

Tauchmann, Harald; Lenz, Silja; Requate, Till; Schmidt, Christoph M.

**Article — Accepted Manuscript (Postprint)**

## Tobacco and alcohol: complements or substitutes?

Empirical Economics

**Provided in Cooperation with:**

RWI – Leibniz-Institut für Wirtschaftsforschung, Essen

*Suggested Citation:* Tauchmann, Harald; Lenz, Silja; Requate, Till; Schmidt, Christoph M. (2013) : Tobacco and alcohol: complements or substitutes?, Empirical Economics, ISSN 1435-8921, Springer, Berlin, New York, Vol. 45, Iss. 1, pp. 539-566,  
<https://doi.org/10.1007/s00181-012-0611-3> ,  
<http://link.springer.com/article/10.1007%2Fs00181-012-0611-3>

This Version is available at:

<https://hdl.handle.net/10419/141886>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

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

PRICE VARIATION IN INDIVIDUAL-LEVEL DATA<sup>\*†</sup>

Harald Tauchmann

*RWI*

Silja Lenz

*RWI*

Till Requate

*University of Kiel*

Christoph M. Schmidt

*RWI, Ruhr-University Bochum, IZA, and CEPR*

## Abstract

The question of whether alcohol and tobacco are consumed as complements or substitutes is crucial for determining the side-effects of anti-smoking policies. 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 is often 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 using German survey data. Our estimation results suggest that a reduction in tobacco consumption results in a moderate reduction in alcohol consumption. It is demonstrated 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. complements or substitutes interdependence in consumption tobacco and alcohol insufficient price-variation instrumental variables approach

**JEL Classification** C31 · D12 · I12

---

<sup>\*</sup>We would like to thank Boris Augurzky and Lars Siemers, 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 für Socialpolitik Annual Conference 2006 for many useful comments.

<sup>†</sup>Address for correspondence: Harald Tauchmann, RWI, Hohenzollernstraße 1-3, 45128 Essen, Germany, phone: +49-201-8149-259, fax: +49-201-8149-200, [harald.tauchmann@rwi-essen.de](mailto:harald.tauchmann@rwi-essen.de)

# 1 Introduction

Tobacco consumption has been identified as a major cause of health problems in industrialized countries (World Health Organization, 2008). Consequently, smoking has been placed under severe restrictions. Even in Germany, one of the last developed countries to consider such restrictions, smoking has been banned from public sector buildings and public transport, and similar bans have recently been introduced for bars and restaurants at the regional level. Yet, 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, conditional on that both drugs are substitutes. This might particularly apply to the socially approved licit drug alcohol.

Using German micro data, this paper analyzes whether tobacco and alcohol are complements or substitutes in consumption. The interdependence between alcohol and tobacco consumption is an important aspect to consider when designing anti-smoking policies: 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, however, a smoking ban could have a decreasing, i.e. a desirable side-effect on the consumption of alcohol. Making matters even more complicated, there might not be an universal answer to this question. Rather, the nature of interdependence in consumption of both drugs might be heterogeneous across cultures and countries (and even individuals). Yet, as in Germany anti-smoking policies – namely smoking bans – have been introduced relatively recently and correspondingly the empirical evidence of side effects of such measures is still limited, the case of Germany appears particularly worth studying.<sup>1</sup>

In technical terms, answering the question of complementarity versus substitutability is far from straightforward. The standard approach of estimating cross-price effects using micro data is often seriously hampered by insufficient variation of prices across observation units. Several approaches were suggested for circumventing this problem. One approach is to generate price variation by calculating weighted averages from prices of numerous specific goods with individual expenditure shares serving as weights (Labeaga, 1999). Another approach is to derive individualized prices – and individualized price indices, respectively – directly from expenditures data (e.g. Cox and Wohlgenant, 1986; Hoderlein and Mihaleva, 2008). However, in many applications individual-level information on expenditures may also be unavailable, as it is the case for the data used for the analysis at hand. To solve the problem of insufficient price variation, we therefore 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 between tobacco and alcohol, thus promising positive side-effects of smoking bans in the form of reduced alcohol consumption.

For other countries than Germany, one can draw on 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. Several – among them some rather early – contributions rely on aggregate data at the regional or national level. Here Jones (1989), using aggregate quarterly expenditures UK data, may serve as an particularly relevant example. Cross-price elasticities for beer, sprits, wine cider, and tobacco are estimated indicating that tobacco is a complement to all considered alcoholic beverages. Other aggregate data based analyses are for instance Goel and Morey (1995), Dee (1999), Bask and Melkersen (2004), and Pierani and Tiezzi (2009).

In contrast, several other studies use survey data at the level of individual consumers. Here Jimenez and Labeaga (1994), Decker and Schwartz (2000), Cameron and Williams (2001), Zhao and Harris (2004), and Picone et al. (2004)<sup>2</sup> may serve as representative examples of the respec-

---

<sup>1</sup>Focussing on revenues of pubs and bars as a rough measure of alcohol consumption, recent empirical evidence for Germany does not point at any significant impact of smoking bans on drinking (Ahlfeldt and Maennig, 2010) or finds just a small decreasing effect on alcohol consumption (Kvasnicka and Tauchmann, 2012). Yet, by completely ignoring home consumption, these analyses might provide a rather incomplete picture.

<sup>2</sup>Chaloupka and Laixuthai (1997), DiNardo and Lemieux (2001), and Williams et al. (2001) address the interdependency of the consumption of alcohol and drugs other than tobacco, such as marijuana. Moreover, several related papers neither use prices as explanatory variables, nor do they pursue a structural model approach similar to the present paper, and,

tive body of literature. In Jimenez and Labeaga (1994), for instance, a demand system including tobacco and alcohol is estimated relying on Spanish household level data, and cross-price elasticities are derived from estimated coefficients. Within this analysis, a special focus is on how to accommodate zero expenditures in the econometric model. Yet, typical for such analyses, identification rests on very limited price variation over time since observed prices do not vary at the level of individual households. This renders disentangling genuine price effects from time or regional effects very difficult. Irrespective of the level of aggregation and the country considered, most of these studies find negative cross-price effects and consequently conclude that alcohol and tobacco are complements. As the only exception Goel and Morey (1995), focussing on liquor rather than alcohol consumption in general, find positive and significant cross-price elasticities.

