Minimum Unit Prices for alcohol

Paul Calcott*

June 24, 2017

Abstract

Minimum Unit Prices (MUPs) have been proposed on the grounds that they can reduce alcohol consumption by the heaviest drinkers, without significantly burdening moderate drinkers. This paper examines the case for MUPs in an optimal tax framework. Conditions are identified under which the optimal policy mix involves a MUP in addition to a corrective tax. The implications that inelastic demand for alcohol may have for this mix are explored with a calibrated numerical example. The final issue to be addressed is the danger that consumers may respond to a specific tax with quality substitution.

JEL CLASSIFICATIONS: H230, I120

KEYWORDS: corrective taxes, Minimum Unit Price, alcohol

*Victoria University of Wellington, paul.calcott@vuw.ac.nz

1 Introduction

A Minimum Unit Price (MUP) for alcohol is a prohibition of alcohol sales for prices less than some minimum level per unit of ethanol content. Related policies have been introduced in Canadian provinces (Stockwell et al. 2012, Sharma et al. 2014). However, they typically set minimum prices per liquid volume by drink type, rather than per ethanol content. The Scottish government passed an Act in 2012 to enable Minimum Unit Pricing (Katikireddi & McLean 2012) but it has not yet been implemented, largely because of legal challenges. This paper examines the microeconomic foundations for such a policy. The optimal mix of a corrective tax and a MUP is characterized analytically, and some determinants of this mix are explored with a numerical example.

Expressions for the socially optimal level of a Minimum Unit Price have not been derived in the previous literature, but there are two strands of research that are relevant. The first strand reports estimated levels for socially optimal alcohol taxes in the absence of a minimum price (Pogue & Sgontz 1989, Saffer & Chaloupka 1994, Kenkel 1996). The second strand models the likely effects of Minimum Unit Pricing (Meier et al. 2010, Sharma et al. 2014, Holmes et al. 2014).

A common approach to calculating optimal tax rates on alcohol is to minimize aggregate deadweight losses across consumer types, where these deadweight losses are inferred from elasticities. This approach was proposed by Pogue & Sgontz (1989), and followed by Saffer & Chaloupka (1994) and then Kenkel (1996). Pogue & Sgontz (1989) assumed two types of drinker (abusers and nonabusers).[1] Saffer & Chaloupka (1994) incorporated cross-price elasticities, and evaluated the case for taxing the alcohol content in each beverage equally. Kenkel (1996) extended the framework, allowing for

[1] with an extension to a disease model of alcoholism.

two types of heavy drinkers (informed and uninformed).

The following analysis draws on this literature, but departs from it in a number of respects. I will address the optimal combination of a corrective tax and a Minimum Unit Price, rather than a corrective tax in isolation. Because a MUP will be allowed for, it will be necessary to allow for quality differences among drinks. After all, a MUP would be a blunt instrument if all alcoholic drinks sold for the same price per unit. I will assume that consumers can choose from a continuous range of alcohol varieties, distinguished by quality and hence price. Given this focus on quality, it is convenient to work directly with utility functions rather than the demand curves on which these earlier studies were focused.[2]

The second strand of literature that is relevant to the current study simulates the effects of a Minimum Unit Price, drawing on estimates of own- and cross-price elasticities of various types of alcohol. A prominent example is the Sheffield study, (Meier et al. 2010, Holmes et al. 2014, Brennan et al. 2015). This work has been used to advocate for introduction of a MUP, by arguing that the burden on moderate drinkers would be relatively small while heavier drinkers would respond with significant reductions in alcohol consumption.

My study differs because it explicitly adopts the normative framework of welfare economics, and examines the determinants of optimal policy instruments in that framework. I will provide a general characterization of policy settings that maximize aggregate welfare, rather than estimate the effects of particular proposals for a MUP that may have been proposed by politicians

[2]Both Parry et al. (2009) and Aronsson & Sjögren (2010) analysed alcohol taxes with reference to utility functions, but assumed a representative consumer and did not allow for variation in product quality. Lockwood & Taubinsky (2017) present an optimal tax approach to soda taxes, although the focus is on redistributive concerns rather than quality changes.

or public health advocates. My study focuses on foundational and conceptual issues, rather than on the selection of a specific level for a Minimum Unit Price. In addition, I deal with variation in beverage quality in a very different way. Instead of representing quality impacts in terms of cross-price elasticities of demand between discrete categories of beverage, I treat quality as an explicit choice variable.[3]

I will assume that drinkers are heterogeneous. If all drinkers were identical, there would be little reason to propose a Minimum Unit Price. Then socially efficient drinking could be implemented with a specific tax on alcohol content, calibrated to match the common magnitude of the externalities plus internalities from marginal consumption. However, drinkers do differ in the amounts that they drink and consequently in the harm caused by marginal consumption.

Heterogeneity among alcohol drinkers reduces the level of social efficiency that can be attained with a single corrective tax. When marginal externalities and internalities differ among drinkers, a single tax rate cannot be calibrated to simultaneously match the distortions for every individual (Crawford et al. 2010). In general, when harms are heterogeneous but the rate of the corrective tax cannot be differentiated, then there may be a case for supplementary regulation (Christiansen & Smith 2012) that targets the relatively undertaxed harms. An application to a MUP for alcohol is suggested by claims that such a policy would have the most impact on excessive drinkers (Holmes et al. 2014, Sharma et al. 2014).

The following section sets out a theoretical model, in which a hetero-

[3]Each approach has advantages and disadvantages. It is convenient to treat quality as an explicit choice variable in more theoretical work, because then quality competition means that a price floor is compatible with zero profits in a long-run competitive equilibrium. However, discrete product types are suitable for more applied work, for then it is straightfoward to incorporate estimated cross-price elasticities of demand.

geneous population of consumers choose quantities and qualities of alcohol to purchase. Section 3 provides an analytic characterization of the optimal policy in this setting, and illustrates with a numerical example calibrated to statistics reported by Kerr & Greenfield (2007). Sections 4 and 5 address some considerations that have been raised in the previous literature on Minimum Unit Prices. Section 4 deals with the overall inelasticity of demand for alcohol, and with variation of that inelasticity between moderate and heavy drinkers. Section 5 addresses the concern that a tax increase might have limited effect, if drinkers respond by switching to lower quality drinks. I argue that this concern does not provide a compelling reason to prefer Minimum Unit Prices.

2 Model

2.1 Payoffs

A population of consumers each choose the quantity and quality of alcohol to consume, $q \in \mathbb{R}_+, a \in \mathbb{R}_+$. I will interpret q as the quantity of pure alcohol contained in drinks, rather than the overall liquid volume. The latter might be viewed as being incorporated into quality.

