuals. Among these, 18,711 individuals drink and 11,272 smoke while 8,675 individuals consume both drugs;[13] cf. Table 1.

3.2 Variables

In our analysis, alcohol consumption is defined as grams of alcohol intake per day which is calculated from the reported glasses of beer, wine and spirits consumed per week.[14] The quantity of tobacco consumed is measured by the average number of smoked cigarettes per day. The variable takes the value zero if the individual answers to be an ex- or never smoker. Numerous consumers do report to be drinkers or smokers but do not report the amount of alcohol or nicotine consumed. In our sample, this applies to 20 percent of all drinkers and to 17 percent of all smokers. In the Tobit specification we do not exclude these observation from our analysis but let the probability to either drink or

[13]In the Tobit model only 25,695 observations are used for estimating the equations explaining alcohol consumption and only 26,353 are used for estimating the equations explaining tobacco consumption, because of missing information on the respective dependent variable. For the conditional linear model the corresponding figures are 6,819 and 6,940 observations. In the latter specification the sample size is considerably reduced because of conditioning on $a_{it} > 0$ and $c_{it} > 0$.

[14]We use standard values for beverages' alcohol content: one glass of beer (0.3l) contains 12 grams of alcohol, one glass of wine (0.25l) 20 grams, and one glass of spirits (0.02l) 5.6 grams.

Table 1: **Descriptive statistics for key variables [in percent]**

	All	**Males**	**Females**
Smoking and Drinking			
drinker	72.9	83.8	62.0
smoker	42.8	46.2	39.4
drinker as well as smoker	34.0	41.2	26.7
Father's drinking habits			
father never drinker	18.6	18.0	19.3
father monthly drinker	23.6	23.6	23.5
father weekly drinker	25.9	26.8	25.0
father daily drinker	31.9	31.6	32.3
Mother's drinking habits			
mother never drinker	50.7	49.9	51.4
mother monthly drinker	28.5	29.2	28.0
mother weekly drinker	13.6	14.1	13.0
mother daily drinker	7.2	6.8	7.6
Father's smoking habits			
father never-smoker	27.6	27.5	27.6
father ex-smoker	35.1	35.7	34.5
father smoker	37.3	36.8	37.9
Mother's smoking habits			
mother never-smoker	67.4	67.6	67.3
mother ex-smoker	12.6	12.8	12.4
mother smoker	20.0	19.7	20.3

Note: See Tables 5 and 6 for a more detailed description.

smoke enter the likelihood function.[15] In contrast, estimation of the conditional linear specification has solely to rely on individuals that report quantities of consumption.

In our empirical analysis, we control for gender, age, age squared, and living in West-Germany. Moreover, the vector x_{it} includes parental education, parental marital status, number of children at parents' home as well as the way individuals have grown up, reflecting the social background of the family. By interacting parental education with dummy variables indicating having grown up with the parent we allow parental education to have an effect only if the respondent has grown up with the parent.

[15]For the univariate Tobit model this can quite easily be implemented by recoding consumers with no information about quantitative consumption as non-consumers and multiplying the explanatory variables by minus one.

Variables often controlled for by other authors – e.g. Chaloupka and Laixuthai (1997), Yen (2005) – like own education, marital and labor market status, number of children, current living situation as well as income are deliberately not used as explanatory variables because of their potential endogeneity. Nevertheless, despite our reservations, we also experimented with including these variables in additional specifications but it turned out that this does not change our main findings.

Most importantly, parental smoking and drinking habits serve as instruments z_{cit} and z_{ait}. Individuals who already have moved out from parental home are retrospectively asked about these variables. For our regression analysis, each parent's smoking behavior is characterized by three categories: (i) smoker, (ii) ex- or (iii) never-smoker, with the latter serving as the reference group. With regard to parents' drinking habits for each parent four categories are distinguished: parent drinks (i) (almost) daily, (ii) several times a week, (iii) several times a month, and (iv) (almost) never. Again, the last category is chosen as reference group. We interact parental consumption habits with the indicator for having grown up with this parent in order to make sure that only those parental habits enter the analysis that could have influenced children's consumption behavior. Table 1 provides the key descriptive statistics. See Tables 5 and 6 in Appendix B for the complete set of descriptive statistics.

4 Estimation Results

Naively estimating equations (1) and (2) by the Tobit or the conditional linear model, ignoring the endogeneity of the right hand side variables c_{it} and a_{it}, respectively, indicates a strong positive correlation between the consumption of both tobacco and alcohol. The estimates for γ_a as well as for γ_c are highly significant and positive. However, these results are certainly biased and do not tell us much about the structural interdependence of the consumption of both drugs. Thus, we now turn to the reduced-form results and to estimating the structural form via instrumental variables.

4.1 Reduced Form Results

The corresponding results for the reduced form equations (3) and (4) are presented in Table 2. In qualitative terms, the main result is that the chosen instruments are highly correlated with the corresponding endogenous variables c_{it} and a_{it}. Thus, the parents' drinking habits exert a significant effect on the drinking behavior of their children and this holds for smoking behavior as well. The inclination to drink increases with the intensity of parental alcohol consumption and the propensity to smoke increases with the intensity of parental smoking. The relevance of these direct links is confirmed by formal tests on instrument relevance (Davis and Kim, 2002)[16] and by tests of joint significance of instruments as well.[17] This holds for both the Tobit and the conditional linear model.

Furthermore, the Tobit results also exhibit distinct "cross-correlations" between parental smoking habits to individuals' drinking habits and vice versa. Remarkably, while the correlation between the propensity to drink and parental smoking behavior, i.e. our cross-relation of primary interest, is positive, we find a significantly negative correlation between the propensity to smoke and parental drinking habits. This correlation raises some doubts whether our identifying assumptions do symmetrically hold for the second equation in our demand system.

With regard to our control variables the Tobit results for the reduced forms exhibit a trend of a decreasing inclination to smoke and drink over time as well as a lower

[16]For the Tobit model the $\chi^2(1)$-statistic of the relevant LR-test takes a value of 716.6 concerning parents' drinking habits and 2,484.9 concerning parents' smoking habits. For the conditional linear model the corresponding values are 121.9 and 122.6. For the former Shea-Partial-R^2's are calculated using Tobit pseudo residuals.

[17]For the Tobit model the F-statistics for the test on joint significance is as high as 272.6 (smoking) and 104.9 (drinking). For the conditional linear model the corresponding values are 30.8 and 20.4.