The remainder of the 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 comprises information on consumed quantities of tobacco and alcohol and several individual level background variables, while expenditures are not reported. For the considered period of time in Germany, the prices of tobacco and alcohol neither display any remarkable variation over time and across regions, nor any variation 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:

$$a_{it} = \gamma_a c_{it} + \beta'_a x_{it} + \delta'_a z_{ait} + \varepsilon_{ait} \quad (1)$$

$$c_{it} = \gamma_c a_{it} + \beta'_c x_{it} + \delta'_c z_{cit} + \varepsilon_{cit} \quad (2)$$

Here,  $\varepsilon_{ait}$  and  $\varepsilon_{cit}$  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 time and region specific variables dummy variables in the vector  $x_{it}$ . Similar structural models have been formulated by Dee (1999) and Bask and Melkersen (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.

The coefficient  $\gamma_a$  in structural equation (1) measures what would happen to the consumption of alcohol if the consumption of tobacco were exogenously reduced by one unit.<sup>3</sup> This interpretation analogously applies to  $\gamma_c$ . We use these coefficients as a measure of complementarity in consumption, since they precisely answer the relevant question pertaining to the possible side-effects of drug related regulation: "Imagine the regulator could manage to reduce individual smoking levels 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.

---

hence, are more concerned with random correlation of drinking and smoking rather than with interdependency, e.g. Su and Yen (2000), Lee and Abdel-Ghany (2004), and Yen (2005).

<sup>3</sup>If feedback-effects are taken into account, one might think of  $(1 - \gamma_a \gamma_c)^{-1} \gamma_a$  as the more appropriate measure.

## 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  $\gamma_a$  and  $\gamma_c$  obtained from naively estimating (1) and (2) are almost certainly biased. Nonetheless, the coefficients  $\theta$  of the corresponding reduced form representation

$$a_{it} = \theta'_{a1}x_{it} + \theta'_{a2}z_{ait} + \theta'_{a3}z_{cit} + v_{ait} \quad (3)$$

$$c_{it} = \theta'_{c1}x_{it} + \theta'_{c2}z_{ait} + \theta'_{c3}z_{cit} + v_{cit} \quad (4)$$

can consistently be estimated by regressing alcohol and tobacco consumption on all exogenous variables.<sup>4</sup> The structural-form coefficients directly translate into reduced form coefficients as follows:

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

The terms for  $\theta_{c1}$ ,  $\theta_{c2}$ ,  $\theta_{c3}$ , and  $v_{cit}$  are defined analogously.

If  $z_{ait}$  and  $z_{cit}$  were empty, that is, if we had no instruments for alcohol or tobacco consumption, respectively, estimates for  $\theta$  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 coefficient including  $\gamma$  from estimates for  $\theta$ , since  $\gamma_a = \theta_{a3k}/\theta_{a2k}$  and  $\gamma_c = \theta_{c3k}/\theta_{c2k}$  hold.<sup>5</sup> 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 includes variables which can be regarded as valid instruments 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.<sup>6</sup> For instance, Bantle and Haisken-DeNew (2002) find significant correlations between parental smoking behavior and children's tobacco consumption in Germany, and – in an analysis addressing the effects of drinking on occupational attainment – MacDonald and Shields (2001) suggest using parental consumption as instruments for filial consumption. In order to use parental consumption habits as instruments in our analysis, 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 direct effect on their drinking habits. Even though parents' tobacco consumption and children's later alcohol use might be correlated, the correlation operates through children's own smoking habits (and other observables). Colloquially speaking, this means that children "learn" to smoke – but not to drink – from watching parents smoke, and vice versa. The suggested approach, hence, requires the corresponding exclusion restriction.

When estimating our coefficient of primary interest  $\gamma_a$  via instrumental variables, we have to

<sup>4</sup>If  $a_{it}$  and  $c_{it}$  were observed for all individuals, the appropriate estimation procedure was OLS. Yet, this does not apply to the data used for the present empirical application. Here  $a_{it}$  and  $c_{it}$  exhibit censoring from the left at zero. See section 2.3 for a detailed discussion about how we deal with this problem.

<sup>5</sup>The subscript  $k$  indicates the  $k$ th 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  $\gamma_a$  and  $\gamma_c$  can be calculated.

<sup>6</sup>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).

acknowledge that this coefficient is not necessarily the same for all individuals. Rather, what we can identify if  $\gamma_a$  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 divergent behavior from what they had (Angrist and Krueger, 2001). One restriction allowing the generalization of 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  $\gamma_a$  in terms of such a structural model parameter, keeping the less restrictive interpretation as a local average effect in mind.

The validity of our exclusion restrictions is decisive for our empirical analysis. As a minimum, we need to justify them, equation by equation, through a priori reasoning. It does not seem implausible that parental smoking behavior may 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.<sup>7</sup> Hence, the structural coefficients  $\gamma_a$  and  $\gamma_c$  are over-identified and one can apply tests for over-identifying restrictions. We apply two different test procedures.

The first approach represents an intuitive test for the validity of over-identifying restrictions. As pointed out, estimating the reduced form model allows the calculation of as many estimates for  $\gamma_a$  and  $\gamma_c$  as instruments are available. If all instruments are valid, one should expect that these different estimates only differ because of sampling error. Hence we test whether the relevant ratios of the reduced form coefficients  $\theta$  coefficients jointly equal each other.<sup>8</sup> Secondly, we employ a formal 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 estimating either the reduced form or the structural form should deviate only because of sampling error. Conventional  $\chi^2$ -tests allow for testing the underlying null-hypotheses.

## 2.3 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. So in our data, the consumption patterns of both alcohol and tobacco are characterized by large shares of zero consumption. To account for this in the econometric analysis, 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 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.<sup>9</sup> The results reported in this paper are based on a particularly straight forward 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  $\theta$ , fitted values  $\hat{a}_{it}^*$  and  $\hat{c}_{it}^*$  are calculated. These serve as regressors in the structural equations, which once again

<sup>7</sup>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.

