THE UNIVERSITY of York

CENTRE FOR HEALTH ECONOMICS

Are Frequency and Intensity of Participation Utility-Bearing Aspects of Consumption? An Analysis of Drinking Behaviour

Fredrik Berggren
Matthew Sutton

CHE Technical Paper Series 1

ARE FREQUENCY AND INTENSITY OF PARTICIPATION UTILITY-BEARING ASPECTS OF CONSUMPTION? AN ANALYSIS OF DRINKING BEHAVIOUR

Fredrik Berggren
Matthew Sutton

October 1996

CENTRE FOR HEALTH ECONOMICS TECHNICAL PAPER SERIES

The Centre for Health Economics has a well established Discussion Paper series which was originally conceived as a means of circulating ideas for discussion and debate amongst a wide readership that included health economists as well as those working within the NHS and pharmaceutical industry.

The introduction of a Technical Paper Series offers a further means by which the Centre's research can be disseminated. The Technical Paper Series publishes papers that are likely to be of specific interest to a relatively specialist audience, for example papers dealing with complex issues that assume high levels of prior knowledge, or those that make extensive use of sophisticated mathematical or statistical techniques.

The series provides a maximum of 20 pages for authors. The content and its format are entirely the responsibility of the author, and papers published in the Technical Paper series are not subject to peer-review or editorial control, unlike those produced in the Discussion Paper series. Offers of further papers, and requests for information should be directed to Frances Sharp in the Publications Office, Centre for Health Economics, University of York.

THE AUTHORS

Fredrik Berggren is a Research Student in Health Economics at the Department of Community Medicine, MAS and in the Department of Economics at Lund University, Sweden. Matthew Sutton is currently a Research Fellow at the Centre for Health Economics, University of York. The research was undertaken whilst Matthew Sutton was a Visiting Research Fellow in Health Economics at the Department of Community Medicine, MAS, Lund University.

ACKNOWLEDGEMENTS

This research was funded by the National Institute of Public Health, Sweden (Folkhalsoinstitutet). We are indebted to Bertil S. Hanson for providing the data set, and to colleagues in the Department of Economics (Lund University) and CHS (Odense University), Andrew Jones (York University) and an internal referee for suggestions on previous drafts. An earlier version of this paper was presented at the *International Health Economics Association* Inaugural Conference in May 1996 and has been published in the *Studies in Health Economics* series from Lund University.

Price £4.00

ABSTRACT

The subject of most economic analyses of drinking behaviour is per-period consumption or expenditure. In this paper, we show that this 'quantity-hypothesis' imposes restrictions on the nature of individuals' choices between how often to drink (frequency) and how much to drink on each occasion (intensity). Under the 'quantity-hypothesis', standard assumptions about the utility function give rise to a convex consumption possibilities frontier along which individuals will position randomly. However, we suggest that the systematic selection of particular optima may be determined by the differential effects of the costs of acquisition on frequency and intensity, differential rates of time preference and demand for health, or varying preferences for the multiple characteristics of alcohol consumption. We propose a test of the 'quantity-hypothesis', based on analysing structural change between the parameters estimated for frequency and intensity, which allows for self-selection, heteroskedastic error terms and dependence between the parameters estimated. Using cross-sectional data, we find evidence to reject the 'quantity-hypothesis' restrictions on the determinants of drinking patterns of the general public. Allowing for simultaneity, our results suggest a positive effect of income on participation and a negative effect on intensity.

INTRODUCTION

Economic analyses of drinking behaviour based on individual-level data have become increasingly common (Leung and Phelps, 1993; Godfrey, 1994) and the main focus has turned towards more complex measures of drinking patterns than per-period consumption or expenditure. Commonly, questions about drinking behaviour relate to two dimensions (Sindelar, 1993): the number of drinking episodes in a defined period - *frequency*; and the average number of drinks consumed in each drinking episode - *intensity*.[1]

In terms of policy analysis, there is a clear rationale for disaggregating consumption along these two dimensions. There is growing evidence of a significant association between episodes of intoxication and adverse effects, such as road traffic accidents and violent crime (Edwards et al, 1994; Stockwell et al, in press). Health advice also relates specifically to intensity and frequency. For example, new guidelines for the UK recommend a 48-hour drink-free period to allow the body to recover following an episode of intense drinking (*Sensible Drinking*, 1995). Therefore, there has emerged a common focus on investigating the determinants of both high-quantity and intense drinking.

However, there has been less agreement on the specification of these variables. Many studies define 'more than *five* drinks' as an occasion of heavy drinking (Manning et al, 1995; Chaloupka and Wechsler, 1995), although Moore and Cook (1995) adopt *six* as the cut-off value. In some analyses, one occasion of heavy drinking is sufficient to be registered as a binge-drinker, in others a threshold number of occasions is required (Moore and Cook, 1995) and in others, the number of binge-episodes is retained as the dependent variable (Manning et al, 1995). Recently, Grossman, Chaloupka and Sirtalan (1995) have analysed a binary indicator of frequent drinking, defined as more than forty drinking occasions in the last year. In another study, drinking levels were specified using the risk-levels adopted for health advice and policy targets (Sutton and Godfrey, 1995).

In studies in which the response variable is dichotomous, there has been little controversy over the appropriate model specification, and there are examples of the use of probit (Chaloupka and Wechsler, 1995), logit (Moore and Cook, 1995) and linear probability model (Grossman et al, 1995) specifications. In contrast, for multiple-response indicators of drinking levels (including continuous consumption measures), various econometric approaches have been used, including: multinomial logit (Grossman, Coate and Arluck, 1987; Coate and Grossman, 1988); ordered probit (Chaloupka and Wechsler, 1995); grouped data (Sutton and Godfrey, 1995); quantile (Manning et al, 1995); and tobit (Atkinson et al, 1990) regression specifications. The aim is to permit responsiveness to economic factors to vary by drinking levels, and in each of these studies, results have been derived for comparison of the relative price-responsiveness of heavy versus moderate drinking. Most studies have found heavy drinking (however defined) to be more price-responsive, but Manning et al (1995) is a recent exception.