Table 2: **Results for the reduced form**

	Tobit Model				Conditional Linear Model			
	Drinking		Smoking		Drinking		Smoking	
Variable	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.
father monthly drinker	1.796**	0.363	-1.542**	0.436	0.101	0.769	-0.334	0.336
father weekly drinker	3.230**	0.365	-1.507**	0.439	2.515**	0.773	-0.055	0.339
father daily drinker	4.146**	0.357	-1.154**	0.426	3.916**	0.751	0.541	0.325
mother monthly drinker	2.903**	0.268	-1.199**	0.329	1.790**	0.550	-0.450	0.243
mother weekly drinker	4.686**	0.355	-0.580	0.437	3.434**	0.722	-0.513	0.324
mother daily drinker	4.557**	0.464	-0.569	0.569	4.185**	0.931	0.180	0.421
father ex-smoker	0.711*	0.288	3.593**	0.366	0.311	0.642	0.422	0.294
father smoker	0.821**	0.297	6.995**	0.369	-0.297	0.639	1.226**	0.288
mother ex-smoker	0.185	0.341	2.997**	0.415	0.360	0.695	0.814**	0.312
mother smoker	0.497	0.296	7.464**	0.345	0.608	0.558	2.136**	0.241
constant	-8.246**	1.978	-35.171**	2.379	2.886	4.662	-9.009**	1.841
year 1986	-2.826**	0.385	-3.893**	0.467	0.595	0.762	-0.267	0.342
year 1990	-3.075**	0.313	-4.983**	0.373	0.596	0.615	-0.778**	0.273
year 1992	-7.622**	0.776	-5.667**	0.768	8.382**	2.516	-0.820	0.564
west	-7.738**	0.654	0.119	0.575	5.890*	2.365	2.940**	0.425
female	-11.231**	0.211	-4.449**	0.257	-11.316**	0.448	-2.787**	0.198
age	2.200**	0.143	3.271**	0.176	0.742*	0.308	1.728**	0.137
$age^2/100$	-3.120**	0.267	-5.536**	0.327	-0.569	0.574	-2.568**	0.253
parents married	-0.215	0.324	-2.086**	0.382	-0.715	0.645	-1.108**	0.276
father has low degree	0.766*	0.364	-0.425	0.436	-0.044	0.747	-0.264	0.326
father has medium degree	0.489	0.455	-0.769	0.553	-0.341	0.932	-0.923*	0.415
father has high degree	0.025	0.736	-1.908*	0.911	-0.021	1.567	-1.557*	0.705
father has univ. degree	0.958	0.503	-2.450**	0.614	-0.081	1.067	-1.209*	0.470
mother has low degree	-0.141	0.280	-0.520	0.338	-0.979	0.571	-0.534*	0.251
mother has medium degree	-0.002	0.359	-0.517	0.439	0.279	0.741	0.010	0.335
mother has high degree	-0.980	0.765	-2.852**	0.968	0.635	1.638	0.391	0.735
mother has univ. degree	-0.094	0.584	-2.121**	0.719	1.576	1.283	-0.860	0.559
grown up with mother	-2.024**	0.716	-4.596**	0.826	-2.947*	1.385	-1.716**	0.601
grown up with father	-4.580**	1.230	-5.387**	1.508	-5.488*	2.729	-1.512	1.115
grown up with both	1.181	1.326	0.873	1.532	4.141	2.735	1.256	1.120
no. children at parents' home	0.104	0.075	0.648**	0.089	0.360*	0.155	0.190**	0.066
number of observations	25,695		26,353		6,819		6,940	
LR-statistic	4,364.43		2,419.64		–		–	
F-statistic	–		–		31.31		34.34	

Note: ** significant at the 1%-level; * significant at the 5% level.

propensity to consume tobacco and alcohol for women compared to men. We also find a significantly positive (but diminishing) correlation with age. Moreover, results indicate a significantly negative correlation of the propensity to drink or smoke with having grown up with at least one parent compared to individuals having grown up with other persons. We further find that parental education has a significantly negative effect on the propensity to smoke. The number of children at parents' home as well as the parental marital status are significant only for the inclination to smoke. Except for the time effect on drinking and those for living in western Germany, the conditional linear model exhibits similar patterns of estimated coefficients. Yet, fewer coefficients are significant.

4.2 Structural Model Results

Table 3 reports the results for the structural equations (1) and (2). For the control variables, the structural estimates by and large confirm the reduced form estimates. Our discussion can therefore concentrate on the parameters of primary interest, γ_a and γ_c. Regarding the effect of smoking on alcohol consumption γ_a, the estimate from the Tobit model is clearly significant while that from the conditional linear model is insignificant, albeit of the same sign and order of magnitude. The lack of significance might most likely be explained by the rather small subsample that is used for estimating the conditional linear model. According to the Tobit results, the estimates exhibit that smoking significantly increases the propensity to drink. Thus smoking and drinking are classified as complements in consumption. By contrast, the Tobit equation for smoking behavior suggests that drinking significantly decreases the propensity to smoke, which would indicate that drinking and smoking are substitutes.

We do now that the true parameters γ_a and γ_c need to bear the same sign, opposite to the sign of the Hicksian cross-price derivatives, which are necessarily symmetric.

