Working Paper

Merger review using online experiments

Farasat A.S. Bokhari [a],
Stefan P. Penczynski [b], and Janithe Siriwardana [a,b]

[a] *School of Economics, Centre for Competition Policy, University of East Anglia*

[b] *School of Economics, Centre for Behavioural and Experimental Social Science, University of East Anglia*

March 19, 2023

Abstract

Merger simulation is a complex exercise and is difficult to implement during merger assessment due to data and time requirements. We use data from an online experiment to estimate demand parameters for beers market and combine this with aggregate national level data on prices, shares and attributes of beers. These allow us to calculate elasticities, markups and marginal costs for a set of real products which compare well to reported estimates in the literature. Our proposed method offers a cheap and fast way of implementing modern IO methods for evaluating cases in real time.

Key words: Merger simulation, demand estimation, preference experiment, alcohol

---

1. Introduction

Several measures that are used to evaluate mergers including price elasticity of demand, diversion ratios and merger simulation exercises require the estimation of demand and demand side parameters. One of the modern workhorse models of demand estimation is the random coefficients mixed logit model (RCML) of Berry et al. (1995, henceforth BLP) and Nevo (2000) which produces more realistic elasticities of demand by allowing for arbitrary correlation between prices and markups, and flexible substitution patterns. Rather than estimating a fixed point for each parameter, we estimate a distribution which represents heterogeneity in preferences between individuals. There are however, several challenges in the estimation procedure such that the models are somewhat underutilised by researchers and agency practitioners. The first of these challenges is in data collection. The models require at a minimum, aggregate level data of purchases obtained from a single market. Data from several markets is advantageous because it results in greater variation in relative prices of the products and/or products offered. However, this can be time consuming and costly to obtain. In many industries there is simply a dearth of information on sales volumes and prices; Moshary et al. (2022) find there is no centralized database that contains information about either individual-level or aggregate gun purchases matched with prices while aggregate proxies for purchases that have been used in previous research are neither detailed to the gun model nor matched with prices. The models also require data on demographic variables which at best can only be approximated by good census data. Finally, prices are often correlated with unobserved variables resulting in endogeneity; this requires a set of relevant and exogenous instrument variables to solve. These issues are challenges for any researcher attempting to estimate demand but particularly for an anti-trust agency evaluating a merger in real time, these data issues represent significant hurdles to a timely analysis.

To circumvent these difficulties with BLP type models, we borrow from the experimental literature on discrete choice experiments, to construct a stated preference (SP) experiment in which subjects are required to make repeated choices on sets of beers. From these repeated choices, we approximate mixed logit choice probabilities and estimate demand parameters to quickly and cheaply enable initial merger simulations. This is in contrast to revealed preference (RP) data used in empirical methods. In our SP experiment, we define a matrix of attributes and levels that combine to form possible alternatives that result in a set of pseudo-products used to estimate demand parameters. Each individual was presented with four alternatives from the possible set of 18 pseudo-products, with each alternative priced at one of 3 values. Individuals were instructed to select their preferred option in each choice set. We combine this micro-data with aggregate level data on real products to obtain a price elasticity of demand matrix for the product set as well as associated price-cost margins, creating an alternative tool for competition economists to use. For industry/regulatory practitioners, the advantage of an SP experiment is that once an appropriate experimental design has been conceived it can be retooled for many different products/situations and implemented quickly [1]. An online survey using existing platforms such as Prolific or Amazon Mechanical Turk can produce thousands of observations from very specific groups of consumers in a matter of days. The researcher also has the ability to randomly set prices in the experiment, so the problem of endogeneity is solved. This enables us to use a mixed logit model to derive choice probabilities and maximum simulated likelihood (MSL) to estimate parameters using a less complex estimation algorithm than BLP models. However, our methodology can also be used in situations where empirical data does not exist or would be very difficult to collect (see the handgun example above) or as an alternative methodology to corroborate empirical estimates.

[1] Regulators often have tight deadlines when conducting merger reviews. The UK Competition Markets Authority (CMA) has 40 working days to complete Phase 1 and a further 24 weeks during Phase 2 to conduct their investigation and submit a final report. In the US, where the Federal Trade Commission (FTC) and Department of Justice (DoJ) are jointly responsible for merger analysis, pre-merger reviews must be completed within 30 days and if necessary the agencies are granted another 30 days to investigate further and take action if required.

Individuals were paid for their participation but rewards were not contingent on their choices. This raised the issue of the role of, or rather lack of, incentivisation in our experiment and how this compares to observed choices garnered through supermarket (RP) data which are necessarily consequential to a consumer because they use funds from their own endowment and actually take home the product. However, we felt that attempting to incorrectly include incentives could take us further away from the real-world ideal.

Following BLP's seminal work on RCML's, a range of papers have sought to improve the performance of these models. Nevo (2000, 2001) attempts to guide practitioners through the model using the ready to eat cereal market as an example. Petrin (2002) uses micro moments obtained using consumer level data to augment market-level data and estimate a demand model for minivans. There is also a related literature on discrete choice models (e.g. Train (2009)) from which we borrow heavily from. Elsewhere Reynaert and Verboven (2014) and Rossi (2014) focus on instrument variables and their role within RCML type models. Others such as Bajari et al. (2007), Fosgerau and Bierlaire (2007), Train (2008), Bastin et al. (2010), and Fosgerau and Mabit (2013) all seek to open up the models beyond the potential misspecification that can occur through the use of inappropriate parametric mixing distributions by introducing more flexible distributions. Train (2016) proposes a logit-mixed logit (LML) model that is a generalised specification that combines much of the previous work in this area; Bansal et al. (2018) extend the LML for a combination of random and fixed parameters. Lee and Seo (2016) work is an example of a third area of research involving the properties of the nested fixed point algorithm at the heart of BLP type models. We place our paper in a fourth strand of the literature devoted to novel, often experimental, methods in demand estimation and the use of demand models in a variety of applied settings.

Conlon and Mortimer (2013) conducted some of the earliest experimental work in merger analysis in response to changes in the DoJ/FTC Horizontal Merger Guidelines that set new standards based on upward pricing pressures (UPP) which in turn rely on diversion ratios. They estimate diversion by exogenously removing products from vending machines and analysing changes in demand, firm profits, diversion ratios and UPP. However, this type of field experiment is both costly and time consuming. In contrast, our experiment is cheap and quick to implement. Conlon and Mortimer (2021) follow up their previous work by establishing a local average treatment effect (LATE) interpretation of diversion ratios and show how diversion ratios can be estimated using different interventions. Although they mention the potential to use a lab (or online) experiment, the paper does not implement any experiments.

Moshary et al. (2022) conduct a similar experiment to ours in that they present subjects with choice sets in an experimental setting in order to elicit demand preferences in the market for firearms. Having obtained substitution patterns for various types of guns, they simulate changes in gun regulations and use the estimated demand model to assess changes in demand and consumer surplus. While the foundation of the experiment is similar, we use our demand model to simulate a merger rather than evaluate regulatory policy.

Magnolfi et al. (2022) take a different approach to experimental demand estimation by using a triplet experiment where subjects are presented with a reference product and are asked to select the two products that are most similar to the reference from a given choice set. They then use a machine learning algorithm to estimate an embedding – a low-dimensional representation of the latent product space. Substitution patterns can be inferred from the distances between product pairs in the embedding. Two other papers also use embeddings in demand estimations. Bajari et al. (2021) use deep neural nets to generate an embedding from products image and text descriptions, useful in cases where the demand relevant information may not easily be defined by a set of measurable characteristics, even though humans are able to process and synthesise the relevant information. However, a key difference between our work is that the embedding serves to augment price and quantity data in a traditional demand estimation model. While we require some data on price and quantity, the requirements are less strenuous (we use readily available national level data) and serve to augment our experimental data. Armona et al. (2021) use search data to estimate consumer preferences for hotels by using a

Bayesian Personalised Ranking to learn products' latent characteristics from consumers web-browsing history. We see these latent attribute methods as complementary to our work using observable product characteristics. As Armona et al. (2021) themselves state 'if the observables are rich, the value add from latent characteristics may be smaller'.

The results of our experiments are promising. Following the estimation of the demand parameters, we use these to estimate substitution effects and markups so it is these that we ultimately compare to previous studies. We calculate elasticities for a set of real products that consists of the 18 most popular beers by market share in 2019. This attempts to place our demand parameters in context by comparing them to results observed by Miller and Weinberg (2017) in work analysing the effects of the Miller-Coors joint venture in 2008. It should be noted that the data set they use is not contemporaneous to ours; the product set is different and the structure of the industry has changed so direct comparisons between our results and those of Miller-Weinberg are not possible. We simply use their results to show that our method can produce what appear to be realistic values for individual product elasticities as well as median own-price elasticities. The median own-price elasticity for our real product set of -4.83 falls close to the range that Miller-Weinberg report of -4.73 – -4.33 for their various random coefficient nested logit specifications, suggesting the methodology can produce realistic substitution patterns. To illustrate our point regarding data requirements, Miller and Weinberg use a dataset that spans 39 geographic regions over years 2001-2011. In a meta-analysis of beer studies globally, Nelson (2014) finds the average market price elasticity is around -0.2 compared to our estimate of -0.13. Our predicted markups in the range of 22.5-23% are lower than Miller and Weinberg's estimated 34%. As has been explicitly stated by Train (2009) and others, the success of these models depends on correctly identifying relevant observable product characteristics and the density of preferences for these characteristics in the population. As we found, the challenge remains correct specification of the indirect utility function particularly in an industry where branding plays a significant role in purchase decisions. Since our model did not include brand effects, we found it did not always accurately predict market shares for some products which had similar observed product characteristics but in reality had markedly different actual market shares. This is the underlying cause of the discrepancy in individual own-price elasticities between our specification and Miller-Weinberg's. However, since this initial model was designed to be the simplest version possible, there are both experimental and econometric changes we could make to improve the accuracy of our elasticity estimates. The trade off is an increase in complexity in both the experiment and the estimation algorithm. We discuss this in greater detail in our concluding remarks.

The rest of the paper is organised as follows. Section 2 describes our model that encompasses elements from various strands of the existing literature. We define indirect utility, choice probabilities, a simulation method, price elasticities and price-cost margins. In section 3, we detail the experimental design guided by Hensher (1994), including testing of experimental features through Monte Carlo simulations, further detailed in Appendix A. Section 4 provides an example of the types of results the estimation procedure can produce and attempts to place them in the context of existing work. Finally, we conclude in section 5 by discussing variations and improvements to the methods discussed previously.

2. Model Specification

The mixed logit model is in the class of random utility models (RUM) derived from assumptions of utility maximisation. Individual n faces a choice between products $j \in J$ over a set of $t \in T$ time periods or choice situations. The utility individual n derives from product j in choice situation t is

$$U_{njt} = \beta_n' x_{njt} + \varepsilon_{njt} \tag{1}$$

An individual will choose product i if and only if $U_{nit} > U_{njt}\ \forall j \neq i$. β_n is a vector of coefficients on the product characteristics that is unobserved for the sample and varies in the population with density $f(\beta|\theta^*)$ where θ^* are the true location and scale parameter of the population distribution. x_{njt} is a vector of observed product characteristics for each alternative in each choice set.