---

[1] The terms *frequency* and *intensity* which we adopt in this paper differ from those used elsewhere. We adopt the term *intensity* for 'quantity per episode' to distinguish this from our use of the term *quantity* to represent the total amount (or *volume*) consumed over the observation period. In these terms, *quantity* equals the product of *frequency* and *intensity* and expenditure is the product of the unit price of alcohol, *frequency* and *intensity*.

This qualitative difference in conclusion may relate to the various restrictions inherent in different econometric specifications.[2] However, the specification of the relationship between participation, frequency and intensity and the nature of the dependent variable may introduce an additional source of inconsistency in results across studies. Investigation of the relationships between participation, frequency and intensity has been proposed as a research priority (Leung and Phelps, 1993), and in this paper we focus on the relationship between frequency and intensity.[3]

The increased popularity of modelling different aspects of drinking patterns seems to be aimed at enhancing the policy relevance of econometric analyses rather than the descriptive validity of economic models of drinking behaviour. Conceptual analysis of the relationships between intensity and frequency can offer guidance on the modelling process. For example, in this paper, we model the benefits of alcohol with a Cobb-Douglas function,[4] which includes frequency and intensity as distinct, utility-bearing aspects of drinking behaviour. Demand functions are derived which suggest that optimal levels of frequency and intensity will be inter-dependent, as they enter the budget constraint as a multiplicative term. Therefore, studies which treat them as independent are at risk of misspecification.

In this framework, analysis of quantity can be expressed as the imposition of a testable restriction on the utility function. Under the restriction, combinations of frequency and intensity will be selected at random. We suggest that there are various conceptual reasons why frequency-intensity bundles will not be selected at random, including differences in acquisition costs. Comparison with data on drinking patterns suggests that this prediction of random choices may be a deficiency in existing economic approaches and that an unrestricted approach may enhance the descriptive ability of economic models of drinking behaviour. In addition, our approach could be used to test whether other combinations of intensity and frequency, such as those that have been used elsewhere in the literature, impose valid restrictions on choices.

A further bi-product of this work relates to the contention over the appropriate time-period for measuring consumption, which may vary from seven days to one year (Sindelar, 1993). As Sindelar (1993) notes, measurement over short time-periods may not reflect 'usual' drinking patterns while recall over longer time-periods is more prone to inaccuracy or bias. To this balance, we would like to add a further condition for the appropriate time-period over which to model *quantity*. According to our definitions, intensity is simply 'quantity per drinking-day'. We show that appropriateness of the period of analysis can be seen as the validity of a restriction of unitary elasticity of substitution between frequency and intensity, and present several methods for testing this restriction. Using this methodology, we show that the restriction implied by modelling quantity over a thirty-day period is rejected for the spirit drinking behaviour of a sample of the Swedish general public.

The paper is organised as follows. In the first section, we emphasise several benefits of considering frequency and intensity as distinct aspects of drinking behaviour. In the second

[2] Manning et al (1995), for example, suggest that Atkinson et al's (1990) findings are driven *completely* by assumed functional form.

[3] We consider participation in our analysis by allowing for self-selection using a Heckman two-step procedure. Participation has been considered more extensively for alcohol by Atkinson et al (1990) and for tobacco by Jones (1989).

[4] The Cobb-Douglas function provides a simple and often used framework for alcohol demand studies in terms of the frequently adopted log-log functional form (Godfrey, 1988; Leung and Phelps, 1993).

section, we derive demand functions for frequency and intensity and demonstrate the restrictions implied by standard modelling of per-period consumption - the 'quantity-hypothesis'. We discuss in the third section the econometric techniques used to test the restriction and to estimate influences on frequency and intensity when they are simultaneously determined. The fourth section contains a description of the data set used in this analysis. The results are presented in the fifth section and discussed in a final section.

FREQUENCY AND INTENSITY: DISTINCT ASPECTS OF CONSUMPTION?

There are several reasons why analysing drinking behaviour disaggregated into the frequency and intensity of participation episodes may be useful: to explore the theory of consumer behaviour applied to alcohol; to enhance the descriptive validity of economic models of drinking behaviour; and to contribute to policy analysis.

Consumer behaviour theory[5]

The conceptual motivations for disaggregating drinking behaviour relate to the range of characteristics of alcohol consumption, the demand for health-detrimental activities and acquisition costs of different drinking patterns.

Alcohol consumption may serve various purposes, such as intoxication, refreshment, relaxation, socialising, and taste. Each of these characteristics of consumption may be produced differentially by the intensity and frequency of drinking episodes. Therefore, consumers of alcohol who drink for different reasons will select different frequency and intensity bundles *systematically*, depending on the motivations for their consumption.

In addition, if the detrimental health effects of drinking are related mainly to intensity then we would expect lower income, younger and less educated individuals, with expected higher rates of time preference and lower demand for health investments (Becker, Grossman and Murphy, 1991), to be more likely to select more intense and less frequent consumption bundles.

Acquisition costs may be significant elements of shadow prices which differ significantly between the frequency and intensity of drinking episodes. For each drinking session we may define three types of costs; transaction costs, drinking-time costs, and recovery-time costs, in which recovery time costs relate to the time period over which an individual's usual activities (including paid work) are impaired following a drinking episode. Each of these types are variable costs in terms of frequency.[6] Transactions costs can be assumed to be fixed in terms of intensity. Drinking-time and recovery-time costs, on the other hand, will *vary* according to intensity. Moreover, we may expect recovery costs to be a non-linear positive function of

[5] As is traditional in economic studies of alcohol consumption, we assume a rational choice framework. Recent developments in economic theories of addiction have shown how addiction may be viewed as a rational process (Becker and Murphy, 1988). Moreover, alcoholics are unlikely to be represented in our data-set, given that it is a postal survey of individuals in private households.

[6] Transactions costs may not be variable in terms of frequency for off-site consumption episodes, where individuals are able to make bulk-frequency purchases in advance. For off-site consumption episodes, therefore, the economies of scale achieved by 'bingeing' are reduced.

intensity.[7] Aggregating these three elements of costs, we would observe economies of scale in intensity in acquisition costs initially, followed by diseconomies of scale.

