Advanced Industrial Organization II
Lecture 4: Cartels. Vertical Mergers.

Måns Söderbom*

14 April 2009

1 Plan

This time we have three main activities:

1. Student presentation of Porter (1983) followed by general discussion.
2. Collusive markers, based on Harrington (2005)
3. Vertical mergers.

*Department of Economics, University of Gothenburg. Office: E526. E-mail: mans.soderbom@economics.gu.se

Next time we meet (April 27th) we will complete our study of mergers. This will involve a student presentation of the paper by Kim & Singal, and a discussion of a policy-oriented paper analyzing merger control in the EU (Neven and Röller, 2002).

2 Porter (1983): Some pointers

Reference: Porter, Robert H., "A Study of Cartel Stability: The Joint Executive Committee, 1880-1886," Bell Journal of Economics, 14 (1983), 301-314.

In the lecture on April 6th I summarized the Porter (1983) paper briefly, and we have now had a student presentation based on it as well. This is an important paper. Here are some pointers that hopefully will make it easier to read.

2.1 Hypothesis

The main null hypothesis in this paper is that *no structural change has occurred*, i.e. that only cooperative or noncooperative (Cournot) behaviour is observed, but not both (p. 311).

2.2 Key behavioral mechanisms

Firms are faced with the problem of detecting and deterring cheating on the cooperative agreement (note also that cartels weren't illegal at the time).

If the market price falls, this could be because of a fall in demand, or because a cartel member has secretly lowered the price (cheating) in order to raise his profits.

2.3 Econometrics

- Demand and supply equations:

$$\begin{aligned} \ln Q_t &= \alpha_0 + \alpha_1 \ln P_t + \alpha_2 L_t + U_{1t} && \text{(Demand)} \\ \ln P_t &= \beta_0 + \beta_1 \ln Q_t + \beta_2 S_t + \beta_3 I_t + U_{2t} && \text{(Supply)}, \end{aligned}$$

where Q is the volume of grain shipped, P is the rate for rail services, L is a dummy variable equal to 1 (0) if the Great Lakes are (not) open for shipping (if they are open, then this provides an alternative way of transportation, thus reducing the demand for rail services; hence we expect $\alpha_2 < 0$); S captures composition of the cartel (changes due to entry etc.).

- Key variable: I_t. This is equal to 1 if firms are in the collusive phase (low quantity, high price), and 0 if firms are in punishment phase (high quantity, low price). I_t is **not observed** by the econometrician.

- Porter considers two solutions to this problem:

  - Create a dummy variable, denoted PO, equal to 0 unless the *Railway Review* (a trade magazine) reported that a price war was occurring in which case it is equal to 1.

– Use a switching regression treating I_t as a stochastic variable, being equal to 1 with probability λ and 0 with probability $\mathbf{1} - \lambda$.

- The residuals U_{1t}, U_{2t} are assumed to follow a **bivariate normal distribution** (no serial correlation):

$$(U_{1t}, U_{2t}) \,\tilde{}\, N(\mathbf{0}, \mathbf{\Sigma}),$$

where $\mathbf{0}$ is a $2 \times \mathbf{1}$ vector of zeros (reflecting zero means) and

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$$

is the variance-covariance matrix. Unless $\sigma_{12} = 0$, this implies that $\ln P_t$ will be an endogenous explanatory variable in the demand equation, and that $\ln Q_t$ will be an endogenous explanatory variable in the supply equation. Why?

- Porter writes the econometric model in compact form (matrix notation) in eq. (4) in the paper. He then presents the **likelihood function** for this model on p.306. Below is a brief summary of the logic underlying likelihood functions, and how they are used in empirical research.

2.3.1 Introduction to likelihood functions (optional)

- We now depart from Porter's paper for a moment, in order to discuss the meaning of likelihood functions in general.

- Suppose that, in the population, there is a variable w which is distributed according to some distribution $f(w;\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is a vector of unknown parameters.