Each individual has their own value of β_n that can be estimated and represents their tastes and preferences over the defined product characteristics. The values of these β_n's are distributed over the population with parameters θ^*. It is these population parameters, θ^*, that we seek to estimate through the mixed logit model. For a known β_n i.e. conditional on the value of β_n, the choice probability of product i for individual n is given by the multinomial logit (MNL) expression

$$L_{nit}(\beta_n) = \frac{\exp(\beta_n' x_{nit})}{\sum_J \exp(\beta_n' x_{njt})} \tag{2}$$

When an individual makes repeated choices over T situations, let $y_n = \langle y_{nt}, ..., y_{nT} \rangle$ represent the chosen alternative in each t. Then the probability of an individual's sequence of choices, y_n is the product of MNL's across each $t \in T$

$$P_n(y_n|\beta_n) = \prod_{t=1}^{T} L_{nt}(y_{nt}|\beta_n) \tag{3}$$

Since each individual's β_n is unobserved, the exact unconditional probability of n's sequence of choices is the integral of the conditional probability over all possible values of β as defined by the true parameters of the distribution of β_n, θ^*

$$P_n(y_n|\theta^*) = \int P_n(y_n|\beta_n) f(\beta|\theta^*) d\beta \tag{4}$$

Having obtained the unconditional probability of each individual's sequence of choices we use maximum likelihood (ML) procedures to derive estimates for θ^*. It can be shown that the asymptotic distribution of the ML estimator, $\hat{\theta}$ is centred on the true values, θ^* and its variance decreases as the sample size increases, where sample size refers to the number of observations, $v = n * t$. Therefore, $\hat{\theta} \xrightarrow{p} \theta^*$ as the sample size rises without bound. Furthermore, the variance of the ML estimator is equal to the inverse of the variance of the scores evaluated at the true parameters, $\mathbf{V}^{-1} = \mathrm{Var}(\partial \ln P_n(\theta^*)/\partial\theta)$ divided by sample size (Self and Liang, 1987). Greene (2000) shows that this value is equal to the lowest possible variance of any estimator. Together, these properties mean that the ML estimator is consistent, efficient and asymptotically normal, $\hat{\theta} \overset{a}{\sim} N(\theta^*, \mathbf{V}^{-1}/N)$.[2]

However, since the integral in (3) does not have a closed form solution, $P_n(y_n|\theta^*)$ must be approximated via simulation. Hereon, we express $P_n(y_n|\theta^*)$ as $P_n(\theta^*)$ for simplicity. We then use maximum simulated likelihood (MSL) estimation which is identical to ML estimation except simulated probabilities are used in place of exact probabilities. The procedure for simulation is as follows: for given values of θ, a β_n is drawn from the density $f(\beta|\theta)$ and used to calculate the statistic $P_n(y_n|\beta_n)$. This is repeated for R number of draws of β_n and the average value of these statistics gives the approximate, or simulated, unconditional choice probability, which we call $\check{P}_n(\theta^*)$, given in equation (4) and specify to be an unbiased estimator of $P_n(\theta^*)$ i.e. $E_r\check{P}_n(\theta^*) = P_n(\theta^*)$.

$$\check{P}_n(\theta) = \frac{1}{R}\sum_{r=1}^{R} P_n(y_n|\beta_n^{r|\theta}) \tag{5}$$

where $\beta_n^{r|\theta}$ is the rth draw from $f(\beta|\theta)$. The MSL estimator then maximises the simulated log-likelihood function, $\text{SLL}(\theta) = \sum_n \ln \check{P}_n(y_n|\theta^*)$. As this is a non-linear transformation of the simulated probabilities, however, $\ln \check{P}_n(y_n|\theta^*)$ is not an unbiased estimator of $\ln P_n(y_n|\theta^*)$ regardless of whether $\check{P}_n(y_n|\theta^*)$ is unbiased for $P_n(y_n|\theta^*)$. This bias in the SLL translates into bias in the MSL estimator. To derive an estimator with the same probability limit as the ML estimator it is sufficient that the sample average of the SLL converges to the same limit as the sample average of the log-likelihood. Since the simulator estimate is composed of the traditional estimator plus simulation bias and simulation noise, to achieve consistent estimates we must have sufficiently large numbers of observations in the sample to reduce asymptotic variance of simulation noise and an R that rises with v to reduce simulation bias. If R rises faster than $\sqrt{v}$ then MSL is asymptotically equivalent to ML (Hajivassiliou and Ruud, 1994). As is increasingly common in mixed logit models we use Halton draws instead of pseudo-random draws in drawing from $f(\beta|\theta)$ (200 repetitions in the simulations and 500 repetitions using real data) and BFGS as

[2] see Train (2009, pp. 247–250) for proofs of consistency, efficiency and asymptotic normality

the numerical maximisation algorithm.[3] Halton draws help avoid the possible issues of coverage and covariance that can arise from taking pseudo-random draws from a distribution. Although there are several maximisation algorithms one could potentially use, a comparison is beyond the scope of this paper. BHRW provide a detailed discussion on the topic.

As discussed earlier, we use the mixed logit because of its flexibility. $\beta_n = b + \gamma_n$ where b is the population mean, represented by the point estimate of the mean within $\hat{\theta}$ and γ_n is an individual's deviation in taste from the mean, represented for the population as the estimate of standard deviation within $\hat{\theta}$. Utility is then composed of a mean component that is common to all members of the population, $b' x_{njt}$ and a stochastic portion for each individual, $\gamma_n' x_{njt} + \varepsilon_{njt}$. This stochastic portion is correlated over alternatives and choice situations because γ_n is a common term so that the model can allow for general models of substitution and is not constrained by IIA. Any RUM model can be approximated by a mixed logit through appropriate selection of product characteristics and distribution for the coefficients (McFadden and Train, 2000); we specify a normal distribution for all non-price characteristics and a log-normal distribution for price such that the coefficient is always negative.

Estimating $\hat{\theta}$ provides a foundation for further analysis. In merger simulations, the demand estimates can be used to calculate price elasticity of demand, which when combined with data on marginal costs and ownership structures can be used to predict the price and welfare effects of a merger. Let $\eta_{jk} = \frac{\partial q_j}{\partial p_k} \frac{p_k}{q_j}$ be the price elasticity of demand where p_j and q_j are the price and quantity of good j in the market. Instead of quantities, in the logit case, we use predicted market shares $s_j = \frac{q_j}{M}$ where M is the total size of the market. Market shares in turn are equivalent to the predicted probabilities,

[3] Unlike Newton-Raphson which uses the actual Hessian at β_z to determine the step to β_{z+1} where β_z is the current estimate of β in the ML/ MSL algorithm, or BHHH which uses the score at β_z to approximate the Hessian, BFGS uses information at several points to obtain a sense of curvature of the log-likelihood function (see Brunner et al. (2017) for a comparison of drawing from densities and optimisation methods)

$$P_{nj} = \int \frac{\exp \beta' x_{nj}}{\sum_J \exp \beta' x_{nj}} f(\beta) d\beta \tag{6}$$

This gives the predicted probability that individual i chooses good j which is a weighted average of the logit formula, evaluated at different values of β and the weights given by the density of $f(\beta)$ as described by the estimated parameters $\hat{\theta}$. Then we have

$$\eta_{jk} = \begin{cases} -\frac{p_j}{s_j} \int \alpha_n P_{nj}(1 - P_{nj}) f(\beta) d\beta & \text{if } j = k, \\ \frac{p_j}{s_k} \int \alpha_n P_{nj} P_{nk} f(\beta) d\beta & \text{otherwise} \end{cases} \tag{7}$$

Note that α is the coefficient on price and is simply one element of the vector of coefficients β. This results in a $J \times J$ (or a $J \times (J+1)$ with an outside good) matrix in which the main diagonals are the own price elasticities of goods $j \in J$ and the off-diagonals are the cross-price elasticities of goods $j, k \in J$. Therefore, by combining our demand estimates with real world observations on price and product characteristics we should be able to obtain measures of price elasticity for real products.

In a monopolistic market obtaining price elasticity is sufficient to infer marginal cost, c. A common measure of market power is the Lerner index, also referred to as the price-cost margin (PCM),

$$\text{Lerner Index} = \frac{p - c}{p} \tag{8}$$

It can be shown that for a profit maximising firm, at the optimum price p^*

$$\frac{p^* - c(q(p^*))}{p} = -\frac{1}{\eta(p^*)} \tag{9}$$

Following our mixed logit specification, in an oligopoly of F firms in which the fth firm produces a subset $\mathcal{F}_f \in J$ products, a firm's joint profit is given by

$$\Pi_f = \sum_{k \in \mathcal{F}_f} (p_k - c_k) s_k(\mathbf{p}; \theta) \tag{10}$$

where c_k is the constant marginal cost of the kth product and $\mathbf{p}$ is a vector of all relevant prices. Assuming Nash-Bertrand competition, the profit maximisation first order condition can be written as

$$\mathbf{p} = \mathbf{c} + \mathbf{\Omega^{-1}s} \qquad \text{where} \qquad \Omega_{jk} = -\phi_{jk} \frac{\partial s_k(\mathbf{p} : \theta)}{\partial p_j}, \tag{11}$$

$\mathbf{s}$ is a vector of market shares and Φ is a 1/0 $J \times J$ matrix where element ϕ_{jk} is 1 if j, k are produced by the same firm and 0 otherwise. Using the matrix of slope coefficients, $\mathbf{B}$ where element j, k is $\frac{\partial s_j(\mathbf{p}:\theta)}{\partial p_k}$ then $\mathbf{\Omega} = \mathbf{\Phi} \circ \mathbf{B}^\top$ where $\circ$ is element by element multiplication. As $\mathbf{B}$ has previously been obtained as the integrals of our elasticity calculations, obtaining the markup, $\mathbf{\Omega^{-1}s}$ is straightforward.

3. Experiment

3.1. **Experimental design.** We chose beer as our primary product because the industry is an oligopolistic differentiated product market that has been studied in the past. It is also an industry that has seen a significant amount of merger activity over the years. A key requirement of the mixed logit model is to obtain data in long form[4]. We find that it is easier to create the experiment with this consideration in mind rather than attempt to switch later. The design process we use is adapted from Hensher (1994). Firstly, we define our set of product characteristics. These must be relevant to the purchase decision as well as observable and measurable. Price is included because marginal utility of income is a key component of the price elasticity of demand function. Based on previous studies including Miller and Weinberg (2017) and Lerro et al. (2020), we chose ABV to represent alcohol content, volume per unit to represent packaging size and can/bottle to represent packaging material as

[4]Each row represents one alternative in a choice set, with either a zero or one to indicate whether that alternative was chosen

our remaining product characteristics. These are identified in the 'attributes' column of Table 1. The levels in column 2 were chosen to balance realism with econometric considerations. The range of values should be believable and large enough to ensure sufficient variability to estimate coefficients but not so large that there are a high number of dominated alternatives. The more levels for each attribute, the more choice tasks are required.