Acquisition costs of frequency and intensity can be used to predict differential drinking patterns between income groups. Given diminishing marginal utility of income and opportunity costs of leisure time increasing in income, those with high income are likely to value the time elements of acquisition costs more highly than those on low income, the opposite being the case for money prices. High income groups are therefore less likely to drink at levels of intensity which create substantial recovery-time. For those on low incomes, money costs will dominate, which are assumed to be a linear function of intensity.[8]

Descriptive validity

A preliminary analysis of the spirit drinking data from *The Malmö Health Survey 1994* (Lindström et al, 1994) demonstrates the potential 'added value' of a frequency-intensity approach. Various aspects of the spirit drinking patterns of different social groups are shown in Table 1. The percentage of respondents reporting consumption of spirits in the last 30 days (the participation rate) is substantially different between males and females but shows no clear relationship with age. In contrast, income and education appear to be positively correlated with the participation rate.

It can be seen from Table 1 that aggregation of frequency and intensity data into a quantity-per-month variable obscures clear clustering of drinking patterns in different age groups. There is no obvious age-related trend in the quantity figures, and yet frequency and intensity appear to exhibit positive and negative correlations with age, respectively. Moreover, examination of the drinking patterns of different income groups lends further support to the hypothesis that frequency and intensity decisions are distinct processes. The J-shaped relationship between income and quantity appears to be generated by the conflicting correlations of income with frequency and intensity. As income increases from a low level, the negative effect of intensity on quantity dominates. At higher levels of income, the positive effect of frequency on quantity is stronger. The resolution of two competing effects in the quantity variable suggests that increased descriptive power may be achieved by modelling frequency and intensity as two distinct but dependent processes.

[7] Here, we consider recovery costs as private costs but Manning et al (1995) have made asimilar hypothesis about the nature of the costs imposed on others as regards intensity: "external costs are believed to be increasing at an increasing rate in the amount consumed in a single setting" (p.139).
[8] Although they may not be perceived as such at the time! A referee has highlighted stylised evidence which suggests that, for some individuals, the marginal utility of income may be decreasing in increasing levels of intensity.

Table 1: Patterns of spirit drinking by social group in Malmö

Group	Participation Rate, % (N)	Quantity, cl per month (N)	Frequency, drinking-days per month (N)	Intensity, cl per drinking-day (N)
Gender:				
Female	29.0 (1655)	32.8 (460)	3.0 (480)	10.4 (460)
Male	57.1 (1675)	69.9 (930)	4.3 (957)	16.4 (930)
Age:				
21 years	44.6 (408)	66.4 (176)	2.4 (182)	22.4 (176)
26 years	42.4 (425)	35.7 (177)	2.2 (180)	15.8 (177)
31 years	40.7 (369)	66.8 (147)	3.0 (150)	18.3 (147)
41 years	39.8 (415)	71.8 (161)	3.5 (165)	14.6 (161)
51 years	52.8 (460)	54.2 (235)	3.7 (243)	12.8 (235)
61 years	45.0 (449)	56.7 (198)	4.7 (202)	12.1 (198)
71 years	42.4 (458)	49.8 (185)	4.7 (194)	10.5 (185)
81 years	34.9 (341)	68.5 (110)	7.4 (119)	9.0 (110)
Income (SEK/month):				
6999 or less	32.7 (671)	100.3 (203)	3.5 (220)	22.8 (203)
7000 - 9999	39.4 (639)	59.6 (243)	3.8 (252)	15.9 (243)
10000 - 13999	42.4 (661)	46.4 (270)	3.8 (280)	12.1 (270)
14000 - 20999	51.6 (754)	42.4 (382)	3.7 (389)	12.5 (382)
21000 or more	57.9 (409)	58.9 (236)	4.6 (237)	11.3 (236)
Education:				
Compulsory	36.6 (1022)	77.2 (358)	4.7 (375)	15.0 (358)
Vocational	44.8 (699)	59.4 (298)	3.9 (313)	14.5 (298)
High School	47.0 (704)	58.7 (321)	3.2 (331)	17.4 (321)
University	46.2 (905)	38.6 (413)	3.5 (418)	11.7 (413)

Policy analysis

Economic research has considered various concerns about policy on the use and abuse of alcoholic beverages, such as minimum drinking age, control of availability, advertising, taxation, evaluation of treatment, demand and social costs (Godfrey, 1994; Berggren and Lindgren, 1995). Most research is conducted on the basis of quantity or expenditure data. Well-known and often cited calculations of social cost based on these sort of data are Berry and Boland (1977), Holtermann and Burchell (1983), Harwood et al (1984) and Rice et al (1990). The use of results from economic research based on frequency and intensity data may add knowledge of policy relevance to society.

Different consumption patterns for different groups of individuals may cause specific problems. Analysis of policy-related changes in social costs from different drinking behaviours could be used in resource allocation. Frequency may be related to both positive and negative health effects, as the body may benefit from having drink-free time to recover between drinking events and frequent consumption of small doses of alcohol may reduce the risks for coronary-heart diseases (*Sensible Drinking*, 1995). High-intensity drinking may be related to short-term harmful behaviours, including violent behaviour, crimes, accidents, and overdoses.

MODEL

Utility stems from *alcohol-satisfaction*, which is produced by the frequency and intensity of drinking episodes. The utility-bearing commodity may *not* be total alcohol consumption. As alcohol-satisfaction is unobservable, we allow frequency and intensity to enter the utility function directly. The optimal combinations of frequency and intensity chosen reflect the technology of producing alcohol-satisfaction.

Two utility-maximisation strategies are considered, both of which are initially affected by the individual's participation decision. The alternatives are then whether the individual (a) decides how much to consume and then produces this level via a combination of frequency and intensity, or (b) makes distinct choices about how much frequency and intensity to consume which are inter-related. Option (a) can be considered a three step procedure: i) self-selection, ii) what quantity to consume, and iii) how to create this consumption level in terms frequency and intensity. Option (b), on the other hand, reduces to a two step procedure: i) self-selection, and ii) choices over patterns of drinking in terms of frequency and intensity.

