1 Market Power in Differentiated Goods Markets

- Basic problem: Want to estimate a $J \times 1$ demand system $q = q(p, \theta)$ defined in J prices. We initially (in this section) attempt to do this in the most general way possible.
- For item/firm j we have
  - demand $q_j(p, \theta)$, and
  - profit $q_j(p, \theta)(p_j - mc_j)$.
- Profit π_f for firm f with products $k \in \mathscr{F}_f$ is $\pi_f \equiv \sum_{k \in \mathscr{F}_f} q_k(p, \theta)(p_k - mc_k)$.
- Firm f's profit maximizing condition for price j assuming multiproduct Nash behaviour[1] is

$$\frac{\partial \pi_f}{\partial p_j} = \sum_{k \in \mathscr{F}_f} \frac{\partial q_k(p, \theta)}{\partial p_j}(p_k - mc_k) + q_j(p, \theta) = 0. \tag{1.1}$$

  Equation 1.1 encapsulates the *two* sources of market power in this model:
  - *product differentiation*: corresponds to element $\frac{\partial q_j(p,\theta)}{\partial p_j}$, present in both single- and multi-product environment, and;
  - *portfolio effect*: corresponding to elements of the sort $\frac{\partial q_k(p,\theta)}{\partial p_j}$, $\forall k \neq j$; present only for multi-product firms.
- Define $\Delta(p, \theta)$ to be a $J \times J$ block-diagonal matrix with zero off-diagonal blocks. Blocks along the diagonal are square matrices, with $card(\mathscr{F}_f)^2$ elements; the number of blocks is equal to the number of firms, and the block indexed by f contains $\frac{\partial q_k(p,\theta)}{\partial p_j}$, $\forall j, k \in \mathscr{F}_f$, with k indexing rows and j indexing columns. Now we can write

$$\Delta(p, \theta)[p - mc] + q(p, \theta) = 0 \tag{1.2}$$

  for the system of J optimality conditions, yielding

$$p = mc - [\Delta(p, \theta)]^{-1} q(p, \theta). \tag{1.3}$$

  The term $[\Delta(p, \theta)]^{-1} q(p, \theta)$ encapsulates both sources of market power in the multiproduct Nash model.
- If we could get data on or estimate mc and estimate θ, then (with functional form assumptions) we can estimate the effect of mergers/demergers by changing elements in $\Delta(p, \theta)$. E.g. under monopoly there are no zero blocks in $\Delta(p, \theta)$ – i.e. portfolio effects are non-zero as the monopolist accounts for all price effects in her profit maximisation; in case of a demerger across all products, $\Delta(p, \theta)$ becomes diagonal and portfolio effects disappear (still left with distortion from product differentiation).
- Important example: collusion. To incorporate collusion, all we need to do is alter $\Delta(p, \theta)$ so that some share of other firms' profit enters profit for each firm f. Denote this by $\Delta(p, \theta, \kappa)$, assuming that share is a constant κ across all firms. Now the block-off-diagonals for the rows belonging to block f in $\Delta(p, \theta, \kappa)$ correspond to $\kappa \frac{\partial q_{k'}(p,\theta)}{\partial p_j}$ for $k' \notin \mathscr{F}_f$ and $j \in \mathscr{F}_f$, and our optimality conditions are

$$p = mc - [\Delta(p, \theta, \kappa)]^{-1} q(p, \theta). \tag{1.4}$$

- In this last example, if we assume mc is constant across products, we have $J(1 + J)$ parameters to estimate. We can attempt to estimate such a model with market-level data $\{s_{jt}, p_{jt}, x_{jt}\}$, where s_{jt}, p_{jt} and x_{jt} are market shares (quantity), prices and product characteristics respectively, with $j = 1, \ldots, J$ and $t = 1, \ldots, T$. Could also user consumer-level data.

These notes are based on Howard Smith's lectures from the 2018-2019 academic year.
Of course, this is our interpretation of the material Howard presented. Good content is his; mistakes are ours.

[1] i.e. prices being set conditioning on other firms' prices and internalising cross-product effects.

