ONLINE APPENDIX

This version: May 20, 2016

This online appendix accompanies the article “The Value of Collusion with Endogenous Capacity and Demand Uncertainty” by Johannes Paha.

A.I . Comments on Section III “The Model”

This section comments on the assumptions made to establish the model presented in Section III of the main text.

Costs and Capacity

The model assumes soft capacity constraints (see equation (5)). From a modeling point of view, this is convenient because it allows us to concentrate on Nash equilibria in pure strategies in the product market game. To see this, consider a situation with hard capacity constraints where demand exceeds a firm's production capacity. This would require to specify a rationing rule (Besanko and Doraszelski [2004]), which may result in a profit function that is not quasi-concave such that pure-strategy equilibria often fail to exist. This is different with soft capacity constraints, where it is not necessary to ration consumers. This results in profit functions that are continuous and quasi-concave. Hence, a unique pure strategy equilibrium exists (Maggi [1996]).

Collusive Conduct

In line with Fershtman and Pakes [2000] the model assumes Nash [1950, 1953] bargaining. The model was also implemented under the assumption of joint profit maximization yielding results that are numerically similar and qualitatively identical in both cases. The assumption of Nash bargaining may however be preferable because it solves certain issues related to price setting and the distribution of collusive profits. As was argued in the main text, it avoids ambiguity by concentrating on a scheme without side payments. It also distributes profits according to the *earnings follows output* principle (Bain [1948]) with firm *j*'s share of collusive outputs under Nash bargaining roughly equaling its capacity-share (i.e., $q_{k,j}/(q_{k,j}+q_{k,-j})\approx s_j/(s_j+s_{-j})$). This is because every colluding firm receives at least its competitive profit, and firms with a higher threat point gain more from collusion.

The assumption of a reversion to competitive prices $\boldsymbol{p}_c$ when at least one firm would want to deviate from collusion at Nash bargaining prices is attributed to the uniqueness of $\boldsymbol{p}_c$. In some cases, deviations could also be prevented if the cartel firms chose a set of prices below Nash bargaining prices but above competitive prices. Determining such a pricing vector is difficult for at least two reasons. First, there is a continuum of constrained semi-collusive equilibria at supra-competitive prices, and the firms have to coordinate on one of them. Especially in industries with asymmetric firms and differentiated products, this poses a selection problem not only to the researcher but also to the firms. Second, this set of prices would have to be calculated *within* the optimization process. This is a highly complex task. To see this, consider that this set of prices would depend on the deviation incentive of the firm, which is a function of its investment strategy, which, again, is a function of prices.

Therefore, assuming collusion at Nash bargaining prices and reversions to competitive prices provides unique pricing strategies that can be calculated independently from the policy functions and the value functions of the firms. This greatly simplifies the task of determining firms' optimal investment policies.

Incremental Value of Collusion

The model assumes the explicit costs of collusion to be zero. It does not specify an explicit rule either that would determine in what capacity and demand states the firms switch from competition to collusion. Alternatively, one might assume that collusion is established once the incremental value of both firms is above the present value of the expected costs of collusion. This section demonstrates that the assumptions that were made in the main text are advantageous when it comes to the computation of the policy and the value functions, and that these assumptions do not impair the main conclusions drawn from the paper.

Not specifying a rule for cartel formation allows calculating the competitive policy and value functions independently from their collusive counterparts. This is done using the algorithm proposed by Pakes and McGuire [1994] whose properties are well-known. Otherwise, if competing firms took into account that collusion will be established once they move in a certain state of capacities and demand, one would have to consider this prospect when determining their competitive investment policies. The collusive value function would have to be treated as an input when calculating the competitive value function, just as the competitive value function is an input when calculating the collusive one. This is because the colluding firms condition their investment policies on the possibility of collusion breaking down in some states, with competition being re-established thereafter. This mutual dependence of the competitive and the collusive policy and value functions would call for a new / modified version of the Pakes and McGuire [1994] algorithm that would have to be capable of determining all four functions at the same time. Such a technical innovation is left for future research because the main conclusion of the article is the same both with or without specifying a rule for cartel formation.