Thus, analysis of individuals' consumption of alcohol must first address individuals' choices of whether to consume alcohol or not. If an individual chooses not to consume alcohol, the problem reduces to the maximisation of utility over all other goods (Z). We can consider the problem as a three-dimensional problem in F, I and Z, in which frequency (F) is the underlying demand for how many days to consume during a month, and intensity (I) is the consumption level in general during such a day.

For non-alcohol-consumers, the maximisation problem relates only to points on the Z-axis and the solution is a single point on this line. If an individual chooses to consume a positive amount of alcohol, the problem is transformed immediately into one in which the consumption possibilities are not well-behaved, as the consumption possibility surface is convex towards the origin in the F-I plane. Despite this convexity, the budget constraint is linear if a cut is made from the maximum point of Z towards the F and I axes.

A general utility function with utility as a function of quantity of alcoholic beverages (Q) and other goods (Z) is

$$U = u(Q, Z)$$

Quantity is defined as the product of frequency and intensity. Therefore, we can substitute to obtain:

$$U = u[(F \cdot I), Z]$$

We may specify this approach by using a Cobb-Douglas function. In a Cobb-Douglas framework, substitution of the product of frequency and intensity for quantity in the utility function suggests that the coefficients on frequency and intensity are identical.

$$U = Q^{\alpha} \cdot Z^{\delta} = (I \cdot F)^{\alpha} \cdot Z^{\delta} = I^{\alpha} \cdot F^{\alpha} \cdot Z^{\delta} = I^{\alpha} \cdot F^{\beta} \cdot Z^{\delta} \quad \text{if } \alpha = \beta$$

If they are not identical, this suggests that frequency and intensity have differential effects on utility.

The budget constraint is given by:

$$p_x \cdot I \cdot F + p_z \cdot Z = M$$

in which: M is the total budget, p_X is the per unit price of alcohol and p_Z is the price on other goods. The contributions of each unit of frequency and intensity to the overall budget for alcohol are dependent on one another:

$$\frac{dM}{dF} = p_x . I \quad , \quad \frac{dM}{dI} = p_x . F$$

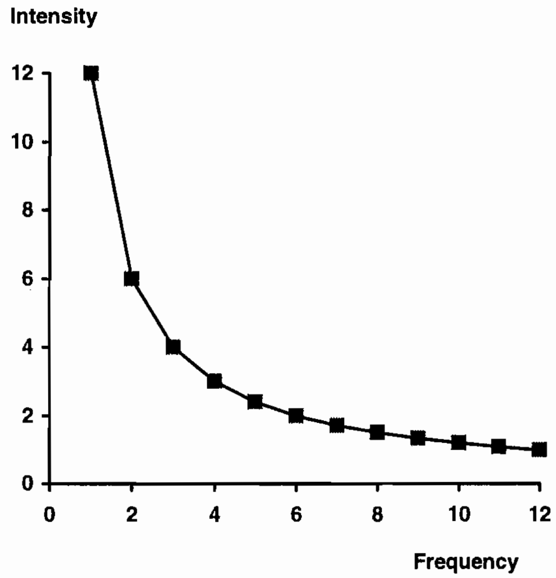
The Marshallian demand functions for frequency and intensity are:

$$I = \frac{\alpha}{(\beta + \delta)} \cdot \frac{M}{p_x \cdot F} \quad , \quad F = \frac{\beta}{(\alpha + \delta)} \cdot \frac{M}{p_x \cdot I}$$

Therefore, the optimal levels of F and I are negatively related and simultaneously determined.

Figure 1 below shows the implications of analysing alcohol consumption in an intensity and frequency framework taking a cut across Z. Figure 1 shows the case where an individual's budget for alcohol, *m*, equals 12 units. The budget line is convex because expenditure on alcohol is the product of frequency, intensity and the unit price of alcohol. The indifference map is parallel to the budget constraint as utility, like expenditure, is a function of total consumption. Therefore, the individual is predicted to be indifferent between all combinations of frequency and intensity along the consumption possibility frontier. In other words, the individual will choose randomly between a) consuming 1 drink in each of *m* drinking episodes, b) consuming *m* drinks in a single episode and abstaining otherwise, and c) consuming $\sqrt{m}$ drinks on $\sqrt{m}$ occasions. The locus of equi-probable optima is a constant-elasticity-of-substitution function with elasticity equal to minus one.

Figure 1: Under the quantity-hypothesis, the budget constraint is parallel to the indifference map



This prediction can be considered in terms of two comparison situations. Firstly, high-quantity consumers are not predicted to select different patterns of frequency and intensity from low-quantity consumers. Secondly, no inter-period stability in the frequency-intensity bundles selected are predicted.[9]

ESTIMATION

Nature of the restrictions

Logarithms of the Marshallian demand functions provides inter-dependent linear expressions for frequency and intensity:

$$\ln I = \ln\alpha - \ln(\beta+\delta) + \ln M - \ln p - \ln F$$

$$\ln F = \ln\beta - \ln(\alpha+\delta) + \ln M - \ln p - \ln I$$

Under the quantity hypothesis and in the absence of money illusion, the utility function is homogenous of degree one and the terms $(\alpha+\delta)$ and $(\beta+\delta)$ equal unity.

[9] Although we would expect relationships between quantities consumed in different periods as a result of the addictive stock (Becker and Murphy, 1988).

There are two restrictions suggested by the quantity-hypothesis: a) the coefficients on the independent variables included in X_k should be the same in the two equations, i.e. $\alpha_k = \beta_k$ and b) the coefficients on frequency in the intensity equation, and on intensity in the frequency equation, equal minus one. Partial analyses of these functions, with α and β determined by individual characteristics, thus provides a testable restriction of whether frequency and intensity represent different aspects of drinking behaviour which could be considered to bear directly on utility.