- Suppose we have a **random sample** $\{w_1, w_2, ..., w_N\}$ drawn from the population distribution $f(w; \boldsymbol{\theta})$ where $\boldsymbol{\theta}$ is unknown.

- Our objective is to estimate $\boldsymbol{\theta}$. Our sample is more likely to have come from a population characterized by one particular set of parameter values, say $\tilde{\boldsymbol{\theta}}$, than from another set of parameter values, say $\breve{\boldsymbol{\theta}}$.

- The maximum likelihood estimate (MLE) of $\boldsymbol{\theta}$ is simply the particular vector $\hat{\boldsymbol{\theta}}^{ML}$ that gives the **greatest likelihood** (or, if you prefer, probability) of observing the sample $\{w_1, w_2, ..., w_N\}$.

- What do I mean by the 'likelihood of observing the sample'? Suppose there are two observations, w_1 and w_2. What is the likelihood of observing this sample?

- Clearly the likelihood of observing w_1 is $f(w_1; \boldsymbol{\theta})$ and the likelihood of observing w_2 is $f(w_2; \boldsymbol{\theta})$. Further, because of the random sampling assumption w_1 and w_2 are independent, hence the likelihood of observing w_1 **and** w_2 (i.e. the sample) is simply

$$L(\boldsymbol{\theta}; w_1, w_2) = f(w_1; \boldsymbol{\theta}) \cdot f(w_2; \boldsymbol{\theta})$$

  i.e. the **product** of the individual likelihoods. The equation just defined is a function of $\boldsymbol{\theta}$: for some values of $\boldsymbol{\theta}$ the resulting L will be relatively high while for other values of $\boldsymbol{\theta}$ it will be low. This is why we refer to equations of this form as **likelihood functions**.

- The value of $\boldsymbol{\theta}$ that gives the maximum value of the likelihood function is the maximum likelihood estimate of θ.

- This idea, of course, is easily generalized to the case where we have N observations on w. In this case we can write the likelihood function as

$$L(\boldsymbol{\theta}; w_1, w_2, ..., w_N) = f(w_1; \boldsymbol{\theta}) f(w_2; \boldsymbol{\theta}) \cdot ... \cdot f(w_N; \boldsymbol{\theta}),$$

or, in more compact notation,

$$L(\boldsymbol{\theta}; w_1, w_2, ..., w_N) = \prod_{i=1}^{N} f(w_i; \boldsymbol{\theta}).$$

Again, the value of $\boldsymbol{\theta}$ that gives the maximum value of the likelihood function is the $\hat{\boldsymbol{\theta}}^{ML}$.

- For computational reasons it is much more convenient to work with the **log-likelihood function**:

$$\ln L(\boldsymbol{\theta}; w_1, w_2, ..., w_N) = \sum_{i=1}^{N} \ln f(w_i; \boldsymbol{\theta}).$$

Clearly, the value of $\boldsymbol{\theta}$ that gives the maximum value of the log likelihood function is the $\hat{\boldsymbol{\theta}}^{ML}$.

Illustration: The log likelihood for the simple linear model Now consider the simple regression model

$$y = \alpha + \beta x + \varepsilon.$$

Suppose that ε follows a normal distribution with mean zero and constant variance:

$$\varepsilon \tilde{} N\left(0, \sigma^2\right).$$

Suppose that x is exogenous. The likelihood of observing ε is thus

$$f(\varepsilon) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\varepsilon^2}{2\sigma^2}\right).$$

Now relate this to data and parameters:

$$f(\varepsilon) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(y-\alpha-\beta x)^2}{2\sigma^2}\right).$$

The sample log likelihood is simply

$$\ln L = \sum_{i=1}^{N} \ln f(y_i, x_i; \beta, \sigma),$$

and the maximum likelihood estimates of β, σ are those that lead to the highest possible sample log likelihood; to find these estimates, we thus maximize the sample log likelihood with respect to β, σ. It turns out that for the present model this results in the OLS estimator. (Optional: can you show this?)

