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 15, 2024

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 the US beer 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 and examines the difference in consumer behaviour in the lab when brands are present versus when they are absent.

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. Yet the challenges associated with estimating models derived from the modern workhorse random coefficients mixed logit (RCML) of Berry et al. (1995, henceforth BLP) and Nevo (2000) are numerous.

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. We then combine these estimates derived from the lab with real world data on prices and product characteristics taken from a single market (in our case US national figures from 2019). This step is crucial in our ability to conduct counterfactual analysis such as merger simulations because it incorporates market clearing equilibrium conditions to the experimental demand parameters such that the merger simulations have real world interpretations. As far as we know, we are the first to apply this combination of experimental and real world data in this setting. Although our methods are not immune to the challenges of data collection, they do not require the highly detailed, multi-time-period, multi-market data sets typically required of empirical demand estimation.

We show that an experiment of this type, under the right circumstances, is quick and cheap to implement; data can be collected and analysed in weeks, if not days, rather than months. Our second contribution involves understanding the issue of incentivisation and the use of brands in lab experiments. Incentivisation in this context has a specific meaning; consumers do not face any consequences for their choices in the form of altered payoffs (see section 2 for more detail). The literature on SP experiments often discusses labelled versus non-labelled (or branded versus non-branded) products in the choice sets consumers see as an option within the experimental design (Louviere et al., 2000). It may be natural to think that since brands play a part in real purchase decisions, brand effects should be included in the experiment. Alternatively, it could be argued that omitting brands from the experiment would lead to unrealistic demand estimates and elasticities. We show that in lab experiments, where incentivisation is not possible, there is no incentive for subjects to *not* engage in cheap talk when brands are present, such that the non-branded experiment is better suited to elasticity calculations. This is an important finding and as far as we are aware none of the other papers in this direct domain make this distinction.

Beyond speed, the methodology addresses other challenges associated with empirical models which 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; for example Moshary et al. (2022) use an SP experiment to estimate demand for handguns because there is no centralized database that contains information about either individual-level or aggregate gun purchases matched with prices. Aggregate proxies for purchases that have been used in previous research are neither detailed to the gun model nor matched with prices and so are not suitable for demand estimation either. 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 present challenges for any researcher attempting to estimate demand but particularly for an anti-trust agency evaluating a merger in real time, they represent significant hurdles to a timely analysis. Imthorn et al. (2016), from the Netherlands Authority for Consumers and Markets (ACM), is the only paper that we have found to have put a similar method in practice during several merger cases including agricultural fertiliser, hospitals and bakery products, preferring them over hypothetical surveys (Imthorn et al., 2016). They do however, consistently apply brands in their methodology which as we show can be problematic. They also specified the conditions for merger simulation only after the experiments were conducted which caused problems including lack of variation in price. We show that by considering the purpose of the demand estimation, we can avoid most of these types of issues through careful design of the experiment. Additionally, given technological improvements since the original Dutch paper, we provide ideas for further adaptation in section 5.

From the repeated choices, we approximate mixed logit choice probabilities and estimate demand parameters to quickly and cheaply enable initial merger simulations. Prices in the experiment are always randomised, eliminating the need for instruments and variation in product characteristics and repeated random assignment of choice sets to consumers enables identification of the demand parameters of the model. To evaluate the effect of branding in the experiment, we conduct two treatments. In the first, choice sets all feature products from a single brand. We call this treatment *intra-brand*. In the second treatment, choices sets are constructed with products from different brands. We call this treatment *inter-brand*. More detail regarding the treatments can be found in section 2.3. In the intra-brand experiment, we define a matrix of attributes and levels that yields a set of pseudo-products used to estimate demand parameters. For the inter-brand treatment, all the product characteristics except price are taken from real products. In each treatment, each individual was presented with four alternatives from the set of 18 possible 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, a further 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. 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 a complementary method to other merger analysis tools such as upward pricing pressures, diversion ratios or qualitative measures. Further still, some competition authorities around the world including the CMA already use surveys and questionnaires in other forms, often qualitative, during merger assessment such that introducing this methodology will not be technically burdensome.

1.1. **Literature.** 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 mini-vans. There is also a related literature on discrete choice models (e.g. Train (2009)) from which we borrow heavily. 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) introduce more flexible distributions to the models to prevent the misspecification that can occur when inappropriate mixing distributions are used. We place our paper in a small but growing strand of literature that uses novel, often experimental, methods to either conduct demand estimation or more generally assess unilateral price effects arising from some change in the market.

[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.

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; it does not solve the problems of RCML's in contrast to our experiment which offers solution to these issues. 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 (although not demand parameters - hence our experiment is more flexible in its use) 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.

Imthorn et al. (2016) advocate the use of conjoint-analysis to overcome biases such as framing effect and those caused by interviewees strategic interests that occur during typical survey methods used by competition authorities.[2] However, Imthorn et al. (2016) themselves state the usage of such methods is limited and we have found no similar implementation by any other competition agency before or since. The authors speculate this may be 'due to a perception that these techniques are complex and time-consuming'. We show in this paper that neither of those limitations hold true. In 2010, the NMa used a choice-based conjoint-analysis (CBC), similar to our inter-brand treatment as part of wider empirical research including interviews and questionnaires to approve the merger between Agrifirm and Cehave, two producers of agricultural products. The resultant merger simulations were used as evidence that the merged entity would not be able to profitably raise prices significantly. Other attempts to use CBCs by the NMa and ACM were less successful in part because of market structure and sometimes unrealistic substitution patterns post

[2] E.g. 'what percentage price increase in product A would it take for you to switch to product B?'.

estimation. We deal with this issue specifically by comparing our inter- and intra-brand treatments.