Testing the restrictions and the exclusion of endogenous variables

Implicit in restriction (a) is the condition that there should be no variables which predict frequency and do not predict intensity and visa versa. This is problematic in practical terms because without equation-specific predetermined variables, the system of equations for F and I is not identified. In other words, if the restriction *is* valid, it is not *possible* to estimate the system of equations.

One approach may be to omit the endogenous variables F and I. This introduces a possible source of omitted variable bias. However, rearranging the equations for F and I:

$$\ln F = \sum_{j=1}^{k-1} \alpha_j X_j + \alpha_k \ln I + \varepsilon_F$$

$$\ln I = \sum_{j=1}^{k-1} \beta_j X_j + \beta_k \ln F + \varepsilon_I$$

we obtain:

$$\ln F = \sum_{j=1}^{k-1} \frac{(\alpha_j + \alpha_k \beta_j)}{(1 - \alpha_k \beta_k)} X_j + \frac{(\varepsilon_F + \alpha_k \varepsilon_I)}{(1 - \alpha_k \beta_k)}$$

$$\ln I = \sum_{j=1}^{k-1} \frac{(\beta_j + \beta_k \alpha_j)}{(1 - \alpha_k \beta_k)} X_j + \frac{(\varepsilon_I + \beta_k \varepsilon_F)}{(1 - \alpha_k \beta_k)}$$

A test of the equivalence of the estimated coefficients is a form of the test for the original coefficients with each of the original coefficients (α_j, β_j) weighted by a function of α_k and β_k. These are the coefficients on the omitted variables, which are also restricted to be equal.

However, with variables omitted from the estimated equations, the sum of squared residuals is overestimated, meaning that s^2 is a biased estimate of σ^2. As the estimate is biased upwards (Greene, 1993), the standard errors of the coefficients are biased upwards and the test of the equivalence of the estimated parameters is a consistent but less powerful test of the restrictions. To summarise, because there is a higher probability of a Type II error, this test is a more conservative test of the restrictions.

Testing restrictions

We wish to make cross-equation comparisons of estimated parameters. The situation is analogous to tests of structural change in time-series analysis. In this case, Greene (1993) suggests stacking the data and estimating the following:

$$\begin{bmatrix} \ln F \\ \ln I \end{bmatrix} = \begin{bmatrix} X_k & 0 \\ 0 & X_k \end{bmatrix} \begin{bmatrix} \alpha_k \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_F \\ \varepsilon_I \end{bmatrix}$$

A variety of tests can then be used to test the restriction $\alpha_k = \beta_k$. A convenient method of estimation is provided by Seemingly Unrelated Regression Equations (SURE). As the frequency and intensity data are measured on different scales, we allow for different constant terms. However, there are two additional complications which we consider using alternative testing approaches: self selection; and heteroskedasticity. Each of our approaches has advantages in some areas and disadvantages in others, and we present a range of possible test results.

Allowing for self-selection

To estimate the influence of the independent variables on frequency and intensity on the sample of spirit-drinkers may introduce sample selection bias. Therefore, the regression equations are estimated allowing for sample selection. The underlying structure of the model is of the following form:

$$d_i^* = \gamma z_i + u_i$$

$$d_i = 1 \quad if\, d_i^* > 0$$
$$d_i = 0 \quad if\, d_i^* \leq 0$$

$$y_i = \beta_k X_{ik} + \varepsilon_i \quad \text{observed iff } d_i = 1$$
$$(u_i, \varepsilon_i) \sim \text{bivariate normal } [0, 0, 1, \sigma_\varepsilon, \rho]$$

in which y_i is a column of stacked data on the log of frequency and intensity. We adopt the Heckman two-step estimation procedure. In particular, since we are interested in allowing for differential effects of incidental truncation on frequency and intensity, we obtain an Inverse Mill's Ratio for each observation through regression of participation in spirit drinking on the independent variables z and estimate y_i using OLS. As is the case for other independent variables in the model, the coefficient on IMR is allowed to vary between frequency and intensity through multiplication by two dummy variables each representing half of the stacked data set.

Allowing for heteroskedasticity

Greene (1993) notes that heteroskedasticity may be a particular problem for tests of structural change. This can cause overestimation of the significance of the test statistic and therefore a higher probability of a Type I error. In our case, it is possible that we have heteroskedastic disturbances along two dimensions: between frequency and intensity; and related to some of the independent variables in the regression. The variances of the error terms are likely to differ between frequency and intensity because of the different scaling of the variables and because OLS regression on the independent variables and the Inverse Mill's Ratios does not take account of the sample selection correction to the estimated variance. Heteroskedasticity which is a function of the variables in the selection equation and their associated coefficients is introduced (Greene, 1993). If the estimated ρ's differ between frequency and intensity, this will create heteroskedastic disturbances between the frequency and intensity equations.

Greene (1993) suggests that a Wald test for structural change is valid in large samples even in the case of heteroskedastic disturbances. However, the validity of the test rests on the assumptions that the parameters are drawn from independent samples and therefore there is zero covariance between the parameters estimated for frequency and intensity. Given that the estimated parameters are functions of common structural parameters, this assumption of the independence of the estimated parameters is unlikely to hold. Nevertheless, as it is less sensitive to heteroskedasticity, we provide Wald-statistics for a test of structural change assuming zero inter-aspect covariance between the parameters estimated using two separate sample-selection models.

LIMDEP provides a Breusch-Pagan Lagrange-Multiplier (BPLM) test for heteroskedasticity. As we have good reasons to suspect heteroskedasticity in the disturbances between the first and second half of the stacked data set, we also calculate a Goldfeld-Quandt (GQ) test for heteroskedasticity. The BPLM test allows the disturbance variance to vary with the set of regressors while the GQ compares the errors in two subsets of the total sample. In our case the GQ test provides a specific test for heteroskedasticity between frequency and intensity whilst BPLM is a more general test.

If the hypothesis of homoskedasticity is rejected by either test, we re-estimate the stacked data regression using White's correction of the standard errors. We also estimate a model of multiplicative heteroskedasticity, in which the set of variables describing the form of heteroskedasticity includes both a dummy variable to distinguish between the two halves of the data set and the variables entered into the selection equation.