2 Estimating θ via discrete choice models (Demand Side)

2.1 Logit

- Let $\theta = (\beta, \alpha)$. Consumer i's utility for product $j \in \mathscr{F} = \{0, \ldots, J\}$ is:

$$u_j^i = \beta x_j - \alpha p_j + \xi_j + \epsilon_j^i \equiv \delta_j + \epsilon_j^i \tag{2.1}$$

  - β: vector of marginal valuations of characteristics x_j (these are common utility function parameters)
  - α: price sensitivity; marginal utility of money
  - ξ_j: mean utility of j's unobserved characteristics
  - $\delta_j \equiv \beta x_j - \alpha p_j + \xi_j$: mean utility of product j (common across consumers – note, assumes ϵ_j^i are mean zero, but not true for Type 1 Extreme Value).
  - ϵ_j^i is consumer i's individual unobserved utility from product j.

- i chooses $j \iff (u_i^j \geq u_i^k \quad \forall k \in \mathscr{F})$ where $\mathscr{F}$ is the choice set.

- Type-1 extreme value (EV(0,1)):

  – $F(\epsilon) = \exp(-\exp(-\epsilon))$; $f(\epsilon) = \exp(-\epsilon)\exp(-\exp(-\epsilon))$; with support $\Re^n$.

  – $\Pr(u_i^j \geq u_i^k \quad \forall k \in \mathscr{F}) = \frac{e^{\delta_j}}{\sum_{k=0}^J e^{\delta_k}} = \frac{e^{\delta_j}}{1+\sum_{k=1}^J e^{\delta_k}}$; ($\delta_0 = 0$ by normalization).[2]

  – First order statistic (expected maximum value) of u_j^i for j in $\mathscr{F}' \subset \mathscr{F}$:

$$CS = \frac{1}{\alpha}\mathbb{E}\left[\max_{j \in \mathscr{F}'}(\delta_j + \epsilon_j^i)\right] = \frac{1}{\alpha}\log\left[\sum_{j \in \mathscr{F}'} \exp(\delta_j)\right] \tag{2.2}$$

  – When $\mathscr{F}' = \mathscr{F}$, CS is a measure of consumer surplus of the choice set (the "log sum" or "inclusive value").

  – Yields linear model:

$$s_j = \frac{e^{\delta_j}}{\sum_{k=0}^J e^{\delta_k}} \implies \log s_j = \delta_j - \log \sum_{k=0}^J e^{\delta_k} \tag{2.3}$$

  and thus

$$\log s_j - \log s_0 = \delta_j = \beta x_j - \alpha p_j + \xi_j \tag{2.4}$$

  which can be used to estimate $\theta = (\beta, \alpha)$ with ξ_j treated as a residual and $j = 1, \ldots, J$ indexing observations.

- For OLS, must assume $\mathbb{E}[\xi_j | (x_j, p_j)] = 0$; ξ observed by seller (possibly by buyer too) so unlikely to hold. IV estimation requires $\mathbb{E}[\xi_j | (x_j, z_j)] = 0$ for some z_j (instrument validity). Candidate z_j's: Marginal cost shifters; product characteristics not in utility; prices of the same product in other cities (Hausman). Might also want instruments for non-price characteristics x_j.

- Logit elasticities:

  – $\frac{\partial s_j}{\partial p_j} = -\alpha s_j(1 - s_j) \implies \frac{\partial s_j}{\partial p_j}\frac{p_j}{s_j} = -\alpha p_j(1 - s_j)$

  – $\frac{\partial s_k}{\partial p_j} = -\alpha s_j s_k \implies \frac{\partial s_k}{\partial p_j}\frac{p_j}{s_k} = -\alpha s_j p_j \quad \forall k \neq j$, which is *independent of* k.

  – Very inflexible; no effect from closeness of the two products observed characteristics.

[2] See Appendix A.

2.2 (Random Coefficients) Mixed Logit

- Used by Berry, Levinsohn and Pakes (1995); β's now vary across consumers:

$$u_j^i = \beta^i x_j - \alpha p_j + \xi_j + \epsilon_j \tag{2.5}$$

  with

$$\beta_a^i = \bar{\beta}_a + \sigma_a \nu_a^i \tag{2.6}$$

  - $\bar{\beta}_a$, mean value of characteristic a.
  - $\nu_a^i \sim F(\nu)$, individual i's taste for a relative to mean taste; distributed $F(\nu)$, density $f(\nu)$.
  - σ_a, scaling term; spread of taste variation of a.