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. As mentioned before, Moshary et al. (2022) illustrate an important use case for experiments where empirical data is simply not available. While the foundation of the experiment is similar, crucially their experiment always shows the gun brand as a product characteristic. They do not conduct an intra-brand equivalent in their study and thus face the same challenges we did when using the brands without an alternative procedure as in this paper. While not experimental as such Qiu et al. (2021) use win/loss data to identify diversion ratios for merger analysis, recognising the need for simple and efficient methodologies to use in real-time. Incidentally, one could generate this data using survey methods; while this elicitation has been criticised by some U.S. courts[3] we believe that certain adaptations can be made to improve their external validity. See section 5 for more detail.

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. (2023) 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

[3]See *U.S. v. H&R Block Inc., et al., D.D.C.* (2011).

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. We consider our methodology to be complimentary to other tools used in merger evaluation both qualitative and quantitative. 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'. Ultimately, the choice of which techniques to apply depends in part on the product(s) of interest.

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 in the US by market share in 2019 using parameters estimated from both treatments. 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 using intra-brand parameters 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. Our predicted markups in the range of 22.5-23% are lower than Miller and Weinberg's estimated 34%. In section 5, we discuss some of the challenges we faced and lessons we learned during the design and implementation of this methodology, including ways in which to improve the accuracy of estimates.

The rest of the paper is organised as follows. In section 2, we detail the experimental design guided by Hensher (1994), including testing of experimental features through Monte Carlo simulations, further detailed in Appendix A. Section 3 describes our model that encompasses elements from various strands of the existing literature. We define indirect utility, choice probabilities, price elasticities and price-cost margins. 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. In section 5 we discuss some of the issues we faced and provide thoughts on how the version of the experiment we conducted can be adapted for real-world use. Finally, we conclude in section 6.

2. Experiment

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 (alcohol by volume) to represent alcohol content, volume per unit to represent packaging size and can/bottle to represent packaging material as our remaining product characteristics. These are identified in the 'attributes' column of Table 1.

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

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

2.1. **Intra- versus inter-brand treatments.** The issue of branding is a key consideration for our experiment. Firms spend heavily on marketing and advertising to increase visibility and recognition of their products and differentiate their brands from competitors in order to reduce the brands own price elasticity of demand. When choices are made in the real-world, consumers consider brand names in their purchase decision because they confer information to the consumer as a result of advertising, 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. Branded 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 brand 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. However, as the brand name itself conveys information to the subjects beyond the attributes specified in the experiment, 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 may be prudent to use unbranded alternatives (Louviere et al., 2000). Here, products have no specific names, and are identified only as option 'A', 'B', 'C' etc. for example.

To balance these issues we settle on two treatments which we call intra-brand and inter-brand. Brands are present in both treatments but are utilised differently.

2.1.1. *Treatment 1: Intra-brands.* In this treatment products are hypothetical and each individual product is constructed as a combination of attributes from the values in 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 identify model parameters 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. In order to increase the salience of the choice in the minds of a subject, we included a randomly chosen brand logo to appear at the top of each choice set. The subject was then asked *"If* [brand name] *launched a new beer, which would you prefer?"*. As a result any brand effects are fixed across all four options. An example screen is shown in Figure 1.

2.1.2. *Treatment 2: Inter-brands.* In this treatment, we use real branded products with real product characteristics except for price which is randomly allocated to a product from the prices in Table 1. An example screen is shown in Figure 2. Now the first row contains the brand name as well as a picture of the product to simulate the choice a consumer might face on a supermarket shelf. Otherwise, the presentation remains the same as in treatment 1.

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

FIGURE 1. Intra-brand treatment example screen

Choice set 1

If STELLA ARTOIS launched a new beer, which would you prefer?

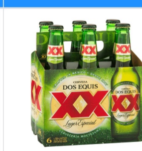
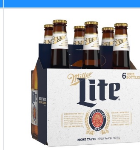
Product name	A	B	C	D
Price/6 pack	$6.49	$6.49	$10.99	$7.99
ABV	3.6%	5.5%	3.6%	4.6%
Container	Bottle	Bottle	Can	Can
Volume per container	12-oz	16-oz	8.4-oz	16-oz
Your choice:	○	○	○	○

Next

Laboratory for Economic and Decision Research, University of East Anglia

FIGURE 2. Inter-brand treatment example screen

Choice set 2

Product name	A	B	C	D
Brand	 Modelo Especial	 Bud Light	 Dos Equis	 Miller Lite
ABV	4.5%	4.2%	4.2%	4.2%
Container	Bottle	Can	Bottle	Bottle
Volume/unit	12-oz	16-oz	12-oz	12-oz
Price/6-pack	$6.49	$6.49	$10.99	$7.99
Your choice:	○	○	○	○

Next

(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 different 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 in both treatments.