Consumer payoff is $u(q, a; \theta) - p(a)q$. Consumers are distinguished by a vector of utility parameters, θ, which is drawn from distribution F. Assume that the consumer's benefit from consumption, $u(q, a; \theta)$, is twice continuously differentiable and strictly concave in q, a. Moreover, consumers prefer higher quality when consuming a strictly positive amount.

$$u_a(q, a; \theta) > 0, \forall q > 0. \tag{1}$$

A long-run competitive equilibrium ensures that the price of alcohol equals its unit cost plus the specific tax rate on units of q, $p(a) = c(a) + \tau$, where unit costs, $c(a)$, are strictly increasing and weakly convex.[4]

Substituting the equilibrium price into the consumer payoff, we obtain:

$$U(q, a; \theta, \tau) = u(q, a; \theta) - [c(a) + \tau]q. \tag{2}$$

Without a binding Minimum Unit Price, a consumer would choose q, a to maximize (2). These choices would satisfy the best responses:

$$q^*(a; \theta, \tau) := \arg\max_q U(q, a; \theta, \tau), \qquad a^*(q; \theta) := \arg\max_a U(q, a; \theta, \tau).$$

The equilibrium choices, $\hat{q}(\tau; \theta), \hat{a}(\tau; \theta)$ can be identified as the intersection of these two best responses.

With a Minimum Unit Price, $\bar{p}$, the consumer obtains no discount from purchasing alcohol with quality below a level, $\bar{a}$. This level is identified by the equality between that price and the tax-inclusive unit cost:

$$c(\bar{a}) + \tau = \bar{p}. \tag{3}$$

Because consumers prefer higher quality by assumption (1), they would not buy alcohol with quality lower than $\bar{a}$. Consequently, the consumer problem can be stated as:

$$\max_{q,a} U(q, a; \theta, \tau) \text{ s.t. } a \geq \bar{a}.$$

When the constraint is not binding, then the consumer's solution is again $q = \hat{q}(\theta, \tau)$ and $a = \hat{a}(\theta, \tau)$. But when it does bind, then quality is $\bar{a} = \tilde{a}(\bar{p} - \tau)$, as implicitly defined by (3), and quantity is at the best response to that level, $q = q^*(\bar{a}; \theta, \tau)$, or equivalently:

$$q = \tilde{q}(p; \theta, \tau) := \arg\max_q \{u(q, c^{-1}(p - \tau)) - pq.\} \tag{4}$$

A tilde will identify choices for which quality is actively bound by the MUP.

[4]These assumptions are not sufficient to ensure that (2) is quasi-concave. However, it will be assumed that consumers' decision rules are continuous.

2.2 The relationship between quality and quantity

A policy to influence the consumption of alcohol will often also have an impact on the quality chosen. This impact depends on whether quality and quantity are complements or substitutes.

Definition 1. *Let quality and quantity be local market complements (substitutes) when the best responses,* $a^*(q;\theta), q^*(a;\theta,\tau)$ *are weakly increasing (decreasing) in* q, a *respectively, in the neighborhood of the equilibrium.*

The next result proposes a sufficient condition for q, a to be local market substitutes.

Lemma 1. *Let* $u_a(q,a)/q$ *be decreasing in* q *and the consumer's solution be interior. Then* q, a *are local market substitutes.*

Proof. First, as the private payoff is strictly concave in q, strictly concave in a and continuous, the best responses are continuous by the Maximum Theorem. Next, for any given value of $q > 0$, maximization of U wrt a is equivalent to maximization of U/q. But the latter satisfies increasing differences in $(a; -q)$ when $u_a(q,a)/q$ is decreasing in q, and so $a^*(q;\theta)$ is decreasing in q by the Monotonicity Theorem of Topkis (1978). The two best responses must have the same sign for their slopes when they go through an interior utility maximum, by Young's Theorem and the Implicit Function Theorem. Therefore, q, a are local market substitutes. □

The scaled marginal benefit of quality, $u_a(q,a)/q$, is decreasing in q when:

$$u_{aq}(q,a) \leq \frac{u_a(q,a)}{q}.$$

This will hold for sure if $u_{qa} < 0$, and otherwise holds when the marginal utility of quality is concave in quantity and the marginal benefit of quality is always nonnegative (so that $u_a(0,a) \geq 0$).

2.3 Slopes of the decision rules

The impacts of a Minimum Unit Price or specific tax will depend on the slopes of consumers' decision rules. The simplest of these impacts may be that on quantity consumed from an incremental change in the specific tax, for fixed beverage quality. In an interior solution, this will be the amount by which the best response for quantity shifts in response to the tax change. For an incremental tax change, this will be $q^*_\tau(a;\theta,\tau)$.

Lemma 2. *The fixed-quality impact of an increase in the specific tax is negative,* $q^*_\tau \leq 0$.

Proof. For fixed a, θ, the private payoff satisfies increasing differences in $(q, -\tau)$. The negative impact of an increase in τ follows from the Monotonicity Theorem. □

More generally, the impact of the specific tax may involve quality changes. Consider this impact for a consumer for whom the MUP does not bind (one who would prefer to drink alcohol with quality greater than $\bar{a}$).

Lemma 3. *Imagine a consumer whose solution is interior, without the MUP binding, and who is free to change the quality of alcohol to purchase. Then the impact on her quantity purchased of an incremental increase in the specific tax on q, is:*

$$\hat{q}_\tau = \frac{q^*_\tau}{1 - q^*_a a^*_q}, \tag{5}$$

for quantity, and:

$$\hat{a}_\tau = \frac{a^*_q q^*_\tau}{1 - q^*_a a^*_q}, \tag{6}$$

for quality. Effect (5) *is negative and is larger in absolute value than the fixed-quality response,* q^*_τ.

Proof. Differentiate $\hat{q}(\tau;\theta) \equiv q^*(a^*(\hat{q}(\tau,\theta));\theta,\tau)$ wrt τ, and solve for $\hat{q}_\tau$ to obtain (5). Follow a similar process for $a = a^*(\hat{q}(\tau,\theta);\theta)$ to obtain (6). To show that (5) is negative, appeal to Lemma 2 plus the second-order condition for maximization of $U(q, a; \theta, \tau)$. To show that it is larger in absolute value than q^*_τ, recall that q^*_a, a^*_q must have the same sign as each other at a point where the two best responses intersect. □

Now consider consumers who would have preferred to buy alcohol with quality lower than $\bar{a}$.