Table 1. Attributes and levels of survey products

Attributes	Levels of features
Beers	
Price/6-pack	\$6.49, \$7.99, \$10.99
ABV	3.6%, 4.6%, 5.5%
Can/Bottle	0 = can, 1 = bottle
Volume/unit	8.4-oz, 12-oz, 16-oz

As it is more practical to ask fewer respondents to make repeated choices rather than ask more respondents to each make a single choice, we use a panel data set. There is some disagreement in the literature regarding an appropriate number of choice sets in the context of subject fatigue. Bradley and Daly (1997) argue that fatigue caused by a large number of choice sets increases the error term variance. Hess et al. (2012) provide evidence that these concerns are overstated. Ultimately, we follow Chung et al.'s (2011) recommendation that difference specifications and functional forms should be pretested in order to identify optimal numbers of products and choice sets. This pretesting is done through a simulation exercise using 'fake' data. The methodology and output of this is described in appendix A. As a result of the simulation, we settle on 4 alternatives in a choice set and 8 choice sets per subject as our final experimental design.

Historically, capacity constraints in the lab meant that the number of observations one could obtain was limited. Therefore, alternatives in a choice set had to be selected in such a way that they extracted the maximum amount of information. For lab studies, orthogonal arrays in which the attribute levels are independent both within and between alternatives, became the preferred experimental design when choosing alternatives for a choice set. The benefit of online experiments is that they are easily scale-able. Random sampling theory guarantees that if we take large enough samples from the complete factorial, we should closely approximate the statistical properties of the factorial itself (Louviere et al., 2000). Since we require a large number of observations to achieve consistent and efficient parameter estimates anyway and our simulation exercise indicates that beyond a few thousand observations the marginal gains in accuracy decrease significantly, we are able to draw on random sampling from the full factorial set as the selection method for alternatives in a choice set, without the need for deriving several complex orthogonal arrays. In fact, for certain cases, Rose and Bliemer (2009) show that an orthogonal design is not the most efficient design and so-called 'efficient' designs are able to produce more efficient data in the sense that more reliable parameter estimates can be achieved with an equal or lower sample size. Random assignment of alternatives to choice sets across a large number of choice sets also achieves attribute level balance which ensures the parameters can be estimated well on the whole range of levels, instead of just having data points at only one or few of the attribute levels.

3.2. **Realism and External Validity.** Of primary concern for any SP type experiment are issues of realism and external validity. By construct, the surveys illicit hypothetical responses and so minimising hypothetical bias, or 'the potential error induced by not confronting the individual with an actual situation' (Schulze et al., 1981) is paramount. It is possible to achieve high levels of realism through complex choice tasks yet this must be balanced with the levels of stress and cognitive burden placed on participants which can reduce the quality of responses (Hensher and Cherchi, 2015).

3.2.1. *Labelled vs unlabelled alternatives.* Also referred to as alternative-specific vs generic, the way products are described in a choice set affects identification and precision, which must be considered alongside complexity and realism of the choice tasks. Labelled (alternative-specific) experiments have specific, real brand or product names attached to each alternative in a choice set and it is the label which is the object of choice. Unlabelled (generic) experiments, on the other hand, have no specific names, and are identified only as option ‘A’, ‘B’, ‘C’ etc. for example. Unlabelled alternatives represent the simplest design type of SP experiments. Choice outcomes are generic as the label attached to each alternative does not convey any information beyond what is described by the given product characteristics in the experiment. With labelled alternatives however, the brand name itself conveys information to the subjects beyond the attributes specified in the experiment and these characteristics are unobserved by the researcher. The key issue is that we, as researchers, have no way of controlling for these unobserved characteristics. Therefore, to avoid problems of endogeneity or omitted variable bias that may arise if unobserved characteristics are correlated with price or the random error term - which in turn can have significant consequences for the magnitude of parameter estimates especially on price where positive associations can lead to underestimating coefficients, while negative associations can lead to overestimating coefficients - it is prudent to use unlabelled alternatives in the simple design we propose (Louviere et al., 2000).

Beyond statistical properties however, the labelling question also has roots in behavioural and realism concerns. When choices are made in the real-world, consumers *do* consider brand names precisely because they confer information to the consumer, particularly in our case, where there is only a limited amount of information conveyed by our product characteristics. Therefore the inclusion of brands would serve to improve external validity. Labelled choices are also more tangible in the minds of subjects and may increase internal validity of the experiment. De Bekker-Grob et al. (2010) find that including labels in the choice of colorectal screening programs does change individual choices and reduces the attention respondents gave to specified attributes. They suggest unlabelled alternatives are more suitable when investigating attribute tastes and associated trade-offs and labelled alternatives may be more appropriate when the goal is to predict real-life choices.

For our first round of data collection, we settled on a hybrid design in which a different brand is specified for each choice set but alternatives within a choice set remain unlabelled such that there is variation in brand-specific unobserved characteristics between choice sets but these unobserved characteristics are constant within a choice set so that order of preferences, and hence the chosen alternative, is not affected by the brand. Including brands at the choice set level maintains some degree of realism by making alternatives appear less abstract without compromising on the statistical properties of the model. We subsequently conducted a second round of data collection on the same population with a set of 18 real products with brand dummies included in the model. This was so we could better understand how brands effect our demand parameter estimates. A discussion of our results is presented in section 4.6.

3.2.2. *Limited attention and Motivation.* One of the biggest challenges for any stated choice experiment is to convince external validity and realism exist when consumers are not making consequential choices (Bergman et al., 2020). If consumers are not spending their own money, they may simplify their decision process for example, always choosing option A. As mentioned earlier, lengthy surveys can result in boredom and cognitive fatigue which increases survey noise and correspondingly reduces the quality of responses. We include attention checks at random points within each round to ensure the participant is not just randomly clicking through choices. However, as of the current experiment we have not devised a satisfactory methodology of incentivising choices which would increase external validity. Experimenting with various incentivisation strategies is an area for further research. No statistical differences in estimates between various methodologies would suggest that incentives do not change respondent's choice. However, this is beyond the scope of the current paper.

We previously alluded to the challenges in incentivising our experiment in such a way that it recreated the experience of consumers in an actual supermarket. One possibility is to give subjects an endowment at the beginning of the experiment so that one of their choices could be randomly chosen to 'purchase' the actual goods. However, we know from the mental accounting literature (see Arkes et al. (1994)) that subjects treat this not as part of their regular endowment but as a windfall and what we observe is how they treat this windfall rather than how they behave with their own money. Additionally, depending on the products in question delivery may be prohibitively expensive or simply infeasible. Regardless, in contrast to supermarket data where the consumer has made a decision about what they might purchase, in an experiment the researcher forces the subject into a decision. We also posit that when faced with a choice as in our experiment consumers default to their past shopping experiences in the absence of any other information thus mimicking those choices closely.

4. Results

4.1. **Data.** We administered the experiment described in section 4 on the online subject recruitment platform Prolific. The subject pool was restricted to US residents aged between 21-30, which gave us the largest geographical market to operate in. Previous work in the US beer market also enabled us to make some comparisons to existing data. The age restriction included the minimum drinking age in the US and an age range most likely to be found on a student campus if we attempted to collect further data through physical participation at the UEA Laboratory for Economic and Decision Research (LEDR).

In total 486 subjects made a choice for each of the eight choice sets presented to them, resulting in 3,888 choice observations. Participants were paid a fixed fee for their time. As a result of an unexpected surge in sign ups to Prolific of young women aged 18-30 around the time of our experiment[5], many studies including ours suffered from a severe gender bias; 79% of subjects were female. We felt the data remained suitable for our methodological purposes but recognise any predictive claims could be weakened by the unrepresentative sample.

[5]The flood of new participants was subsequently attributed to a viral TikTok in which a teenager promoted Prolific as a 'side-hustle'; an easy way to make a few extra dollars. The video garnered 4.1 million views within a month (Letzter, 2021).

On top of their product choices, data on subjects demographics including age, gender, income and location by state was collected.

4.2. **Mixed Logit.** We first estimate a mixed logit model on the data. The model parameters, $\hat{\theta}$, refer to the mean and standard deviation of each of the elements of the vector β. Each product characteristic is specified to have a random component such that there is heterogeneity in preferences and we do not include any demographic variables. The random parameter on price, which is commonly referred to as α is an element of β and is specified as log-normal for two reasons. Firstly, prior studies have shown that this is typically the shape for the distribution of preferences on price. Secondly, it ensures all parameter estimates have the same sign so that the parameter estimate on price α is negative for all n. All other random parameters are specified to be normally distributed. This is of course, an a priori assumption but it is straightforward to estimate the parameters of any parametric distribution including a uniform or triangular distribution where appropriate. Estimating non-parametric distributions is possible; as McFadden and Train (2000) state, it is possible to estimate any RUM model to any degree of accuracy by a mixed logit with appropriate observed product characteristics and mixing distribution. However, as the number of parameters to estimate per characteristic increases, the estimation becomes computationally complex. Although the likes of Fosgerau and Mabit (2013) and Train (2016) have detailed methods to navigate these estimations, we have no reason to believe preferences on our chosen characteristics are distributed in such fashion. Column 1 of Table **??** shows the results of the unconditional mixed logit (random parameter) estimation for beers.

We can see that consumers prefer a higher ABV, and volume per unit but a negative coefficient on container indicates that subjects prefer cans to bottles. The standard errors on these non-price product characteristics are all small and the estimates are statistically significant. Similarly, the standard deviations are all statistically significant which suggests the presence of unobserved heterogeneity in preferences and that a random specification is appropriate.

For the parameter on price, the log-normal coefficients m and s are estimated such that the reported mean is equal to $\exp(m + (s^2/2))$ and the reported standard deviation is equal to $m * \sqrt{\exp(s^2) - 1}$. The sign is negative, indicating utility goes down as price goes up, but of course this is as result of the log-normal specification we used.

4.3. **Conditional Distributions.** Using the the point estimates of $\hat{\theta}$ we can calculate each subjects tastes conditional on the sequence of choices they made, presented in column 2 of Table **??**, which shows the mean and standard deviation of the 486 individual coefficients, β_n. The means of β_n are very close to the population mean in all cases. This similarity is expected for a correctly specified and consistently estimated model. The standard deviations are considerably greater than zero and are also similar to their population counterparts. For example, the conditional mean of ABV has a standard deviation of 1.116, and the population estimate of the standard deviation is 1.430. Thus, variation in $\bar{\beta}_n$ captures more than 78% pf the total estimated variation in the coefficient.