Table 3: **Results for the structural model**

		Tobit Model				Conditional Linear Model			
		Drinking		**Smoking**		**Drinking**		**Smoking**	
Parameter / Exp. Variable		Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.
γ_a	fitted smoking	0.089**	0.024	–	–	0.107	0.209	–	–
γ_c	fitted drinking	–	–	-0.224**	0.049	–	–	0.057	0.042
δ_a	father monthly drinker	1.947**	0.299	–	–	0.126	0.706	–	–
	father weekly drinker	3.399**	0.323	–	–	2.493**	0.734	–	–
	father daily drinker	4.293**	0.371	–	–	3.789**	0.712	–	–
	mother monthly drinker	2.994**	0.276	–	–	1.860**	0.521	–	–
	mother weekly drinker	4.707**	0.404	–	–	3.553**	0.698	–	–
	mother daily drinker	4.587**	0.513	–	–	4.248**	1.036	–	–
δ_c	father ex-smoker	–	–	3.775**	0.440	–	–	0.461	0.267
	father smoker	–	–	7.222**	0.408	–	–	1.348**	0.267
	mother ex-smoker	–	–	3.063**	0.420	–	–	0.778**	0.298
	mother smoker	–	–	7.604**	0.353	–	–	2.064**	0.248
β	constant	-5.234**	1.947	-37.053**	2.387	3.890	4.450	-9.234**	1.883
	year 1986	-2.491**	0.358	-4.515**	0.482	0.636	0.753	-0.306	0.305
	year 1990	-2.655**	0.340	-5.649**	0.453	0.722	0.612	-0.774**	0.251
	year 1992	-7.136**	0.590	-7.366**	0.892	8.524**	1.962	-1.240*	0.578
	west	-7.751**	0.478	-1.618*	0.753	5.628**	1.837	2.572**	0.434
	female	-10.841**	0.253	-6.957**	0.639	-11.030**	0.668	-2.152**	0.486
	age	1.915**	0.155	3.763**	0.202	0.548	0.458	1.688**	0.143
	age^2/100	-2.631**	0.281	-6.230**	0.361	-0.280	0.764	-2.533**	0.264
	parents married	0.016	0.271	-2.171**	0.388	-0.520	0.698	-1.107**	0.289
	father has low degree	0.808*	0.324	-0.325	0.463	-0.001	0.802	-0.311	0.353
	father has medium degree	0.559	0.379	-0.758	0.647	-0.183	0.964	-1.003*	0.423
	father has high degree	0.191	0.601	-1.951	1.001	0.207	1.608	-1.593*	0.734
	father has univ. degree	1.176**	0.432	-2.296**	0.664	0.126	1.121	-1.330*	0.517
	mother has low degree	-0.102	0.270	-0.575	0.333	-0.906	0.588	-0.509*	0.243
	mother has medium degree	0.039	0.341	-0.547	0.491	0.292	0.791	-0.074	0.331
	mother has high degree	-0.732	0.860	-3.028**	0.957	0.671	1.646	0.321	0.808
	mother has univ. degree	0.085	0.560	-2.144*	0.884	1.687	1.411	-0.968	0.556
	grown up with mother	-1.653*	0.674	-5.111**	0.854	-2.584	1.666	-1.763**	0.635
	grown up with father	-3.898**	1.232	-6.899**	1.537	-5.451*	2.626	-1.286	1.113
	grown up with both	1.078	1.150	1.128	1.515	3.930	2.734	1.080	1.157
	no. children at parents' home	0.045	0.084	0.681**	0.093	0.329*	0.164	0.182**	0.070
number of observations		25,695		26,353		6,819		6,940	
LR-statistic		4,362.32		2,400.04		–		–	
F-statistic		–		–		34.72		40.50	
tests for over-identifying restrictions (p-values):									
intuitive		0.555		0.012		0.874		0.242	
Hausman		–		–		0.449		0.009	
Newey & McFadden		0.489		0.000		0.647		0.023	

Notes: ** significant at the the 1% level; * significant at the the 5% level; bootstrapped standard errors reported.

Thus this asymmetry in estimation results reveals that our identifying assumptions do not apply to both of our equations. In order to gain more insights, we turn to the tests on over-identifying restrictions. According to these tests, the exclusion restrictions are warranted in the equation for alcohol consumption (1). In contrast, for the smoking equation (2) all over-identification tests but one presented in Table 3 clearly reject the null-hypothesis of parental drinking habits having no direct effects on children's later tobacco consumption. Drinking at the parental home accordingly seems to affect children's future lives in a more general way than parental smoking habits. This is quite plausible in the case of severe alcohol abuse that is likely to damage family life in general and therefore might affect children through various channels. Excessive smoking – though harmful to health – is not likely to have comparable effects. Yet, the asymmetry may even apply to moderate consumption. Unlike smoking, drinking often is a social activity and possibly even a reflection of competence in the controlled consumption of psychoactive substances. Thus, we can be confident that parental smoking behavior constitutes a valid instrument in the equation explaining alcohol consumption. Drinking and smoking seem to be complements in consumption.

While this result confirms the main body of previous literature, the insignificant estimate for γ_a obtained from the conditional linear specification neither confirms nor contradicts this result. We regard the substantially smaller sample size as the most likely explanation for this insignificance. Correspondingly, our conclusions are based on the Tobit model. In quantitative terms, our preferred estimate for γ_a indicates that one cigarette less per day results in roughly the consumption of a tenth of a gram alcohol less per day.[18] This represents merely one-hundredth of a half-pint. Thus, the reduction of drinking levels that will result from successful anti-smoking policies

[18]The implicit assumption that the size of this effect does not depend on consumption levels is quite strong. We therefore estimate extensions to the basic model that allow for consumption-dependent effects. Yet, due to the lack of additional valid instruments, the corresponding estimation results do not provide further insights.

is likely to be rather moderate. But still, the estimate for γ_a clearly argues against an unintended side-effect. That is, effective anti-smoking policies will not result in an increase of the consumption of alcohol, but rather tend to improve population health on several margins simultaneously.

4.3 Separate Models for Males and Females

Our analysis reveals pronounced gender-effects on the consumption of tobacco as well as the consumption of alcohol, see Tables 2 and 3. In order to analyze whether gender does not only matter for the level of consumption but also for the interdependence in consumption, the model is estimated separately for males and females. Table 4 displays our preferred estimates for the structural coefficients γ; see Tables 7 to 10 in Appendix B for a comprehensive list of estimation results.

For both the Tobit as well as the conditional linear specification, results of corresponding LR-tests argue in favor of separate models for males and females. Yet, in qualitative terms the results are similar to those obtained from the pooled model for men and women. In the case of the conditional linear model the estimated coefficients γ are insignificant for either gender, confirming the results for the pooled model. Tests for over-identification do not reject the validity of instruments in either demand equations. Yet, the p-values are still considerable smaller for the smoking equation. Apparently, the power of these tests is considerably reduced by the smaller sample size.

Regarding the gender-specific variants of the Tobit specification for smoking, for both men and women $\hat{\gamma}_c$ is negative, yet – as in the pooled model – over-identification tests reject the identifying assumptions. In contrast, $\hat{\gamma}_a$ takes positive values for both genders and our identification strategy is supported by the relevant test-statistics. The main differences to the results from Table 3 are quantitative in nature. While for

Table 4: **Results for separate models for males and females**

	Tobit Model				Conditional Linear Model			
	Drinking		**Smoking**		**Drinking**		**Smoking**	
Parameter	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.
Males								
γ_a	0.135**	0.041	–	–	0.115	0.314	–	–
γ_c	–	–	-0.159**	0.055	–	–	0.056	0.042
number of obs.	12,922		13,064		4,212		4,313	
tests for over-identifying restrictions (p-values):								
intuitive	0.383		0.080		0.434		0.408	
Hausman	–		–		0.410		0.122	
Newey & McFadden	0.379		0.000		0.297		0.214	
Females								
γ_a	0.037	0.025	–	–	-0.008	0.205	–	–
γ_c	–	–	-0.337**	0.102	–	–	0.037	0.144
number of obs.	12,773		13,291		2,607		2,627	
tests for over-identifying restrictions (p-values):								
intuitive	0.936		0.207		0.938		0.989	
Hausman	–		–		0.638		0.225	
Newey & McFadden	0.920		0.061		0.676		0.263	

Notes: ** significant at the the 1% level; * significant at the the 5% level; bootstrapped standard errors reported.

males $\hat{\gamma}_a$ is of a substantially larger magnitude than in the pooled model, the parameter takes a much smaller value for females and even becomes insignificant. Therefore, the complementarity between smoking and drinking seems to be a predominately male phenomenon.