2.2. **Identification.** From Table 1 we can see that if we ignore price, there are 18 different products that can be constructed from the levels and features shown. Since we choose 4 products to appear in each of our choice sets with replacement there are $C^R(18, 4) = 5,985$ different combinations of products that could appear. Given the 8 choice sets per subject and the average cost per subject we were unable to recruit sufficient participants to ensure that every combination ought to appear at least once, such that there were instances where the product characteristics were held constant while price was varied enabling identification of price.

As Holmes et al. (2017) state 'an experimental design must contain sufficient independent variation among attribute levels within and across alternatives so that each preference parameter can be identified. For example, if the levels of an attribute are always identical across alternatives, it will not be possible to identify the effect of that attribute on responses.' In panel A of Table 2, we attempt to illustrate this independent variation both within a single choice set and between all the choice sets a particular subject observes. The mean within choice set variation details how much on average each attribute varies between the 4 alternatives presented in a choice set. The units are the units of measurement of each particular attribute.[5] In the mean subject variation, we observe (the average of) the variation in each attribute among all alternatives in all the choices a subject sees. Since the parameter on price is

[5]Price is measured in dollars; ABV in percentage; container is a 0/1 dummy and volume is measured in fluid ounces.

our primary coefficient of interest, in panel B we illustrate that in around 4% of all choice sets in the intra-brand and 2% in the inter-brand, all the non-price product characteristics were held constant while price varied. These figures are admittedly small and are consequence of the size of our sample population. Scaling the experiment ought to increase these numbers.[6]

Historically, capacity constraints in the lab meant that the number of observations one could obtain was limited. Therefore, alternatives in each choice set had to be selected in such a way that they extracted the maximum amount of information so that the model could be correctly identified. This is particularly true of the intra-brand treatment as it consists entirely of hypothetical products. 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. Identification is then achieved because we have (1) variation in our product characteristics by construction from Table 1 across

[6] With a much larger budget, Moshary et al. (2022) were able to survey over 22,000 respondents.

and within subjects, alongside attribute level balance and variation in choice sets between and within subjects. In panel A of Table 2 we can see that there is less variation in the ABV and volume when using real products in the inter-brand treatment. When we randomise these product characteristics in the intra-brand treatment we get greater variation. Price is randomised in both treatments which is why the mean within choice set variation is similar and container only has two options so its variation is similar as well. In part this explains why we find smaller (in magnitude) parameter estimates in the inter-brand treatment; there is simply far less variation in these specific product characteristics in the set of real products that parsing out preferences is difficult. This lack of variation in product characteristics is not uncommon in empirical data. Panel B shows the number of times we held the non-price product characteristics constant and varied price for individuals and the mean number of times this occurred for each individual.

2.3. **Realism and External Validity.** Of primary concern for any SP type experiment are issues of realism and external validity. By construct, the surveys elicit 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).

2.3.1. *Incentivisation.* 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

Table 2. Summary Statistics

Panel A: Independent variation of product characteristics

	Price	ABV	C'tainer	Volume
Intra-brand				
Choice sets where x is held constant	3.11%	1.20%	11.3%	3.11%
Mean within choice set variation†	1.539	0.644	0.409	2.587
Mean within subject variation‡	1.840	0.767	0.494	3.066
Between subject variation	0.111	0.048	0.009	0.179
Inter-brand				
Choice sets where x is held constant	314	36	800	2661
Mean within choice set variation†	1.518	0.290	0.422	0.950
Mean within subject variation‡	1.840	0.338	0.492	1.369
Between subject variation	0.114	0.024	0.010	0.202

†We take the standard deviation of a characteristic from the four alternatives in a choice set for each choice set and report the mean of these values.

‡We take the standard deviation of a characteristic among all the alternatives an individual subject see across all his choice sets and report the mean of these values.

Panel B: Holding non-price X constant while varying price

	Intra-brand	Inter-brand
Choice sets where price is varied, X is constant	4.60%	1.77%
Mean per person choice sets where X is constant§	0.363	0.142

§ The average number of choice sets each subject sees that has fixed non-price product characteristics and varying prices.

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, but beyond the scope of this paper.

The challenge in our experiment is to provide incentives in such a way that it recreates 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. On the goods side, depending on the products in question, an actual provision or delivery may be prohibitively expensive or simply infeasible. Finally, the close recreation of incentives involves an outside option of no purchase in the experiment. But then the actual outside option – outside of the experiment – becomes relevant and is difficult to control or observe.

Ultimately, we posit that when faced with a choice in our experiment consumers default to their past shopping experiences in the absence of any other information and thus mimic those choices closely.

2.4. **Data.** We administered the treatments described above in June/July 2021 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.

In total 1,000 subjects, split equally between treatments, made a choice for each of the eight choice sets presented to them in each treatment resulting 4,000 observations per treatment.[7] Participants were paid a fixed fee for their time. As a result of an unexpected surge in sign ups to Prolific of young women

[7] A small number of participants in each treatment were excluded from the final analysis because they failed one of the random attention checks during the experiment or or for missing or inappropriately answering the demographic questions. For example several participants stated their age was outside of the specified range.

aged 18-30 around the time of our experiment[8], 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. On top of their product choices, data on subjects demographics including age, gender, income and location by state was collected.