<sup>8</sup>Standard errors for the ratios of coefficients are derived through bootstrapping.

<sup>9</sup>Hard-coded procedures for the FIML as well as the two-step approach are available in econometric software packages like Stata.

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  $\text{var}(\varepsilon_c)$ .<sup>10</sup> Thus, marginal effects on expected actual consumption are also not identified. Yet, since we are primarily interested in the signs of the parameters  $\gamma_a$  and  $\gamma_c$  this limitation is of marginal relevance.

The Tobit approach to the model is quite restrictive, though. Besides normality, it assumes that the decisions on drinking and smoking participation are determined by exactly the same mechanism that determines the intensity of consumption. For this reason, we also try a specification that interprets the equations (1) and (2) as conditional on  $a_{it} > 0$  and  $c_{it} > 0$ . Thus, this alternative approach separates the participation decisions from the demand equations. It can be estimated as a generalized Heckman (1976) selection model (Poirier, 1980) or in the fashion of a two part model; see Jones (2000) for a discussion of both econometric approaches. In the former models (Heckman, 1976; Poirier, 1980) identification critically relies on valid exclusion restrictions, i.e. explanatory variables that only affect the decision whether or not to consume but not the actual consumed quantity. Yet, in our data we cannot identify any variable that may legitimately be excluded only from the continuous model. For this reason we prefer the latter specification, to which we refer as the conditional linear model; see section 4.3 for estimation results.

The 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 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”<sup>11</sup> collected by IFT<sup>12</sup> Munich; see Kraus and Augustin (2001) for a detailed description. The data is originally comprised of nine 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, 2003, and 2006. 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. Up until 1992, only German citizens were interviewed, immigrants not holding the German citizenship were disregarded. Later on, the entire German-speaking population was included in the survey, irrespective of citizenship. The data provides comprehensive information with respect to various legal and illicit drugs regarding prevalence, frequency and intensity of consumption, consumption habits, and age at the time of first use. Additionally, detailed information on socioeconomic characteristics is provided along with information on attitudes towards several drug-related issues. Any information on family background and, in particular, on parental characteristics and parental behavior is reported by respondents in retrospect rather than by parents themselves who were not personally interviewed.

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 subsequent waves, the upper age limit was successively raised up to 39 in 1990. Since 1995, the target population has solely consisted of adults aged 18 to 59. As a consequence, consumers’ family background has increasingly become a minor issue. The result is that neither smoking nor drinking habits at the parental home are reported in waves after 1992. This also means that the recent waves lack those instrumental variables that are decisive for our econometric model. Thus, our analysis has to rely on data collected in 1980, 1986, 1990, and 1992. As the empirical analysis, hence, has to rest on rather old data some caution is required in interpreting the results. Though it is unlikely that

<sup>10</sup>Cf. Rivers and Vuong (1988) for an equivalent problem in the case of the Probit model.

<sup>11</sup>Bundesstudie “Repräsentativerhebung zum Gebrauch psychoaktiver Substanzen in Deutschland”

<sup>12</sup>Institute for Therapy Research (Institut für Therapieforschung)

the relationship of consuming alcohol and tobacco has fundamentally changed, the ‘culture’ of smoking and drinking may still not be the same as in the 1980s and the early 1990s.

We do not consider individuals younger than 16 years for estimating the model. The rationalization behind this is that although 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;<sup>13</sup> see 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.<sup>14</sup> 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. Since the questionnaire asks about average consumption per week and per day, respectively, infrequency of purchase and consumption is not observed in the data. 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. We do not exclude these observations from our analysis, but let the probability to either drink or smoke enter the likelihood function.<sup>15</sup>

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. Here two dummies indicate whether the respondent grew up with father and mother, respectively, while the interaction of both, which is additionally included, indicates a complete nuclear family. By interacting parental education with dummy variables indicating that the respondent grew up with the parent and letting only this interaction enter the regression equations, we allow parental education to have an effect only if the respondent has grown up with the parent. We also include a set of time dummies that implicitly control for variables that do not vary at the level of individual consumers but only over time. This first of all applies to prices. Variables often controlled for by other authors – e.g. Chaloupka and Laixuthai (1997), Yen (2005) – such as own education, marital status, labor market status, number of children, current living situation, and income are deliberately not used as explanatory variables because of their potential endogeneity. Nevertheless, despite our reservations, we estimated alternative model specifications that include such variables, but it turned out that this does not change our main findings.<sup>16</sup>

Most importantly, parental smoking and drinking habits serve as instruments  $z_{cit}$  and  $z_{ait}$ . Individuals who have already moved out from the 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, four categories are distinguished for each parent: parent drinks (i) (almost) daily, (ii) several times a week, (iii) several times a month, and (iv)

<sup>13</sup>Some 25 695 observations are used for estimating the equations explaining alcohol consumption and 26 353 are used for estimating the equations explaining tobacco consumption. That is, 821 observations cannot be used for estimating the alcohol equation because of missing information on this variable. Yet, as information on tobacco consumption is available for them, they can still be used for estimating the tobacco equation. The opposite applies to 163 observations that can only be used for estimating the alcohol equation.

<sup>14</sup>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; where for Germany 10 grams of alcohol is typically defined a standard drink (DHS, 2003).

<sup>15</sup>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.

<sup>16</sup>In detail the additional controls are the respondent’s own education (four categories, with ‘no educational attainment [yet]’ serving as reference), marital status (indicator) and labor market status (four categories, with ‘not participating’ serving as reference), number of children and monthly net-income (five categories, with ‘income < 1 000 DM’ serving as reference; at survey time the ‘Deutsche Mark’ (DM) was the currency in Germany). Results for this extended model are available upon request.



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

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

(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

Estimation results indicate that consumption of tobacco and alcohol clearly differs between men and women. This does not only refer to the levels of consumption but also to the influence of exogenous variables. In accordance with this, LR-test<sup>17</sup> strongly reject a pooled model. Thus in this section we only report results for separate models for males and females.