5 Conclusions

This paper proposes a new approach for analyzing the interdependence in the consumption of alcohol and tobacco and applies this idea to German survey data. We use an alternative measure of complementarity which – in qualitative terms – is shown to be equivalent to conventional Hicksian cross-price derivatives, yet it is not based on the estimation of cross-price effects. In fact, the proposed instrumental variable approach mimics an experimental study and therefore does not rely on high-quality price data which often may not be available. This makes it particularly well-suited to the German case where price variation for both goods is extremely limited. Moreover, the lack of price variation is a frequent obstacle to survey data-based analyses of consumer behavior irrespective of the specific goods under scrutiny. Instrumental variables approaches, similar to the one proposed here, might therefore serve as a promising modeling strategy for gathering evidence on interdependencies in consumption.

Our estimation results suggest that tobacco and alcohol are consumed as complements. This result rests on a positive effect from the consumption of tobacco to the consumption of alcohol that is found in the data using a Tobit specification for estimation. Less restrictive specifications of the model neither confirm nor challenge this finding but suffer from smaller number of observations. From a policy perspective, complementarity can be interpreted as follows: if the government could achieve a reduction in smoking or in the inclination to smoke by any anti-drug policy, this would also decrease the propensity to consume alcohol. Thus, there would be no unintended side-effects in form of an increased (ab)use of alcohol to compensate for the reduced level of nicotine intake. Even the reverse, i.e. a moderate reduction in the consumption of alcohol, seems to be the consequence. Yet, this result seems only to be relevant for males.

References

Angrist, J. D. and Krueger, A. B. (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments, *Journal of Economic Perspectives* **15**: 69–85.

Bantle, C. and Haisken-DeNew, J. P. (2002). Smoke signals: The intergenerational transmission of smoking behavior, *DIW Discussion Paper* **227**.

Bask, M. and Melkerson, M. (2004). Rationally addicted to drinking and smoking?, *Applied Economics* **36**: 373–381.

Cameron, L. and Williams, J. (2001). Cannabis, alcohol and cigarettes: Substitutes or complements?, *The Economic Record* **77**: 19–34.

Chaloupka, F. J. and Laixuthai, A. (1997). Do youths substitute alcohol and marijuana? Some econometric evidence, *Eastern Economic Journal* **23**: 253–276.

Davis, G. C. and Kim, S. Y. (2002). Measuring instrument relevance in the single endogenous regressor - multiple instrument case: a simplifying procedure, *Economics Letters* **74**: 321–325.

Decker, S. L. and Schwartz, A. E. (2000). Cigarettes and alcohol: Substitutes or complements, *NBER Working Paper Series* **7535**.

Dee, T. S. (1999). The complementarity of teen smoking and drinking, *Journal of Health Economics* **18**: 769–793.

DiNardo, J. and Lemieux, T. (2001). Alcohol, marijuana, and American youth: the unintended effects of government regulation, *Journal of Health Economics* **20**: 991–1010.

Dow, W. H. and Norton, E. C. (2003). Choosing between and interpreting the heckit and two-part models for corner solutions, *Health Services and Outcomes Research Methodology* **4**: 5–18.

Florkowski, W. J. and McNamara, K. T. (1992). Policy implications of alcohol and tobacco demand in Poland, *Journal of Policy Modelling* **14**: 93–99.

Goel, R. K. and Morey, M. J. (1995). The interdependence of cigarette and liquor demand, *Southern Economics Journal* **62**: 451–459.

Greene, W. H. (2002). *LIMDEP 8.0 - Econometric Modeling Guide, Vol.2*, Plainview, New York.

Heckman, J. J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models, *Annals of Economics and Social Measurement* **5**: 475–492.

Imbens, G. W. and Angrist, J. D. (1994). Identification and estimation of local average treatment effects, *Econometrica* **62**: 467–476.

Jimenez, S. and Labeaga, J. M. (1994). It is possible to reduce tobacco consumption via alcohol taxation, *Health Economics* **3**: 231–241.

Jones, A. M. (1989). A system approach to the demand for alcohol and tobacco, *Bulletin of Economic Research* **41**: 85–106.

Jones, A. M. (2000). Health econometrics, *in* A. Culyer and J. Newhouse (eds), *Handbook of Health Economics, Vol. 1A*, pp. 265–344.

Kraus, L. and Augustin, R. (2001). Population survey on the consumption of psychoactive substances in the German adult population 2000, *Sucht – Zeitschrift für Wissenschaft und Praxis* **47**: 1–88.

Lee, Y. G. and Abdel-Ghany, M. (2004). American youth consumption of licit and illicit substances, *International Journal of Consumer Studies* **28**: 454–465.

Maddala, G. S. (1983). Limited-dependent and qualitative variables in econometrics, Cambridge University Press, Cambridge U.K.

Murphy, K. M. and Topel, R. H. (1985). Estimation and inference in two-step econometric models, *Journal of Business & Economic Statistics* **3**: 370–379.

Nelson, F. D. and Olsen, L. (1978). Specification and estimation of a simultaneous equation model with limited dependent variables, *International Economic Review* **19**: 695–710.

Newey, W. K. (1986). Efficient estimation of limited dependent variable models with endogenous explanatory variables, *Journal of Econometrics* **36**: 231–250.

Newey, W. K. and McFadden, D. (1994). Large sample estimation and hypothesis testing, *in* R. Engle and D. McFadden (eds), *Handbook of Econometrics*, Vol. 4, Elsevier, pp. 2111–2245.

Picone, G. A., Sloan, F. and Trogdon, J. G. (2004). The effect of the tobacco settlement and smoking bans on alcohol consumption, *Health Economics* **13**: 1063–1080.

Poirier, D. J. (1980). Partial observability in bivariate probit models, *Journal of Econometrics* **12**: 209–217.

Rivers, D. and Vuong, H. Q. (1988). Limited information estimators and exogeneity tests for simultaneous probit models, *Journal of Econometrics* **39**: 347–366.

Smith, R. and Blundell, R. (1986). An exogeneity test for the simultaneous equation tobit model with an application to labor supply, *Econometrica* **54**: 679–685.

Su, S. B. and Yen, S. T. (2000). A censored system of cigarette and alcohol consumption, *Applied Economics* **32**: 729–737.

Williams, J., Pacula, R. L., Chaloupka, F. J. and Wechsler, H. (2001). Alcohol and marijuana use among college students: economic complements or substitutes?, *NBER Working Paper Series* **8401**.

Winter-Ebmer, R. (2006). Coping with a structural crisis: Evaluating an innovative redundancy-retraining project, *International Journal of Manpower* **27**: 700–721.

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

Yen, S. T. (2005). A multivariate sample-selection model: estimating cigarette and alcohol demands with zero observations, *American Journal of Agricultural Economics* **87**: 452–466.