3. 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 $t \in T$ choice situations in the experiment. 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 shown in the experiment 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 beer 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.

Since each individual's β_n is unobserved, the exact unconditional probability of n's sequence of choices made during the experiment is the integral of the conditional probability over all possible values of β as defined by the true

[8]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).

parameters of the distribution of β_n, θ^*,

$$P_n(y_n|\theta^*) = \int P_n(y_n|\beta_n) f(\beta|\theta^*) d\beta. \tag{2}$$

However, since the integral in (2) does not have a closed form solution, $P_n(y_n|\theta^*)$ must be approximated via simulation by taking R draws of β_n for a given θ, calculating the statistic $P_n(y_n|\beta_n)$ for each draw and averaging.[9]

As discussed earlier, we use the mixed logit because of its flexibility. We have $\beta_n = b + \gamma_n$ 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 such that

[9] We do not cover these derivations as they are covered in detail elsewhere. For an excellent presentation see Train (2009), chapters 3 and 8.

$$\eta_{jk} = \begin{cases} -\frac{p_j}{s_j} \int \alpha_n P_{nj}(1 - P_{nj}) f(\beta) d\beta & \text{if } j = k, \\ \frac{p_j}{s_k} \int \alpha_n P_{nj} P_{nk} f(\beta) d\beta & \text{otherwise.} \end{cases} \tag{3}$$

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 because at the profit-maximising price, the price-cost margin is equal to the negative reciprocal of the price elasticity of demand.

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

$$\Pi_f = \sum_{k \in \mathcal{F}_f} (p_k - c_k) s_k(\mathbf{p}; \theta), \tag{4}$$

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

$$\mathbf{p} = \mathbf{c} + \mathbf{\Omega^{-1}s}, \qquad \text{where} \qquad \Omega_{jk} = -\phi_{jk} \frac{\partial s_k(\mathbf{p} : \theta)}{\partial p_j}, \tag{5}$$

$\mathbf{s}$ 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.

4. Results

We first estimate a mixed logit model on the data from each treatment using PyBLP (Conlon and Gortmaker, 2020). The results are presented in Table 3. 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.

4.1. **Treatment 1: Intra-brand.** Column 1 of Table 3 shows the results of the intra-brand treatment. 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 a result of the log-normal specification we used.

4.1.1. *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 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 2 of Table 3.

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

Table 3. Mixed logit estimates

	Intra-brand			Inter-brand	
	(1)	(2)	(3)	(4)	(5)
Price					
Mean (μ_α)	-1.027*	-1.025*	-1.017*	-0.275*	-0.220*
	(0.077)	(0.068)	(0.068)	(0.047)	(0.030)
SD (σ_α)	1.148*	1.112*	1.133*	0.836*	0.652*
	(0.185)	(0.154)	(0.161)	(0.354)	(0.224)
ABV					
Mean (μ_{β_1})	1.444*	1.439*	2.864*	1.605*	0.202*
	(0.089)	(0.090)	(0.752)	(0.104)	(0.027)
SD (σ_{β_1})	1.430*	1.455*	1.442*	1.741*	
	(0.089)	(0.088)	(0.089)	(0.109)	
Container					
Mean (μ_{β_2})	-0.686*	-0.701*	-0.276*	1.215*	0.589*
	(0.099)	(0.101)	(0.869)	(0.081)	(0.037)
SD (σ_{β_2})	1.675*	1.710*	1.699*	1.411*	
	(0.111)	(0.114)	(0.113)	(0.081)	
Unit volume					
Mean (μ_{β_3})	0.256*	0.261*	0.260*	-0.001	-0.147*
	(0.016)	(0.017)	(0.017)	(0.026)	(0.009)
SD (σ_{β_3})	0.257*	0.249*	0.250*	0.369*	
	(0.019)	(0.018)	(0.018)	(0.032)	
Price × income		0.001			
		(0.001)			
ABV × age			-0.059		
			(0.031)		
ABV × gender			-0.017		
			(0.035)		
Observations	3888	3840	3728	3944	3944

Column (1) shows results from base intra-brand specification; columns (2) and (3) add demographic interaction terms for price and ABV respectively. Column (4) shows estimates from the inter-brand experiments without brand dummies included the model, whereas column (5) includes brand dummies in the specification

Standard errors in parentheses: $^*p < 0.05$

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 3, that includes an interaction between age and ABV, and gender and ABV. The results suggest that younger people and women prefer a stronger beer, although neither estimate is statistically 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.

4.2. **Treatment 2: Inter-brand.** The results of the same mixed logit specification as column 1 but using the data from the inter-brand treatments are shown in column 4; although brands were included in the experiment, no brand fixed effects are included in the model. In column 5 we add brand dummies to account for brand fixed effects. 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 $J*1$ vector of brand dummy coefficients, $d = (d_1, ..., d_j)'$ using the previously described mixed logit procedure. From the original indirect utility equation 1 it follows that

$$d = X\beta + \xi, \tag{6}$$

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

$$\hat{\beta} = (X'{V_d}^{-1}X)^{-1}(X'{V_d}^{-1}\hat{d}), \qquad \hat{\xi} = \hat{d} - X\hat{\beta} \tag{7}$$

The difference between columns 1 and 4/5 are stark. The magnitude of the mean value of price is much smaller at -0.275 and -0.220 compared to -1.027. Without brand dummies the magnitude of the coefficients on ABV and container are much larger than without. The mean coefficients for container and unit volume also reverse signs between the inter and intra-branded experiments. Finally the standard deviations are all smaller in the inter-brand experiment save for unit volume. As per Nevo (2000) we consider the specification with brand dummies going forward for the inter-brand treatment.