Porter uses this framework, tailored to his model in which there are two residuals. It can be shown that if I_t is observed, the estimates maximizing the likelihood function

$$L(I_1, ..., I_T) = \prod_{t=1}^{T} h(y_t | I_t)$$

are the two-stage least squares (2SLS) estimates. Hence he uses 2SLS rather than maximum likelihood for his model in which I_t is constructed based on auxiliary information on price wars.

The switching regression model is more complicated however. Here he treats I_t as a **stochastic** variable, being equal to 1 with probability λ and 0 with probability $1 - \lambda$. Basically, we change the likelihood of each observation, from $h(y_t | I_t)$ to

$$f(y_t) = \lambda \times h(y_t | I_t = \Delta) + (1 - \lambda) \times h(y_t | I_t = 0),$$

and write the sample likelihood as

$$L = \prod_{t=1}^{T} f\left(y_t\right). \tag{1}$$

Thus, rather than conditioning on the regime variable - which we can't do since it is unobservable - we "integrate out" the stochastic regime variable using λ as the weight. $\mathbf{\Delta}$ and λ are thus estimated parameters.

The likelihood function (1) can be maximized using a wide range of algorithms - Porter uses a so called E-M (expectation maximization) algorithm but we don't have to worry about that here.

2.4 Questions

- Given that the focus of the paper is on the supply relationship, why does Porter estimate a demand function as well?

- How, exactly, are the 2SLS results in Table 3 obtained? What does the first-stage equation look like? What does the second-stage equation look like?

- Why does Porter argue that the econometric results in Table 3 are "not consistent with single-period profit maximization"?

- How does Porter test whether the hypothesis that no structural change has occurred? What does he conclude?

3 Collusive Markers

Reference: Section 3 in Harrington (2005).

- What **behavioral patterns** are indicative of collusion? Behavior that distinguishes collusion from competition are often referred to as a **collusive marker.** If we know what constitutes a collusive marker, we know what to look for.

- **Theory** plays an important role in providing collusive markers. Basically, empirical findings need to be explainable by theory, otherwise your empirical analysis will not be convincing. Thus, interpretation should be done in light of the underlying theoretical model. When we find evidence of collusion, there is always the possibility that there actually is no collusion and the problem is we've misspecified the non-collusive model, for example. Still, collusive markers can serve to **screen** industries to determine whether they are worthy of more intense investigation.

- Harrington discusses distinguishing features of collusion and competition, focusing on patterns for **prices** and **market shares.** In particular, he discusses how collusive markers can be based on:
  - Relationship between the firm's prices and demand movements
  - Stability of price & market shares
  - Relationship between firms' prices.

Basically, theories exist indicating how these are **affected** by collusion.

- The general point is that various theories offer various predictions. From our point of view, it is more important to understand how these collusive markers are derived, rather than memorizing what they actually are.

- I will focus on the relationship between the firm's prices and demand movements, because I think the discussion in the paper is more intuitive than for price stability and the relationship between firms' prices; for the latter, I find the discussion a little too brief for it to be possible to fully understand the main mechanisms. Read if you are interested.

3.1 Predictions on price

3.1.1 Relationship between the firm's prices and demand movements

- Repeated game model (dynamic): Is collusion sustainable? This depends on the rewards and punishments in the model.

- Example - consider static game in which firms compete according to the Cournot model. Firms do have an incentive to collude (so as to mimic a monopolist), but coordination problems imply it will be optimal for each firm to cheat, i.e. increase output over and above the agreed level, since this will lead to higher profits (the reward).

- Firms can only be deterred from cheating if they experience a future loss, in the form of intensified competition in response to cheating (the punishment).