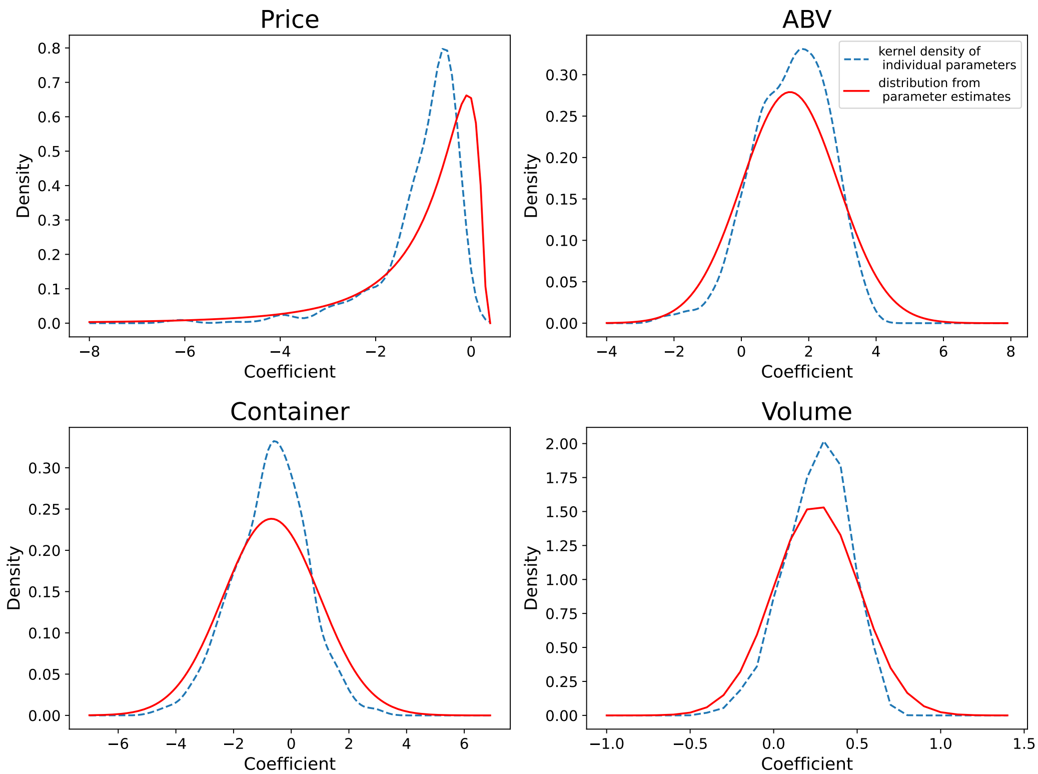
Figure 1 shows a similar pattern for all the other coefficients. The dashed line shows the kernel density of the individual parameters and the standard deviation of this distribution is marginally less than the standard deviation of the equivalent population distribution. This shows that the mean of a subjects conditional distribution captures a large share of the variation in coefficients across subjects and has the potential to be meaningful in distinguishing customers.

4.4. **Interaction Effects.** Where available individual-specific demographic data can be included in the model as a source of observed heterogeneity through an interaction with relevant product characteristics. Some of the unobserved heterogeneity in the population model can then be 'explained' by the observed demographic characteristics of sampled individuals. Although it

Table 2. Beer demand estimates

	Population	Individual	Interactions	
	β	$\bar{\beta}_n$	w/ income	w/ age
Price				
Mean	-1.027	-1.041	-1.025	-1.017
	(0.077)		(0.068)	(0.068)
Std. dev.	1.148	0.876	1.112	1.133
	(0.185)		(0.154)	(0.161)
ABV				
Mean	1.444	1.468	1.439	2.864
	(0.089)		(0.090)	(0.752)
Std. dev.	1.430	1.116	1.455	1.442
	(0.089)		(0.088)	(0.089)
Container				
Mean	-0.686	-0.720	-0.701	-0.276
	(0.099)		(0.101)	(0.869)
Std. dev.	1.675	1.266	1.710	1.699
	(0.111)		(0.114)	(0.113)
Unit volume				
Mean	0.256	0.260	0.261	0.260
	(0.016)		(0.017)	(0.017)
Std. dev.	0.257	0.190	0.249	0.250
	(0.019)		(0.018)	(0.018)
Demographic Interactions				
Price × income			0.001	
			(0.001)	
ABV × age				-0.059
				(0.031)
ABV × container				-0.017
				(0.035)
Observations	3888	3888	3840	3728
Draws	500	500	500	500

FIGURE 1. Comparison of population and individual parameters



may be tempting to add pairwise interactions between each demographic variable and each product characteristic, the larger the number of interactions, the greater the number of moment restrictions required. Hence a researcher must decide which demographic and product characteristics interact in reality.

We collected data on income, age, ethnicity and home state for each individual. As Hensher and Greene (2003) state, these demographic effects can be included in the model by interacting the variable with the random parameter and adding it in as a fixed parameter. In this specification, $U_{nj} = \beta_n' x_{nj} + \kappa(z_n x_{nj}) + \varepsilon_{nj}$ where z_n is a vector of demographic characteristics, and κ is a fixed parameter (we drop t for notational simplicity). A common and plausible interaction is between price and income. Of the 486 subjects, six declined to provide information on their income so they were dropped from the sample for this specification. The results are presented in column 3 of Table **??**.

The results show that there is a small interaction effect and the positive sign suggests that as income rises subjects are slightly less sensitive to price. However, this effect is not statistically significant which means that there is absence of heterogeneity around the mean on the basis of observed income. This is not to say that income has no effect on the distribution of preferences on price, simply that we have failed to discover its presence. It must be noted at this point that there is an issue with our data with regards to income. Participants were asked, 'What is your monthly income in dollars?'. Some subjects clearly stated their annual income but more importantly around 13% of subjects responded with 0. This is likely to be students not in any form of employment. Of course, these subjects still have a monthly budget and it is this including all loans, stipends and allowances that was required. As a result, we question the non significance of the income interaction. To further illustrate the point we include a second specification, in column 4, that only includes an interaction between gender and ABV. The results suggest that women prefer a stronger beer and the estimate this time is significant. Again, we do not place too much emphasis on the result itself because the gender bias in the sample means that female preferences drive the estimates. Nevertheless, it serves to illustrate the mechanism of the interaction.

4.5. **Branded versus non-branded products.** A second round of data collection on the same population (US beer drinkers aged 21-30 on Prolific), was conducted that include brand dummies on a set of the 18 most popular beers by market share in the US in 2019. The experiment was in every other way the same as the non-branded iteration. The addition of brand dummies captures a large proportion of unobserved (to the researcher) effects. The difficulty is that the number of parameters to estimate increases in proportion to the number of brands, and characteristics that are fixed across choice situations are difficult to identify. This second problem requires the use of a minimum distance procedure (Chamberlain (1982), Nevo (2000)) to estimate taste coefficients β. We first estimate a $J*1$ of brand dummy coefficients, $d = (d_1, ..., d_j)'$ using the previously describes mixed logit procedure. From the original indirect utility equation 1 it follows that

$$d = X\beta + \xi \tag{12}$$

where X is a $J * K$ matrix of product characteristics that are fixed and ξ' is a vector of $J * 1$ unobserved product characteristics. Assuming that $E(\xi|X) = 0$ then

$$\hat{\beta} = (X'{V_d}^{-1}X)^{-1}(X'{V_d}^{-1}\hat{d}), \qquad \hat{\xi} = \hat{d} - X\hat{\beta} \tag{13}$$

Demand parameter estimates are presented in Table 3. Column 1 shows the unbranded results from column 1 in Table 2. Column 2 shows the results from the branded experiment. The differences are stark. The magnitude of the mean value of price is much smaller at -0.220 compared to -1.027. In fact, the absolute value of the mean coefficients are all smaller except for container. The mean coefficients for container and unit volume also reverse signs between the branded and non-branded experiments. Finally the standard deviations are all smaller in the branded experiments.

4.6. **Substitution patterns and markups.** We then use the various demand estimates to calculate price elasticity matrices and price cost margins for a our set of pseudo-products and a set of real products (1) using the non-branded estimates and (2) using the brand dummy estimates and compare them to existing estimates from previous studies. The real product set contained 18 of the most popular beers in the US plus an outside good matching the size of the pseudo-product set. Ownership of the brands was split between five firms; AB InBev, Molson Coors, Constellation Brands, Heineken and Blue Ribbon specified in the ownership matrix Φ. It must be noted that studies on aggregate data use observations from the entire population while our sample was restricted to ages 21-30. Table 4 presents a sample of the estimated elasticity matrix for the real products using the unbranded experiment estimates. Tables B.1 and B.2, in the appendix presents the same for the set of pseudo-products and the real products using brand dummies. Each entry i, j, where i indexes the row and j indexes the column, gives the elasticity of brand

TABLE 3. Branded versus unbranded demand parameter estimates

Variable	Parameter	Unbranded	Branded
Price	Mean (μ_α)	-1.027	-0.220
		(0.077)	(0.030)
	Standard deviation (σ_α)	1.148	0.652
		(0.185)	(0.224)
ABV	Mean (μ_{β_1})	1.444	0.202
		(0.089)	(0.026)
	Standard deviation (σ_{β_1})	1.430	0.244
		(0.089)	
Container	Mean (μ_{β_2})	-0.686	0.589
		(0.099)	(0.037)
	Standard deviation (σ_{β_2})	1.675	0.341
		(0.111)	
Unit Volume	Mean (μ_{β_3})	0.256	-0.147
		(0.016)	(0.009)
	Standard deviation (σ_{β_3})	0.257	0.082
		(0.019)	

i with respect to a change in the price of j. As the full matrix is too large to include here, only columns of brands owned by the two largest manufacturers ABInBev (green) and Molson Coors (orange) are shown in the table as these products were most scrutinised following the joint-venture between Miller and Coors investigated by Miller and Weinberg[6]. We can see evidence of the flexibility of the mixed logit in the heterogeneity in cross-price elasticities that exists within a single column. We also compare our estimates with those achieved by Miller and Weinberg (2017) in a study that uses a random coefficient nested logit model to compare predictions from demand estimation to ex-post merger price effects. Own-price elasticities for a selection of products that appear in both studies as well as summary statistics are presented in Table 5.

[6]The brands in red, blue and pink are owned by Constellations Brands, Heineken and Pabst respectively

We do not present this comparison as a benchmarking exercise. As we have mentioned before, the product sets, sample populations, time periods and product characteristics are all too different between our study and that of Miller-Weinberg to make direct comparisons and hypothesis testing is not possible. We include these here to show our estimates are *broadly* in line with previous studies as an illustration that our methodology produces what appears to be realistic estimates of elasticities.