The article shows in line with established literature that, on the one hand, collusive strategies can only be sustained when the firms discount profits at a low rate and when their capacities are sufficiently symmetric. On the other hand and as a new result, a low value of the discount rate induces competing firms to go for a preemption race strategy that yields fairly asymmetric capacities. Hence, collusion can only be sustained when firms' discount rate is sufficiently low, but not so low that the firms would invest in fairly asymmetric capacities prior to the establishment of collusion. This challenges the common notion that a low value of the discount rate always facilitates collusion.

This conclusion can be made as long as a firm's investment policy in *competition* is independent of the costs of collusion. These costs may affect firms' investment policies in *collusion*. The firms would likely invest in lower collusive capacities because of the diminished return of collusive investments. As these costs are incurred only in collusion they do not affect firms' competitive investment strategies and, thus, they do not affect the main result of this article.

This conclusion would neither be challenged if one specified an explicit rule for cartel formation. In this case, the prospect of colluding might motivate the competing firms to choose an investment policy that generates symmetric capacities more frequently, where collusion can be sustained and will be established in the future. However, given the small firms' incentive to make higher investments in semi-collusion (Property 9) a dominant firm may still find it value-maximizing to defend its dominant position rather than entering a collusive agreement. Hence, firms with a low discount rate may still have an incentive to engage in preemption races and become dominant in competition. Endogenizing cartel formation would only have an effect on the exact values of δ and r where preemption races can be observed without invalidating the economic rationale presented in this article. Not specifying a rule for cartel formation thus keeps the model parsimonious and helps to establish its main results most transparently, while using algorithms whose properties are well-known.

A.II . Comments on Section IV "The Stage Game"

The main text states Property 1.

Property 1: $\Delta\pi_{c,j}/\Delta s_{-j}<0$

$$\Delta\pi_{c,j}/\Delta s_j \begin{cases} \geq 0 & \text{for} \quad s_j \leq s_{c,j}* \\ <0 & \text{for} \quad s_j > s_{c,j}* \end{cases}$$

To explain this property, when firm *-j* moves in a state with higher capacity s_{-j} it reduces its marginal costs c_{-j} and, thus, optimally lowers its price $p_{c,-j}$. This causes a business stealing effect which decreases the competitive profit of firm *j* (i.e., $\Delta\pi_{c,j}/\Delta s_{-j}<0$). The inverse u-shape of profits $\pi_{c,j}$ in firm *j*'s own capacity state s_j has been discussed in Besanko et al. [2010.2] and can also be shown in models of product market competition with homogeneous goods and hard capacity constraints (Kreps and Scheinkman [1983], Deneckere and Kovenock [1996], Allen et al. [2000]). Here, it results from the assumption of soft capacity constraints that have both a direct and an indirect impact on $\pi_{c,j}$. As to the direct impact, an expansion of s_j reduces firm *j*'s marginal costs c_j, resulting in lower equilibrium prices

$p_{c,j}$, and a higher output $\tilde{\tilde{q}}_{c,j}$. As to the indirect impact, setting lower prices $p_{c,j}$ causes business stealing from firm $-j$, lowering both the rival's output $\tilde{\tilde{q}}_{c,-j}$ and (by reversing the effect of decreasing returns to scale) also its marginal costs c_{-j}. This makes the rival $-j$ a fiercer competitor resulting in lower prices. For $s_j<s_{c,j}*$, the direct impact dominates the indirect impact and causes $\Delta\pi_{c,j}/\Delta s_j>0$. For $s_j>s_{c,j}*$, the indirect impact dominates the direct one implying $\Delta\pi_{c,j}/\Delta s_j<0$. The inverse u-shape vanishes when $s_{c,j}*$ exceeds the maximum number M of capacity states. This is the case, first, when business stealing is less pronounced (e.g., when demand d is high, or when products are more differentiated, i.e., lower θ), or, second, when changes in output $\tilde{\tilde{q}}_{c,-j}$ have a small impact on firm $-j$'s marginal costs (e.g., when its capacity s_{-j} is large, or when capacity constraints are soft, i.e., lower η). This is shown by (A24).

$$\begin{array}{lll} s_{c,j}*(s_{-j})<s_{c,j}*(s'_{-j}) & \text{if} & s_{-j}<s'_{-j} \\ s_{c,j}*(\theta)<s_{c,j}*(\theta') & \text{if} & \theta>\theta' \\ s_{c,j}*(\eta)<s_{c,j}*(\eta') & \text{if} & \eta>\eta' \\ s_{c,j}*(d)<s_{c,j}*(d+1) & & \end{array} \tag{A24}$$

The main text also states Property 2.