Lemma 4. *Imagine a consumer who chooses interior values of q, a, but for whom the MUP is binding, $\hat{a}(\tau; \theta) < \bar{a}$. Then the marginal impacts of a specific tax on q are:*

$$\tilde{q}_\tau(\bar{p}; \theta, \tau) = -\frac{q^*_\tau}{c'(a)}, \quad \tilde{a}_\tau(\bar{p}; \theta, \tau) = -\frac{1}{c'(a)}, \tag{7}$$

and the marginal impacts of an increase in the MUP are:

$$\tilde{q}_{\bar{p}}(\bar{p}; \theta, \tau) = \frac{q^*_a}{c'(a)}, \quad \tilde{a}_{\bar{p}}(\bar{p}; \theta, \tau) = \frac{1}{c'(a)}.$$

Proof. Let $g(a, p) := \arg\max_q\{u(q, a) - pq\}$. Then $q^* = g(a, c(a) + \tau)$ and $\tilde{q} = g(c^{-1}(p - \tau), p)$. Differentiate the former to find that $q^*_a = g_a + g_p c'(a)$ and $q^*_\tau = g_p$. Differentiate the latter to determine that $\tilde{q}_p = g_a/c'(a) + g_p$ and $\tilde{q}_\tau = -g_a/c'(a)$. Divide the expressions for q^*_a and q^*_τ by $c'(a)$ to obtain $\tilde{q}_p$ and $-\tilde{q}_\tau$. To find $\tilde{a}_\tau, \tilde{a}_{\bar{p}}$, totally differentiate $c(a) + \tau = \bar{p}$. □

According to Lemma 4, it is not safe to assume that a price increase due to a higher tax would have the same impact as the same price change due to a MUP. This is because they involve quite different effects on the quality of alcohol purchased. Moreover, the impact of a specific tax is very different on consumers, depending on whether their choices are actively bound by the constraint implied by the Minimum Unit Price. If the MUP binds before and after a tax increase, then the increase has no impact on the price that consumers pay. Consequently, the conventional mechanism by which taxes affect consumption would have been deactivated. Instead, the effect of the tax is mediated through the minimum quality, $\bar{a}$.

3 Optimal policy

The social planner sets levels for two policy instruments, a linear specific tax on the pure alcohol content of beverages, and a Minimum Unit Price on the same.[5] He maximizes aggregate wellbeing, private utility net of distortions:

$$W = \int_{\Omega} [u(q, \bar{a}; \theta) - c(\bar{a})q - D(q)]dF + \int_{\Omega^c} [u(q, a; \theta) - c(a)q - D(q)]dF.$$

In this expression, $D(q)$ is the distortion from alcohol consumption, the portion of social harm that would not be incorporated into private decisions. It reflects externalities, and possibly internalities as well. The set of consumers whose quality choices are actively bound by the MUP is denoted Ω, and its complement is Ω^c.

With $\bar{p}, \tau$ as the policy variables for the social planner, consumption will be $q = \hat{q}(\tau, \theta), a = \hat{a}(\tau, \theta)$ for unconstrained consumers, and $q = \tilde{q}(\bar{p}; \theta, \tau), \bar{a} = \tilde{a}(\bar{p} - \tau)$ for constrained consumers.

3.1 Social welfare maximization

Take the planner's first-order condition with respect to $\bar{p}$, and substitute in consumers' private first-order conditions with respect to q:

$$\frac{dW}{d\bar{p}} = \int_{\Omega} \left([\tau - D'(q)]\tilde{q}_{\bar{p}} + MSB_a \tilde{a}_{\bar{p}} \right) dF = 0, \tag{8}$$

where MSB_a is the marginal social benefit of a consumer's chosen beverage quality:

$$MSB_a = u_a(q, a; \theta) - c'(a)q,$$

which is decreasing in a.

[5] Theoretically, a more flexible set of instruments might be considered, that would determine a menu of combinations of price and quality for drinkers to choose from.

Substitute in expressions for $\tilde{q}_{\bar{p}}, \tilde{a}_{\bar{p}}$ from Lemma 4, then multiply through by $c'(\bar{a})$ to deliver the marginal social benefit of $\bar{a}$:

$$\frac{dW}{d\bar{a}} = \int_{\Omega} \left([\tau - D'(q)] q_a^* + MSB_a \right) dF = 0. \tag{9}$$

The planner's second FOC is wrt τ, and it implies (after substituting in consumers' private FOCs wrt q):

$$0 = \int_{\Omega} \left([\tau - D'(q)] \tilde{q}_\tau + MSB_a \tilde{a}_\tau \right) dF + \int_{\Omega^c} \left([\tau - D'(q)] \hat{q}_\tau + MSB_a \hat{a}_\tau \right) dF. \tag{10}$$

This expression can also be rewritten. Note that the quality choices of unconstrained drinkers are not directly distorted. Appeal to the consumers' private first-order conditions with respect to a to show that $MSB_a = 0, \forall \theta \in \Omega^c$.

If there was no MUP, then $\Omega = \emptyset$, and so (10) would simplify to $0 = \int [\tau - D'(q)] \hat{q}_\tau dF$ Consequently, the required corrective tax would be a weighted average of marginal distortions:

$$\tau = \int \alpha(\theta) D'(q) dF, \tag{11}$$

evaluated at the constrained-optimal outcome. The weights represent the share of the overall quantity response that each drinker type is responsible for. With a continuous distribution of θ, the weights are:

$$\alpha(\theta) = \frac{\hat{q}_\tau f(\theta)}{\int \hat{q}_\tau dF}. \tag{12}$$

Lemma 4 implies that $\tilde{q}_\tau = q_\tau^* - \tilde{q}_{\bar{p}}$ and $\tilde{a}_\tau = -\tilde{a}_{\bar{p}}$, which can be substituted in to obtain:

$$\frac{dW}{d\tau} + \frac{dW}{d\bar{p}} = \int_{\Omega} [\tau - D'(q)] q_\tau^* dF + \int_{\Omega^c} [\tau - D'(q)] \hat{q}_\tau dF. \tag{13}$$

If $\bar{p}$ and τ are both set optimally, then $dW/d\tau + dW/d\bar{p} = 0$. Again, the corrective tax should be set to a weighted average of marginal distortions as in (11), but now the weights are:

$$\alpha(\theta) = \begin{cases} \frac{q^*_\tau f(\theta)}{\int_\Omega q^*_\tau dF + \int_{\Omega^c} \hat{q}_\tau dF} & \text{if } \theta \in \Omega \\ \frac{\hat{q}_\tau f(\theta)}{\int_\Omega q^*_\tau dF + \int_{\Omega^c} \hat{q}_\tau dF} & \text{if } \theta \notin \Omega. \end{cases} \tag{14}$$

This generalizes the characterization of corrective taxes for heterogeneous harms proposed by Diamond (1973).