### 4.1 Reduced Form Results for the Tobit Specification

The 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  $a_{it}$ . This holds for both, males and females. 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 if one of the parents smokes and – albeit less prominent – even if one of the parents had smoked in the past. The relevance of these direct links is confirmed by formal tests on instrument relevance (Davis and Kim, 2002)<sup>18</sup> and by

<sup>17</sup>For the Tobit model the  $\chi^2(58)$ -statistic for the reduced form takes a value as high as 2978. If smoking and drinking are treated separately, pooled models are rejected for either drug.

<sup>18</sup>For males, the  $\chi^2(1)$ -statistic of the relevant LR-test takes a value of 576.3 concerning parents' drinking habits and 986.6 concerning parents' smoking habits. For these tests Shea-Partial- $R^2$ s are calculated using Tobit pseudo residuals. For females, the corresponding values are 536.6 and 1536.5.

Table 2: Tobit reduced form estimates

	Males				Females			
	Drinking		Smoking		Drinking		Smoking	
	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.
constant	-18.087**	3.310	-40.064**	3.433	-7.900**	1.881	-34.242**	3.239
year 1986	-1.564*	0.628	-3.967**	0.663	-3.698**	0.377	-3.740**	0.647
year 1990	-2.241**	0.515	-5.763**	0.538	-3.372**	0.303	-4.165**	0.509
year 1992	-5.609**	1.363	-7.064**	1.117	-6.943**	0.708	-4.336**	1.044
west	-8.271**	1.169	-2.124*	0.840	-4.853**	0.586	2.278**	0.779
age	2.952**	0.237	3.666**	0.254	1.216**	0.136	2.853**	0.239
age <sup>2</sup> /100	-4.278**	0.445	-6.026**	0.473	-1.606**	0.255	-5.010**	0.445
parents married	-0.407	0.542	-2.001**	0.556	0.015	0.309	-2.139**	0.516
father has low degree	0.760	0.606	-1.244*	0.630	0.650	0.348	0.382	0.593
father has medium degree	-0.126	0.754	-2.017*	0.797	0.945*	0.438	0.505	0.754
father has high degree	-1.307	1.222	-2.440	1.314	1.071	0.708	-1.439	1.241
father has univ. degree	-0.077	0.846	-3.692**	0.892	1.749**	0.478	-1.124	0.832
mother has low degree	-0.591	0.464	-0.449	0.488	0.286	0.270	-0.546	0.460
mother has medium degree	-0.178	0.602	-0.193	0.639	0.250	0.342	-0.878	0.593
mother has high degree	-1.596	1.242	-2.441	1.346	-0.146	0.751	-3.232*	1.376
mother has univ. degree	-1.293	0.986	-2.037	1.051	0.905	0.552	-2.120*	0.967
grown up with mother	-4.071**	1.200	-4.971**	1.216	-0.123	0.682	-4.151**	1.103
grown up with father	-8.252**	2.111	-7.333**	2.169	-0.026	1.279	-3.031	2.068
grown up with both	4.409*	2.154	4.214	2.206	-2.453	1.304	-2.730	2.098
# of children at parents' home	0.338**	0.128	0.764**	0.131	-0.116	0.071	0.538**	0.118
father monthly drinker	1.820**	0.604	-2.007**	0.634	1.413**	0.346	-1.061	0.590
father weekly drinker	4.200**	0.606	-2.080**	0.635	1.863**	0.350	-0.924	0.597
father daily drinker	5.914**	0.597	-1.154	0.620	2.021**	0.339	-1.158*	0.575
mother monthly drinker	2.689**	0.443	-1.093*	0.471	2.505**	0.260	-1.351**	0.452
mother weekly drinker	4.784**	0.584	-0.719	0.627	3.903**	0.344	-0.400	0.599
mother daily drinker	4.291**	0.795	-0.305	0.840	4.230**	0.433	-1.322	0.760
father ex-smoker	1.257**	0.476	3.956**	0.524	0.051	0.277	3.159**	0.502
father smoker	1.209*	0.496	7.411**	0.536	0.296	0.283	6.402**	0.501
mother ex-smoker	0.235	0.566	2.759**	0.598	0.236	0.327	3.241**	0.567
mother smoker	0.764	0.491	6.935**	0.501	0.214	0.285	7.858**	0.468
number of observations	12 922		13 063		12 773		13 290	
LR-statistic	1 085.0		1 177.3		966.0		1 098.5	

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

tests of joint significance of instruments as well.<sup>19</sup>

Furthermore, for both genders, the 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 indeed symmetrically hold for the second equation in our demand system.

With regard to the control variables, the results for the reduced forms exhibit a trend of a decreasing inclination to smoke and drink over time. We also find a significantly positive (but diminishing) correlation with age. Both, women and men living in West Germany drink significantly less than their counterparts from the eastern part of the country. Yet, with respect to tobacco consumption the results offer a more heterogenous picture. While the west-east differential in the inclination to smoke is negative for males, the opposite applies to females. Moreover, for males the estimation results indicate a significantly negative correlation between the propensity to drink or smoke and having grown up with at least one parent. That is males having grown up with other persons exhibit higher consumption levels for both considered psychoactive substances. In terms of point estimates, this pattern is less pronounced for females but it is still statistically significant.

Parental education seems to have a less prominent effect on the offspring’s consumption of psychoactive substances than one may hypothesize. Maternal education is (jointly) never significant at the 5 percent level, while a (jointly) significant effect of paternal education is only found

<sup>19</sup>For males, the F-statistics for the test on joint significance is as high as 126.0 (smoking) and 53.5 (drinking). For females, the corresponding values are 147.2 and 59.3.

for sons' propensity to smoke and daughters' propensity to drink. The number of children at parents' home significantly and positively affects the propensity to consume psychoactive substances, with the exception of female drinking. Parental marital status is significant only for the inclination to smoke, yet for both genders.