The Structural Unrestricted Model

The structural unrestricted model is a simultaneous equation system with selectivity, which can be estimated by Two-Stage Least-Squares (Greene, 1991). The model is assumed to be of the form:

$$d_i^* = \gamma z_i + u_i$$

$$d_i = 1 \quad if\; d_i^* > 0 \quad ;\ 0 \text{ otherwise.}$$

$$\left.\begin{aligned} \ln F_i &= \beta_m X_{im} + \beta_n W_n + \delta_I \ln I + \varepsilon_{iF} \\ \ln I_i &= \beta_m X_{im} + \beta_p V_p + \delta_F \ln F + \varepsilon_{iI} \end{aligned}\right\} \quad \text{observed iff } d_i = 1$$

in which X_m is a vector of common determinants of frequency and intensity, W_n is a set of independent variables determining frequency and V_p contains influences on intensity. Initially, we include gender, age and income in X_m, education and economic status in V_p, and family composition variables, to reflect the transaction costs of socialising, in W_n. Education reflects health-related knowledge and economic status captures differences in the opportunity costs of recovery time.

DATA

The Malmö Health Survey 1994 (Hälsoläget i Malmö) is a mail survey of 5600 individuals in private households registered with the municipal database, *MAGDA*. The sample was randomly selected from eight age groups. The survey was conducted anonymously so that no further information could be linked to the questionnaire. The survey included questions about health and health risks, with an aim to create a cross-sectional data set which would provide a picture of the prevalence of health risks in the local population. There was 178 'technical' drop-outs, and 3861 of the remaining 5422 individuals answered the questionnaire. This gives a total drop-out rate of 29 percent. The drop-outs were greater for the younger group than for the older groups. Of those who answered the questionnaire, 50.5 percent were female. Variables included in this analysis are age, gender, whether born in Sweden, education, income, marital status, perceived level of social support, economic status, and family composition. The dependent variables relate to self-reported spirit drinking in the last 30 days.[10] 3087 individuals gave information on all variables which are given in Table 2.

RESULTS

The results of the probit regression analysis of whether an individual has consumed spirits in the last 30 days are shown in Table 3. Using a cut-off probability of 0.5, two-thirds of the cases are correctly predicted into participating and non-participating groups. All of the variables take the expected sign, and most are significantly different from zero at the 5% level. In particular, females and those with children under five years-old in the household are found to be significantly less likely to report drinking spirits, and the participation rate is negatively related to age. Participation in the drinking of spirits is significantly and positively related to income and education. Given Sweden's location in the 'vodka-belt' (Bruun et al, 1975; Davies and Walsh, 1983), the variable representing immigration (NONSWEDE) is expected to be negative.

[10] Our analysis relates only to spirit-drinking and we have not allowed for cross-drink effects. There is no common agreement as to the significance and nature of these influences empirically (Leung and Phelps, 1993). Moreover, we do not consider addiction effects. The data are taken from a postal questionnaire of individuals in private households and it is unlikely that heavily-addicted drinkers are included. Furthermore, it is believed that the addictive effects for alcohol are not as strong as they are for tobacco (Grossman et al, 1995).

Table 2. Variable definitions and sample means.

Variable	Definition	Mean Full-sample (3087)	Mean Drinkers (1333)
FREQUENCY	Log of number of days drunk spirits in the last 30 days	-	1.330
INTENSITY	Log of how much spirits (cl.), in general, drunk on each occasion	-	2.460
AGE	log of age in years	3.764	3.762
FEMALE	1 if female	0.497	0.329
INCOME2	1 if income[11] 7000-9999 SEK	0.204	0.182
INCOME3	1 if income 10000-13999 SEK	0.211	0.203
INCOME4	1 if income 14000-20999 SEK	0.241	0.287
INCOME5	1 if income 21000- or more	0.132	0.177
VOCATION	1 if 1+ year of vocational education	0.209	0.216
HIGHSCHOOL	1 if 1+ year of high school (gymnasieskola)	0.213	0.231
UNIVERSITY	1 if 1+ year university/higher education	0.283	0.302
NONSWEDE	1 if not born in Sweden	0.258	0.202
DIVORCED	1 if divorced	0.101	0.078
MARRIED	1 if married/cohabiting, widow/widowed	0.653	0.672
CHILD<5	1 if child below age 5 at home	0.117	0.081
NCHILD<15	Log of number of children below age 15 at home	0.181	0.140
NOSUPPORT	1 if perceive no social support for handling problems	0.587	0.590
UNEMP	1 if currently unemployed	0.100	0.095
STUDENT	1 if currently full-time student	0.107	0.089

Regression results have been obtained for the various reduced-form regression models for frequency and intensity and are available from the authors on request. Although the equations suffer from omitted variable bias, and therefore the coefficients represent the combined effects of several factors, the point estimates of the parameters are stable across different estimation procedures and there is little change in the estimated standard errors. The GQ test of heteroskedastic errors in the OLS regression on the stacked data on frequency and intensity suggest no evidence on which to reject homoskedasticity (F(1322,1322) = 1.01; crit. value = 1.06). However, the BPLM test statistic suggests the presence of general heteroskedasticity (χ^2 (21) = 208,782; p<0.001). Therefore, it may be important to undertake the test of the equivalence of the estimated parameters with adjustment for heteroskedasticity.

[11] Household disposal income after tax including social benefits.

Table 3. Probit equation for whether drank spirits in last 30 days.