Zhao, X. and Harris, M. N. (2004). Demand for marijuana, alcohol and tobacco: participation, levels of consumption and cross-equation correlation, *The Economic Record* **80**: 394–410.

Appendix

A Equivalence of Measures of Complementarity

The measure of complementarity γ_a that is used in this analysis is defined in terms of observed changes in consumption, i.e. in terms of Marshallian demand. It represents the derivative of the Marshallian demand for alcohol with respect to the exogenously given consumption of tobacco. This analogously applies to γ_c. In micro-economic theory however complementarity is defined in terms of cross-price effects on Hicksian, i.e. compensated, demand. Though Hicksian demand is an theoretical concept that cannot directly be observed, it allows for disentangling pure substitution effects from income effects.[19] In this appendix we show that the cross-price effect of increasing the price of tobacco on the Hicksian demand for alcohol has always the opposite sign of the effect resulting from increasing the consumption of tobacco on the Marshallian demand for alcohol. For this reason, in qualitative terms the measure of complementarity that is used in this analysis corresponds with the standard definition of complementarity.

To see this, we write the consumer's direct utility as $U(a, c, w)$, where we denote by a, c, and w the amounts of consumed alcohol, tobacco and a compound good consisting of all other goods, respectively. For simplicity, any subscripts i and t denoting specific individuals and periods are skipped. The corresponding prices are p_a, p_c, and p_w. Hicksian demand for alcohol is written as $a^H(p_a, p_c, p_w, \overline{U})$, for some fixed utility level $\overline{U}$. Accordingly, the restricted Marshallian demand for alcohol, if the consumption of tobacco $\overline{c}$ is given, is denoted by $a^M(p_a, p_c, p_w, \overline{c}, y)$ where y is income. We now state the following result:

[19] Cross-price effects on Marshallian demand capture both, substitution and price-induced income effects and therefore their sign may differ from those on Hicksian demand.

Proposition: If U is strictly quasi-concave, and both the Marshallian and the Hicksian demand is characterized by interior solutions in a, c, and w, then

$$\operatorname{sign}\left[\frac{\partial a^H(p_a, p_c, p_w, \overline{U})}{\partial p_c}\right] = -\operatorname{sign}\left[\frac{\partial a^M(p_a, p_c, p_w, \overline{c}, y)}{\partial \overline{c}}\right]. \tag{5}$$

Proof: By definition $a^H(p_a, p_c, p_w, \overline{U})$ is the solution of $\min_{a,c,w} \{p_a a + p_c c + p_w w\}$ subject to

$$U(a, c, w) = \overline{U}. \tag{6}$$

The first-order necessary conditions for the expenditure minimum are given by

$$\begin{aligned} U_a(a, c, w) &= \lambda^{-1} p_a \equiv \mu p_a && (7)\\ U_c(a, c, w) &= \lambda^{-1} p_c \equiv \mu p_c && (8)\\ U_w(a, c, w) &= \lambda^{-1} p_w \equiv \mu p_w && (9) \end{aligned}$$

where U_a, U_c, and U_w are partial derivatives of $U(\cdot)$ and λ, with $\lambda > 0$, is the Langrange multiplier with respect to (6) and $\mu \equiv \lambda^{-1}$. In order to obtain $\partial a^H/\partial p_c$ we differentiate the equation system (7) trough (9) and (6) totally with respect to p_c to obtain:

$$\begin{bmatrix} U_{aa} & U_{ac} & U_{aw} & -p_a \\ U_{ac} & U_{cc} & U_{cw} & -p_c \\ U_{aw} & U_{cw} & U_{ww} & -p_w \\ \mu p_a & \mu p_c & \mu p_w & 0 \end{bmatrix} \times \begin{bmatrix} \partial a^H/\partial p_c \\ \partial c^H/\partial p_c \\ \partial w^H/\partial p_c \\ \partial \mu/\partial p_c \end{bmatrix} = \begin{bmatrix} 0 \\ \mu \\ 0 \\ 0 \end{bmatrix} \tag{10}$$

where we have made use of (7) trough (9) in the last row of the matrix. Solving (10) we obtain for $\partial a^H/\partial p_c$ (we omit the expressions of the other effects being of no further interest here):

$$\frac{\partial a^H}{\partial p_c} = -\mu \frac{p_w^2 U_{ac} - p_c p_w U_{aw} - p_a p_w U_{cw} + p_a p_c U_{ww}}{D} \tag{11}$$

where the denominator D is given by

$$\begin{aligned} D &= p_a^2[U_{cc}U_{ww} - U_{cw}^2] + p_c^2[U_{aa}U_{ww} - U_{aw}^2] + p_w^2[U_{aa}U_{cc} - U_{ac}^2] \\ &\quad +2p_a p_c[U_{aw}U_{cw} - U_{ac}U_{ww}] + 2p_a p_w[U_{ac}U_{cw} - U_{cc}U_{aw}] \\ &\quad +2p_c p_w[U_{ac}U_{aw} - U_{aa}U_{cw}] \end{aligned}$$

and is greater than zero by strict quasi-concavity and the resulting second-order condition of the consumer's expenditure minimization problem.

We now look at the restricted Marshallian demand $a^M(p_a, p_c, p_w, \overline{c}, y)$ which by definition is the solution of $\max\limits_{a,c,w} U(a, c, w)$ subject to

$$p_a a + p_c c + p_w w = y \tag{12}$$

and $c \leq \overline{c}$. The Lagrange function is then given by $\mathcal{L}(a, c, w, \mu, \nu) = U(a, c, w) + \mu[y - p_a a - p_c c - p_w w] + \nu[\overline{c} - c]$. Assuming that the constraint $c \leq \overline{c}$ holds with equality, the first-order necessary conditions for the utility maximum are given by

$$U_a(a, \overline{c}, w) = \mu p_a \tag{13}$$

$$U_w(a, \overline{c}, w) = \mu p_w. \tag{14}$$

Differentiating (13), (14) and (12) with respect to $\overline{c}$ we obtain:

$$\begin{bmatrix} U_{aa} & U_{aw} & -p_a \\ U_{aw} & U_{ww} & -p_w \\ p_a & p_w & 0 \end{bmatrix} \times \begin{bmatrix} \partial a^M/\partial \overline{c} \\ \partial w^M/\partial \overline{c} \\ \partial \mu/\partial \overline{c} \end{bmatrix} = \begin{bmatrix} -U_{ac} \\ -U_{cw} \\ -p_c \end{bmatrix}.$$

Solving this system for $\partial a^M/\partial \overline{c}$ (again omitting the other expressions) we obtain:

$$\frac{\partial a^M}{\partial \overline{c}} = -\frac{p_w^2 U_{ac} - p_c p_w U_{aw} - p_a p_w U_{cw} + p_a p_c U_{ww}}{D'} \tag{15}$$

where the denominator $D' = p_w^2 U_{aa} + p_a^2 U_{ww} - 2 p_a p_w U_{aw}$ is negative by strict quasi-concavity.