4.3. **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 (a) using the intra-brand estimates from and (b) using the inter-brand with 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 intra-branded experiment estimates. Tables B.2 and B.3, in the appendix presents the same for the set of pseudo-products and the real products using parameters estimates from the inter-brand specification with brand dummies. Each entry i, j, where i indexes the row and j indexes the column, gives the elasticity of brand 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.[10] We can see evidence

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

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.

TABLE 4. Intra-brand parameters applied to 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
	Median X-PeD	0.046	0.177	0.032	0.352	0.064	0.139	0.028	0.045	0.054	0.551	0.158	0.136	0.177
	Mean X-PeD	0.049	0.187	0.038	0.464	0.071	0.124	0.043	0.047	0.077	0.816	0.198	0.150	0.187

Table 5. Comparison of beer elasticity estimates

	(1)	(2)	(3)	(4)
		Intra-	Inter-	
	Pseudo	brand	brand	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

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.

Indeed, when we look at the summary statistics for our intra-brand real set, the median own-price elasticity falls close to 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 experiment we conducted in treatment 1 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 taste are most likely to be the cause. When we use the estimates from the inter-brand experiment with brand dummies 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. At the inter-brand level, when 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 inter-brand experiment as seen by the smaller absolute values of the taste coefficients. Therefore, for our purposes we prefer the intra-brand experiment as it focuses subjects on the observed product characteristics, especially price, which is crucial for downstream estimation of elasticities and the merger simulation.

While there exist many papers that estimate elasticities on the market for beer as a whole[11], there are fewer paper that estimate elasticities for a set of differentiated products. Figure 3 shows mean and median own-price elasticity estimates from studies that have previously estimated a differentiated demand system in the beer industry in the US or UK. 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 inter-brand results with brand dummies results appear somewhat of an outlier which supports our preference to use the intra-brand estimates for the proceeding merger simulation.

[11] see Fogarty (2010) and Nelson (2014) for two meta-studies that compare market elasticites across countries and time-periods.

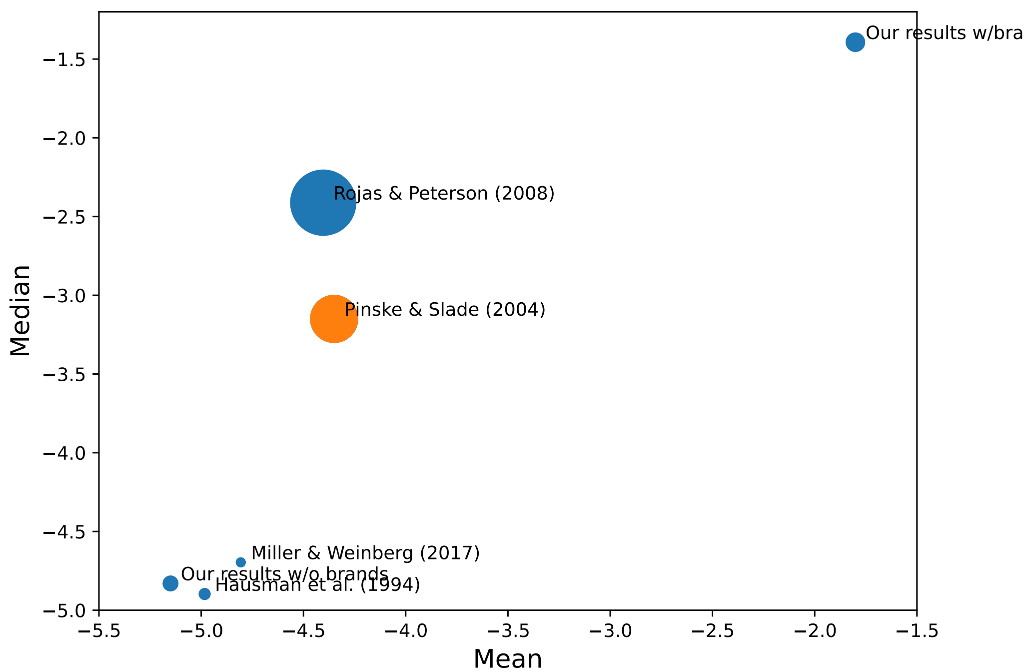
Finally, we use the elasticity matrix to calculate marginal costs using equation 5 for the real product sets. (It is not possible to do this for the pseudo-set as there is no ownership matrix). We obtain median and mean price cost margins of around 22-23% in the intra-brand version compared to 34% in 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 inter-brand experiment with brand dummies, this increases to 60% further reinforcing our idea that for our purposes the non-branded experiment is preferable.