Variable	Coefficient	Std. error	t-ratio
CONSTANT	0.523	0.267	1.960
AGE	-0.163	0.066	-2.485
NONSWEDE	-0.523	0.057	-5.053
INCOME2	0.203	0.075	2.690
INCOME3	0.257	0.075	3.419
INCOME4	0.501	0.073	6.822
INCOME5	0.588	0.088	6.712
VOCATION	0.094	0.069	1.373
HIGHSCHOOL	0.162	0.079	2.058
UNIVERSITY	0.127	0.069	1.849
FEMALE	-0.734	0.048	-15.331
CHILD<5	-0.423	0.080	-5.257

Likelihood-Ratio Index = 9.9%
Correct predictions = 66.7%

Table 4 shows the results from the range of tests of the null 'quantity-hypothesis'. The tests are conducted on the parameter estimates from: the separate sample selection equations for frequency and intensity; the stacked data regression of frequency and intensity with adjustments for heteroskedasticity using White's correction for standard errors and a multiplicative heteroskedastic regression model (MHR); and a Seemingly Unrelated Regression Equation system (SURE). The test statistic from tests on all the parameter estimates (ALL in table 4) are highly significant, suggesting a rejection of the restriction of equal impact on utility from frequency and intensity. The restrictions fail particularly for age and education. Three of the four tests for gender fail at the 5-percent level, and all income tests fail at the 10-percent level.

Table 4. Results of the tests of the 'quantity-hypothesis' restriction.

Variables	Sample-selection	OLS, White's correction	MHR	SURE
GENDER	p=0.056	p=0.038	p=0.085	p=0.042
AGE	p<0.001	p<0.001	p<0.001	p<0.001
EDUCATION	p<0.001	p<0.001	p<0.001	p<0.001
INCOME	p=0.100	p=0.094	p=0.077	p=0.065
ALL	p<0.001	p<0.001	p<0.001	p<0.001

The results from the structural equations for frequency and intensity are shown in Table 5. The income variables are excluded from the frequency equation because they are not found to be significantly different from zero. Age is the only variable which is significantly different from zero in the frequency equation, although the difference between males and females approaches significance at the 5% level. As expected, frequency is increasing in age. The coefficients on the other variables entered into the equation, which represent family circumstances, are close to zero. The coefficient on the instrumental variable for intensity is positive and not significantly different from zero. We find few determinants of frequency, *given positive consumption.*

The intensity equation, on the other hand, is well-defined. *Conditional on the level of frequency*, we find significantly negative relationships for both education and income. Individuals in the highest-income group are estimated to drink 28% less on each drinking occasion than those in the lowest-income group. University-educated individuals are predicted to drink 24% less intensely than those who completed only compulsory education. Unemployed individuals are estimated to drink 23% more on each occasion than their employed counterparts. The coefficient on the instrumental variable for frequency is positive but not significantly different from zero at the 5% level.

DISCUSSION

In this paper we have shown that traditional economic analyses of alcohol consumption place a restriction on individuals' choices of how often to consume and how much to consume on each occasion. With standard assumptions about the nature of the utility function, it has been shown that the restriction implies that individuals will locate randomly along a consumption possibilities frontier for frequency and intensity with unitary constant-elasticity-of-substitution. Descriptive analysis of data from *The Malmö Health Survey 1994* suggested that there is systematic clustering of particular drinking patterns in different social groups. Differences in acquisition costs of frequency and intensity, the demand for health and preferences for different types of drinking patterns may all contribute to observed similarities in intra-group behaviour.

Based on Marshallian demand functions for frequency and intensity, tests are made of the validity of the 'quantity-hypothesis' restriction using cross-sectional data on individuals' spirit-drinking behaviour. The implication of the restriction of random-positioning along the consumption possibility frontier is that independent variables have equal impacts on frequency and intensity. This hypothesis is rejected for age, gender, education level and income. Moreover, the rejection is robust to adjustments for heteroskedasticity and self-selection. An unrestricted simultaneous equation system with selectivity for frequency and intensity indicates that economic factors significantly influence intensity and participation but not frequency. The independent effects of lower income, unemployment and a lower-education level are to intensify drinking behaviour.

Table 5. Simultaneous equation model for frequency and intensity.

Variable	Coefficient (t-ratio) Frequency	Coefficient (t-ratio) Intensity
CONSTANT	-0,300 (-0,616)	4,891 (18,563)
AGE	0,443 (6,334)	-0,768 (-3,295)
FEMALE	-0,182 (-1,904)	-0,252 (-1,913)
CHILD<5	-0,137 (-1,412)	-
NCHILD<15	0,073 (1,060)	-
MARRIED	-0,040 (-0,770)	-
DIVORCED	-0,102 (-1,404)	-
NOSUPPORT	-0,003 (-0,087)	-
INTENSITY	0,030 (0,273)	-
IMR[12]	-0,017 (-0,138)	-0,012 (-0,068)
VOCATION	-	0,015 (0,206)
HIGHSCHOOL	-	-0,119 (-1,903)
UNIVERSITY	-	-0,274 (-4,228)
INCOME2	-	-0,163 (-2,222)
INCOME3	-	-0,208 (-2,138)
INCOME4	-	-0,227 (-2,499)
INCOME5	-	-0,329 (-3,417)
UNEMP	-	0,210 (2,944)
STUDENT	-	-0,132 (-1,760)
FREQUENCY	-	0,636 (1,201)

[12] Inverse Mill's Ratio from probit regression in Table 3.

The data set which we have used to both test the quantity-hypothesis restriction and to model frequency and intensity is cross-sectional in design. It is not possible, therefore, to incorporate the effects of price changes on consumption and, more importantly, to investigate whether there are differential effects of prices on frequency and intensity. The approach used, however, provides the tools for analysing participation, frequency and intensity and expressing these results in terms of elasticities.

In general, the results presented indicate that utility can be derived differentially from frequency and intensity of spirit consumption. Just as explicit consideration of the joint impact of frequency and intensity on utility has proved useful, it may be fruitful to model the impact of drinking behaviour on social costs in a similarly unrestricted manner. It is clear that intensity is not the sole link to accidents and violence and that health effects are not only determined by frequency.

How frequently and intensely individuals participate in the consumption of alcohol are relevant pieces of information on individuals' preferences for various characteristics of drinking and shadow prices of various aspects of drinking. This has been shown using average frequency and average intensity over a period of 30 days. Even this may be a restriction, but to test this would require information from *daily* 'drinking diaries' or measures of the *variability* of the frequency and intensity of drinking episodes (Sindelar, 1993).