- **Incentive Compatibility Constraint** (ICC): The discounted **foregone** future profit stream associated with cheating is at least as large as the current gain from cheating. If the punishment is reversion to the non-collusive outcome for T periods, then the ICC is

$$\sum_{\tau=1}^{T} \delta^{\tau} \left(\pi^{c} - \pi^{nc}\right) \geq \pi^{d} - \pi^{c},$$

  or, equivalently,

$$\pi^{c} + \sum_{\tau=1}^{T} \delta^{\tau} \pi^{c} \geq \pi^{d} + \sum_{\tau=1}^{T} \delta^{\tau} \pi^{nc}$$

where π^c is profits under collusion, π^c is profits under non-collusion (e.g. Cournot competition) and π^d is profits from deviating (cheating). In equilibrium, the ICC holds and so nobody cheats.

- What's the effect of a decrease in the discount factor δ on the nature of the punishment, for the ICC to hold?
- What's the effect of a decrease in the punishment duration T on the nature of the punishment, for the ICC to hold?

- Now suppose that firms are **not** colluding (competitive benchmark), and suppose demand is stochastic. In such a framework, what is the relationship between prices and demand?

- Basically, you expect demand and prices to be positively correlated - demand curve shifts outward, price increases.

- Other extreme: perfect collusion - i.e. firms are colluding to mimic a monopolist - then also price and demand are positively correlated.

- So - can't distinguish between competition and perfect collusion. Are we making any progress? Well no, if we suspect it's either perfect collusion or competition. But maybe the type of collusion we're looking for is unlikely to be perfect - e.g. because of information imperfections (cf. Porter, 1983).

- Suppose now we are pretty confident firms behave either competitively or according to imperfect collusion; by the latter I mean a scenario in which firms have an incentive to deviate. This means the ICC binds - relax it, and firms would change their behavior (i.e. deviate). Suppose demand shocks are iid - notably **serially uncorrelated.** Now ask yourself: what kind of relationship between demand shocks and prices would you expect to see in this scenario?

- Return to the (binding) ICC:

$$\sum_{\tau=1}^{T} \delta^{\tau} \left(\pi^{c} - \pi^{nc}\right) = \pi^{d} - \pi^{c}.$$

Suppose demand is atypically high in the current period. How does this affect the ICC? Well, since the demand shock is assumed iid - serially uncorrelated, in particular- it follows that the **future** loss from cheating is not affected. The cartel board understands this, of course, and also understands that it therefore needs to keep $\pi^{d} - \pi^{c}$ constant in order to preserve the equality of the ICC. So suppose the cartel board does nothing - what would happen to $\pi^{d} - \pi^{c}$?

- If the other cartel members produce and supply the market according to the agreement, then clearly firm i can make a big profit by (say) increasing the quantity supplied to the market over and above its quota. Indeed, the current gain from deviating is higher in periods where demand is strong, than in periods of weak demand. Hence, if demand is strong, and the cartel does nothing, then $\pi^d - \pi^c$ will increase, thus violating the ICC.

- What is the cartel board to do? One option is to increase the supply to the market (thus tending to **reduce** the price), since this will weaken the incentive to deviate. Think of this as an outward shift of the supply curve.

- Depending on the size of the cartel and the slope of demand & supply curves, this mechanism may result in a negative correlation between prices and demand - high demand $\longrightarrow$ stronger incentive to cheat$\longrightarrow$ cartel reduces the price in order to weaken the incentive to cheat.

- We now have our first collusive marker: in a competitive equilibrium (and under perfect collusion), prices and demand are **positively** correlated, whereas under imperfect collusion they can be **negatively** correlated.

- If demand is cyclical, we can adopt this way of thinking. For example, at the peak of the cycle, everyone knows demand in the near future will be lower than at present. If there is collusion, this means the punishment is weaker **at** the peak than at points **just before** the peak. Thus preserving the ICC is more difficult at the peak than just before the peak, so the cartel board has to consider lowering the price at the peak. Hence, the price path will **lead** the demand cycle.

- In contrast, under non-collusion, the price just follows demand.

- Collusive marker: if demand is highly cyclical - collusion, and price will lead the demand cycle; competition, and price will move simultaneously with demand cycle.

- Collusive marker related to Porter (1983): large changes in price & quantity in the absence of demand and cost changes (regime switching).