- Realised utility is then given as

$$u_j^i = \sum_a [\bar{\beta} + \sigma_a \nu_a^i] x_{ja} - \alpha p_j + \xi_j + \epsilon_j^i \tag{2.7}$$

$$\equiv \delta_j + \sum_a \sigma_a \nu_a^i x_{ja} + \epsilon_j^i. \tag{2.8}$$

  with $\delta_j \equiv \bar{\beta} x_j - \alpha p_j + \xi_j$ (i.e. product characteristics).

- Conditioning on i's draw of ν^i we have standard closed form for the individual's choice probability:

$$\Pr(u_j^i > u_l^i \ \forall\ k \neq l \mid \nu^i) = \frac{\exp(\delta_j + \sum_a \sigma_a \nu_a^i x_{ja})}{1 + \sum_{k=1}^J \exp(\delta_k + \sigma_a \nu_a^i x_{ka})} \tag{2.9}$$

  i.e. we have logit conditional on the ν^i's; σ_a adds flexibility/heterogeneity in preferences for x; this means that person with taste for specific characteristics will substitute more readily to alternatives with similar characteristics.

- Obtaining s_j:
  - Integrate over density $f(\nu)$ in the population of consumers:

$$s_j(\xi, \theta) \overset{(2.9)}{=} \int_v \left(\frac{\exp(\delta_j + \sum_a \sigma_a \nu_a^i x_{ja})}{1 + \sum_{k=1}^J \exp(\delta_k + \sigma_a \nu_a^i x_{ka})} \right) f(\nu) d\nu \tag{2.10}$$

  - We do not know $\theta = (\beta, \alpha, \sigma)$ or ξ, and there is no analytic form for this integral, but for a given (ξ, θ) we can evaluate the integral via numerical (Monte Carlo) integration: assuming a density $f(\nu)$ for ν, draw ns values of ν^i from $f(\nu)$ and compute

$$s_j(\xi, \theta)^{sim} = \frac{1}{ns} \sum_{i=1}^{ns} \frac{\exp(\delta_j + \sum_a \sigma_a \nu_a^i x_{ja})}{1 + \sum_{k=1}^J \exp(\delta_k + \sigma_a \nu_a^i x_{ka})} \tag{2.11}$$

- Estimation:
  1. Assume that at true $\theta = (\alpha, \beta, \sigma)$ the unobserved utility ξ is mean-independent of product characteristics x and other instruments z, i.e. $\mathbb{E}[\xi(\theta)|x, z] = 0$, and estimate using these moment conditions in a GMM estimator.
  2. For observations $(\hat{p}, \hat{x}, \hat{s})$ and trial parameter values θ_t there is a unique J-vector ξ implied by inverting the market share equation 2.11: $\hat{s} = s^{sim}(\xi, \theta_t) \implies \xi = \xi(\theta_t, \hat{s})$. This inversion is possible because the market share function is one-to-one: for each j, $\hat{s}$ is a smooth function monotonically increasing in ξ_j and decreasing in ξ_k for all $k \neq j$ (Berry, 1994 conditions).
  3. No closed form solution for $\xi = \xi(\theta_t, \hat{s})$; must be solved numerically at each trial θ_t. This is achieved via the contraction mapping:

$$\xi_n = \xi_{n-1} + \ln \boldsymbol{s} - \ln \boldsymbol{s}^{sim}(\xi_{n-1}, \theta) \tag{2.12}$$

     until some level of tolerance is achieved.