We might expect that a specific tax would be less effective at restraining consumption for consumers that are actively bound by the MUP (according to Lemmas 3, 4. Then, if those consumers whose choices are bound by the Minimum Unit Price are also those whose drinking refects the greatest distortion, then (14) suggests that introduction of a MUP might call for a lower tax rate.

3.2 The case for introducing a MUP

In order to investigate conditions under which a MUP would be beneficial, imagine a benchmark policy setting in which there is no MUP, but in which the corrective tax is otherwise optimal. Now introduce a MUP that is initially just so low that it does not affect any purchases, but is then increased incrementally. The impact on aggregate welfare of this increase would be (by analogy to (9)):

$$\int_{\mathcal{A}} \left([\tau - D'(q)]q^*_a + MSB_a\right) dF,$$

where $\theta \in \mathcal{A}$ are the drinkers buying the cheapest drinks. Without a MUP, consumers' private first-order conditions imply that $MSB_a = 0$, and this will still be approximately true for a levels of $\bar{a}$ close to the minimum level that any drinker would like to buy. It follows that an incremental increase in $\bar{a}$ will be beneficial if the affected drinkers, $\theta \in \mathcal{A}$, should be forced to buy slightly more expensive drinks:

$$\int_{\mathcal{A}} [\tau - D'(q)]q^*_a dF > 0. \tag{15}$$

Consequently, (15) is a sufficient condition for introduction of a MUP to be beneficial. It will be satisfied if the least discriminating drinkers, $\theta \in \mathcal{A}$, are (i) currently undertaxed:

$$\tau < D'(\hat{q}(\tau, \theta)), \forall \theta \in \mathcal{A}$$

and also (ii) treating q, a as local market substitutes, $q_a^* < 0$.

Previous studies have not focused on the extent to which the cheapest varieties of alcohol would be undertaxed without a MUP. However, some suggestive information can be assembled from statistics presented for other purposes. For example, Table 1.1 from Meng et al. (2013) implies that only about 26% of alcohol in their dataset is consumed by "harmful drinkers", but 46% of cheap alcohol is consumed by this category of drinker.

Nor has there been much empirical attention to whether alcohol quality and quantity are substitutes or complements. However, this should be reflected in the sign of the impact of a tax on quality chosen (by Lemma 3) and in the sign of the impact of a MUP on quantity consumed (by Lemma 4). Some research findings relevant to these impacts will be noted in Section 5.

3.3 Baseline example

Next, the optimal policy mix will be illustrated with a numerical example.[6] In order to construct a setting in which a Minimum Unit Price would be beneficial, the joint sufficient condition from Section 3.2 will be imposed.

To ensure that quality and quantity are market substitutes, a suitable functional form for utility will be chosen:

$$u(q, a; \theta) = -\frac{\beta + \mu/a}{q^{\rho}}, \ \beta > 0, \mu > 0, \rho > 0$$

where $\theta = \{\beta, \mu, \rho\}$. As u_a/q is decreasing in q, Lemma 1 applies and so q, a will be substitutes. This functional form will also be convenient when

[6] See the Appendix for a description of the computational procedure.

we explore the consequences of elasticity, because it implies a determinate (fixed-quality) elasticity of demand equal to $-1/(1+\rho)$.

To ensure that purchases of the cheapest drinks are associated with higher marginal distortions, I assume that the distortion function is convex in consumption, and calibrate the utility parameters to data in which cheap drinks are disproportionately bought by heavy drinkers. This consumption data is summarized in the second panel of Table 1. It is drawn from Table 4 of Kerr & Greenfield (2007), and adjusted for underreporting. They present average consumption levels and prices paid for five categories of consumer, distinguished by how heavily they drink. Their data is drawn from the 2000 National Alcohol Survey conducted in the U.S.A.

The utility function should be consistent with inelastic demand, to conform with the empirical literature. Wagenaar et al. (2009) conduct a meta-analysis of elasticity estimates, and report a simple mean for elasticity estimates of general alcohol demand equal to -0.51. The value of ρ is set to 0.95 which delivers a (fixed quality) elasticity of demand equal to -0.513.

Values for the other two utility parameters β, μ are calibrated so that the model predicts Kerr & Greenfield's consumption and price data, assuming a benchmark specific tax of $\tau = 0.1$ and no MUP. The resulting parameter values are presented in the third panel of Table 1.

Unit costs are assumed to be linear in quality, $c(a) = 0.3+0.1a$. Following Manning et al. (1989), it is assumed that there is no distortion to alcohol consumption under 3.0 ounces a day, but a distortion of 1.19 in 1986 dollars for each ounce in excess of this level.[7] After adjusting this value to reflect the date of the survey used by Kerr and Greenfield, the distortion function becomes $D(q) = \max\{0, q-3\} \times 1.87$.

The fourth panel of Table 1 deals with the constrained welfare optimum under the assumption that the policymaker does not have the capacity to

[7] Their figure does not include internalities.

Consumer category	A	B	C	D	E
Proportion of consumers	5%	5%	15%	25%	50%
Initial consumption					
ozs of ethanol per day	10.975	3.977	2.122	0.841	0.153
price paid per oz	1.32	1.30	2.23	3.37	7.92
Utility parameters					
implied values of β	50.16	6.91	2.24	0.41	0.02
implied values of μ	904.4	119.6	145.2	62.9	14.5
$\tau = 0.77, \bar{p} = 0$					
ozs of ethanol per day	6.78	2.45	1.34	0.54	0.11
price paid per oz	2.54	2.51	3.94	5.63	11.90
$\tau = 0.21, \bar{p} = 2.52$					
ozs of ethanol per day	6.27	2.25	1.89	0.75	0.14
price paid per oz	2.52	2.52	2.57	3.83	8.73

Table 1: Benchmark case, $\rho = 0.95$

impose a MUP. In this case, the optimal tax would be $\tau = 0.77$. This may seem reasonably high. After all, the only consumption that requires correction is by the heaviest drinking type of consumer, which only makes up five percent of the population.[8] If we failed to account for differences of the slopes of decision rules among consumer types, we might expect the tax rate to be equal to the marginal distortion on this type (1.87) multiplied by the proportion of the population whose alcohol consumption is distorted ($\pi_1 = 0.05$), giving a tax rate equal to only $\tau = 0.0935$. However, if we interpret equation (13) in terms of discrete types, the optimal tax rate would