Finally, comparing (11) and (15) we obtain

$$\frac{\partial a^H}{\partial p_c} = \frac{\partial a^M}{\partial \overline{c}} \frac{\mu D'}{D} \tag{16}$$

establishing (5).

B Supplementary Tables

Table 5: **Description of dependent variables**

Variable	Mean	Std. Dev.	Number of obs.
All			
drinker	0.729	0.444	25,654
smoker	0.428	0.495	26,353
drinker as well as smoker	0.340	0.474	25,532
grams of alcohol consumed by drinkers	15.112	16.418	15,505
number of cigarettes smoked by smokers	16.043	8.547	9,372
drinker without quantitative information	0.203	0.402	18,711
smoker without quantitative information	0.169	0.374	11,272
Males			
drinker	0.838	0.369	12,916
smoker	0.462	0.499	13,063
drinker as well as smoker	0.412	0.492	12,817
grams of alcohol consumed by drinkers	19.076	18.763	9,056
number of cigarettes smoked by smokers	17.433	8.801	5,061
drinker without quantitative information	0.185	0.389	10,817
smoker without quantitative information	0.161	0.368	6,034
Females			
drinker	0.620	0.486	12,738
smoker	0.394	0.489	13,290
drinker as well as smoker	0.267	0.442	12,715
grams of alcohol consumed by drinkers	9.546	10.037	6,449
number of cigarettes smoked by smokers	14.412	7.936	4,311
drinker without quantitative information	0.228	0.419	7,894
smoker without quantitative information	0.177	0.382	5,238

Table 6: **Description of explanatory variables**

	All		**Males**		**Females**	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
father never drinker	**0.186**	0.389	**0.180**	0.384	**0.193**	0.395
father monthly drinker	**0.236**	0.424	**0.236**	0.425	**0.235**	0.424
father weekly drinker	**0.259**	0.438	**0.268**	0.443	**0.250**	0.433
father daily drinker	**0.319**	0.466	**0.316**	0.465	**0.323**	0.468
mother never drinker	**0.507**	0.500	**0.499**	0.500	**0.514**	0.500
mother monthly drinker	**0.285**	0.452	**0.292**	0.455	**0.280**	0.449
mother weekly drinker	**0.136**	0.342	**0.141**	0.348	**0.130**	0.336
mother daily drinker	**0.072**	0.259	**0.068**	0.251	**0.076**	0.266
father never-smoker	**0.276**	0.447	**0.275**	0.447	**0.276**	0.447
father ex-smoker	**0.351**	0.477	**0.357**	0.479	**0.345**	0.475
father smoker	**0.373**	0.484	**0.368**	0.482	**0.379**	0.485
mother never-smoker	**0.674**	0.469	**0.676**	0.468	**0.673**	0.469
mother ex-smoker	**0.126**	0.331	**0.128**	0.334	**0.124**	0.330
mother smoker	**0.200**	0.400	**0.197**	0.397	**0.203**	0.402
year 1980	**0.187**	0.390	**0.193**	0.395	**0.180**	0.384
year 1986	**0.128**	0.334	**0.133**	0.340	**0.123**	0.328
year 1990	**0.581**	0.493	**0.573**	0.495	**0.590**	0.492
year 1992	**0.104**	0.305	**0.101**	0.302	**0.107**	0.309
west	**0.838**	0.368	**0.844**	0.362	**0.832**	0.374
female	**0.503**	0.500	–	–	–	–
age	**24.31**	6.297	**24.213**	6.275	**24.421**	6.317
$\text{age}^2/100$	**6.310**	3.351	**6.256**	0.064	**6.363**	3.366
parents married	**0.807**	0.395	**0.811**	0.392	**0.805**	0.396
father has no school degree	**0.139**	0.346	**0.138**	0.344	**0.140**	0.347
father has a low degree	**0.531**	0.499	**0.533**	0.499	**0.529**	0.499
father has a medium degree	**0.156**	0.363	**0.159**	0.366	**0.151**	0.358
father has a high degree	**0.032**	0.175	**0.032**	0.176	**0.032**	0.176
father has a university degree	**0.142**	0.349	**0.138**	0.345	**0.148**	0.355
mother has no school degree	**0.309**	0.462	**0.304**	0.460	**0.315**	0.464
mother has a low degree	**0.416**	0.493	**0.424**	0.494	**0.408**	0.491
mother has a medium degree	**0.191**	0.393	**0.189**	0.391	**0.193**	0.395
mother has a high degree	**0.024**	0.154	**0.025**	0.157	**0.023**	0.150
mother has a university degree	**0.059**	0.236	**0.058**	0.234	**0.061**	0.240
grown up with mother	**0.953**	0.212	**0.954**	0.211	**0.952**	0.214
grown up with father	**0.891**	0.312	**0.891**	0.312	**0.890**	0.312
grown up with both	**0.881**	0.324	**0.881**	0.324	**0.882**	0.323
no. children at parents' home	**2.752**	1.462	**2.732**	1.434	**2.772**	1.488

Notes: Descriptive statistics for those 26,516 observations that are included in at least one of the reduced form Tobit regressions; statistics are constructed for all variables prior to interacting with dummies indicating having grown up with the parent; reference-categories italicized.