3 Ackerberg, Benkard, Berry, and Pakes (2007, Section 1)

3.1 Demand Side Model

- Let individual i's utility from product j at time/in market t[3] be $u_{ijt} = U(\tilde{x}_{jt}, \xi_{jt}, z_{it}, \nu_{it}, y_{it} - p_{jt}, \theta)$ where:
  - $\tilde{x}_{jt}$ is a K-dimensional vector of observed product characteristics other than price;
  - p_{jt} is the price of the product;
  - ξ_{jt} are the unobserved characteristics of the product (allowed to vary over t);
  - z_{it} and ν_{it} are differences in taste of consumer i, respectively observed and unobserved by the econometrician;
  - y_{it} is consumer i's income;
  - θ is the vector of parameters to be estimates.
- The partial equilibrium nature of the problem is incorporated into the model by letting utility depend on the money available to spend outside of this market $(y_i - p_j)$ (quasi-linear preferences assumed implicitly); sometimes expenditure in other markets is not "explicitly" modelled (as it is here). Then utility is of the form $u_{ij} = U(\tilde{x}_j, \xi_j, z_i, \nu_i, p_j, \theta)$; this is the form they seem to actually work with.
- The model incorporates an option "purchase nothing" as the outside good $j = 0$ with utility $u_{i0} = U(\tilde{x}_0, \xi_0, z_i, \nu_i, \theta)$.
- Product choice is determined by $u_{ij} > u_{ir}, \ \forall\ r \neq j$.
- Assuming linearity and letting $x_j := (\tilde{x}_j, p_j)$ we have $U_{ij} = \sum_k x_{jk}\theta_{ik} + \xi_j + \epsilon_{ij}$, with $\theta_{ik} = \bar{\theta}_k + \theta_j^{o\prime} z_i + \theta_j^{u\prime} \nu_i$; i.e. the observed and unobserved taste parameters interact with the product characteristics. $U_{i0} = 0$ by normalization.
- Unobserved product characteristics ξ_j **must be characterisable as a scalar**; taste for this characteristic is assumed not to vary across consumers (this is a real assumption – not a normalisation – that implies ξ_j can be thought of as a residual).
- ϵ_{ij}'s are a bit awkward; taste variation for product j of individual i that is independent across products and individuals. Can do away with this term; leads to model with nice properties, but it is computationally convenient to keep ϵ_{ij}'s.
- Substituting in the expression for θ_{ik} (remember, i, k is an individual-characteristic pair) we obtain

$$U_{ij} = \delta_j + \sum_{kr} \theta^u_{rk} x_{jk} z_{ir} + \sum_{kl} \theta^u_{lk} x_{jk} \nu_{il} + \epsilon_{ij} \tag{3.1}$$

  where r and l index the elements in z_i and ν_i respectively, and

$$\delta_j = \sum_k \bar{\theta}_k x_{jk} + \xi_j. \tag{3.2}$$

- The interactions of x_{jk} with z_i and ν_i are what generate flexible own and cross price elasticities.

3.2 Estimation with Product-Level Data

- With product level data, $z_i \equiv 0$ for all i (no individuals are observed). For elements in ν_i we might have some data sources (e.g. census data) yielding the relevant distribution; if we don't, we assume some parametric distribution for ν_i (preferably one we can draw from) and include the parameters of that distribution in our vector of parameters we wish to estimate – for example, we might assume a normal distribution for ν_i with mean and standard deviation vectors $\bar{\theta}$ and θ^u respectively.

[3] t dropped in the following for simplicity.

- Since ξ_j's are likely correlated with elements in x_j, we require a set of instruments w_j such that $\mathbb{E}(\xi_j|w_j) = 0$.[4]
- Steps in estimation are actually quite simple:
  1. Get an estimate of $\xi(\cdot)$ as a function of θ – i.e. we need an estimate of a function that varies in the parameters of the model; this is really conceptually the same as OLS, wherein the errors are also functions of the model parameters;
  2. Make some identifying assumptions on $\xi(\cdot;\theta)$, specifying that at the true value of the parameters ($\theta = \theta_0$) it obeys some restrictions – this is where we use our instruments;
  3. Use a standard method of moments estimator to impose the identifying restrictions and back out the unique value of θ such that $\xi(\cdot;\theta)$ satisfies our restrictions.
- Step I: Evaluate, for candidate parameter vector, the integral