Property 2: $$\begin{array}{l} \Delta\pi_{k,j}/\Delta s_{-j}<0 \\ \Delta\pi_{k,j}/\Delta s_j \begin{cases} \geq 0 & \text{for} \quad s_j\leq s_{k,j}* \\ <0 & \text{for} \quad s_j>s_{k,j}* \end{cases} \\ s_{k,j}*>s_{c,j}* \end{array}$$

The collusive profits $\pi_{k,j}$ of firm j fall in s_{-j} because the higher capacity s_{-j} enables firm $-j$ to claim a larger share of aggregate profits and puts firm j at a bargaining disadvantage. Collusive profits hardly exhibit an inverse u-shape ($s_{k,j}*>s_{c,j}*$) for two reasons. First, because of the lower collusive output the firms are affected by diminishing returns to scale to a lesser extent than in competition: Costs that are already low cannot be lowered much more by building further capacity. Second, the collusive agreement implies that for given marginal costs the firms keep prices constant when capacity changes.

A.III. Comments on Section V(i) "Technical Details" – The Algorithms

Pseudocode 1 presents the implementation of the Pakes-McGuire [1994] algorithm used for the optimization in the competitive model. The main characteristics of the algorithm are the use of a Gauss-Jacobi scheme (14:) (i.e., the policy function in iteration l is calculated by use of information from the value and policy function obtained in iteration l-1 only). Moreover, the value function (15:) is updated using a policy iteration scheme (i.e., the algorithm iterates on the value function for (typically) 3 steps while using the same candidate policy function). The optimization is stopped when both sup norm distance measures *tolV* and *tolx* fall below a tolerance of *tol*=5e-8 (13:), i.e., when the modification in the entries of the value function and the policy function between any two iterations is very small.

The policy function $\boldsymbol{x0}_c$ is initialized (4:) by multiplying the profits $\pi_{c,j}(\boldsymbol{s},d)$ of firm j that correspond to the entries in the policy function by a pseudo-random number $\zeta(\boldsymbol{s},d,j)$ drawn from a uniform distribution in the interval [0,1]. The value function $\boldsymbol{V0}_c$ is initialized (7:) in a similar fashion by multiplying a naïve estimate of firm j's value (i.e., $\pi_{c,j}(\boldsymbol{s},d)\cdot(1+r)/r$) with a pseudo-random number.

A dampening scheme prevents the algorithm from visiting a sequence of policy and value functions all over again. Hence, the value and policy functions used in iteration l+1 are generated as weighted averages of the functions from iterations l and l-1 (25: and 26:). When using a fixed dampening factor λ, the circling behavior of the algorithm is not always avoided completely. Therefore, a new value of $\lambda \in [0,1]$ is drawn from a uniform distribution in every iteration. This dampening scheme is only applied in iterations where one of the distance measures, *tolV* or *tolx*, exceeds its respective value from the previous iteration (22:). Otherwise, a value of λ=1 (23:) is used. This helps the algorithm to stay on a convergence path while avoiding jumps away from the previous candidate solutions. Such jumps often indicate that the algorithm oscillates between different types of candidate solutions rather than converging to an equilibrium of the game. The occasional use of undampened updating (λ=1) results in a faster convergence than in the case of continuously employed dampening.

The randomness in the initialization of the value function $\boldsymbol{V0}_c$ and the policy function $\boldsymbol{x0}_c$ as well as the randomness in the choice of the dampening factor λ ensures that the algorithm generates a unique sequence of candidate solutions in every restart with otherwise identical industry parameter values. When a game has multiple equilibria, the algorithm possesses the ability to converge towards them. This multiplicity is explored by running the algorithm several times on the same set of parameter values.