4.4. **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 5, 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

$$\begin{bmatrix} \mathbf{p} \\ \mathbf{q} \end{bmatrix} = \begin{bmatrix} (\mathbf{\Phi} \circ \mathbf{B}') & \mathbf{I} \\ -\mathbf{B} & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} (\mathbf{\Phi} \circ \mathbf{B}') & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{c} \\ \mathbf{a} \end{bmatrix}, \tag{8}$$

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 8 to predict the new prices and quantities. Table 6 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%.

Figure 3. 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.

Table B.4 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 can 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

5. Discussion

As we have discovered, brand effects and the issue of incentivising the experiment go hand in hand. In our inter-brand treatment we saw that subjects appear to undervalue the importance of price when brands are present, a result that is also mentioned in Moshary et al. (2022). The reason for this is because the experiment was entirely hypothetical and subjects were not required to make any purchases based on the choices they made. This, of course is in contrast to what they would experience purchasing beer in the supermarket where there is a trade-off between preferences for a particular brand and the associated price. For example, an individual's all else equal 'favourite' brand may also be the most expensive, while their close second favourite brand is significantly cheaper such that they usually purchase the second favourite. This type of behaviour is not captured by a non-incentivised branded experiment because there is no consequential trade-off between price and brand; the hypothetical individual in the previous example would choose his favourite brand because the price is not relevant to him - the brand effect supersedes all other product characteristics. Branding provides signals to consumers about product quality and enables firms to change higher prices. In the real-world we are able to calculate the willingness to pay for these brands through revealed preference data but this is not transferable to the current non-incentivised experiment. Typical structural empirical work includes prices on the right hand side, but does not usually have advertising information, so that advertising information then sits in the error term. These studies almost always use instrument variables which makes price orthogonal to the error term, such that they able to capture the price effect very effectively. In contrast, to understand the coefficient on the price more accurately, we must focus on the interbrand treatment because in the intrabrand treatment people are focused on the wrong aspect of the purchase decision. A question then for further investigation is whether there is a design that allows for brands to be included alongside some form of incentivisation? Regardless, the success of these models depends on correctly identifying relevant observable product characteristics and the density of preferences for these characteristics in the population (Train, 2009). We were limited to measurable characteristics but in future the use of machine learning techniques and deep neural nets could allow the inclusion of qualitative characteristics to synthesise substitution patterns.

The use of similar techniques by agencies around the world is varied. Survey methods, questionnaires and experiments have been used more often in the UK by the Competition Markets Authority, (CMA, 2018) than in the US.[12] As Imthorn et al. (2016) argue, the specific use of discrete choice experiments avoids many of the biases inherent in more general survey methods. In their guidelines, the CMA discuss best practices for recruitment of subjects and questionnaire-type survey design including identifying choice attributes by either asking consumers to identify the most important reason(s) for their purchase or how important each attribute is to the customer, using a categorical scale.[13] However, the guidelines suggest discrete choice experiments of the type we use are not extensively used because of time constraints. We show that once an effective framework is designed, an experiment can be quickly deployed for goods or markets that exhibit similar characteristics.

The guidelines also express concerns with finding representative samples using online surveys. There are however, several ways in which our current methodology can be easily adapted. For certain demographics/products current online platforms including Prolific and Amazon MTurk will allow for representative samples - particularly as the numbers on these platforms are growing. But competition authorities that already use questionnaires in their analysis can utilise their existing recruitment methods. As the experiment only needs a mobile device and internet connection to administer and takes between 10-15 minutes to complete, customers can even be intercepted outside of stores if an agency so wishes. Else, participants can be drawn from third-party marketing firms such as YouGov in the UK or Harris Poll in the US that have large subject lists with considerable geographic and demographic coverage. The

[12]In *U.S. v. H&R Block Inc., et al., D.D.C.* (2011), Judge Howell criticises several elements of a survey used by defendant in the merger case including the leading nature of the questions and potential biases in recruitment.

[13]For example 'essential', 'very important', 'fairly important' and so on.

proliferation of video calling and online meetings even allows for the adaptation of interviewer led telephone surveys to our methodology which enables researchers to sample potentially more representative populations than are available through online recruitment platforms. However, this is beyond the scope of this paper. Agencies can even recruit participants to a physical lab(s) across the country - the only requirement is that there are sufficient devices on which subjects can make choices, although this is considerably more expensive and time-consuming. The limited success of the ACM and NMa shows that experiments can be useful, certainly in various types of merger cases, but also in answering other policy questions. Beyond understanding incentives and the role of brands in experimental preferences many of the underlying capabilities required for implementation already exist within mature competition agencies, particularly subject recruitment infrastructure and coding abilities. Nevertheless, there has not been a broader uptake elsewhere in the European Commision, or further afield suggesting that perhaps the biggest challenge for experiments is to convince institutions to begin using them. We hope that this paper, alongside others advocating similar methods, can begin to do just that.

6. 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.

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 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 initial 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.

The addition of brands saw our calculated elasticities move significantly away from both our non-branded results and those from Miller-Weinberg. It appears some factor in the presentation of the choice sets leads to consumers only considering the brand such that the other product characteristics, including price are less salient in the choice. An area for further research is to explore alternative experimental designs to solve this issue. 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, then a possible solution might be that only the largest brands in a market need be considered. Therefore it should be possible to present subjects with a choice of, for example, 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.