Disaggregating alcohol consumption in frequency and intensity highlights more commonality between policy concerns for tobacco and alcohol. Because both tobacco consumption and the intensity of drinking episodes cluster in particular social groups and are monotonically related to health damage, they may be explained by similar theories of the determinants of health-detrimental behaviours. From a public health perspective, these results are important because they explain observed patterns of binge drinking as a rational response to the cost of consumption periods. Furthermore, because availability affects acquisition costs, different responses would be expected from frequency and intensity. If transaction costs are raised, this may increase the intensity of consumption, particularly for those who do not have high recovery-time costs, such as young individuals and low income groups.

When we allow frequency and intensity to both enter the utility function separately, the multiplicative form of the budget constraint suggests optimal levels of frequency and intensity will be dependent on one other. The estimates from the derived demand functions do not support the modelling of quantity but suggest that frequency and intensity are utility-bearing aspects of drinking behaviour. The quantity and the frequency-intensity approaches are equivalent only when there is constant-elasticity-of-substitution between frequency and intensity.

REFERENCES

Atkinson, A.B., Gomulka, J., & Stern, N.H. (1990) Spending on alcohol: Evidence from the Family Expenditure Survey 1970-1983, *Economic Journal*, 100, 808-27.

Becker, G.S. & Murphy, K.M. (1988) A theory of rational addiction, *Journal of Political Economy*, 96, 675-700.

Becker, G.S., Grossman, M. & Murphy, K.M. (1991) Rational Addiction and the Effect of Price on Consumption, *American Economic Review*, 81, 237-241.

Berggren, F. & Lindgren, B. (1995), Alkohol - är det något för ekonomer? (Alcohol - something for economists?), *Socialvetenskaplig tidskrift*, 4, 298-315.

Berry, R.E. & Boland, J.P. (1977) *The Economic Cost of Alcohol Abuse*, New York: Free Press.

Bruun, K., Edwards, G., Lumio, M. et al (1975) *Alcohol Control Policies in Public Health Perspective*. (The Finnish Foundation for Alcohol Studies, Vol. 25, Forssa).

Chaloupka, F.J. & Wechsler, H. (1995) The impact of price, availability, and alcohol control policies on binge drinking in college, NBER, Working paper no. 5319.

Coate, D. & Grossman, M. (1988) Effects of alcoholic beverage prices and legal drinking ages on youth alcohol use, *Journal of Law and Economics*, 31, 145-171.

Davies, P. & Walsh, D. (1983) *Alcohol Problems and Alcohol Control in Europe*. Crom Helm: London.

Edwards, G. et al (1994) *Alcohol Policy and the Public Good*. Oxford: Oxford University Press.

Godfrey, C. (1988) Licensing and the demand for alcohol, *Applied Economics*, 20, 1541-58.

Godfrey, C. (1994) Economic influences on changes in population and personal substance behaviour, in Edwards, G. & Lader, M. (eds), Addiction: processes of change, Society for the study of Addiction Monograph No. 3 Oxford: Oxford University Press,

Greene, W.H. (1991) *LIMDEP User's Manual and Reference Guide Version 6.0.* Econometric Software, Inc.: New York.

Greene, W.H. (1993) *Econometric Analysis*. Second edition. New York: MacMillan.

Grossman, M., Coate, D. & Arluck, G. (1987) Price sensitivity of alcoholic beverages in the United States: Youth alcohol consumption, in Holder, H.D., ed. *Control Issues in Alcohol Abuse Prevention: Strategies for States and Communities*, Greenwich, CT, JAI, 169-98.

Grossman, M., Chaloupka, F.J. & Sirtalan, I. (1995) An empirical analysis of alcohol addiction: results from the Monitoring the Future Panels, NBER, Working paper no. 5200.

Harwood, H.J., Napolitano, D.M., Kristiansen, P.L., & Colins, J.J. (1984) Economic Costs to Society of Alcohol and Drug Abuse and Mental Illness: 1980, Research Triangle Park, N.C.: Research Triangle Institute.

Holtermann, S. & Burchell, A., (1983) The Costs of Alcohol Misuse, Government Economic Service Working Paper no 37, London, Department of Health and Social Security.

Jones, A.M. (1989) A double-hurdle model of cigarette consumption. *Journal of Applied Econometrics*. 4: 23-39.

Leung, S.I. & Phelps, C.E. (1993) "My Kingdom for a Drink...?":A Review of Estimates of the Price Sensitivity of Demand for Alcoholic Beverages, in *Economics and the Prevention of Alcohol-related Problems*, eds. Hilton, M.E. & Bloss, G., U.S. Department of Health and Human Services, 1-31.

Lindström, M., Bexell, A., Hanson, B.S. & Isacsson, S-O. (1994) Hälsoläget i Malmö, Rapport från postenkätundersökningen våren 1994 (The Malmö Health Survey 1994), Socialmedicinska avdelningen, Samhällsmedicinska institutionen, Malmö, Lunds universitet.

Manning, W.G., Blumberg, L. & Moulton, L.H. (1995) The demand for alcohol: The differential response to price, *Journal of Health Economics*, !4, 123-48.

Moore, M.J. & Cook, P.J. (1995) Habit and heterogeneity in the youthful demand for alcohol, NBER, Working paper no. 5152.

Rice, D.P., Kelman, S., Miller, L.S. & Dunmeyer, S. (1990) The Economic Costs of Alcohol and Drug Abuse and Mental Illness: 1985, U.S. Department of Health and human Services.

Sensible Drinking (1995), The Report of an Inter-Departmental Working group, Department of Health.

Sindelar, J.L. (1993), Measurement Issues in Alcohol Survey Data, in *Economics and the Prevention of Alcohol-related Problems*, eds. Hilton, M.E. & Bloss, G., U.S. Department of Health and Human Services, 201-228.

Stockwell, T. et al (in press) Total versus hazardous per capita consumption as predictors of acute and chronic alcohol related harm, *Contemporary Drug Problems*.

Sutton, M. & Godfrey, C. (1995) A grouped data regression approach to modelling the economic and social determinants of individual drinking behaviour, *Health Economics*, 4, 237-47.