Pseudocode 2 presents the implementation of the Pakes and McGuire algorithm [1994] used for the collusive model. Its main structure mimics that for the optimization of the competitive model. However, it is in some instances more complex. One needs to calculate (or load) the product market equilibria in competition, collusion and deviation periods (3: to 6:). The collusive policy and value functions $\boldsymbol{x0}_k$ and $\boldsymbol{V0}_k$ are initialized (8: to 14:) by multiplying each entry of their competitive counterparts by a number $\zeta(\boldsymbol{s},d,j)$ drawn pseudo-randomly from a uniform distribution in the interval [0.5,1.5]. The algorithm assesses in every iteration whether some firm would want to deviate from a collusive equilibrium at Nash bargaining prices and, if yes, set prices, quantities and profits at competitive, (i.e., price war) levels or at Nash bargaining levels otherwise (21: to 22:). Additionally (24:), if for some combination of $\boldsymbol{s}$ and d the algorithm finds $V0_k(\boldsymbol{s},d)<V1_c(\boldsymbol{s},d)$, i.e., collusion cannot be stabilized even at competitive prices, the collusive agreement would be terminated and the algorithm sets $V0_k(\boldsymbol{s},d)=V1_c(\boldsymbol{s},d)$. The stopping criterion of the algorithm (19:) is determined by the convergence of the collusive policy function and value function ($\boldsymbol{x0}_k$ and $\boldsymbol{V0}_k$) only. This is because the deviation policy function and value function ($\boldsymbol{x0}_d$ and $\boldsymbol{V0}_d$) closely resemble their competitive versions, so that convergence of $\boldsymbol{x0}_d$ and $\boldsymbol{V0}_d$ is achieved more quickly than convergence of $\boldsymbol{x0}_k$ and $\boldsymbol{V0}_k$.

The competitive and collusive equilibria of the dynamic game were calculated for 3,600 combinations of the discount rate $r \in \{0.01, 0.02, ..., 0.1\}$, the depreciation probability $\delta \in \{0.01, 0.02,$

..., 0.1}, the persistence of demand $\rho \in \{0.5, 0.7, 0.9\}$, the hardness of capacity constraints $\eta \in \{2.5, 10, 40\}$, the product differentiation parameter $\theta \in \{0.75, 0.85, 0.90, 0.95\}$, and the investment success α=0.0625. Selected parameter combinations were also evaluated for $\rho \in \{0.1, 0.3, 0.35, 0.40, 0.45\}$ and $\alpha \in$ {0.09375, 0.125, 0.25. 0.375, 0.5, 0.625, 6.25}.

Pseudocode 1: Competitive optimization

Initialization
 Calculate competitive equilibrium $\boldsymbol{p}_c(\boldsymbol{s},d)$, $\boldsymbol{q}_c(\boldsymbol{s},d)$, $\boldsymbol{\pi}_c(\boldsymbol{s},d)$ $\forall$ $\boldsymbol{s}$, d
 Policy and value function
  Policy function: $x0_{c,j}(\boldsymbol{s},d) = \zeta(\boldsymbol{s},d,j)\cdot \pi_{c,j}(\boldsymbol{s},d)$ $\forall$ $\boldsymbol{s}$, d, j
   $x1_{c,j}(\boldsymbol{s},d) = x0_{c,j}(\boldsymbol{s},d)$
  Value function: $V0_{c,j}(\boldsymbol{s},d) = \zeta(\boldsymbol{s},d,j)\cdot \pi_{c,j}(\boldsymbol{s},d)\cdot(1+r)/r$ $\forall$ $\boldsymbol{s}$, d, j
   $V1_{c,j}(\boldsymbol{s},d) = V0_{c,j}(\boldsymbol{s},d)$
  With values of $\zeta(\boldsymbol{s},d,j)$ being drawn pseudo-randomly from a uniform distribution $\in$ [0,1]
 Program controls
  $tolV = tolV\text{-}1 = tolx = tolx\text{-}1 = 20$
  $\lambda = 1$
Optimization
 while ($tolV > tol$=5e-8) and ($tolx > tol$==5e-8)
  Use $\boldsymbol{x0}_c$ and $\boldsymbol{V0}_c$ (Gauss-Jacobi scheme) to calculate optimal policy function $\boldsymbol{x1}_c$ according to equations (11) and (12)
  Update value function $\boldsymbol{V1}_c$ with values of $\boldsymbol{x1}_c$ and perform a policy iteration on $\boldsymbol{V1}_c$
  Update distance measures
   $tolV\text{-}1 = tolV$
   $tolx\text{-}1 = tolx$
   $tolV = \max \mid (\boldsymbol{V1}_c\text{-}\boldsymbol{V0}_c)/(1+|\boldsymbol{V0}_c|) \mid$
   $tolx = \max \mid (\boldsymbol{x1}_c\text{-}\boldsymbol{x0}_c)/(1+|\boldsymbol{x0}_c|) \mid$
  Determine dampening factor λ
   if ($tolV > tolV\text{-}1$) or ($tolx > tolx\text{-}1$),
   draw λ pseudo-randomly from a uniform distribution $\in$ [0,1]
   else, set $\lambda = 1$
  Update value and policy function with dampening
   $\boldsymbol{V0}_c = \boldsymbol{V1}_c$
   $\boldsymbol{x0}_c = \lambda\cdot\boldsymbol{x1}_c + (1\text{-}\lambda)\cdot\boldsymbol{x0}_c$
 end
 Return $\boldsymbol{V1}_c$ and $\boldsymbol{x1}_c$