Table 7: **Males: reduced form estimates**

	Tobit Model				Conditional Linear Model			
	Drinking		**Smoking**		**Drinking**		**Smoking**	
Variable	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.
father monthly drinker	1.820**	0.604	-2.006**	0.634	0.131	1.114	-0.084	0.426
father weekly drinker	4.200**	0.606	-2.076**	0.439	3.827**	1.117	0.042	0.428
father daily drinker	5.914**	0.597	-1.154	0.620	5.869**	1.092	0.797	0.414
mother monthly drinker	2.689**	0.443	-1.095**	0.471	2.310**	0.804	-0.393	0.312
mother weekly drinker	4.784**	0.584	-0.722	0.627	4.091**	1.067	-0.422	0.426
mother daily drinker	4.291**	0.795	-0.305	0.840	4.213**	1.427	0.072	0.556
father ex-smoker	1.257**	0.476	3.957**	0.523	1.107	0.941	0.015	0.375
father smoker	1.209*	0.497	7.411**	0.535	-0.203	0.949	0.980**	0.372
mother ex-smoker	0.235	0.566	2.758**	0.598	0.918	1.036	1.126**	0.405
mother smoker	0.764	0.491	6.934**	0.501	0.857	0.826	2.042**	0.312
constant	-18.087**	3.310	-40.047**	3.433	-1.397	6.910	-12.838**	2.36
year 1986	-1.564*	0.628	-3.968**	0.663	1.370	1.103	-0.724	0.436
year 1990	-2.241**	0.515	-5.763**	0.538	1.044	0.912	-1.054**	0.357
year 1992	-5.609**	1.363	-7.078**	1.116	8.519*	3.756	-0.834	0.716
west	-8.271**	1.169	-2.138**	0.839	3.880	3.540	2.451**	0.537
age	2.952**	0.237	3.666**	0.254	1.129*	0.455	1.975**	0.176
$age^2/100$	-4.278**	0.445	-6.026**	0.473	-1.122	0.848	-2.897**	0.326
parents married	-0.407	0.421	-2.000**	0.556	-1.261	0.942	-1.097**	0.352
father has low degree	0.760	0.606	-1.243*	0.630	-0.182	1.087	-0.179	0.414
father has medium degree	-0.126	0.754	-2.018*	0.797	-1.031	1.372	-0.614	0.537
father has high degree	-1.307	1.222	-2.439	1.314	-0.606	2.258	-0.955	0.911
father has univ. degree	-0.077	0.846	-3.693**	0.892	-1.092	1.578	-1.220*	0.605
mother has low degree	-0.591	0.464	-0.447	0.488	-1.355	0.832	-0.557	0.323
mother has medium degree	-0.179	0.602	-0.193	0.639	0.689	1.091	-0.154	0.432
mother has high degree	-1.596	1.242	-2.441	1.346	-0.939	2.357	-1.502	0.929
mother has univ. degree	-1.293	0.986	-2.038	1.051	0.932	1.917	-1.071	0.726
grown up with mother	-4.071**	1.199	-4.970**	1.216	-4.799*	2.021	-0.950	0.772
grown up with father	-8.252**	2.111	-7.333**	2.168	-10.072*	3.981	-1.250	1.433
grown up with both	4.409*	2.154	4.213	2.206	8.755*	3.984	0.568	1.439
no. children at parents' home	0.338**	0.128	0.764**	0.131	0.657**	0.227	0.240**	0.084
number of observations	12,922		13,064		4,212		4,313	
LR-statistic	1,084.95		1,177.50		–		–	
F-statistic	–		–		8.81		21.32	

Note: ** significant at the 1%-level; * significant at the 5% level.

Table 8: **Females: reduced form estimates**

	Tobit Model				Conditional Linear Model			
	Drinking		Smoking		Drinking		Smoking	
Variable	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.
father monthly drinker	1.413**	0.346	-1.059	0.590	-0.170	0.859	-0.746	0.542
father weekly drinker	1.863**	0.350	-0.914	0.596	0.102	0.867	-0.245	0.551
father daily drinker	2.021**	0.339	-1.156*	0.574	0.598	0.833	0.147	0.522
mother monthly drinker	2.505**	0.260	-1.355**	0.452	0.855	0.603	-0.534	0.384
mother weekly drinker	3.903**	0.344	-0.405	0.599	2.408**	0.783	-0.584	0.497
mother daily drinker	4.230**	0.433	-1.324	0.760	3.768**	0.96	0.077	0.640
father ex-smoker	0.051	0.277	3.163**	0.502	-0.910	0.705	1.222**	0.468
father smoker	0.296	0.283	6.402**	0.501	-0.510	0.686	1.601**	0.452
mother ex-smoker	0.236	0.327	3.239**	0.567	-0.119	0.743	0.451	0.484
mother smoker	0.214	0.285	7.856**	0.468	0.264	0.601	2.339**	0.378
constant	-7.900**	1.881	-34.204**	3.238	-0.320	5.009	-5.812*	2.922
year 1986	-3.698**	0.377	-3.741**	0.647	-0.865	0.86	0.650	0.553
year 1990	-3.372**	0.303	-4.166**	0.509	-0.202	0.661	-0.297	0.421
year 1992	-6.943**	0.709	-4.368**	1.043	6.312*	2.687	-0.894	0.918
west	-4.853**	0.586	2.248**	0.778	7.496**	2.514	3.884**	0.692
age	1.216**	0.136	2.852**	0.239	0.124	0.332	1.344**	0.215
$\text{age}^2/100$	-1.606**	0.255	-5.008**	0.445	0.307	0.62	-2.074**	0.398
parents married	0.015	0.309	-2.137**	0.516	-0.022	0.711	-1.198**	0.440
father has low degree	0.650	0.348	0.383	0.592	0.317	0.828	-0.308	0.524
father has medium degree	0.945*	0.438	0.503	0.754	0.606	1.015	-1.300*	0.652
father has high degree	1.071	0.708	-1.439	1.241	1.522	1.777	-2.303*	1.108
father has univ. degree	1.749**	0.478	-1.127	0.832	1.801	1.155	-1.193	0.743
mother has low degree	0.286	0.270	-2.122*	0.967	-0.305	0.632	-0.401	0.398
mother has medium degree	0.250	0.342	-0.542	0.460	-0.432	0.808	0.211	0.527
mother has high degree	-0.146	0.751	-0.876	0.593	2.910	1.863	3.646**	1.197
mother has univ. degree	0.905	0.552	-3.229*	1.376	2.146	1.368	-0.388	0.872
grown up with mother	-0.123	0.682	-4.149**	1.103	-0.056	1.527	-3.097**	0.947
grown up with father	-0.026	1.279	-3.035	2.068	1.019	3.012	-2.005	1.761
grown up with both	-2.453	1.304	-2.730	2.098	-2.199	3.024	2.456	1.768
no. children at parents' home	-0.116	0.071	0.536**	0.118	-0.166	0.172	0.085	0.103
number of observations	12,773		13,291		2,607		2,627	
LR-statistic	966.00		1,098.15		–		–	
F-statistic	–		–		3.43		10.42	

Note: ** significant at the 1%-level; * significant at the 5% level.