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., Z. Cen, V. Chernozhukov, M. Manukonda, S. Vijaykunar, J. Wang, R. Huerta, J. Li, L. Leng, G. Monokroussos, et al. (2023). "Hedonic prices and quality adjusted price indices powered by AI". *arXiv preprint arXiv:2305.00044*.

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.

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.

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.

CMA (May 2018). *Good practice in the design and presentation of customer survey evidence in merger cases*. URL: https://www.gov.uk/government/publications/mergers-consumer-survey-evidence-design-and-presentation/good-practice-in-the-design-and-presentation-of-customer-survey-evidence-in-merger-cases.

Conlon, C. and J. Gortmaker (2020). “Best practices for differentiated products demand estimation with PyBLP”. *The RAND Journal of Economics* 51(4), pp. 1108–1161. DOI: https://doi.org/10.1111/1756-2171.12352. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1756-2171.12352. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/1756-2171.12352.

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.

Fogarty, J. (2010). “The demand for beer, wine and spirits: a survey of the literature”. *Journal of Economic Surveys* 24(3), pp. 428–478.

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.

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.

Holmes, T. P., W. L. Adamowicz, and F. Carlsson (2017). “Choice experiments”. *A primer on nonmarket valuation*, pp. 133–186.

Imthorn, M., R. Kemp, and I. Nobel (Apr. 2016). *Using Conjoing Analysis in Merger Control*. URL: https://www.acm.nl/sites/default/files/old_publication/publicaties/15747_acm-working-paper-conjoint-analysis-2016-04-25.pdf.

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).

Nelson, J. P. (2014). "Estimating the price elasticity of beer: Meta-analysis of data with heterogeneity, dependence, and publication bias". *Journal of Health Economics* 33(C), pp. 180–187. DOI: 10.1016/j.jhealeco.2013.1.

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.

— (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.

Qiu, J., M. Sawada, and G. Sheu (2021). "Win/Loss Data and Consumer Switching Costs: Measuring Diversion Ratios and the Impact of Mergers". *Masayuki and Sheu, Gloria, Win/Loss Data and Consumer Switching Costs: Measuring Diversion Ratios and the Impact of Mergers (November 6, 2021).*

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.

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.

U.S. v. H&R Block Inc., et al., D.D.C. (2011). URL: https://www.justice.gov/atr/case-document/memorandum-opinion-2.

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.[14] 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.

[14] 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

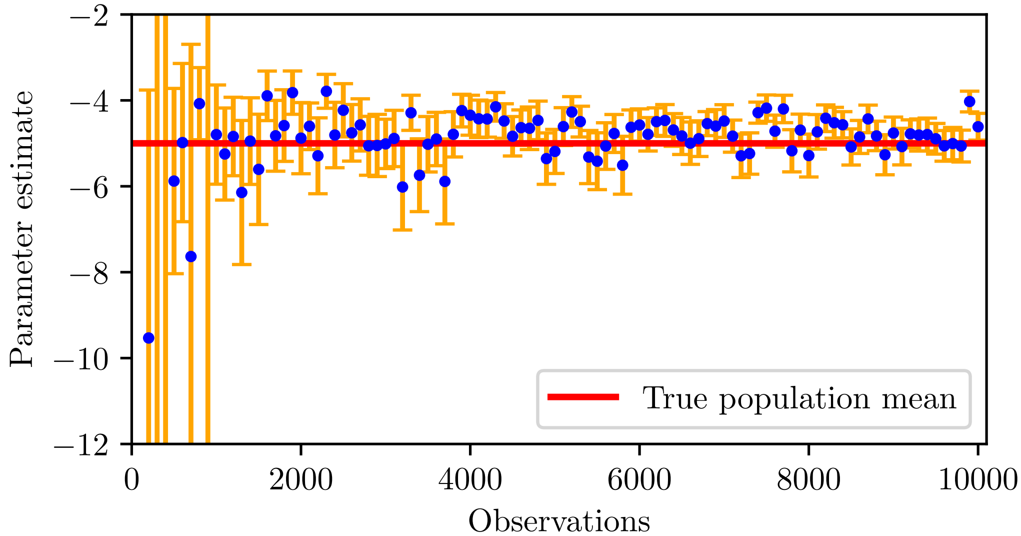
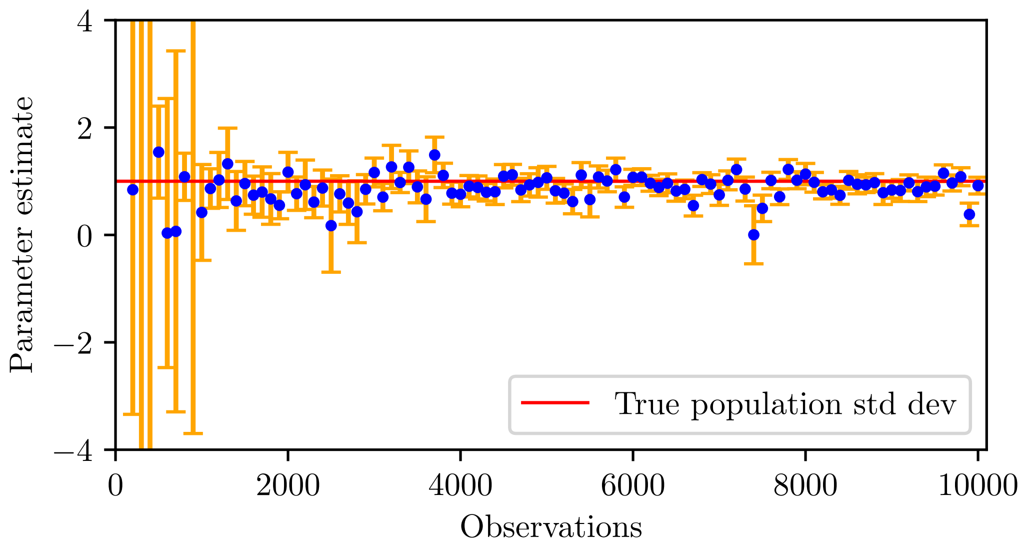