$$\sigma_j(\delta,\theta) = \int_v \left(\frac{\exp(\delta_j + \sum_{kl} x_{jk}\nu_{lr}\theta^u_{kl})}{1 + \sum_{q=1}^J \exp(\delta_j + \sum_{kl} x_{jq}\nu_{lr}\theta^u_{ql})} \right) f(\nu)d\nu. \tag{3.3}$$

numerically as

$$\sigma_j(\delta,\theta,P^{ns}) = \sum_{r=1}^{ns} \frac{\exp(\delta_j + \sum_{kl} x_{jk}\nu_{lr}\theta^u_{kl})}{1 + \sum_{q=1}^J \exp(\delta_q + \sum_{kl} x_{jq}\nu_{lr}\theta^u_{ql})}. \tag{3.4}$$

The key thing to note here is that once you have your draws of ν_{lr} (i.e. you have P^{ns}, the empirical distribution of your simulation draws) you can evaluate this thing at different values of ξ,θ.

- Step II: From the empirical market shares $s^n = [s^n_1,\dots,s^n_J]$, where n denotes the size of the sample from which these shares are calculated (often large – but if we have market level data we don't have this sample), find the unique vector δ that makes the predicted shares $\sigma_j(\xi,\theta,P^{ns})$ equal the actual shares s^n for a given value of θ and simulation draws P^{ns}. This is achieved via the contraction mapping (as shown by BLP 1995):

$$\delta^k_j(\theta) = \delta^{k-1}_j + ln[s^n] - ln[\sigma_j(\delta^{k-1}_j,\theta,P^{ns})]. \tag{3.5}$$

where k indexes loop iterations. Recall that δ is a function of some elements in θ; but of course we're not allowing theta to vary here – only δ, and by implication $\sigma_j(\delta^{k-1}_j,\theta,P^{ns})$ is changing. We stop this process when $\sigma_j(\delta^{k-1}_j,\theta,P^{ns})$ and s^n_j are "close enough". Call the fixed point obtained by this process $\delta(\theta,s^n,P^{ns})$. The model 3.1 then implies that

$$\xi_j(\theta,s^n,P^{ns}) = \delta(\theta,s^n,P^{ns}) - \sum_k x_{jk}\bar{\theta}_k. \tag{3.6}$$

I.e. we've managed to obtain an estimate of the ξ_j's as a function of the model parameters θ; this thing varies as we vary θ, so now we need to specify the restrictions on ξ_j and vary θ until we find the set of parameters that most closely satisfies those restrictions.

Side note on *identification*: As mentioned, identification here follows from restrictions on the true distribution of ξ; for identification, as n and ns go to infinity, $\xi(\theta,s^\infty,P^\infty)$ will/must uniquely satisfy these restrictions at $\theta = \theta_0$. You could use different restrictions to achieve this result; we work with $\mathbb{E}[\xi(\theta_0)|w] = 0$, which leads to the third and final step of the algorithm:

- Step III: Minimize $||G_{j,n,ns}(\theta)||$ over θ, where

$$G_{j,n,ns}(\theta) = \sum_j^J \xi(\theta,s^n,P^{ns})f(w_j). \tag{3.7}$$

Note that given θ^u there is a closed form solution for $\bar{\theta}$; so the nonlinear search is over θ^u only.

[4] Is this maybe a stronger assumption than necessary? Could we merely have assumed orthogonality?

3.3 Improving Efficiency

- ABBP (2007) also discuss improving model efficiency via jointly estimating the pricing equation outlined in section 1;
- requires assumptions regarding the functional form of the marginal cost function, as well as assumptions on the market structure (as outlined above).
- Alternatively, micro data can be used to improve efficiency.
- Especially useful when second choice data or panel data are available, but details on individual characteristics also useful (allows us to shift some elements out of ν and into z).
- If we don't have data on individual choice behaviour, we can still use micro data to draw values of ν for simulating the model with market level data.
- Almost nothing changes when we have micro data on consumers – all we need do is distinguish between the observed and unobserved taste shifters in the market share equation:

$$\sigma_j(\delta, \theta) = \int_v \left(\frac{\exp(\delta_j + \sum_{kl} x_{jk} z_{lr} \theta^o_{kl} + \sum_{kl} x_{jk} \nu_{lr} \theta^u_{kl})}{1 + \sum_{q=1}^{J} \exp(\delta_q + \sum_{kl} x_{jk} z_{lr} \theta^o_{kl} + \sum_{kl} x_{jq} \nu_{lr} \theta^u_{ql})} \right) f(\nu) d\nu. \tag{3.8}$$

4 Application: Miller and Weinberg (2017)

The authors estimate a differentiated-goods pricing model with a random coefficients mixed logit demand side to explore the implications of the Miller-Coors joint venture for the beer market. Analysis based on market level data from 2001-2010 for multiple bundles of beer and multiple producers.