- Summary:

Under certain conditions, price and quantity are **negatively** *correlated, and price* **leads** *a demand cycle; while price and quantity are* **positively** *correlated and price* **follows** *a demand cycle.*

3.1.2 Discussion

- As mentioned above, Harrington also discusses collusive markers for price stability (some models show that prices and market shares are **more stable** under collusion than under competition) and the relationship between firms' prices (some models show that firms prices tend to be **more strongly correlated** under collusion).

- Important: None of the markers discussed in the paper are universal or guaranteed to work.

- For example, we saw above how one collusive marker is based on the ICC binding. For very strong cartels this may not be the case, in which case they will not be discovered using the approach above.

- Also, collusive markers are to some extent model specific, and it is not so clear always what would happen if you relax certain theoretical assumptions.

- Also, existing models do not distinguish between tacit and explicit collusion - it's the latter form of collusion that the antitrust authorities are concerned with. Need to find evidence of direct communication among firms - e.g. through "dawn raids" by antitrust lawyers.

- The current methods available can find a smoking gun, but cannot deliver convincing and complete evidence of collusion.

4 Vertical Mergers

Johan Stennek has discussed horizontal mergers earlier in this course. We now look at **vertical mergers** - i.e. mergers of firms that produce different products that get combined to yield the final good or service. One example would be firms operating at different levels of the production chain, e.g. a wholesaler & a retailer.

- In the decades prior to 1980, vertical mergers were often seen as anticompetitive because they would facilitate foreclosure (e.g. downstream firms not part of the merger will not be able to find a supplier after the merger)

- In the 1980s, it was realized that vertical mergers may generate **efficiency gains** and improvement in consumer welfare. This could happen, for example, if, prior to the merger, both the upstream and the downstream firms were monopolies. As we shall see below, such a chain of monopolies may well lead to a higher price for the final product than if the two firms were to merge and act as a single monopolist.

- Research in the mid 1990s, however, challenged this view, highlighting conditions under which vertical mergers would harm consumers.

- We begin by looking at the argument that vertical mergers are **procompetitive**, and good for consumers.

4.1 Procompetitive vertical mergers

- Consider the case where there is a single upstream supplier (the manufacturer; firm U) how sells a unique product to a single downstream firm (the retailer; firm D). That is, we have two monopolists.

- Firm U has marginal cost c, and sells the product to D at price r.

- Firm D has no retailing cost (for simplicity; easy to relax), and sells the product to the consumer at price P.

- Consumer demand according to linear inverse demand function

$$P = A - BQ.$$

- D's profit:

$$\begin{aligned}\Pi^D &= (P - r)\,Q \\ \Pi^D &= (A - BQ - r)\,Q.\end{aligned}$$

- D's profit maximizing level of Q satisfies $\frac{\partial \Pi^D}{\partial Q} = 0$, i.e.

$$\begin{aligned}(A - r) - 2BQ^D &= 0 \\ Q^D &= \frac{(A - r)}{2B}.\end{aligned}$$

  As we shall see, this effectively is the demand curve facing the upstream firm.

- Plug this into the demand function and you get the market clearing price:

$$
\begin{aligned}
P &= A - BQ^D \\
P &= A - B\frac{(A-r)}{2B} \\
P &= \frac{2A-(A-r)}{2} \\
P &= \frac{(A+r)}{2}.
\end{aligned}
$$

- D's profit:

$$\begin{aligned}\Pi^D &= (A - BQ - r)\,Q \\ \Pi^D &= (A - r)\,Q - BQ^2 \\ \Pi^D &= (A - r)\,\frac{(A - r)}{2B} - B\left[\frac{(A - r)}{2B}\right]^2 \\ \Pi^D &= \frac{2 \times (A - r)^2}{2 \times 2B} - \frac{(A - r)^2}{4B} \\ \Pi^D &= \frac{(A - r)^2}{4B}.\end{aligned}$$