## 4.2 Structural Model Results

Table 3 reports the results for the structural equations (1) and (2). For the control variables, the structural estimates confirm by and large the reduced form estimates. Our discussion can therefore concentrate on the parameters of primary interest,  $\gamma_a$  and  $\gamma_c$ . Regarding the effect of smoking on alcohol consumption  $\gamma_a$ , the estimated coefficient is positive for both, men and women. That is, smoking increases the propensity to drink. Thus smoking and drinking are classified as complements in consumption. Yet, in terms of magnitude, the estimate is much larger for men and becomes statistically insignificant for women, while it is highly significant for men. Nevertheless, based on a one-sided test, a negative value for  $\gamma_a$  is still rejected for females at fairly small level of significance ( $p$ -value: 0.70).<sup>20</sup> Turning to the equation explaining smoking behavior, the estimates for  $\gamma_c$  is negative and statistically significant for either gender. This suggests that drinking significantly decreases the propensity to smoke, which would indicate that drinking and smoking are substitutes.

However, we do know that the true parameters  $\gamma_a$  and  $\gamma_c$  need to bear the same sign, opposite to the sign of the Hicksian cross-price derivatives, which are necessarily symmetric. Thus, this asymmetry in estimation results reveals that our identifying assumptions do not apply to both of our equations. In order to gain more insight in which estimate one can trust, we turn to the tests on over-identifying restrictions. According to these tests, the exclusion restrictions are warranted in the equation (1) explaining alcohol consumption for both, men and women. In contrast, for the smoking equation (2) the Newey and McFadden (1994) test rejects the null of valid instruments for both genders at fairly small levels of significance. The intuitive test is less clear. Yet, the  $p$ -values are still rather small.

This casts substantial doubts on the identifying assumption 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.

Turning to the magnitude of the estimates for  $\gamma_a$ , the results indicate substantial heterogeneity across genders. For women, the estimated effect is not only statistically insignificant at the five percent level, it is also negligibly small in economic terms. The point estimate of 0.038 states that one cigarette less per day reduces female daily alcohol consumption by less than 40 milligrams. Keeping in mind that a standard drink is considered to contain 10 grams, the estimated effect is of almost no consequence. For males, the estimated effect is almost four times larger than its counterpart for females, indicating that one cigarette less per day results in a reduction in the inclination to drink by roughly 140 milligram of alcohol per day.<sup>21</sup> Therefore, the complementarity between smoking and drinking seems to be a predominately male phenomenon. But still, even for males the estimated effect of one cigarette less represents merely one-hundredth of a half-pint of beer. Thus, the reduction of drinking levels that will result from successful anti-smoking policies is likely to be rather moderate. Nevertheless, the estimates for  $\gamma_a$  argue against an unin-

<sup>20</sup>For the specification that uses an extended list of covariates the estimate for  $\gamma_a$  is positive and statistically significant for both, men and women.

<sup>21</sup>The implicit assumption that the size of this effect does not depend on consumption levels is quite strong. We therefore estimated extensions to the basic model that allow for consumption-dependent effects. Yet, due to the lack of additional valid instruments, the corresponding estimation results did not provide further insights.

Table 3: Tobit structural form estimates

	Males				Females			
	Drinking		Smoking		Drinking		Smoking	
	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.
$\gamma_a$ fitted smoking	0.135**	0.040	-	-	0.037	0.025	-	-
$\gamma_c$ fitted drinking	-	-	-0.159**	0.061	-	-	-0.337**	0.093
constant	-12.868**	3.397	-42.953**	3.761	-6.608**	1.926	-37.011**	3.757
year 1986	-1.048	0.605	-4.218**	0.678	-3.557**	0.388	-4.963**	0.741
year 1990	-1.494**	0.523	-6.118**	0.573	-3.213**	0.370	-5.264**	0.619
year 1992	-4.696**	1.180	-7.956**	1.176	-6.773**	0.673	-6.637**	1.171
west	-8.003**	0.928	-3.468**	0.972	-4.933**	0.485	0.674	0.886
age	2.467**	0.260	4.134**	0.318	1.109**	0.152	3.265**	0.293
age <sup>2</sup> /100	-3.472**	0.493	-6.697**	0.555	-1.418**	0.279	-5.553**	0.526
parents married	-0.034	0.599	-2.143**	0.554	0.087	0.310	-2.121**	0.549
father has low degree	0.940	0.647	-1.265	0.705	0.634	0.330	0.583	0.615
$\beta$ father has medium degree	0.150	0.826	-2.233*	0.920	0.921*	0.423	0.809	0.794
father has high degree	-0.972	1.101	-2.771	1.472	1.119	0.672	-1.077	1.413
father has univ. degree	0.432	0.950	-3.901**	0.991	1.787**	0.456	-0.504	0.899
mother has low degree	-0.536	0.497	-0.595	0.488	0.307	0.289	-0.457	0.471
mother has medium degree	-0.159	0.612	-0.287	0.701	0.286	0.339	-0.796	0.634
mother has high degree	-1.249	1.242	-2.696	1.422	-0.015	0.861	-3.159	1.695
mother has univ. degree	-1.020	0.993	-2.240	1.157	0.993	0.566	-1.846*	0.905
grown up with mother	-3.429*	1.487	-5.771**	1.259	0.017	0.693	-4.177**	1.186
grown up with father	-6.940**	1.993	-9.330**	2.460	0.087	1.724	-3.346	2.060
grown up with both	3.775	2.137	4.878	2.567	-2.335	1.741	-3.616	2.038
# of children at parents' home	0.232	0.151	0.836**	0.150	-0.136	0.071	0.499**	0.125
father monthly drinker	2.108**	0.511	-	-	1.454**	0.326	-	-
father weekly drinker	4.529**	0.522	-	-	1.898**	0.355	-	-
$\delta_a$ father daily drinker	6.122**	0.524	-	-	2.070**	0.337	-	-
mother monthly drinker	2.813**	0.450	-	-	2.556**	0.266	-	-
mother weekly drinker	4.855**	0.577	-	-	3.918**	0.359	-	-
mother daily drinker	4.278**	0.917	-	-	4.272**	0.486	-	-
father ex-smoker	-	-	4.207**	0.580	-	-	3.139**	0.522
$\delta_c$ father smoker	-	-	7.697**	0.568	-	-	6.458**	0.513
mother ex-smoker	-	-	2.825**	0.614	-	-	3.345**	0.590
mother smoker	-	-	7.092**	0.561	-	-	7.962**	0.445
number of observations	12 922		13 063		12 773		13 290	
LR-statistic	1 082.2		1 157.9		965.6		1 091.2	
tests for over-identifying restrictions (p-values):								
intuitive	0.447		0.155		0.939		0.397	
Newey & McFadden	0.379		0.000		0.921		0.062	

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

tended side-effect. That is, effective anti-smoking policies will most likely not result in an increase of the consumption of alcohol, but rather tend to improve population health on several margins simultaneously.