[8] This is comparable to the 5.67% of the British population classified as "harmful drinkers" by Ludbrook et al. (2012) and the 5.3% given this classification by Meng et al. (2013). However, these figures do not include drinkers classified as "hazardous".

actually be:

$$\tau = \frac{\hat{q}_\tau(\tau;\theta_1) \times 0.05}{\sum_{j=1}^{5} \hat{q}_\tau(\tau;\theta_j)\pi_j} = 0.77.$$

The tax has a much larger marginal impact on consumption by heavy drinkers than on lighter drinkers, and consequently it is effective at reducing harmful consumption without creating a substantial distortion on moderate drinkers. The reason is that every category of consumer has the same elasticity of demand, but some categories have much higher levels of consumption and pay significantly lower prices. As a result, these categories have much larger absolute impacts on their drinking from higher taxes.

The final panel of Table 1 pertains to the optimal combination of a specific tax and a MUP. The policy combination that maximizes aggregate payoffs net of distortions involves a lower tax on the alcohol content of beverages equal to $\tau = 0.50$, but a hefty minimum unit price of 2.18 per ounce. The optimal MUP is above the average price that would otherwise be paid by the heaviest two categories of drinker.

It is no surprise that the corrective tax is lower when a MUP is available. The two policy instruments are alternative approaches to restraining demand with higher prices. More formally, the optimal tax, (11), can be expected to be lower when weights are given by (14) rather than (12). Because q^*_τ is less than $\hat{q}_\tau$ in absolute value by Lemma 3, there will be reduced weights on the largest distortions when a MUP is applied. When the MUP is binding on the most distorted drinking, then that drinking will be less influenced by a corrective tax.

It is also notable that the price paid by the heaviest drinkers is no higher when a MUP is used than when only a specific tax is available. Whether this price is increased or decreased by the availability of the second policy instrument depends on the specific calibration. Although a MUP can help the policymaker to target heavy drinkers more accurately, it also introduces a second mechnanism for restraining heavy consumption. That is, consumption by category A is dampened not only by higher prices but also because (by assumption) the marginal utility of quantity is decreasing in the quality of alcohol.

4 Elasticity

In the previous literature, the effects of taxes and MUPs have typically been analyzed in terms of price elasticities. However, a variety of such measures can be distinguished in the current model, based on various slopes of decision rules identified in Section 2.3. The following discussion will deal with the simplest, the fixed-quality price elasticity of demand for alcohol. This is the proportional change in quantity demanded purchased over the proportional change in price, assuming that the quality of alcohol is fixed. In the example presented in Section 3.3, the (absolute value of) this elasticity would be $\varepsilon = 1/(1+\rho)$.

4.1 Inelastic demand

One concern is whether the overall elasticity of demand has sufficient magnitude for a price instrument such as a MUP or tax to be effective. This concern contrasts with the literature on corrective taxes in first-best environments, where the corrective tax should be set equal to the marginal distortion, irrespective of the magnitude of the elasticity. The current discussion deals with heterogeneous harms, and so the marginal distortion differs among consumers. However, the implications of elasticity for the optimal policy choice may still not be obvious.[9]

[9] Elasticities can be important when policymakers have redistributive motives. For example, if those with lower incomes have higher consumption (due to preference heterogeneity) and if elasticity is low, then a tax may impose considerable burden on low income

As in Section 3.3, the model is calibrated to match summary statistics from Kerr & Greenfield (2007). The only difference is that a range of values of ρ will now be considered.

Elasticity of demand is determined by parameters of the utility function, and as a result is difficult to vary without changing other features of the model. For example, if an increase in ρ was not combined with other parameter changes, it would result in increased consumption of alcohol and hence a larger distortion. But the increase in the distortion would call for a higher corrective tax, independent of the elasticity of demand. In order to isolate this effect, the other two utility parameters, β, μ, are recalibrated for every value of ρ considered, so that baseline consumption of alcohol is unchanged. However, as will be discussed below, changes in ρ will still have other implications.

The results are summarized in Figure 1. As in the benchmark case of $\rho = 0.95$, the optimal corrective tax rate is smaller when it is combined with a MUP than when it is used in isolation. In addition, it is more sensitive to changes in elasticity when it is the only policy instrument used. For example, comparing $\rho = 0.5$ with $\rho = 1.0$ (elasticities of 2/3 versus 1/2) shows that the corrective tax at $\rho = 1.0$ is only about 37% of the level with $\rho = 0.5$ when a MUP is also available. But when the tax is the only policy instrument, this ratio is about 94%.

One way to think about the insensitivity of the optimal tax rate to changes in elasticity is to consider the special case in which product quality is fixed (or equivalently where $q_a^* \equiv 0$) so the tax has no impact on a. Then $\hat{q}_\tau = q_\tau^*$ by Lemma 3, and the variable-quality elasticity of demand is the same as the fixed-quality elasticity, $\varepsilon = q_\tau^* p/q$. The expression for the optimal tax

consumers without much behavioral consequence (Lockwood & Taubinsky 2017).

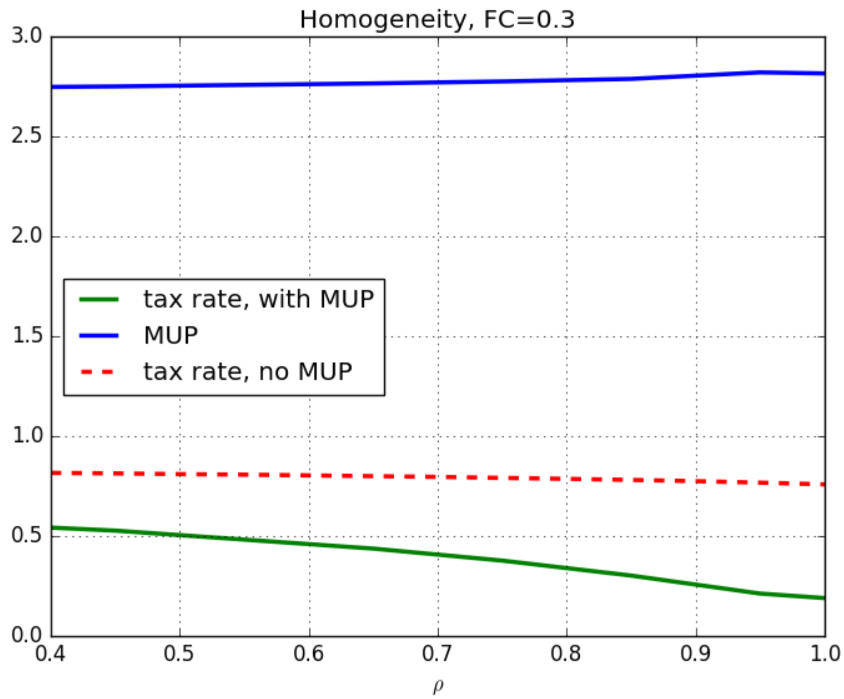

Figure 1: Changes in ρ

rate, (11) would simplify (still assuming that no MUP is in place) to:

$$\tau = \frac{\int \hat{q}_\tau D'(q) dF}{\int \hat{q}_\tau dF} = \frac{\int \varepsilon D'(q) q/p dF}{\int \varepsilon q/p dF}.$$