- Now consider firm U. What price should this firm charge? We write down U's profit function,

$$\Pi^U = (r - c)\,Q.$$

What is the relevant demand curve? Since the buyer of Q from U is the downstream firm D, it must be that the demand for U's product is given by the retailer's demand function; recall

$$Q^D = \frac{(A - r)}{2B},$$

thus inverse demand is

$$r = A - 2BQ$$

Hence U's profit is

$$\begin{aligned} \Pi^U &= (r - c)\, Q \\ \Pi^U &= ((A - 2BQ) - c)\, Q \\ \Pi^U &= AQ - 2BQ^2 - cQ, \end{aligned}$$

and so $\frac{\partial \Pi^U}{\partial Q} = 0$ implies

$$
\begin{aligned}
A - c &= 4BQ \\
Q &= \frac{A-c}{4B}
\end{aligned}
$$

is the optimal quantity supplied by U; at price

$$
\begin{aligned}
r &= A - 2BQ \\
r &= A - 2B\frac{A-c}{4B} \\
r &= A - \frac{A-c}{2} \\
r &= \frac{A+c}{2}
\end{aligned}
$$

- The intuition is very straightforward: Firm U adds a mark-up to its production cost c, resulting in

$$r = \frac{A+c}{2} > c,$$

where the inequality follows from the assumption that $A > c$ which has to hold (Why? Check inverse demand). Firm D adds a mark-up to its purchase cost r resulting in

$$\begin{aligned} P &= \frac{A+r}{2} \\ P &= \frac{A+\frac{A+c}{2}}{2} \\ P &= \frac{3A+c}{4}. \end{aligned}$$

- So there's markup twice; the price is thus higher, and the quantity supplied lower, than what would be the case if the two firms merged and became one monopolist.

[Figure 17.2 here]

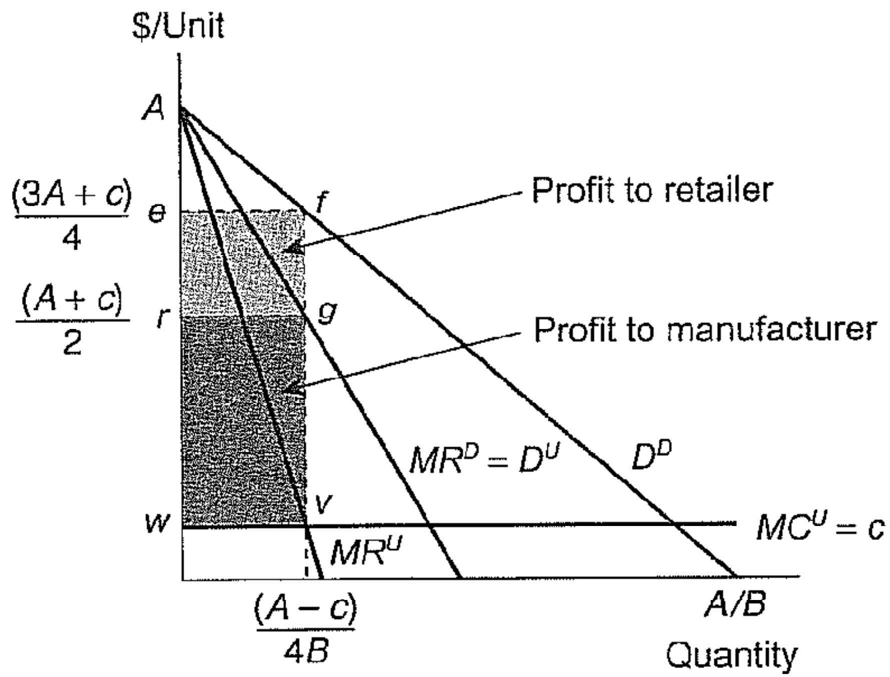


Figure 17.2 Upstream and downstream profit maximization without vertical integration
The retailer's marginal revenue curve MR^D is the manufacturer's demand curve D^U. Double marginalization results when the manufacturer sets its optimal wholesale price $r = (A + c)/2$ above marginal cost c, after which the retailer adds a further markup by setting retail price $P = (3A + c)/4$. Retail profit is area *refg*. The manufacturer's profit is area *wrgv*.