TABLE 4. Unbranded real product set elasticity matrix

	Brand	1	2	3	4	5	6	7	8	9	10	11	12	13
1	Bud Light	**-3.941**	0.144	0.057	0.136	0.069	0.089	0.019	0.053	0.034	0.451	0.062	0.062	0.144
2	Budweiser	0.050	**-4.882**	0.014	0.352	0.100	0.148	0.028	0.054	0.037	0.551	0.135	0.136	0.374
3	Michelob Ultra	0.111	0.080	**-4.047**	0.073	0.041	0.051	0.051	0.033	0.093	0.262	0.155	0.153	0.080
4	Natural Light	0.019	0.138	0.005	**-8.757**	0.057	0.193	0.008	0.041	0.038	3.701	0.135	0.030	0.138
5	Busch Light	0.063	0.260	0.019	0.381	**-4.884**	0.163	0.026	0.068	0.047	0.618	0.124	0.090	0.260
6	Busch	0.047	0.227	0.014	0.749	0.096	**-5.751**	0.019	0.069	0.054	1.144	0.151	0.066	0.227
7	Stella Artois	0.042	0.173	0.056	0.130	0.061	0.078	**-4.594**	0.039	0.106	0.205	0.268	0.300	0.173
8	Coors Light	0.070	0.203	0.022	0.395	0.099	0.170	0.024	**-4.771**	0.054	0.657	0.116	0.068	0.203
9	Miller Lite	0.044	0.137	0.062	0.366	0.068	0.133	0.064	0.054	**-5.276**	0.63	0.457	0.192	0.137
10	Keystone Light	0.032	0.114	0.01	1.941	0.049	0.154	0.007	0.036	0.035	**-6.635**	0.113	0.028	0.114
11	Miller High Life	0.032	0.198	0.041	0.504	0.070	0.145	0.064	0.045	0.179	0.801	**-5.663**	0.264	0.198
12	Blue Moon	0.038	0.235	0.047	0.134	0.060	0.074	0.084	0.031	0.089	0.232	0.311	**-4.591**	0.235
13	Coors Banquet	0.050	0.374	0.014	0.352	0.100	0.148	0.028	0.054	0.037	0.551	0.135	0.136	**-4.882**
14	Corono Extra	0.054	0.138	0.077	0.136	0.060	0.081	0.076	0.043	0.122	0.251	0.258	0.235	0.138
15	Modelo Especial	0.054	0.138	0.077	0.136	0.060	0.081	0.076	0.043	0.122	0.251	0.258	0.235	0.138
16	Heineken	0.044	0.193	0.058	0.116	0.058	0.070	0.083	0.034	0.091	0.206	0.265	0.347	0.193
17	Dos Equis	0.057	0.117	0.084	0.178	0.062	0.095	0.07	0.051	0.154	0.328	0.285	0.186	0.117
18	Pabst Blue Ribbon	0.040	0.315	0.011	0.754	0.098	0.198	0.021	0.058	0.045	1.104	0.169	0.093	0.315
19	Outside	0.037	0.181	0.016	1.510	0.066	0.161	0.020	0.044	0.050	2.737	0.160	0.080	0.181

Table 5. Comparison of beer elasticity estimates

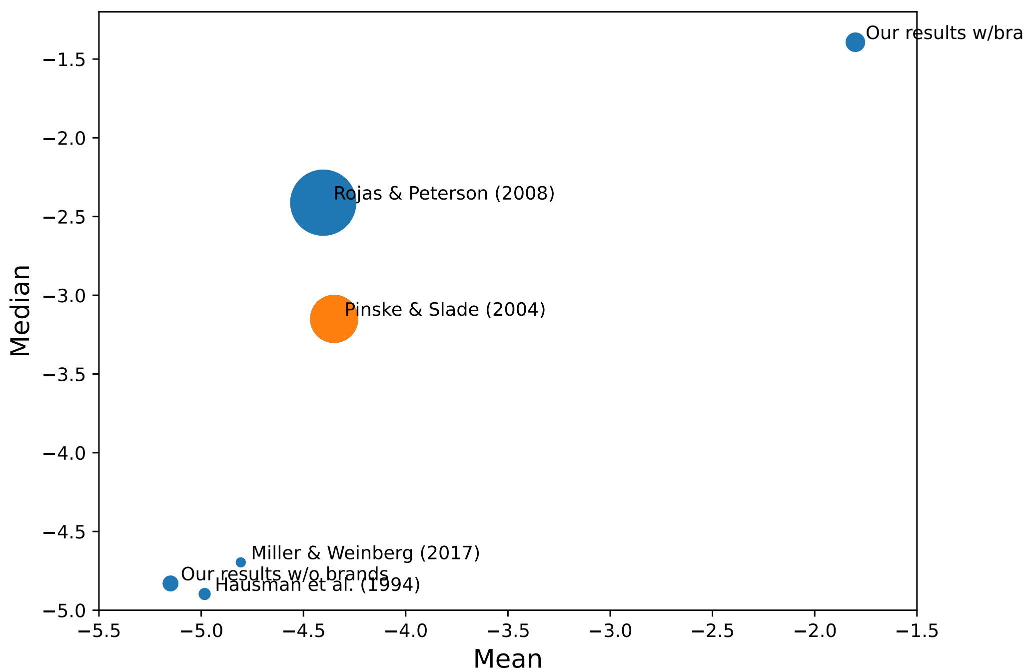
	(1) Pseudo	(2) Real unbranded	(3) Real branded	Miller-Weinberg
Own-price elasticities				
Bud Light		-3.941	-1.116	-4.389
Coors Light		-4.771	-1.487	-4.628
Miller Lite		-5.276	-4.081	-4.517
Budweiser		-4.882	-1.468	-4.272
Michelob Ultra		-4.047	-1.025	-4.970
Corona Extra		-4.529	-1.086	-5.178
Heineken		-4.577	-1.158	-5.147
Miller High Life		-5.662	-1.148	-3.495
Coors Banquet		-4.882	-1.084	-4.371
Summary Statistics				
Median Own-PED	-4.71	-4.83	-1.39	-4.73 – -4.33
Mean PCM		22.5%	81.5%	34%
Median PCM		23.0%	91.8%	

Abbreviations: PED is price-elasticity of demand; PCM is price cost margin

Indeed, when we look at the summary statistics for our unbranded real set, the median own-price elasticity falls close Miller-Weinberg's range which suggests that while there may be individual discrepancies between the results the overall industry outcomes are similar. The median own price elasticity in column 2 is -4.83. This is in line with Miller and Weinberg (2017) who achieve median own price elasticities of between -4.74 and -4.33 for three of their specifications. The pseudo-product set has a very similar median-own price elasticity. This suggests that if the aim to get a general understanding of a market rather than make predictions about specific products the kind of unlabeled experiment we conducted can be useful. Despite this, our model struggled to accurately predict market shares of beers because some beers with similar observed characteristics had markedly different actual market shares suggesting factors other than our observed characteristics were driving choices. Unobserved characteristics such as branding are most likely to be the cause. When we use the estimates from the branded experiment to calculate elasticities we can see the impact that these differences have. The median own-price elasticity is now -1.39 compared to -4.83. We attribute these changes to the non-incentivisation of our experiment. When the real brands are involved and there is no consequence to a subjects wealth, it appears they pick their favourite brand regardless of price, and for reasons not captured by our observed product characteristics. This is supported by the fact that in general subjects are less sensitive to changes in the observed characteristics in the branded experiment as seen by the smaller absolute values of the taste coefficients. Therefore, for our purposes we prefer the non-branded experiment as it focuses subjects on the observed product characteristics, especially price, which is crucial for downstream estimation of elasticities and the merger simulation. Figure 2 shows mean and median own-price elasticity estimates from studies that have previously estimated a differentiated demand system in the beer industry. Markers in blue are from the US whereas markers in orange are from the UK. The size of the marker represents the standard deviation of own-price elasticities. Our branded results appear somewhat of an outlier which supports our preference to use the non-branded estimates for the proceeding merger simulation.

Finally, we use the elasticity matrix to calculate marginal costs using equation 11 for the real product sets. (It is not possible to do this for the pseudo-set as there is no ownership matrix). We obtain median and mean price cost margins of around 22-23% in the unbranded version compared to 34% in Miller and Weinberg. This is equivalent to a markup of around \$3.00 compared to an average markup of \$3.60 in Miller and Weinberg. Significant changes in the structure of the market and in preferences towards craft beers could explain the differences. Results also vary depending on the model of competition that is used. In the branded experiment, this increases to 60% further reinforcing our idea that for our purposes the non-branded experiment is preferable.

FIGURE 2. Comparison of differentiated demand estimates with previous studies



Markers in blue are from the US; markers in orange are from the UK. The size of the marker represents the standard deviation of own-price elasticities.

4.7. **Merger Simulation.** With all the ingredients in place we are able to simulate the effects of a potential merger between firms in the industry using the unbranded elasticities. As an illustrative exercise, we choose to observe the effects of a merger between the two largest parent companies; ABInBev and Molson-Coors. The set of J demand equations $\mathbf{q} = \mathbf{a} + \mathbf{B}\mathbf{p}$ and J prices from equation 11, derived from the first order conditions specific to this demand system, jointly determine price and quantity (market share) in any market. Stacking and rearranging gives

$$\begin{bmatrix} \mathbf{p} \\ \mathbf{q} \end{bmatrix} = \begin{bmatrix} (\mathbf{\Phi} \circ \mathbf{B}') & \mathbf{I} \\ -\mathbf{B} & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} (\mathbf{\Phi} \circ \mathbf{B}') & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{c} \\ \mathbf{a} \end{bmatrix} \tag{14}$$

where I and 0 are $J * J$ identity and zero matrices respectively. To simulate a merger, we change the ownership matrix, $\mathbf{\Phi}$ to reflect the brands that would be under common ownership, and solve equation 14 to predict the new prices and quantities. Table B.3 shows our predictions where the merged entity has the same marginal costs as pre-merger. The new entity, referred to as AM in the table, now owns 13 of the top 18 brands in the market. All prices rise, with an average price increase of 3.41%. As a result mean PCMs rise from 22.5% to 25%. The total market share of the top brands falls from 63.6% to 53.5%.

Table 7 is the same as above, but now the merged entity benefits from an 25% saving in marginal costs. While this is extremely unrealistic, the analysis shows that if the merger was to elicit such cost savings then some of that saving would be passed on to the consumer in the form of lower prices. The average price reduction in this case would be 7.6%. At the same time, PCMs increase to 37%. The point here is to illustrate the flexibility of the model. A merger between any combination of incumbent firms through the appropriate ownership matrix and any marginal cost savings and be simulated easily.