When the elasticity of demand, ε, is common to all drinkers it cancels out of this expression. Moreover, as we recalibrate the other parameters to keep each drinker's q/p constant as the elasticity changes, there is no impact of a change in elasticity on the optimal tax. As with first-best environments, changes to a common elasticity of demand would have little relevance for the optimal tax rate.

This account needs to be amended to acknowledge that q_a^* will typically be nonzero. The optimal tax rate does change a little with ρ because a

is endogenous. More importantly, if q^*_a is negative then an increase in the corrective tax (which moves a consumer along $a = a^*(q; \theta)$) and an increase in the MUP (which moves a consumer along $q^*(a; \theta, \tau)$) can be viewed as substitutes. Both policy changes would reduce consumption and also increase the quality of alcohol consumed.

Changing ρ has implications for how close a substitute the MUP is for the specific tax, i.e., how similar the slope of $q^*(a; \theta)$ is to the inverse of the slope of $a^*(q; \theta)$). In the current example, a higher value of ρ means that the two policy instruments are more closely substitutable. Then a given decrease in consumption would be accompanied by a similar change in quality, irrespective of whether that decrease was due to a tax change or a change to the MUP.

As ρ gets higher, the policymaker increasingly relies on the MUP. Its effect on heavy drinkers is increasingly similar to that of a corrective tax, but without distorting the decisions of moderate drinkers who pay more than the MUP for their alcohol. As a result, the optimal policy involves a higher MUP, and consequently a larger difference between $\bar{p}$ and τ, $c(\bar{a}) = \bar{p} - \tau$.

4.2 Lower elasticities for heavy drinkers

Another concern about the use of Minimum Unit Prices relates to heterogeneity in elasticities of demand. Although there is some dispute in the empirical literature, it is sometimes claimed that heavy drinkers may have lower elasticity of demand than moderate drinkers (Kenkel 1996, Wagenaar et al. 2009).

Unlike the the low common elasticity of demand considered in Section 4.1, this is well-established as potentially undermining the benefits of a corrective tax. Kenkel (1996) derives expressions for the deadweight losses from excessive alcohol consumption and from specific taxes. These expressions are phrased in terms of the ratio of elasticities between different categories of drinker. Similarly, Bernheim & Rangel (2004) show that inelasticity of compulsive demand can undermine the rationale for a corrective tax on addictive substances. When consumption of moderate drinkers is not distorted but the consumption of heavy drinkers is totally inelastic, a specific tax will be ineffective in restraining heavy drinkers and only serve to distort consumption of moderate drinkers.

A simple way to incorporate heterogeneity of elasticities in the model of Section 3.3 is to introduce a parameter γ, so that:

$$u = -\frac{\beta + \mu/a}{q^\rho} - \gamma q.$$

Effectively, Section 3.3 assumed that $\gamma = 0$. With this new parameter, the (absolute value of) fixed-quality elasticity of demand is:

$$\varepsilon = \frac{1}{1+\rho} \cdot \frac{p}{p+\gamma}.$$

Consequently, the consumers who pay higher prices for alcohol (predominantly light and moderate drinkers in the data reported by Kerr & Greenfield) will have more elastic demand.

An increase in γ will make demand less elastic, and will also increase the extent to which demand by heavy drinkers is less elastic than demand by moderate drinkers. Figure 2 illustrates the results for a range of values of γ. The value of ρ is kept equal to 0.95. As with the previous analysis, the values of β, μ are recalibrated to match the consumption figures reported by Kerr & Greenfield.

Some aspects of these results are similar to Figure 1. For example, the optimal tax rate does not change as much when used in isolation, as when it is accompanied by a MUP. However, there is also a striking difference from Figure 1. This time decreases in elasticity are associated with an *increased* corrective tax even though the difference between the elasticities of heavy

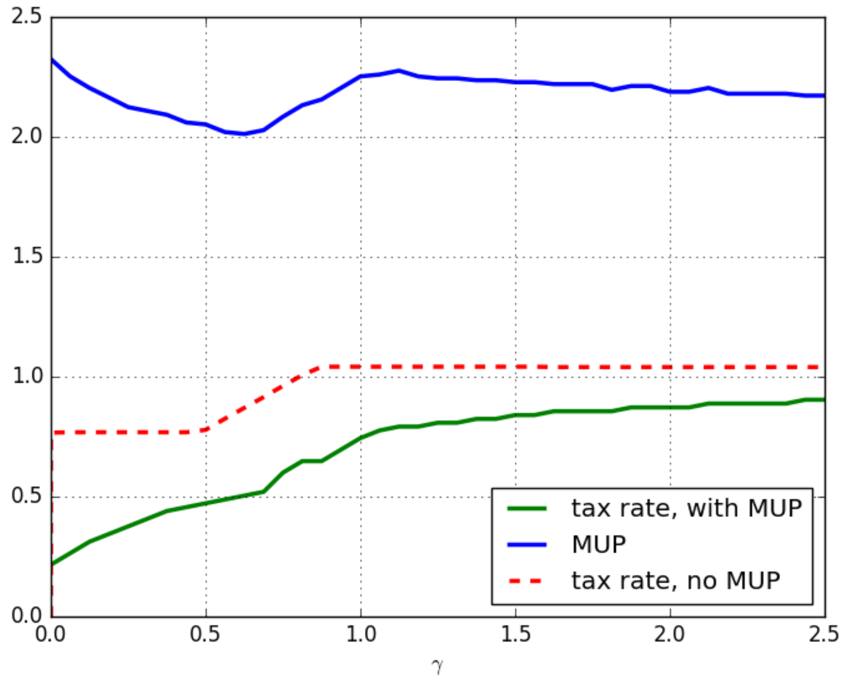

Figure 2: Changes in γ

and moderate drinkers has also increased. The reason is that decreases in elasticity (which now result from increases in γ) are associated with a greater discrepancy between $q_a^\star$ and $(a_q^\star)^{-1}$, and consequently with the tax and the MUP becoming less substitutable.