### 4.3 Results for the Conditional Linear Model

As an alternative to the restrictive Tobit approach, in this section we shortly discuss results for the conditional linear model; see section 2. Indeed, Vuong (1989) tests favor this less restrictive model against the Tobit specification. Table 4 displays estimated coefficients for the structural model. For the reduced form estimates, see Appendix B, Table 7. As for the Tobit specification, the Newey and McFadden (1994) test for over-identifying restrictions argues against parental drinking habits representing valid instruments for alcohol consumption, while warranting the exclusion restrictions for the other structural equation. That is the test supports the identifying assumptions with regard to the effect of smoking on drinking i.e.  $\gamma_a$ . Moreover, for males the point estimate for  $\gamma_a$  is very close to its counterpart obtained from Tobit estimation. Yet, the associated standard error is much larger. This does not come as a surprise as effectively a much smaller sample is used for estimating this model. This is because only individuals who drink and smoke and for which quantitative information on the consumption of either drug is available enter the estimation sample. For females the point estimate for  $\gamma_a$  is virtually zero and also associated by a large standard

Table 4: Structural form estimates for the conditional linear model

	Males				Females			
	Drinking		Smoking		Drinking		Smoking	
	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.
$\gamma_a$ fitted smoking	0.115	0.321	-	-	-0.008	0.205	-	-
$\gamma_c$ fitted drinking	-	-	0.056	0.042	-	-	0.037	0.144
constant	0.135	7.687	-12.844**	2.245	0.478	5.208	-5.767*	2.784
year 1986	1.441	1.111	-0.799	0.422	-0.912	0.859	0.616	0.629
year 1990	1.231	0.891	-1.071**	0.323	-0.206	0.730	-0.280	0.447
year 1992	8.736**	3.107	-1.264	0.709	6.346**	1.380	-1.089	1.267
west	3.739	3.125	2.187**	0.499	7.524**	1.278	3.570**	1.288
age	0.883	0.771	1.916**	0.179	0.128	0.471	1.334**	0.222
age <sup>2</sup> /100	-0.757	1.266	-2.837**	0.332	0.291	0.810	-2.073**	0.421
parents married	-0.934	1.027	-1.049**	0.357	-0.141	0.670	-1.260**	0.443
father has low degree	-0.105	1.320	-0.220	0.432	0.287	0.806	-0.357	0.596
$\beta$ father has medium degree	-0.821	1.498	-0.646	0.535	0.531	1.050	-1.435*	0.662
father has high degree	-0.431	1.912	-0.998	0.956	1.384	2.060	-2.335*	1.136
father has univ. degree	-0.798	1.697	-1.308*	0.575	1.736	1.212	-1.334	0.873
mother has low degree	-1.248	0.811	-0.515	0.288	-0.271	0.664	-0.427	0.400
mother has medium degree	0.760	1.141	-0.277	0.431	-0.377	0.724	0.178	0.494
mother has high degree	-0.506	2.189	-1.512	0.995	2.964	2.276	3.577*	1.480
mother has univ. degree	1.068	1.953	-1.162	0.679	2.139	1.943	-0.480	0.969
grown up with mother	-4.339	2.496	-0.862	0.867	-0.060	1.478	-3.338**	1.009
grown up with father	-9.610**	3.288	-0.617	1.463	0.641	4.646	-2.404	1.995
grown up with both	8.372*	3.479	0.074	1.507	-2.229	4.524	2.710	1.988
# of children at parents' home	0.605*	0.240	0.212*	0.091	-0.162	0.164	0.109	0.124
father monthly drinker	0.124	0.976	-	-	-0.175	0.955	-	-
father weekly drinker	3.761**	1.005	-	-	0.030	0.951	-	-
$\delta_a$ father daily drinker	5.686**	1.079	-	-	0.511	0.907	-	-
mother monthly drinker	2.373**	0.737	-	-	0.897	0.607	-	-
mother weekly drinker	4.266**	1.003	-	-	2.471**	0.872	-	-
mother daily drinker	4.391*	1.860	-	-	3.846**	0.964	-	-
father ex-smoker	-	-	0.012	0.386	-	-	1.318**	0.449
$\delta_c$ father smoker	-	-	1.099**	0.350	-	-	1.733**	0.450
mother ex-smoker	-	-	1.043*	0.416	-	-	0.475	0.452
mother smoker	-	-	1.956**	0.319	-	-	2.282**	0.392
number of observations	4212		4313		2607		2627	
F-statistic	9.65		25.39		3.75		12.23	
tests for over-identifying restrictions (p-values):								
intuitive	0.907		0.101		0.821		0.709	
Newey & McFadden	0.410		0.122		0.638		0.225	

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

error. In consequence, the structural form coefficients  $\gamma$  are all statistically insignificant and even one-sided tests are inconclusive with respect to the direction of the effect.<sup>22</sup> In essence, while the results for the conditional linear model do not contradict those obtained from Tobit estimation, due to statistical insignificance, they either cannot confirm them. Thus our main finding still rests on a rather restrictive specification.

## 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 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 behav-

<sup>22</sup>As a single exception, in the males' sample one may reject a negative value for  $\gamma_c$  with a  $p$ -value of 0.089. Yet, the test of over-identifying restrictions argues against  $\gamma_c$  being identified by exclusion restrictions.

ior 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. Yet, this result seems only to be relevant for males, while for females the estimated effect is very small and statistically insignificant. 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 through 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. In fact, the reverse, i.e. a moderate reduction in the consumption of alcohol, seems to be the consequence. Yet, this result has to be interpreted with some caution, though, as it rests on relatively old data and a restrictive empirical approach.