Table 6. Simulated merger between ABInBev and Molson-Coors with constant marginal cost

	Pre-merger values				Post-merger values				
	Price	Mkt Shr	MC	PCM	New Firm	Price	Mkt Shr	PCM	% Δ Price
Bud Lgt	15.99	0.90	11.51	28.0	AM	16.57	0.83	30.6	3.64
Budweiser	11.99	3.44	9.23	23.0	AM	12.56	2.99	26.5	4.75
Michelob	18.99	0.39	13.87	27.0	AM	19.63	0.36	29.4	3.39
Natural Lgt	7.99	13.15	6.99	12.5	AM	8.51	8.62	17.8	6.45
Busch Lgt	11.99	1.32	9.12	24.0	AM	12.55	1.16	27.3	4.66
Busch	9.99	2.70	7.92	20.7	AM	10.53	2.27	24.8	5.37
Stella Art	15.99	0.42	12.11	24.2	AM	16.68	0.37	27.4	4.30
Coors Lgt	11.99	0.91	8.94	25.4	AM	12.49	0.83	28.4	4.14
Miller Lte	11.99	0.92	9.03	24.7	AM	12.43	0.85	27.4	3.70
Keyst. Lgt	7.99	25.06	6.69	16.3	AM	8.38	21.03	20.1	4.83
Miller HL	10.99	2.56	8.46	23.0	AM	11.42	2.32	25.9	3.90
Blue Moon	14.99	1.59	11.09	26.1	AM	15.47	1.48	28.3	3.17
Coors Bnqt	11.99	3.44	9.10	24.1	AM	12.49	3.09	27.2	4.19
Corona	15.99	0.31	12.41	22.4	-	16.13	0.32	23.0	0.84
Modelo Esp	15.99	0.31	12.41	22.4	-	16.13	0.32	23.0	0.84
Heineken	15.99	0.73	12.47	22.0	-	16.13	0.76	22.7	0.85
Dos Equis	14.99	0.30	11.65	22.3	-	15.13	0.31	23.0	0.91
Pabst BR	9.99	5.13	8.27	17.2	-	10.13	5.54	18.4	1.40

Firms are colour-coded as follows: ABInBev; Molson-Coors; Constellation Brands; Heineken; Pabst. AM is the new firm arising through the merger of ABInBev and Molson-Coors

PCM is price-cost margin = $(p-c)/p$ expressed as a percentage

TABLE 7. Simulated merger between ABInBev and Molson-Coors with 25% MC savings

	Pre-merger values			Post-merger values				
	P($)	Mkt Shr	PCM	New Firm	P($)	Mkt Shr	PCM	% Δ Price
Bud Lgt	15.99	0.90	28.04	AM	15.03	1.02	42.59	-5.99
Budweiser	11.99	3.44	23.01	AM	11.29	3.81	38.67	-5.85
Michelob	18.99	0.39	26.98	AM	17.76	0.45	41.44	-6.47
Natural Lgt	7.99	13.15	12.49	AM	7.58	14.75	30.82	-5.14
Busch Lgt	11.99	1.32	23.96	AM	11.30	1.46	39.49	-5.74
Busch	9.99	2.70	20.70	AM	9.45	2.94	37.12	-5.42
Stella Art	15.99	0.42	24.24	AM	15.01	0.47	39.48	-6.11
Coors Lgt	11.99	0.91	25.44	AM	11.27	1.02	40.50	-6.02
Miller Lte	11.99	0.92	24.72	AM	11.19	1.06	39.52	-6.65
Keyst. Lgt	7.99	25.06	16.25	AM	7.49	30.94	33.01	-6.24
Miller HL	10.99	2.56	23.00	AM	10.25	2.97	38.10	-6.70
Blue Moon	14.99	1.59	26.05	AM	13.93	1.86	40.32	-7.06
Coors Bnqt	11.99	3.44	24.14	AM	11.24	3.89	39.30	-6.28
Corona	15.99	0.31	22.37	Con	14.18	0.42	34.37	-11.29
Modelo Esp	15.99	0.31	22.37	Con	14.18	0.42	34.37	-11.29
Heineken	15.99	0.73	22.04	H	14.18	1.00	34.07	-11.32
Dos Equis	14.99	0.30	22.25	H	13.30	0.41	34.26	-11.3
Pabst BR	9.99	5.13	17.22	P	8.78	7.70	29.39	-12.07

5. Conclusions

So far we have presented a background and methodology that can be used to estimate demand parameters and utilise these estimates in further analysis relevant to merger evaluation. Among our primary goals was to simplify the process so that it could be easily adapted and replicated. As a result, there are several possible alternatives and extensions we have so far failed to discuss. The following section highlights some of the most salient of these, augmented by lessons learned during the administering of the experiment and the subsequent data analysis.

Although we were able to obtain estimates of taste parameters, elasticities and PCMs the observed characteristics we used for beer did not always accurately predict market shares, even when we used the set of real products. While the characteristics we used were guided by previous studies and a survey published by the Craft Brewing Business, it was apparent that many products in the real world were very similar in these characteristics. Despite these apparent similarities however, the products enjoyed different market shares. These differences must be as a result of unobserved factors such as taste and branding. Although it is difficult to measure taste, information about this may be conferred through the brand for well known brands. Since we used unlabelled alternatives in our choice sets, we have no information about specific brand fixed effects. This may be sufficient if the aim is to simply obtain demand estimates which could be confounded by brand effects but if the aim is to predict elasticities and supply side factors then it would appear brand fixed effects are required, particular in an industry where consumers consider branding to be important.

The addition of brands would require an alternative experimental design; exploring these is an area for further research. It is infeasible and inadvisable to provide subjects with a choice set of all brands in the market. If however, we consider the purpose of the demand estimation to be evaluating a merger and that a merger will only come to the attention of regulators when there are competition implications, only the largest brands in a market need be considered. Therefore it should be possible to present subjects with a choice of, for example, all the top 10 brands in a single choice set, while the other product characteristics are allocated randomly as before. This would allow brand fixed effects for each of these 10 products to be estimated and lead to more accurate predicted market shares.

We are currently in the process of designing and implementing a version of this experiment with brands on the same group of people to observe any differences in the estimation of taste parameters, elasticities and marginal costs compared to without brands. Depending on the outcome of that analysis, results may be added to this working paper or included in a separate study.

One problem with using brand dummies that we mentioned earlier was that it confounds identification of demand parameters. However, Nevo (2000) provides an elegant solution to this using a two stage projection method. First, the brand dummy coefficients and their variance-covariance matrix is estimated. Then a GLS regression is used to retrieve the taste parameters where the brand dummies are the independent variable and the number of observations is the number of brands used. Even in an empirical model, however, this restricts the number of observations. Where we suggest using only the top J brands the ability of this method to identify taste parameters must be examined further. The requirement of brands may vary by industry such that where observable product characteristics are more salient in consumer decisions, correct specification of these characteristics may result in a sufficiently identified model.

From an estimation perspective, there are several alternative methods we could explore. Rather than maximum likelihood, hierarchical Bayes estimation can be used and should achieve the same results if the model is correctly specified and identified. Even with an SLL estimation there are a number of different algorithms and methods for drawing from sample in simulation that can be tested. However, in our experience the marginal gains can often be small if there are sufficient observations.

In general we have presented a method that uses experimental data to estimate demand parameters and useful measures in merger evaluation quickly, with some degree of success. We managed to obtain estimates of elasticities and markups that appear to be realistic when compared to previous studies. However, there are several areas in the very simple experiment we conducted that could be improved to enhance the accuracy of estimates further. The precise experimental requirements are likely to be industry dependant, and indeed the model only suited to consumer goods, but once a satisfactory experiment has been designed it can be easily reworked to the specific products in question to provide guidance in initial merger evaluations.

References

Arkes, H. R., C. A. Joyner, M. V. Pezzo, J. G. Nash, K. Siegel-Jacobs, and E. Stone (1994). "The psychology of windfall gains". *Organizational behavior and human decision processes* 59(3), pp. 331–347.

Armona, L., G. Lewis, and G. Zervas (2021). "Learning Product Characteristics and Consumer Preferences from Search Data". *Proceedings of the 22nd ACM Conference on Economics and Computation*, pp. 98–99.

Bajari, P., J. T. Fox, and S. P. Ryan (2007). "Linear regression estimation of discrete choice models with nonparametric distributions of random coefficients". *American Economic Review* 97(2), pp. 459–463.

Bajari, P. L., Z. Cen, V. Chernozhukov, M. Manukonda, J. Wang, R. Huerta, J. Li, L. Leng, G. Monokroussos, S. Vijaykunar, et al. (2021). *Hedonic prices and quality adjusted price indices powered by AI*. Tech. rep. cemmap working paper.

Bansal, P., R. A. Daziano, and M. Achtnicht (2018). "Extending the logit-mixed logit model for a combination of random and fixed parameters". *Journal of choice modelling* 27, pp. 88–96.

Bastin, F., C. Cirillo, and P. L. Toint (2010). "Estimating nonparametric random utility models with an application to the value of time in heterogeneous populations". *Transportation science* 44(4), pp. 537–549.

Bergman, A., A. Chinco, S. M. Hartzmark, and A. B. Sussman (2020). *Survey Curious? Start-Up Guide and Best Practices For Running Surveys and Experiments Online*. Available at SSRN: https://ssrn.com/abstract=3701330. [Accessed 15-April-2021].

Berry, S., J. Levinsohn, and A. Pakes (1995). "Automobile prices in market equilibrium". *Econometrica: Journal of the Econometric Society*, pp. 841–890.

Bradley, M. A. and A. J. Daly (1997). "Estimation of logit choice models using mixed stated preference and revealed preference information". *Understanding travel behaviour in an era of change*. Ed. by P. Stopher and M. Lee-Gosselin. Oxford: Pergamon, pp. 209–232.

Brunner, D., F. Heiss, A. Romahn, and C. Weiser (2017). *Reliable estimation of random coefficient logit demand models*. Available at https://EconPapers.repec.org/RePEc:zbw:dicedp:267. DICE Discussion Papers.

Chamberlain, G. (1982). "Multivariate regression models for panel data". *Journal of econometrics* 18(1), pp. 5–46.

Chung, C., T. Boyer, and S. Han (2011). "How many choice sets and alternatives are optimal? Consistency in choice experiments". *Agribusiness* 27(1), pp. 114–125.

Conlon, C. and J. H. Mortimer (2021). "Empirical properties of diversion ratios". *The RAND Journal of Economics* 52(4), pp. 693–726.

Conlon, C. T. and J. H. Mortimer (2013). "Demand estimation under incomplete product availability". *American Economic Journal: Microeconomics* 5(4), pp. 1–30.

De Bekker-Grob, E. W., L. Hol, B. Donkers, L. Van Dam, J. D. F. Habbema, M. E. Van Leerdam, E. J. Kuipers, M.-L. Essink-Bot, and E. W. Steyerberg (2010). "Labeled versus unlabeled discrete choice experiments in health economics: an application to colorectal cancer screening". *Value in Health* 13(2), pp. 315–323.

Fosgerau, M. and M. Bierlaire (2007). "A practical test for the choice of mixing distribution in discrete choice models". *Transportation Research Part B: Methodological* 41(7), pp. 784–794.