5 A flight from quality?

The final consideration to be addressed is whether a corrective tax would induce compensating behavior, in the form of quality substitution. The possibility has been raised that MUPs may have an advantage over excise taxes, if the latter encourage consumers to substitute to lower quality alcohol, and possibly as a result be less effective at reducing alcohol consumption (Babor 2010, Sharma et al. 2014).

5.1 Empirical literature

Gruenewald et al. (2006) suggest that alcohol price increases may lead consumers to substitute to lower quality, on the basis of an investigation of Swedish data. There have also been claims that consumers substitute to lower quality for other goods. For example, Gibson & Kim (2013) report that increases in the price of rice resulted in a shift to low quality rice.

However, there is also a literature that suggests the opposite effect for a range of goods. Evidence that higher prices might induce consumers to substitute to *higher* quality has been presented for the markets for cigarettes (Sobel & Garrett 1997, Chiou & Muehlegger 2014) and gasoline (Nesbit et al. 2007).

Additional evidence on whether quality and quantity are substitutes might be found in estimates of the effect of a Minimum Unit Price on quantity consumed. Lemma 4 proposes that a MUP would not be locally effective at reducing consumption unless q, a were substitutes. Stockwell et al. (2012) report that Canadian price floors for alcohol have been effective in decreasing the quantity of alcohol consumed.

5.2 Theoretical analysis

Early theoretical studies emphasized substitution to higher quality, as in the "flight to quality" thesis of Barzel (1976). Bohanon & Van Cott (1991) argue that the magnitude of such an effect depends on the degree of substitutability between quality and quantity. However, Keen (1998) argues that while a flight to quality is more plausible than a flight from quality, the latter is also possible.

The analysis in Section 2 above supports Keen's conclusion, and moreover provides a condition under which substitution will be toward rather than away from quality. For unconstrained consumers, an incremental increase in the specific tax will induce purchases with higher quality if q, a are local market substitutes, but with lower quality if they are complements. This is an immediate consequence of Lemma 3.

For consumers whose decisions are actively bound by the Minimum Unit Price, an incremental increase in the MUP induces the consumer to choose higher quality, but an increase in the specific tax would induce lower quality. This is a consequence of Lemma 4.

These findings do not suggest that there is an advantage for a MUP, due to specific taxes being compromised by downward quality substitution. The first reason is that the decline in alcohol consumption as a result of a specific tax would be greater with quality substitution than without. This follows from Lemma 3. The second reason is that we can expect downward quality substitution when q, a are local market complements. But this is just when a Minimum Unit Price would be (locally) ineffective at reducing consumption, by Lemma 4.

The core of the preceding argument is that the direction of the effect of a small change in the MUP on quantity consumed, or of the effect of a small change in the specific tax on quality chosen, depends on whether q, a are local complements or substitutes. But it is worth noting that this is only a local result. Even if q_a^* was positive at the initial equilibrium, a large change in the MUP might move the equilibrium to a section of $q^*(a; \theta, \tau)$ that has the opposite slope. Consequently, large changes in the MUP may sometimes reduce consumption, even when the best responses had positive slopes in an equilibrium without a MUP.

6 Conclusion

The agenda of the current paper is to formalize the rationale for Minimum Unit Prices in an optimal tax framework, and to examine the implications of inelastic demand for alcohol and quality substitution for the optimal policy.

The first conclusion is that there is a case for a Minimum Unit price when both (i) purchases of the cheapest forms of alcohol are relatively undertaxed, and (ii) quality and quantity of alcohol are local market substitutes.

The second conclusion is that a lower elasticity of demand does not have a generally determinate implication for the magnitude of corrective taxes and Minimum Unit Prices. It depends on the cause of this lower elasticity, and whether it increases or decreases the degree to which corrective taxes and Minimum Unit Prices are close substitutes.

The third conclusion is that the danger of downward quality substitution in response to a specific tax does not provide a compelling reason to rely on a Minimum Unit Price rather than a corrective tax.

The preceding analysis neglects a range of concerns that have been raised in the existing literature. One such concern is with equity, and the burden on low-income drinkers (Holmes et al. 2014). Another is loss-leading, imperfect competition and other sources of variation in the passthrough of taxes (Ally et al. 2014). A third is the revenue motivation for corrective taxes (Parry et al. 2009), and a fourth is the degree to which alcohol is a complement or substitute to other intoxicants (Moore 2010).

Appendix: the numerical example

A Utility maximization

Consider the example set out in Section 3.3. It is possible to derive closed-form expressions for some of the decision rules. Let $c(a) = \kappa_0 + \kappa_1 a$. Then the two private first-order conditions are:

$$\frac{(\beta + \mu/a)\rho}{q^{1+\rho}} - [\tau + \kappa_0 + \kappa_1 a] = 0, \qquad \frac{\mu}{a^2 q^\rho} - \kappa_1 q = 0.$$

Consequently:

$$\frac{1}{q^{1+\rho}} = \frac{\tau + \kappa_0 + \kappa_1 a}{(\beta + \mu/a)\rho} = \frac{\kappa_1 a^2}{\mu},$$

which can be solved for a:

$$\hat{a} = \frac{(1-\rho)\kappa_1\mu \pm \sqrt{(1-\rho)^2\kappa_1^2\mu^2 + (\kappa_0 + \tau)4\beta\rho\kappa_1\mu}}{2\beta\rho\kappa_1}.$$

The second term must be added rather than subtracted, as we know that (i) $\hat{a}_\tau \geq 0$ for local market substitutes, and (ii) a must be positive.

The best response for q can be found from the consumer's first-order condition with respect to q:

$$q^* = \left(\frac{\beta + \mu/a}{\tau + c(a)}\right)^{\frac{1}{1+\rho}}.$$

B Computational procedure

Benchmark values of ρ, q, a are chosen from the elasticity and consumption data. Given these values, the consumers' first-order conditions are inverted to identify values of β, μ. Then for each combination of the two policy instruments $\tau, \bar{p}$, the utility maximizing levels of q, a are derived for each of the five consumer types. These values are substituted into total welfare, the sum of private utilities plus tax revenue and less any distortion. Then we identify the policy setting that generates the highest level of total welfare.