- Now suppose that U and D merge, becoming a single monopolist. The demand curve of the producer is now given by the consumer demand curve, and so the profit of the monopolist is given by

$$\begin{aligned} \Pi^I &= (P-c)\,Q \\ \Pi^I &= AQ - cQ - BQ^2, \end{aligned}$$

using

$$P = A - BQ.$$

I can quickly get to the main point - profit maximization implies

$$Q = \frac{A-c}{2B},$$

and so the price of the product is

$$P = A - B\frac{A-c}{2B}$$
$$P = \frac{A+c}{2}.$$

It's straightforward to show that the merger has resulted in:

- An increase in production
- A decrease in price
- An increase in the combined profits of U & D
- An increase in consumer welfare.

• Thus, from a social welfare point of view, the merger has benefitted everyone!

- See Figure 17.3 in Pepall et al. (p. 434) for a graphical illustration.

- Sounds too good to be true? Maybe. Caveats:

  – As the firm becomes big, it may be more difficult to preserve the right **incentives** for people in the new organization. You add more tiers in the organization, and so you may have to spend more resources managing the organization (supervise the supervisors etc.). This may reduce efficiency.

  – We've seen a lot of outsourcing - why?

– **Technology**: If there are several inputs in the downstream production function, the D firm will respond to the high price charged by U by using less of that input (e.g. capital; K) and more of a different input (e.g. labour; L). This would mitigate the inefficiencies posed by double markup, and thus reduce the gains from the merger. Also note that this is an area where it might be useful to estimate the elasticity of substitution between inputs, e.g. by estimating a production function (why might it not be a good idea to adopt the Cobb-Douglas specification?).

4.2 Possible anticompetitive effects of vertical mergers

- Now let's go beyond the simple setting in which there is a single monopolist upstream and a single monopolist downstream. After all, this really is a special case.

4.2.1 Vertical merger to facilitate price discrimination

- Suppose the U firm is a monopolist who sells to lots of downstream firms. If the demand for U's product differs across the D firms, then the U firm would like to **price discriminate**, as this would lead to higher profits than if a single price were charged of everyone.

- In order to be able to price discriminate, the U firm must stop D firms from reselling the product (i.e. the D firms charged a low price sell the product to D firms that would be charged a high price). That is, U must stop arbitrage opportunities.