## References

- Ahlfeldt, G. and Maenning, W. (2010). Impact of non-smoking ordinances on hospitality revenues: the case of Germany, *Jahrbücher für Nationalökonomie und Statistik – Journal of Economics and Statistics* **230**: 506–521.
- 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 Melkersen, 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.
- Cox, T. L. and Wohlgenant, M. K. (1986). Prices and quality effects in cross-sectional demand analysis, *American Journal of Agricultural Economics* **68**: 908–919.
- 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.
- DHS (ed.) (2003). *Alkoholabhängigkeit*, Suchtmedizinische Reihe Band 1, Wissenschaftliches Kuratorium der Deutschen Hauptstelle für Suchtfragen (DHS) e.V., Hamm.
- DiNardo, J. and Lemieux, T. (2001). Alcohol, marijuana, and American youth: the unintended effects of government regulation, *Journal of Health Economics* **20**: 991–1010.
- Goel, R. K. and Morey, M. J. (1995). The interdependence of cigarette and liquor demand, *Southern Economic 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.
- Hoderlein, S. and Mihaleva, S. (2008). Increasing the price variation in a repeated cross section, *Journal of Econometrics* **147**: 316–325.
- 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.
- Kvasnicka, M. and Tauchmann, H. (2012). Much Ado About Nothing? Smoking Bans and Germany's Hospitality Industry, *Applied Economics* **44**: 4539–4551.
- Labeaga, J. M. (1999). A double-hurdle rational addiction model with heterogeneity: estimating the demand for tobacco, *Journal of Econometrics* **93**: 49–72.
- Lee, Y. G. and Abdel-Ghany, M. (2004). American youth consumption of licit and illicit substances, *International Journal of Consumer Studies* **28**: 454–465.
- MacDonald, Z. and Shields, M. A. (2001). The impact of alcohol consumption on occupational attainment in England, *Economica* **68**: 427–453.
- 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. and McFadden, D. (1994). Large sample estimation and hypothesis testing, Vol. 4 of *Handbook of Econometrics*, 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.
- Pierani, P. and Tiezzi, S. (2009). Addiction and interaction between alcohol and tobacco consumption, *Empirical Economics* **37**: 1–23.
- Poirier, D. J. (1980). Partial observability in bivariate probit models, *Journal of Econometrics* **12**: 209–217.
- Rivers, D. and Vuong, 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.-J. and Yen, S. T. (2000). A censored system of cigarette and alcohol consumption, *Applied Economics* **32**: 729–737.
- Vuong, Q. H. (1989). Likelihood ratio test for model selection and non-nested hypothesis, *Econometrica* **57**: 307–333.
- 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**.
- Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge Massachusetts.
- World Health Organization (ed.) (2008). *WHO report on the global tobacco epidemic, 2008: The MPOWER package.*, WHO Press, Geneva.
- 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  $\gamma_a$  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  $\gamma_c$ . However, in micro-economic theory, complementarity is defined in terms of cross-price effects on Hicksian, i.e. compensated, demand. Though Hicksian demand is a theoretical concept that cannot be observed directly, it allows for disentangling pure substitution effects from income effects.<sup>23</sup> In this appendix, we show that the cross-price effect of increasing the price of tobacco on the Hicksian demand for alcohol always has the opposite sign of the effect resulting from increasing the consumption of tobacco on the Marshallian demand for alcohol. For this reason, the measure of complementarity that is used in this analysis corresponds with the standard definition of complementarity in qualitative terms.

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  $t$  that denote specific individuals and periods are skipped. The corresponding prices are  $p_a$ ,  $p_c$ , and  $p_w$ . Hicksian demand for alcohol is written as  $a^H(p_a, p_c, p_w, \bar{U})$ , for some fixed utility level  $\bar{U}$ . Accordingly, the restricted Marshallian demand for alcohol, if the consumption of tobacco  $\bar{c}$  is given, is denoted by  $a^M(p_a, p_c, p_w, \bar{c}, y)$  where  $y$  is income. We now state the following result:

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

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

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

$$U(a, c, w) = \bar{U}. \quad (6)$$

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

$$U_a(a, c, w) = \lambda^{-1} p_a \equiv \mu p_a \quad (7)$$

$$U_c(a, c, w) = \lambda^{-1} p_c \equiv \mu p_c \quad (8)$$

$$U_w(a, c, w) = \lambda^{-1} p_w \equiv \mu p_w \quad (9)$$

where  $U_a$ ,  $U_c$ , and  $U_w$  are partial derivatives of  $U(\cdot)$  and  $\lambda$ , with  $\lambda > 0$ , is the Lagrange 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) through (9) and (6) totally with respect to  $p_c$  to obtain:

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

where we have made use of (7) through (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, since they are of no further interest here):

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

<sup>23</sup>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.

where the denominator  $D$  is given by

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

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  $a^M(p_a, p_c, p_w, \bar{c}, y)$  which by definition is the solution of  $\max_{a, c, w} U(a, c, w)$  subject to

$$p_aa + p_cc + p_w w = y \quad (12)$$

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

$$U_a(a, \bar{c}, w) = \mu p_a \quad (13)$$

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

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

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

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

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

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

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

$$\frac{\partial a^H}{\partial p_c} = \frac{\partial a^M}{\partial \bar{c}} \frac{\mu D'}{D} \quad (16)$$

establishing (5).