Fosgerau, M. and S. L. Mabit (2013). "Easy and flexible mixture distributions". *Economics Letters* 120(2), pp. 206–210.

Greene, W. H. (2000). *Econometric Analysis 4th edition*. Upper Saddle River, NJ: Prentice Hall, pp. 201–215.

Hajivassiliou, V. A. and P. A. Ruud (1994). "Classical estimation methods for LDV models using simulation". *Handbook of Econometrics*. Ed. by R. Engle and D. McFadden. Vol. 4. Elsevier, pp. 2383–2441.

Hensher, D. A. (1994). "Stated preference analysis of travel choices: the state of practice". *Transportation* 21(2), pp. 107–133.

Hensher, D. A. and W. H. Greene (2003). "The mixed logit model: the state of practice". *Transportation* 30(2), pp. 133–176.

Hensher, D. A. and E. Cherchi (2015). "Stated preference surveys and experimental design: an audit of the journey so far and future research perspectives".

Hess, S., D. A. Hensher, and A. Daly (2012). "Not bored yet – revisiting respondent fatigue in stated choice experiments". *Transportation Research Part A: Policy and Practice* 46(3), pp. 626–644.

Lee, J. and K. Seo (2016). "Revisiting the nested fixed-point algorithm in BLP random coefficients demand estimation". *Economics Letters* 149, pp. 67–70.

Lerro, M., G. Marotta, and C. Nazzaro (2020). "Measuring consumers' preferences for craft beer attributes through Best-Worst Scaling". *Agricultural and Food Economics* 8(1), pp. 1–13.

Letzter, R. (Sept. 2021). *A teenager on TikTok disrupted thousands of scientific studies with a single video*. URL: https://www.theverge.com/2021/9/24/22688278/tiktok-science-study-survey-prolific.

Louviere, J. J., D. A. Hensher, and J. D. Swait (2000). *Stated Choice Methods: Analysis and Application*. Cambridge: Cambridge University Press.

Magnolfi, L., J. McClure, and A. T. Sorensen (2022). "Embeddings and Distance-based Demand for Differentiated Products". *Available at SSRN 4113399*.

McFadden, D. and K. Train (2000). "Mixed MNL models for discrete response". *Journal of applied Econometrics* 15(5), pp. 447–470.

Miller, N. H. and M. C. Weinberg (2017). "Understanding the price effects of the MillerCoors joint venture". *Econometrica* 85(6), pp. 1763–1791.

Moshary, S., B. Shapiro, and S. Drango (2022). "Preferences for Firearms and Their Implications for Regulation". *University of Chicago, Becker Friedman Institute for Economics Working Paper* (115).

Nevo, A. (2000). "A practitioner's guide to estimating random-coefficient logit models of demand". *Journal of Economics and Management Strategy* 9(4), pp. 513–548.

Nevo, A. (2001). "Measuring market power in the ready-to-eat cereal industry". *Econometrica* 69(2), pp. 307–342.

Petrin, A. (2002). "Quantifying the benefits of new products: The case of the minivan". *Journal of political Economy* 110(4), pp. 705–729.

Reynaert, M. and F. Verboven (2014). "Improving the performance of random coefficients demand models: the role of optimal instruments". *Journal of Econometrics* 179(1), pp. 83–98.

Rose, J. M. and M. C. Bliemer (2009). "Constructing efficient stated choice experimental designs". *Transport Reviews* 29(5), pp. 587–617.

Rossi, P. E. (2014). "Even the rich can make themselves poor: A critical examination of IV methods in marketing applications". *Marketing Science* 33(5), pp. 655–672.

Schulze, W. D., R. C. d'Arge, and D. S. Brookshire (1981). "Valuing environmental commodities: some recent experiments". *Land Economics* 57(2), pp. 151–172.

Self, S. G. and K.-Y. Liang (1987). "Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions". *Journal of the American Statistical Association* 82(398), pp. 605–610.

Train, K. (2016). "Mixed logit with a flexible mixing distribution". *Journal of choice modelling* 19, pp. 40–53.

Train, K. E. (2008). "EM algorithms for nonparametric estimation of mixing distributions". *Journal of Choice Modelling* 1(1), pp. 40–69.

— (2009). *Discrete Choice Methods with Simulation*. 2nd Edition. New York: Cambridge University Press.

Appendix A. Simulation

While this will not help in selecting appropriate product characteristics it will help choosing the number and combination of attributes and choice sets as well as an idea of the required number of observations. ‘Participants’ are generated and assigned values of tastes and preferences for the observed characteristics, drawn from distributions specified by the researcher. Additional noise, drawn from a standard Type 1 extreme value (Gumbel) distribution is assigned per participant per alternative per choice set. This ‘birthing’ of respondents can be repeated as many times as required for the sample size. Each participant is presented with repeated choice situations as in the real experiment. Each choice situation contains four alternatives and the program chooses the alternative that has the highest utility based on the preferences of each individual in the sample.[7] This process is repeated over n participants and t repetitions per person to obtain observations $= n * t$. Once the data is obtained, analysis is via the process outlined in section 3; we use SLL of mixed logit probabilities to estimate mean and standard deviation of the distribution in the population with the aim of estimating parameters from the previously specified distribution as consistently and efficiently as possible

In order to test for consistency and efficiency of parameter estimates, for a given seed, we presented increasing numbers of computer generated participants with a single choice and estimated the value of the mean and standard deviation of $f(\beta)$ in the population and the associated standard errors, presented in figures A.1 and A.2. In each graph, the red line represents the true parameter values (-5 and 1 respectively). We can clearly see that as the number of observations increases, the parameter estimates converge quickly to the true values, for both mean and standard deviation. We can also see that as the number of observations increase, the standard errors of the estimate, denoted by the gold bars, reduce significantly. Together, these results indicate that we can achieve consistency and efficiency using this model and data collected in a similar fashion.

[7] Choice behaviour need not be utility maximisation - the model simply describes the relation of explanatory variables to the outcome of a choice, without reference to how the choice is made.

The next step was to ensure that these results were not as a result of peculiar phenomenon occurring within the particular seed we had randomly chosen. In order to test this, we repeated the experiment 50 times each for specified combinations of participants, n and choice sets, t and then reported mean values for parameter estimates and standard errors. The results of this exercise are presented in table A.1. The pre-specified, 'true' values are given in parentheses next to the name of the attribute. The first 4 columns show the results for increasing numbers of participants each making a single choice. This is essentially the same as presented in figures A.1 and A.2 except the experiments have been repeated 50 times with different samples. The key thing to note is that as we move from column 1 to 4, the mean value of the mean price coefficient approaches -5, the mean value of the standard deviation of the price coefficient approaches 1, and the standard errors, denoted in parentheses for each parameter, drop significantly. Of course, as we mentioned earlier it is impractical to only ask one choice of each participant, so we conduct our 50 repetitions for different combinations of n and t, shown in columns 5-8. What we can see is that if we can obtain at least 1000 observations then the point estimates for mean and standard deviation are very close to the true values in the population. Increasing the observations to 5000 serves to improve the standard errors. We focus on the price coefficient as the observed characteristic of most interest, however, it can be seen that the mean point estimates for μ and σ all converge to their true values as the number of observations increases for all observed attributes. Similarly, the standard errors all decrease significantly as we move from left to right from column 1 to column 8. This suggests the mixed logit of the experimental data is able to accurately derive the true population parameters, θ^*.

Figure A.1. Estimates of population mean of price coefficient for increasing sample sizes

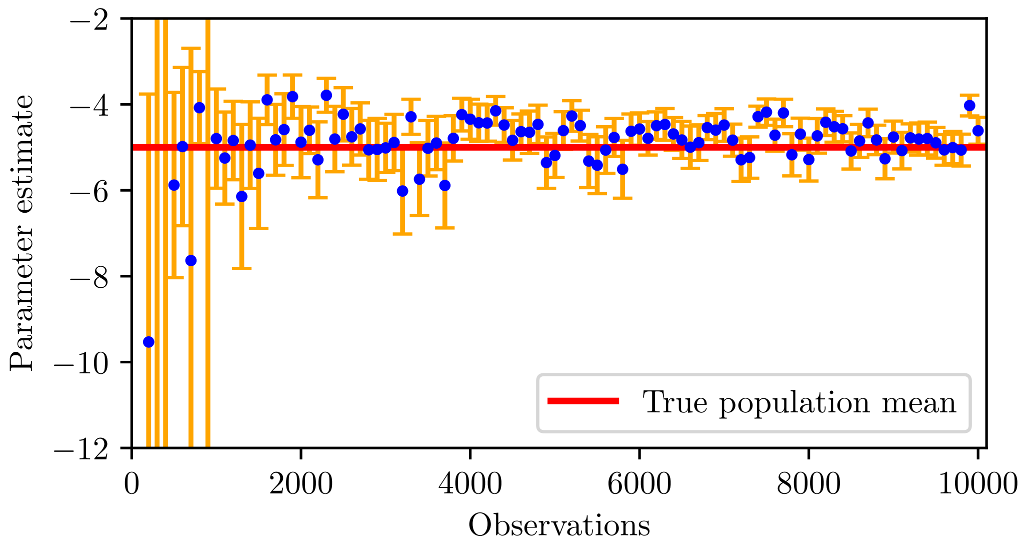
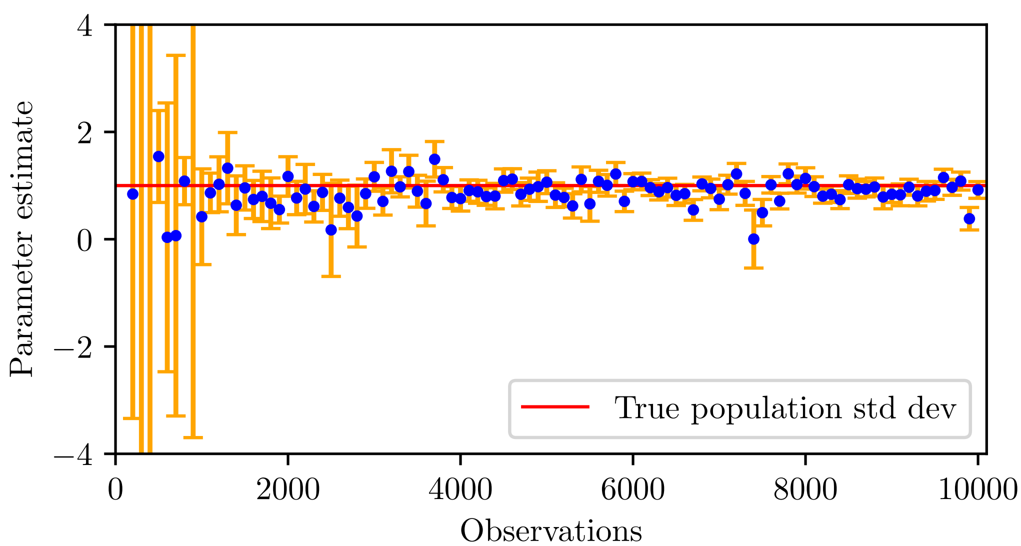