Figure A.2. Estimates of population standard deviation of price coefficient for increasing sample sizes



**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. Additional Tables

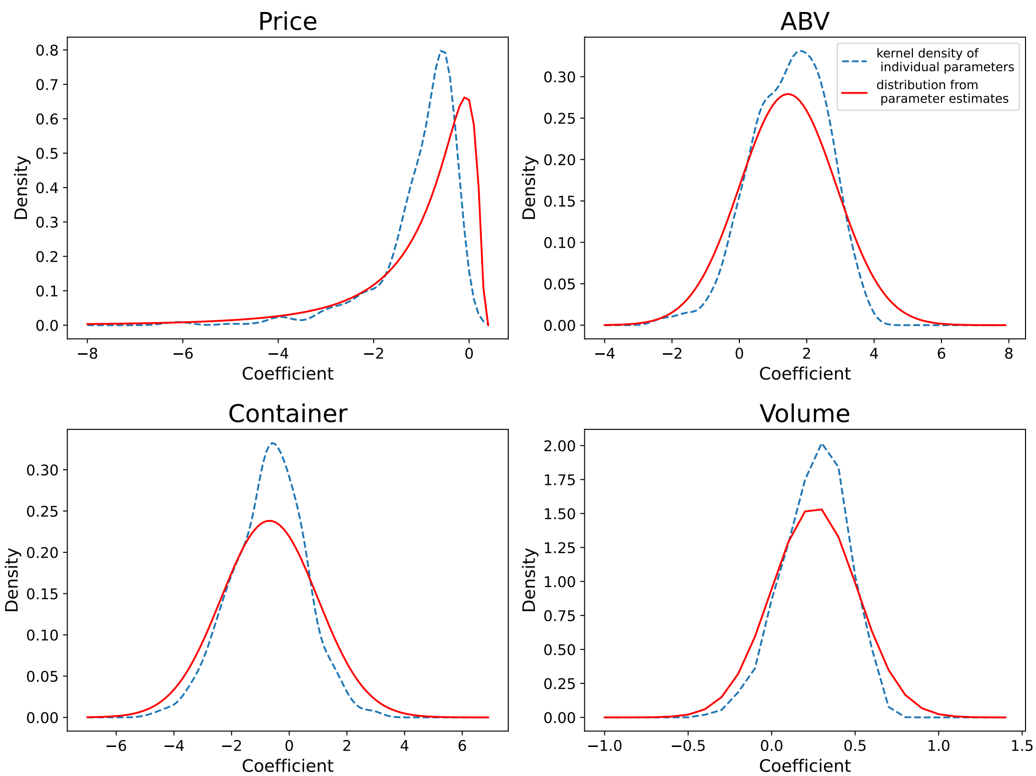
B.1. **Conditional Distributions.** Using the point estimates of $\hat{\theta}$ from the intra-brand treatment we can calculate each subject's tastes conditional on the sequence of choices they made, presented in column 2 of Table B.1, which shows the mean and standard deviation of the 486 individual coefficients, β_n. Column 1 includes the base intra-brand results from table 3 for easy comparisons. 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 estimate of the standard deviation on ABV is 1.116, and the population estimate of the standard deviation is 1.430. Thus, variation in $\bar{\beta}_n$ captures more than 78% of the total estimated variation in the coefficient.

Table B.1. Mixed logit estimates

		Population (1)	Individual (2)
Price	Mean (μ_α)	-1.027	-1.041
	SD (σ_α)	1.148	0.876
ABV	Mean (μ_{β_1})	1.444	1.468
	SD (σ_{β_1})	1.430	1.116
Container	Mean (μ_{β_2})	-0.686	-0.720
	SD (σ_{β_2})	1.675	1.266
Unit volume	Mean (μ_{β_3})	0.256	0.260
	SD (σ_{β_3})	0.257	0.190
Observations		3888	3888

Figure B.1 shows a similar pattern for all the other coefficients. The dashed line shows the kernel density of the individual parameters and the standard

Figure B.1. Comparison of population and individual parameters



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.

Table B.2. 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.3. 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.4. 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
Nat. 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
Bl. Moon	14.99	1.59	11.09	26.1	AM	13.93	1.86	40.3	-7.06
Coors Bqt	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	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