References

Ally, A. K., Meng, Y., Chakraborty, R., Dobson, P. W., Seaton, J. S., Holmes, J., Angus, C., Guo, Y., Hill-McManus, D., Brennan, A. & Meier, P. S. (2014), 'Alcohol tax pass-through across the product and price range: do retailers treat cheap alcohol differently?', *Addiction* **109**(12), 1994–2002.

Aronsson, T. & Sjögren, T. (2010), 'An optimal-tax approach to alcohol policy', *FinanzArchiv: Public Finance Analysis* **66**(2), 153–169.

Babor, T. (2010), *Alcohol: No Ordinary Commodity: Research and Public Policy*, Oxford medical publications, OUP Oxford.

Barzel, Y. (1976), 'An alternative approach to the analysis of taxation', *The Journal of Political Economy* pp. 1177–1197.

Bernheim, B. D. & Rangel, A. (2004), 'Addiction and cue-triggered decision processes', *The American Economic Review* **94**(5), 1558–1590.

Bohanon, C. E. & Van Cott, T. N. (1991), 'Product quality and taxation: A reconciliation', *Public Finance Review* **19**(2), 233–237.

Brennan, A., Meier, P., Purshouse, R., Rafia, R., Meng, Y., Hill-Macmanus, D., Angus, C. & Holmes, J. (2015), 'The Sheffield Alcohol Policy Model – a mathematical description', *Health Economics* **24**(10), 1368–1388.

Chiou, L. & Muehlegger, E. (2014), 'Consumer response to cigarette excise tax changes', *Available at SSRN 1693263* .

Christiansen, V. & Smith, S. (2012), 'Externality-correcting taxes and regulation', *The Scandinavian Journal of Economics* **114**(2), 358–383.

Crawford, I., Keen, M. & Smith, S. (2010), 'Value added tax and excises', *Dimensions of tax design: the Mirrlees review* pp. 275–362.

Diamond, P. A. (1973), 'Consumption externalities and imperfect corrective pricing', *The Bell Journal of Economics and Management Science* pp. 526–538.

Gibson, J. & Kim, B. (2013), 'Quality, quantity, and nutritional impacts of rice price changes in Vietnam', *World Development* **43**, 329–340.

Gruenewald, P. J., Ponicki, W. R., Holder, H. D. & Romelsjö, A. (2006), 'Alcohol prices, beverage quality, and the demand for alcohol: quality substitutions and price elasticities', *Alcoholism: Clinical and Experimental Research* **30**(1), 96–105.

Holmes, J., Meng, Y., Meier, P. S., Brennan, A., Angus, C., Campbell-Burton, A., Guo, Y., Hill-McManus, D. & Purshouse, R. C. (2014), 'Effects of minimum unit pricing for alcohol on different income and socioeconomic groups: a modelling study', *The Lancet* **383**(9929), 1655–1664.

Katikireddi, S. V. & McLean, J. A. (2012), 'Introducing a minimum unit price for alcohol in Scotland: considerations under European law and the implications for European public health', *The European Journal of Public Health* **22**(4), 457–458.

Keen, M. (1998), 'The balance between specific and ad valorem taxation', *Fiscal Studies* **19**(1), 1–37.

Kenkel, D. S. (1996), 'New estimates of the optimal tax on alcohol', *Economic Inquiry* **34**(2), 296–319.

Kerr, W. C. & Greenfield, T. K. (2007), 'Distribution of alcohol consumption and expenditures and the impact of improved measurement on coverage of alcohol sales in the 2000 National Alcohol Survey', *Alcoholism: Clinical and Experimental Research* **31**(10), 1714–1722.

Lockwood, B. B. & Taubinsky, D. (2017), Regressive sin taxes, Working Paper 23085, National Bureau of Economic Research.

Ludbrook, A., Petrie, D. & Farrrar, S. (2012), 'Tackling alcohol misuse', *Appl Health Econ Health Policy* **10**(1), 51–63.

Manning, W. G., Keeler, E. B., Newhouse, J. P., Sloss, E. M. & Wasserman, J. (1989), 'The taxes of sin: do smokers and drinkers pay their way?', *JAMA* **261**(11), 1604–1609.

Meier, P., Purshouse, R. & Brennan, A. (2010), 'Policy options for alcohol price regulation: The importance of modelling population heterogeneity', *Addiction* **105**(3), 383–393.

Meng, Y., Brennan, A., Holmes, J., Hill-McManus, D., Angus, C., Purshouse, R. & Meier, P. (2013), 'Modelled income group-specific impacts of alcohol minimum unit pricing in England 2014/15: policy appraisals using new developments to the Sheffield Alcohol Policy Model (v2. 5)'.

Moore, S. C. (2010), 'Substitution and complementarity in the face of alcohol-specific policy interventions', *Alcohol and Alcoholism* **45**(5), 403–408.

Nesbit, T. et al. (2007), 'Excise taxation and product quality: the gasoline market', *Economic Issues* **12**(2), 1–14.

Parry, I. W., West, S. E., Laxminarayan, R. et al. (2009), 'Fiscal and externality rationales for alcohol policies', *The BE Journal of Economic Analysis & Policy* **9**(1), 1–29.

Pogue, T. F. & Sgontz, L. G. (1989), 'Taxing to control social costs: the case of alcohol', *The American Economic Review* **79**(1), 235–243.

Saffer, H. & Chaloupka, F. (1994), 'Alcohol tax equalization and social costs', *Eastern Economic Journal* **20**(1), 33–43.

Sharma, A., Vandenberg, B. & Hollingsworth, B. (2014), 'Minimum pricing of alcohol versus volumetric taxation: which policy will reduce heavy consumption without adversely affecting light and moderate consumers?', *PloS One* **9**(1), e80936.

Sobel, R. S. & Garrett, T. A. (1997), 'Taxation and product quality: new evidence from generic cigarettes', *Journal of Political Economy* **105**(4), 880–887.

Stockwell, T., Zhao, J., Giesbrecht, N., Macdonald, S., Thomas, G. & Wettlaufer, A. (2012), 'The raising of minimum alcohol prices in Saskatchewan, Canada: impacts on consumption and implications for public health', *American Journal of Public Health* **102**(12), e103–e110.

Topkis, D. M. (1978), 'Minimizing a submodular function on a lattice', *Operations Research* **26**(2), 305–321.

Wagenaar, A. C., Salois, M. J. & Komro, K. A. (2009), 'Effects of beverage alcohol price and tax levels on drinking: a meta-analysis of 1003 estimates from 112 studies', *Addiction* **104**(2), 179–190.