Figure A.2. Estimates of population standard deviation of price coefficient for increasing sample sizes



*Note: Dots represent point estimates of parameters and gold bars represent standard errors

Table A.1. Mean values of the estimated parameters on all coefficients over 50 samples

	1	2	3	4	5	6	7	8
n	200	500	1000	5000	200	1000	100	500
t	1	1	1	1	5	5	10	10
observations $= n * t$	200	500	1000	5000	1000	5000	1000	5000
Price (true values: $\mu = -5; \sigma = 1$)								
mean	-20.34	-8.866	-5.501	-4.729	-5.195	-4.928	-5.407	-5.003
	(124.6)	(17.99)	(1.634)	(0.485)	(0.832)	(0.312)	(0.746)	(0.262)
standard deviation	2.865	1.088	0.928	0.863	0.924	1.024	1.164	1.127
	(24.96)	(2.542)	(0.858)	(0.232)	(0.645)	(0.180)	(0.561)	(0.157)
ABV (true values: $\mu = 2; \sigma = 1.5$)								
mean	8.822	3.689	2.296	2.028	2.216	2.054	2.194	2.074
	(53.70)	(7.791)	(0.656)	(0.198)	(0.365)	(0.136)	(0.338)	(0.124)
standard deviation	6.712	2.835	1.711	1.434	1.615	1.518	1.598	1.514
	(42.39)	(6.653)	(0.750)	(0.234)	(0.379)	(0.145)	(0.345)	(0.124)
Can (true values: $\mu = 1.5; \sigma = 0.8$)								
mean	6.979	2.653	1.698	1.458	1.563	1.461	1.632	1.499
	(43.74)	(5.620)	(0.512)	(0.160)	(0.315)	(0.121)	(0.308)	(0.114)
standard deviation	6.760	2.504	1.232	0.766	0.996	0.853	1.064	0.881
	(2459)	(389.8)	(65.14)	(23.75)	(33.45)	(12.46)	(25.60)	(9.243)
Volume(true values: $\mu = 4; \sigma = 2.5$)								
mean	16.71	6.954	4.376	3.814	4.175	3.916	4.247	3.913
	(102.9)	(14.23)	(1.250)	(0.370)	(0.656)	(0.243)	(0.606)	(0.216)
standard deviation	10.99	4.692	2.827	2.444	2.684	2.549	2.653	2.565
	(64.41)	(11.07)	(0.992)	(0.302)	(0.537)	(0.199)	(0.489)	(0.173)

Appendix B.

Table B.1. Psuedo product set elasticity matrix

Name	1	2	3	4	6	7	9	10	11	13	15	16	17	18
1	**-3.868**	0.077	0.166	0.013	0.043	0.097	0.208	0.025	0.030	0.230	0.395	0.065	0.070	0.120
2	0.042	**-4.144**	0.284	0.010	0.078	0.062	0.310	0.017	0.030	0.131	0.535	0.038	0.061	0.167
3	0.025	0.076	**-4.185**	0.006	0.134	0.031	0.435	0.009	0.025	0.059	0.679	0.018	0.045	0.218
4	0.025	0.036	0.083	**-3.752**	0.097	0.048	0.113	0.056	0.069	0.125	0.230	0.155	0.169	0.296
5	0.019	0.039	0.143	0.021	0.174	0.031	0.168	0.037	0.067	0.071	0.307	0.089	0.144	0.403
6	0.011	0.037	0.236	0.013	**-4.175**	0.015	0.236	0.019	0.055	0.031	0.384	0.04	0.104	0.513
7	0.041	0.047	0.087	0.010	0.025	**-4.629**	0.183	0.031	0.033	0.437	0.617	0.135	0.130	0.198
8	0.030	0.051	0.150	0.008	0.044	0.071	0.284	0.021	0.033	0.261	0.887	0.082	0.118	0.290
9	0.017	0.047	0.243	0.005	0.075	0.036	**-5.038**	0.011	0.028	0.122	1.190	0.039	0.090	0.397
10	0.018	0.022	0.043	0.021	0.053	0.055	0.098	**-4.511**	0.073	0.237	0.347	0.309	0.299	0.464
11	0.014	0.025	0.076	0.016	0.096	0.036	0.152	0.045	**-4.808**	0.141	0.496	0.186	0.270	0.670
12	0.008	0.023	0.125	0.010	0.163	0.018	0.222	0.024	0.061	0.065	0.661	0.088	0.203	0.904
13	0.020	0.021	0.035	0.006	0.01	0.091	0.129	0.028	0.027	**-4.961**	0.857	0.228	0.202	0.287
14	0.014	0.022	0.059	0.004	0.018	0.059	0.201	0.019	0.027	0.422	1.265	0.140	0.187	0.428
15	0.008	0.020	0.093	0.002	0.030	0.030	0.291	0.010	0.022	0.199	**-4.537**	0.067	0.144	0.591
16	0.009	0.010	0.017	0.012	0.022	0.047	0.069	0.061	0.059	0.377	0.475	**-4.992**	0.456	0.651
17	0.007	0.011	0.030	0.009	0.039	0.031	0.108	0.040	0.059	0.229	0.701	0.312	**-5.361**	0.966
18	0.004	0.010	0.048	0.005	0.065	0.016	0.157	0.021	0.048	0.108	0.957	0.148	0.321	**-4.803**

TABLE B.2. Brand dummies elasticity matrix

	Brand	1	2	3	4	5	6	7	8	9	10	11	12	13
1	Bud Light	**-1.116**	0.017	0.015	0.005	0.007	0.005	0.016	0.006	0.032	0.159	0.029	0.048	0.044
2	Budweiser	0.009	**-1.468**	0.014	0.017	0.011	0.011	0.025	0.009	0.095	0.290	0.038	0.079	0.053
3	Michelob Ultra	0.007	0.012	**-1.025**	0.003	0.005	0.003	0.011	0.005	0.021	0.124	0.026	0.035	0.040
4	Natural Light	0.007	0.043	0.009	**-6.698**	0.017	0.039	0.039	0.014	3.008	0.641	0.032	0.122	0.040
5	Busch Light	0.009	0.028	0.014	0.017	**-1.483**	0.011	0.025	0.009	0.095	0.289	0.037	0.078	0.053
6	Busch	0.009	0.037	0.014	0.055	0.015	**-2.140**	0.034	0.012	0.294	0.447	0.040	0.106	0.055
7	Stella Artois	0.009	0.027	0.014	0.017	0.011	0.011	**-1.467**	0.009	0.096	0.289	0.037	0.078	0.053
8	Coors Light	0.008	0.025	0.015	0.016	0.010	0.010	0.023	**-1.487**	0.099	0.300	0.038	0.072	0.055
9	Miller Lite	0.007	0.039	0.010	0.501	0.015	0.034	0.036	0.015	**-4.081**	0.651	0.033	0.112	0.042
10	Keystone Light	0.008	0.030	0.014	0.026	0.012	0.013	0.027	0.011	0.161	**-1.370**	0.041	0.085	0.056
11	Miller High Life	0.007	0.018	0.014	0.006	0.007	0.005	0.016	0.006	0.037	0.189	**-1.148**	0.051	0.047
12	Blue Moon	0.009	0.028	0.014	0.017	0.011	0.011	0.025	0.009	0.095	0.290	0.038	**-1.416**	0.053
13	Coors Banquet	0.007	0.016	0.014	0.005	0.006	0.005	0.015	0.006	0.031	0.167	0.031	0.046	**-1.084**
14	Corona Extra	0.007	0.016	0.014	0.005	0.006	0.005	0.015	0.006	0.031	0.167	0.031	0.046	0.046
15	Modelo Especial	0.007	0.016	0.014	0.005	0.006	0.005	0.014	0.006	0.030	0.167	0.031	0.046	0.045
16	Heineken	0.007	0.018	0.014	0.006	0.007	0.006	0.016	0.007	0.038	0.188	0.032	0.050	0.048
17	Dos Equis	0.009	0.038	0.013	0.055	0.015	0.019	0.034	0.012	0.294	0.448	0.041	0.107	0.055
18	Pabst Blue Ribbon	0.007	0.016	0.013	0.005	0.006	0.005	0.014	0.006	0.030	0.168	0.031	0.046	0.045
19	Outside	0.008	0.027	0.013	0.080	0.011	0.013	0.024	0.010	0.465	0.335	0.036	0.077	0.050

BLANK PAGE

Table B.3. Simulated merger between ABInBev and Molson-Coors with constant marginal cost

	Pre-merger values				Post-merger values				
	Price	Mkt Shr	MC	PCM	New Firm	Price	Mkt Shr	PCM	% Δ Price
Bud Lgt	15.99	0.90	11.51	28.0	AM	15.03	1.02	42.6	-5.99
Budweiser	11.99	3.44	9.23	23.0	AM	11.29	3.81	38.7	-5.85
Michelob	18.99	0.39	13.87	27.0	AM	11.76	0.45	41.4	-6.47
Natural Lgt	7.99	13.15	6.99	12.5	AM	7.58	14.75	30.8	-5.14
Busch Lgt	11.99	1.32	9.12	24.0	AM	11.30	1.46	39.5	-5.74
Busch	9.99	2.70	7.92	20.7	AM	9.45	2.94	37.1	-5.42
Stella Art	15.99	0.42	12.11	24.2	AM	15.01	0.47	39.5	-6.11
Coors Lgt	11.99	0.91	8.94	25.4	AM	11.27	1.02	40.5	-6.02
Miller Lte	11.99	0.92	9.03	24.7	AM	11.19	1.06	39.5	-6.65
Keyst. Lgt	7.99	25.06	6.69	16.3	AM	7.49	30.94	33.0	-6.24
Miller HL	10.99	2.56	8.46	23.0	AM	10.25	2.97	38.1	-6.70
Blue Moon	14.99	1.59	11.09	26.1	AM	13.93	1.86	40.3	-7.06
Coors Bnqt	11.99	3.44	9.10	24.1	AM	11.24	3.89	39.3	-6.28
Corona	15.99	0.31	12.41	22.4	-	14.18	0.42	34.4	-11.29
Modelo Esp	15.99	0.31	12.41	22.4	-	14.18	0.42	34.4	-11.29
Heineken	15.99	0.73	12.47	22.0	-	14.18	1.00	34.1	-11.32
Dos Equis	14.99	0.30	11.65	22.3	-	13.30	0.41	34.3	-11.30
Pabst BR	9.99	5.13	8.27	17.2	-	8.78	7.70	29.39	-12.07

Firms are colour-coded as follows: ABInBev; Molson-Coors; Constellation Brands; Heineken; Pabst. AM is the new firm arising through the merger of ABInBev and Molson-Coors

PCM is price-cost margin = $(p - c)/p$ expressed as a percentage