The basic idea is that the merger could have moved the market to a new equilibrium with facilitated price coordination with ABI, a large competitor. The table below shows that the post-merger price changes for Miller-Coors and ABI are impressively close regardless of specification, which in principle is consistent with the idea of a coordination equilibrium.

TABLE II
CHANGES IN RETAIL PRICES BY FIRM[a]

	(i)	(ii)	(iii)	(iv)
$\mathbb{1}$\{MillerCoors\} × $\mathbb{1}$\{Post-Merger\}	0.098	0.050	0.047	0.069
	(0.007)	(0.004)	(0.005)	(0.007)
$\mathbb{1}$\{ABI\} × $\mathbb{1}$\{Post-Merger\}	0.087	0.040	0.038	0.062
	(0.007)	(0.005)	(0.005)	(0.007)
$\mathbb{1}$\{Post-Merger\}	−0.031	−0.007	−0.002	0.010
	(0.005)	(0.004)	(0.004)	(0.009)
log(Employment)	–	–	−0.051	0.131
	–	–	(0.080)	(0.081)
log(Earnings)	–	–	0.156	0.152
	–	–	(0.029)	(0.035)
Pre-Merger Average Price	11.75	11.14	11.14	11.14
Product Trends	No	No	Yes	Yes
Covariates	No	No	Yes	Yes
# Observations	25,740	167,695	167,695	151,525

The demand side is standard random coefficients mixed logit with the sole purpose of providing a way of quantifying cross-price elasticities before and after the merger. Refer to previous sections for estimation; the only thing worth noting is that the set of instruments used include cost shifters (distance from factory) and a post-merger indicator – under the assumption of exogeneity of the merger with respect to the competitive environment.

The supply model allows for post-merger departures from Nash-Bertrand play for Miller-Coors and ABI, while forcing other competitors to play Nash-Bertrand. The assumption that changes in post-merge unobserved demand for ABI are not different from changes for other competitors allows to

identify collusion as price response in excess of what can be explained by unilateral effects.

The demand side of the model is estimated first; these estimates are then used to populate the cross-price effects matrix and estimate the supply side in search for collusion. The demand side is **not** estimated before and after the merger: instead, the authors use estimated cross-price elasticities from the pre-merger period to populate the post-merger $\Delta(\kappa, p, \hat{\theta})$ matrix and to estimate κ (assumed constant). In particular, remember that for the supply side we have

$$p = mc - [\Delta(\kappa, p, \hat{\theta})]^{-1} q(p, \hat{\theta})$$

where $\hat{\theta}$ are demand parameters; for $mc = w\gamma + \eta$, this implies

$$p = w\gamma + \eta - [\Delta(\kappa, p, \hat{\theta})]^{-1} s \tag{4.1}$$

clearly, endogeneity of Δ prevents direct estimation of this equation; however, post-merger dummies again save the day as instruments allowing to identify (κ, γ). The table below clearly shows that **the null of no collusion is rejected**.