## B Supplementary Tables

Table 5: Description of dependent variables

	Mean	Std. Dev.	Number of obs.
All observations			
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	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
<i>year 1980</i>	0.187	0.390	0.193	0.395	0.180	0.384
<i>year 1986</i>	0.128	0.334	0.133	0.340	0.123	0.328
<i>year 1990</i>	0.581	0.493	0.573	0.495	0.590	0.492
<i>year 1992</i>	0.104	0.305	0.101	0.302	0.107	0.309
<i>west</i>	0.838	0.368	0.844	0.362	0.832	0.374
<i>female</i>	0.503	0.500	-	-	-	-
<i>age</i>	24.310	6.297	24.213	6.275	24.421	6.317
<i>age<sup>2</sup> / 100</i>	6.310	3.351	6.256	0.064	6.363	3.366
<i>parents married</i>	0.807	0.395	0.811	0.392	0.805	0.396
<i>father has no school degree</i>	0.139	0.346	0.138	0.344	0.140	0.347
<i>father has a low degree</i>	0.531	0.499	0.533	0.499	0.529	0.499
<i>father has a medium degree</i>	0.156	0.363	0.159	0.366	0.151	0.358
<i>father has a high degree</i>	0.032	0.175	0.032	0.176	0.032	0.176
<i>father has a university degree</i>	0.142	0.349	0.138	0.345	0.148	0.355
<i>mother has no school degree</i>	0.309	0.462	0.304	0.460	0.315	0.464
<i>mother has a low degree</i>	0.416	0.493	0.424	0.494	0.408	0.491
<i>mother has a medium degree</i>	0.191	0.393	0.189	0.391	0.193	0.395
<i>mother has a high degree</i>	0.024	0.154	0.025	0.157	0.023	0.150
<i>mother has a university degree</i>	0.059	0.236	0.058	0.234	0.061	0.240
<i>grown up with mother</i>	0.953	0.212	0.954	0.211	0.952	0.214
<i>grown up with father</i>	0.891	0.312	0.891	0.312	0.890	0.312
<i>grown up with both</i>	0.881	0.324	0.881	0.324	0.882	0.323
<i># of children at parents' home</i>	2.752	1.462	2.732	1.434	2.772	1.488
<i>father never drinker</i>	0.186	0.389	0.180	0.384	0.193	0.395
<i>father monthly drinker</i>	0.236	0.424	0.236	0.425	0.235	0.424
<i>father weekly drinker</i>	0.259	0.438	0.268	0.443	0.250	0.433
<i>father daily drinker</i>	0.319	0.466	0.316	0.465	0.323	0.468
<i>mother never drinker</i>	0.507	0.500	0.499	0.500	0.514	0.500
<i>mother monthly drinker</i>	0.285	0.452	0.292	0.455	0.280	0.449
<i>mother weekly drinker</i>	0.136	0.342	0.141	0.348	0.130	0.336
<i>mother daily drinker</i>	0.072	0.259	0.068	0.251	0.076	0.266
<i>father never smoker</i>	0.276	0.447	0.275	0.447	0.276	0.447
<i>father ex-smoker</i>	0.351	0.477	0.357	0.479	0.345	0.475
<i>father smoker</i>	0.373	0.484	0.368	0.482	0.379	0.485
<i>mother never smoker</i>	0.674	0.469	0.676	0.468	0.673	0.469
<i>mother ex-smoker</i>	0.126	0.331	0.128	0.334	0.124	0.330
<i>mother smoker</i>	0.200	0.400	0.197	0.397	0.203	0.402

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

Table 7: Reduced form estimates for the conditional linear model

	Males				Females			
	Drinking		Smoking		Drinking		Smoking	
	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.
constant	-1.397	6.910	-12.838**	2.360	0.320	5.009	-5.812*	2.922
year 1986	1.370	1.103	-0.724	0.436	-0.865	0.860	0.650	0.553
year 1990	1.044	0.912	-1.054**	0.357	-0.202	0.661	-0.297	0.421
year 1992	8.519*	3.756	-0.834	0.716	6.312*	2.687	-0.894	0.918
west	3.880	3.540	2.451**	0.537	7.496**	2.514	3.884**	0.692
age	1.129*	0.455	1.975**	0.176	0.124	0.332	1.344**	0.215
age <sup>2</sup> /100	-1.122	0.848	-2.897**	0.326	0.307	0.620	-2.074**	0.398
parents married	-1.261	0.942	-1.097**	0.352	-0.022	0.711	-1.198**	0.440
father has low degree	-0.182	1.087	-0.179	0.414	0.317	0.828	-0.308	0.524
father has medium degree	-1.031	1.372	-0.614	0.537	0.606	1.015	-1.300*	0.652
father has high degree	-0.606	2.258	-0.955	0.911	1.522	1.777	-2.303*	1.108
father has univ. degree	-1.092	1.578	-1.220*	0.605	1.801	1.155	-1.193	0.743
mother has low degree	-1.355	0.832	-0.557	0.323	-0.305	0.632	-0.401	0.398
mother has medium degree	0.689	1.091	-0.154	0.432	-0.432	0.808	0.211	0.527
mother has high degree	-0.939	2.357	-1.502	0.929	2.910	1.863	3.646**	1.197
mother has univ. degree	0.932	1.917	-1.071	0.726	2.146	1.368	-0.388	0.872
grown up with mother	-4.799*	2.021	-0.950	0.772	-0.056	1.527	-3.097**	0.947
grown up with father	-10.072*	3.981	-1.250	1.433	1.019	3.012	-2.005	1.761
grown up with both	8.755*	3.984	0.568	1.439	-2.199	3.024	2.456	1.768
# of children at parents' home	0.657**	0.227	0.240**	0.084	-0.166	0.172	0.085	0.103
father monthly drinker	0.131	1.114	-0.084	0.426	-0.170	0.859	-0.746	0.542
father weekly drinker	3.827**	1.117	0.042	0.428	0.102	0.867	-0.245	0.551
father daily drinker	5.869**	1.092	0.797	0.414	0.598	0.833	0.147	0.522
mother monthly drinker	2.310**	0.804	-0.393	0.312	0.855	0.603	-0.534	0.384
mother weekly drinker	4.091**	1.067	-0.422	0.426	2.408**	0.783	-0.584	0.497
mother daily drinker	4.213**	1.427	0.072	0.556	3.768**	0.960	0.077	0.640
father ex-smoker	1.107	0.941	0.015	0.375	-0.910	0.705	1.222**	0.468
father smoker	-0.203	0.949	0.980**	0.372	-0.510	0.686	1.601**	0.452
mother ex-smoker	0.918	1.036	1.126**	0.405	-0.119	0.743	0.451	0.484
mother smoker	0.857	0.826	2.042**	0.312	0.264	0.601	2.339**	0.378
number of observations	4 212		4 313		2 607		2 627	
F-statistic	8.81		21.32		3.43		10.42	

Notes: \*\* significant at the 1% level; \* significant at the 5% level.