Table 9: **Males: structural form estimates**

		Tobit Model				Conditional Linear Model			
		Drinking		**Smoking**		**Drinking**		**Smoking**	
Parameter / Exp. Variable		Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.
γ_a	fitted smoking	0.135**	0.041	–	–	0.115	0.314	–	–
γ_c	fitted drinking	–	–	-0.159**	0.055	–	–	0.056	0.042
δ_a	father monthly drinker	2.108**	0.458	–	–	0.124	0.937	–	–
	father weekly drinker	4.529**	0.533	–	–	3.761**	1.050	–	–
	father daily drinker	6.122**	0.547	–	–	5.686**	1.050	–	–
	mother monthly drinker	2.813**	0.414	–	–	2.373**	0.809	–	–
	mother weekly drinker	4.855**	0.563	–	–	4.266**	0.988	–	–
	mother daily drinker	4.279**	0.885	–	–	4.391**	1.676	–	–
δ_c	father ex-smoker	–	–	4.209**	0.596	–	–	0.012	0.374
	father smoker	–	–	7.697**	0.557	–	–	1.099**	0.371
	mother ex-smoker	–	–	2.824**	0.592	–	–	1.043*	0.407
	mother smoker	–	–	7.092**	0.544	–	–	1.956**	0.312
β	constant	-12.869**	3.167	-42.937**	3.797	0.135	6.986	-12.844**	2.140
	year 1986	-1.048	0.666	-4.218**	0.727	1.441	1.119	-0.799*	0.389
	year 1990	-1.494*	0.593	-6.119**	0.603	1.231	0.909	-1.071**	0.321
	year 1992	-4.694**	1.218	-7.970**	1.220	8.736**	3.133	-1.264	0.736
	west	-8.001**	0.922	-3.481**	0.954	3.739	3.007	2.187**	0.535
	age	2.467**	0.242	4.134**	0.311	0.883	0.732	1.916**	0.166
	$\text{age}^2/100$	-3.472**	0.441	-6.696**	0.548	-0.757	1.201	-2.837**	0.309
	parents married	-0.034	0.608	-2.142**	0.615	-0.934	1.094	-1.049**	0.363
	father has low degree	0.940	0.656	-1.264	0.671	-0.105	1.174	-0.220	0.429
	father has medium degree	0.150	0.779	-2.234**	0.768	-0.821	1.408	-0.646	0.542
	father has high degree	-0.972	1.225	-2.771*	1.375	-0.431	2.141	-0.998	0.901
	father has univ. degree	0.432	0.839	-3.903**	0.957	-0.798	1.689	-1.308*	0.587
	mother has low degree	-0.536	0.397	-0.593	0.465	-1.248	0.803	-0.515	0.319
	mother has medium degree	-0.159	0.549	-0.287	0.659	0.760	1.126	-0.277	0.421
	mother has high degree	-1.249	1.057	-2.695*	1.307	-0.506	2.124	-1.512	0.940
	mother has univ. degree	-1.020	0.950	-2.240*	1.044	1.068	2.304	-1.162	0.735
	grown up with mother	-3.429*	1.565	-5.770**	1.255	-4.339	2.577	-0.862	0.849
	grown up with father	-6.939**	1.968	-9.329**	2.041	-9.610**	3.389	-0.617	1.428
	grown up with both	3.775	2.154	4.877*	2.063	8.372*	3.624	0.074	1.519
	no. children at parents' home	0.232	0.144	0.835**	0.132	0.605*	0.260	0.212*	0.087
number of observations		12,922		13,064		4,212		4,313	
LR-statistic		1,082.24		1,158.07		–		–	
F-statistic		–		–		9.65		25.39	
tests for over-identifying restrictions: (p-values):									
intuitive		0.383		0.080		0.434		0.408	
Hausman		–		–		0.410		0.122	
Newey & McFadden		0.379		0.000		0.297		0.214	

Notes: ** significant at the the 1% level; * significant at the the 5% level; bootstrapped standard errors reported.

Table 10: **Females: structural form estimates**

		Tobit Model				Conditional Linear Model			
		Drinking		Smoking		Drinking		Smoking	
Parameter / Exp. Variable		Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.
γ_a	fitted smoking	0.038	0.025	–	–	-0.008	0.205	–	–
γ_c	fitted drinking	–	–	-0.337**	0.102	–	–	0.037	0.144
δ_a	father monthly drinker	1.453**	0.370	–	–	-0.175	0.951	–	–
	father weekly drinker	1.898**	0.411	–	–	0.030	0.926	–	–
	father daily drinker	2.070**	0.386	–	–	0.511	0.910	–	–
	mother monthly drinker	2.556**	0.253	–	–	0.897	0.635	–	–
	mother weekly drinker	3.918**	0.311	–	–	2.471**	0.854	–	–
	mother daily drinker	4.272**	0.449	–	–	3.846**	0.954	–	–
δ_c	father ex-smoker	–	–	3.145**	0.513	–	–	1.318**	0.460
	father smoker	–	–	6.458**	0.488	–	–	1.733**	0.434
	mother ex-smoker	–	–	3.343**	0.588	–	–	0.475	0.485
	mother smoker	–	–	7.960**	0.487	–	–	2.282**	0.398
β	constant	-6.610**	1.830	-36.975**	3.659	0.478	4.708	-5.767	3.030
	year 1986	-3.556**	0.367	-4.965**	0.879	-0.912	0.751	0.616	0.613
	year 1990	-3.213**	0.355	-5.267**	0.655	-0.206	0.715	-0.280	0.425
	year 1992	-6.772**	0.626	-6.672**	1.376	6.346**	1.328	-1.089	1.105
	west	-4.932**	0.466	0.642	1.039	7.524**	1.209	3.570**	1.220
	age	1.109**	0.139	3.265**	0.276	0.128	0.419	1.334**	0.227
	$\text{age}^2/100$	-1.418**	0.253	-5.551**	0.517	0.291	0.723	-2.073**	0.423
	parents married	0.087	0.353	-2.119**	0.581	-0.141	0.728	-1.260**	0.460
	father has low degree	0.634*	0.314	0.584	0.598	0.287	0.839	-0.357	0.555
	father has medium degree	0.921*	0.398	0.807	0.842	0.531	1.087	-1.435*	0.657
	father has high degree	1.119	0.697	-1.077	1.452	1.384	1.953	-2.335*	1.128
	father has univ. degree	1.787**	0.471	-0.506	0.885	1.736	1.257	-1.334	0.857
	mother has low degree	0.306	0.279	-0.453	0.497	-0.271	0.655	-0.427	0.397
	mother has medium degree	0.286	0.330	-0.795	0.692	-0.377	0.785	0.178	0.505
	mother has high degree	-0.015 -	0.842	-3.155	1.758	2.964	2.397	3.577*	1.627
	mother has univ. degree	0.993	0.580	-1.848	1.061	2.139	1.678	-0.480	0.927
	grown up with mother	0.017	0.649	-4.177**	1.218	-0.060	1.542	-3.338**	0.983
	grown up with father	0.087	1.896	-3.346	2.191	0.641	4.669	-2.404	1.932
	grown up with both	-2.335	1.820	-3.617	2.416	-2.229	4.477	2.710	1.944
	no. children at parents' home	-0.136	0.073	0.498**	0.119	-0.162	0.166	0.109	0.122
number of observations		12,773		13,291		2,607		2,627	
LR-statistic		965.58		1,090.84		–		–	
F-statistic		–		–		3.75		12.23	
tests for over-identifying restrictions (p-values):									
intuitive		0.936		0.207		0.938		0.989	
Hausman		–		–		0.638		0.225	
Newey & McFadden		0.920		0.061		0.676		0.263	

Notes: ** significant at the the 1% level; * significant at the the 5% level; bootstrapped standard errors reported.