TABLE VI

BASELINE SUPPLY ESTIMATES[a]

Demand Model:		NL-1	RCNL-1	RCNL-2	RCNL-3	RCNL-4
Data Frequency:		Monthly	Monthly	Quarterly	Monthly	Quarterly
Variable	Parameter	(i)	(ii)	(iii)	(iv)	(v)
Post-Merger Internalization of Coalition Pricing Externalities	κ	0.374 (0.034)	0.264 (0.073)	0.249 (0.087)	0.286 (0.042)	0.342 (0.054)
Marginal Cost Parameters						
MillerCoors × PostMerger	γ_1	−0.608 (0.039)	−0.654 (0.050)	−0.649 (0.060)	−0.722 (0.042)	−0.526 (0.040)
Distance	γ_2	0.142 (0.046)	0.168 (0.059)	0.163 (0.059)	0.169 (0.060)	0.148 (0.049)

[a]This table shows the baseline supply results. We use the method of moments for estimation. There are 94,656 observations at the brand–size–region–month–year level in columns (i), (ii), and (iv) and 31,784 observations at the brand–size–region–year–quarter level in columns (iii) and (v). The samples exclude the months/quarters between June 2008 and May 2009. All regressions include product (brand × size), period (month or quarter), and region fixed effects. Standard errors clustered by region and shown in parentheses.

TABLE X

RESULTS FROM COUNTERFACTUAL ANALYSIS[a]

Coordinated Effects:	yes	yes	no	no	no
Unilateral Effects:	yes	yes	yes	yes	no
Efficiencies:	yes	no	yes	no	no
	(i)	(ii)	(iii)	(iv)	(v)
			Retail Prices		
ABI	10.03	10.14	9.38	9.55	9.43
Miller	8.94	9.37	8.28	8.72	8.19
Coors	10.18	10.85	9.56	10.22	9.26
			Brewer Markups		
ABI	4.45	4.56	3.81	3.97	3.84
Miller	4.52	4.32	3.83	3.63	3.05
Coors	4.25	4.06	3.61	3.41	2.68
			Welfare Statistics		
Producer Surplus	22.1%	19.1%	10.3%	8.2%	–
ABI	10.3%	19.8%	−0.08%	9.3%	–
Miller	37.8%	20.2%	24.6%	9.1%	–
Coors	47.8%	12.9%	34.7%	3.5%	–
Consumer Surplus	−3.7%	−5.3%	−0.2%	−2.1%	–
Total Surplus	1.3%	−0.6%	1.8%	−0.1%	–

The counterfactual analysis exercise shows that the merger increases producer surplus by more than 20%, with around half of these gains resulting from coordination; consumer surplus is however reduced. In total, the merger increases welfare by around 1%, but this increase would have been higher under no coordination.

A Derivation of Logit functional form

$$\Pr(u_j^i \geq u_k^i) = \Pr(\epsilon_k^i \leq \delta_j^i - \delta_k^i + \epsilon_j^i) \Rightarrow \Pr(\epsilon_k^i \leq \delta_j^i - \delta_k^i + \epsilon_j^i \mid \epsilon_j^i) = \Pi_{k \neq j} e^{e^{-(\epsilon_j^i + \delta_j^i - \delta_k^i)}}$$

$$\Rightarrow \Pr(\epsilon_k^i \leq \delta_j^i - \delta_k^i + \epsilon_j^i) = \int_\epsilon (\Pi_{k \neq j} e^{-e^{-(\epsilon_k^i + \delta_j^i - \delta_k^i)}}) e^{-\epsilon_k^i} e^{-e^{-\epsilon_k^i}} d\epsilon_k^i = \int_\epsilon (\Pi_k e^{-e^{-(\epsilon_k^i + \delta_j^i - \delta_k^i)}}) e^{-\epsilon_k^i} d\epsilon_k^i =$$

$$= \int_\epsilon e^{-\sum_k e^{-(\epsilon_k^i + \delta_j^i - \delta_k^i)}} \underbrace{e^{-\epsilon_k^i} d\epsilon_k^i}_{c.o.v. \to dt} = \int_0^\infty e^{-t \sum_k e^{-(\delta_j^i - \delta_k^i)}} dt = \left[\frac{e^{-t \sum_k e^{-(\delta_j^i - \delta_k^i)}}}{-\sum_k e^{-(\delta_j^i - \delta_k^i)}} \right]_0^\infty = 0 - \frac{1}{-\sum_k e^{-(\delta_j^i - \delta_k^i)}} =$$

$$\frac{1}{\sum_k e^{-(\delta_j^i - \delta_k^i)}} = \frac{e^{\delta_j^i}}{\sum_k e^{\delta_k^i}} \quad \square$$