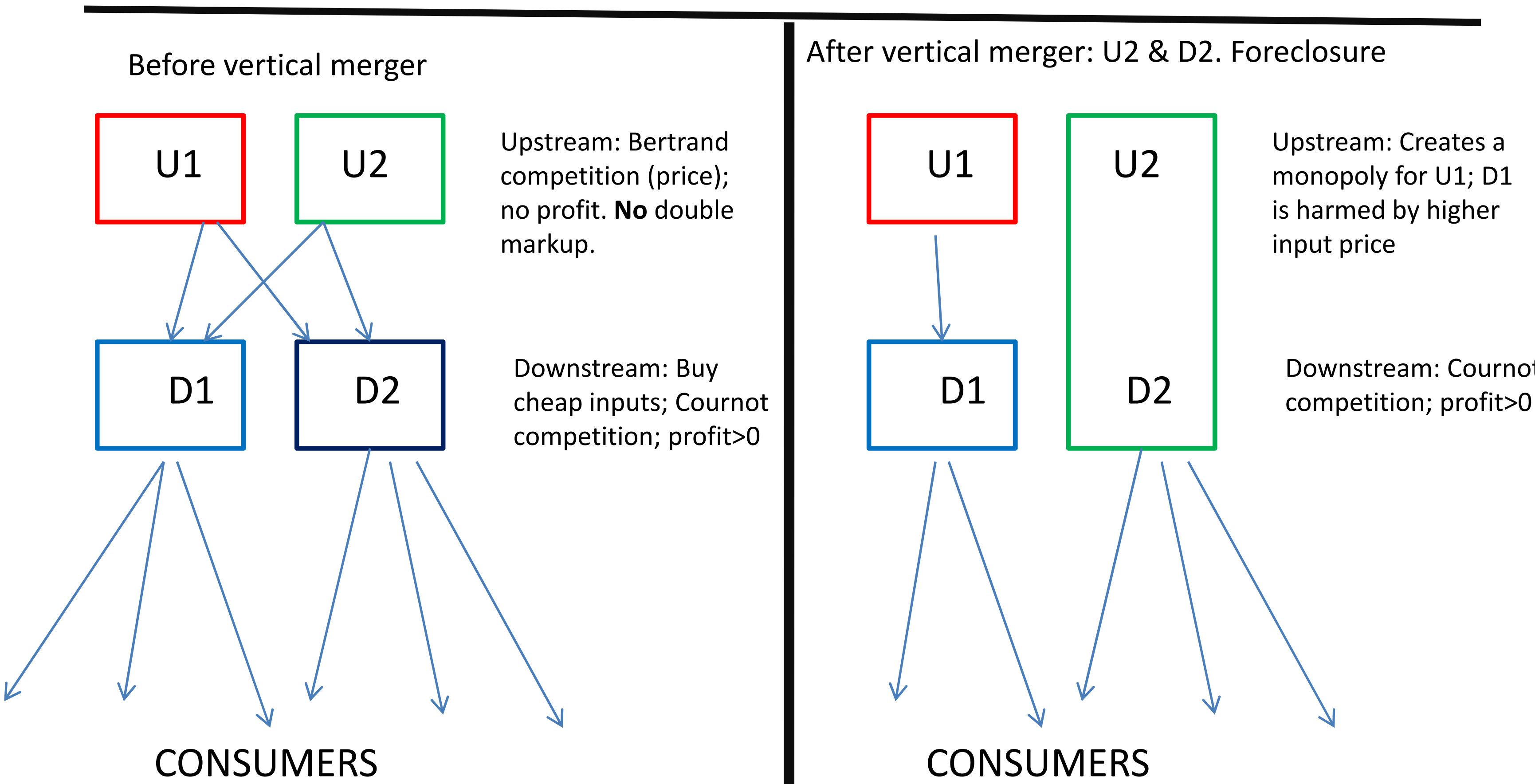
- One potentially effective way of preventing arbitrage is to merge with one or several of the D firms - these firms cannot then sell the product to other D firms as they are being controlled by the U firm.

- So a **reason** for wanting to merge might be to facilitate price discrimination. The effect on consumer welfare is not necessarily positive (it's ambiguous - double markup disappears within the new U+D firm, but prices go up amongst D firms not part of the merger).

[Foreclosure illustration here]

Vertical Merger is anti-competitive if it leads to foreclosure

Suppose the upstream and downstream markets are in fact oligopolies. What will the effect be of a vertical merger? Illustration:



- D1 less able to compete
- U2D2 will raise its price slightly

Next time we will model the effects of vertical mergers in an oligopoly setting more formally. After all, some important questions remain unanswered at this point, for example:

- Why will the integrated firm stop selling to independent downstream firms? Can it be shown that this is rational?

- The obvious response will be more vertical mergers - in the previous example, it would seem obvious that U1 and D1 should merge. How can we go about modelling this formally?

- The student presentation next time is based on the following paper:

Kim, E. Han and Vijay Singal, "Mergers and Market Power: Evidence from the Airline Industry," American Economic Review, 83 (1993), 549-569.

This paper looks at the effects of (horizontal) airline mergers in the 1980s in the US on airfares for domestic flights The results show that prices increased on routes served by the merging firms relative to a control group of routes unaffected by the merger. The authors argue that, while mergers may lead to more efficient operations, these are more than offset by exercise of increased market power.