Pseudocode 2: Collusive optimization

Initialization
  Product market equilibria
    Load competitive equilibrium $\boldsymbol{p}_c(\boldsymbol{s},d)$, $\boldsymbol{q}_c(\boldsymbol{s},d)$, $\boldsymbol{\pi}_c(\boldsymbol{s},d)$ $\forall$ $\boldsymbol{s}$, d, $\boldsymbol{V1}_c$ and $\boldsymbol{x1}_c$
    Calculate Nash bargaining equilibrium $\boldsymbol{p}_{NBS}(\boldsymbol{s},d)$, $\boldsymbol{q}_{NBS}(\boldsymbol{s},d)$, $\boldsymbol{\pi}_{NBS}(\boldsymbol{s},d)$ $\forall$ $\boldsymbol{s}$, d
    Calculate deviation equilibrium $\boldsymbol{p}_d(\boldsymbol{s},d)$, $\boldsymbol{q}_d(\boldsymbol{s},d)$, $\boldsymbol{\pi}_d(\boldsymbol{s},d)$ $\forall$ $\boldsymbol{s}$, d
    Initialize collusive equilibrium $\boldsymbol{p}_k(\boldsymbol{s},d)=\boldsymbol{p}_{NBS}(\boldsymbol{s},d)$, $\boldsymbol{q}_k(\boldsymbol{s},d)=\boldsymbol{q}_{NBS}(\boldsymbol{s},d)$, $\boldsymbol{\pi}_k(\boldsymbol{s},d)=\boldsymbol{\pi}_{NBS}(\boldsymbol{s},d)$ $\forall$ $\boldsymbol{s}$, d
  Policy and value function
    Policy function: $x0_{k,j}(\boldsymbol{s},d) = \zeta(\boldsymbol{s},d,j)\cdot x1_{c,j}(\boldsymbol{s},d)$ $\forall$ $\boldsymbol{s}$, d, j
      $x1_{k,j}(\boldsymbol{s},d) = x0_{k,j}(\boldsymbol{s},d)$
      $x1_{d,j}(\boldsymbol{s},d) = x0_{d,j}(\boldsymbol{s},d) = x0_{c,j}(\boldsymbol{s},d)$
    Value function: $V0_{k,j}(\boldsymbol{s},d) = \zeta(\boldsymbol{s},d,j)\cdot V1_{c,j}(\boldsymbol{s},d)$ $\forall$ $\boldsymbol{s}$, d, j
      $V1_{k,j}(\boldsymbol{s},d) = V0_{k,j}(\boldsymbol{s},d)$
      $V1_{d,j}(\boldsymbol{s},d) = V0_{d,j}(\boldsymbol{s},d) = V0_{c,j}(\boldsymbol{s},d)$
    With values of $\zeta(\boldsymbol{s},d,j)$ being drawn pseudo-randomly from a uniform distribution $\in$ [0.5,1.5]
  Program controls
    $tolV = tolV\text{-}1 = tolx = tolx\text{-}1 = 20$
    $\lambda = 1$
Optimization
  while ($tolV > tol$) and ($tolx > tol$)
    Price war assessment
      If for some combination of $\boldsymbol{s}$ and d a deviation is profitable for at least one firm, set $\boldsymbol{p}_k(\boldsymbol{s},d)=\boldsymbol{p}_c(\boldsymbol{s},d)$, $\boldsymbol{q}_k(\boldsymbol{s},d)=\boldsymbol{q}_c(\boldsymbol{s},d)$, $\boldsymbol{\pi}_k(\boldsymbol{s},d)=\boldsymbol{\pi}_c(\boldsymbol{s},d)$
      Otherwise set $\boldsymbol{p}_k(\boldsymbol{s},d)=\boldsymbol{p}_{NBS}(\boldsymbol{s},d)$, $\boldsymbol{q}_k(\boldsymbol{s},d)=\boldsymbol{q}_{NBS}(\boldsymbol{s},d)$, $\boldsymbol{\pi}_k(\boldsymbol{s},d)=\boldsymbol{\pi}_{NBS}(\boldsymbol{s},d)$
    Profitability assessment
      If for some combination of $\boldsymbol{s}$ and d collusion is not value-maximizing for at least one firm (i.e., $V0_k(\boldsymbol{s},d)<V1_c(\boldsymbol{s},d)$), set $V0_k(\boldsymbol{s},d)=V1_c(\boldsymbol{s},d)$
    Use $\boldsymbol{x0}_k$ and $\boldsymbol{V0}_k$ (Gauss-Jacobi scheme) to calculate optimal collusive policy function $\boldsymbol{x1}_k$
    Use $\boldsymbol{x0}_d$ and $\boldsymbol{V0}_d$ (Gauss-Jacobi scheme) to calculate optimal deviant policy function $\boldsymbol{x1}_d$
    Update value function $\boldsymbol{V1}_k$ with values of $\boldsymbol{x1}_k$ and perform a policy iteration on $\boldsymbol{V1}_k$
    Update value function $\boldsymbol{V1}_d$ with values of $\boldsymbol{x1}_d$ and perform a policy iteration on $\boldsymbol{V1}_d$
    Update distance measures
      $tolV\text{-}1 = tolV$
      $tolx\text{-}1 = tolx$
      $tolV = \max |(\boldsymbol{V1}_k - \boldsymbol{V0}_k)/(1+|\boldsymbol{V0}_k|)|$
      $tolx = \max |(\boldsymbol{x1}_k - \boldsymbol{x0}_k)/(1+|\boldsymbol{x0}_k|)|$
    Determine dampening factor λ
      if ($tolV > tolV\text{-}1$) or ($tolx > tolx\text{-}1$), draw λ pseudo-randomly from a uniform distribution $\in$ [0,1]
      else, set $\lambda = 1$
    Update value and policy function with dampening
      $\boldsymbol{V0}_k = \lambda\cdot\boldsymbol{V1}_k + (1-\lambda)\cdot\boldsymbol{V0}_k$ and $\boldsymbol{x0}_k = \lambda\cdot\boldsymbol{x1}_k + (1-\lambda)\cdot\boldsymbol{x0}_k$
      $\boldsymbol{V0}_d = \lambda\cdot\boldsymbol{V1}_d + (1-\lambda)\cdot\boldsymbol{V0}_d$ and $\boldsymbol{x0}_d = \lambda\cdot\boldsymbol{x1}_d + (1-\lambda)\cdot\boldsymbol{x0}_d$
  end
  Return $\boldsymbol{V1}_k$, $\boldsymbol{x1}_k$, $\boldsymbol{V1}_d$ and $\boldsymbol{x1}_d$

A.IV. Comments on Section V(i) "Technical Details" – Limiting Distributions

Given the policy function $x_j(\boldsymbol{s},d)$ of optimal investments the limiting distribution $prob(\boldsymbol{s},d)$ is simulated as follows. The algorithm chooses a random starting state $(\boldsymbol{s},d)$ and lets the industry evolve over 10^7 periods with firms investing in capacities according to their policy functions. Recording the frequency with which each state was visited provides the unconditional limiting distribution $prob(\boldsymbol{s},d)$.

Section V of the article shows that collusion cannot always be sustained despite the reversion to competitive pricing. In this case, the cartel is assumed to dissolve and there is a permanent transition to the competitive regime. This suggests that competition is an absorbing state and that – over infinitely many periods – the collusive limiting distribution would equal the competitive limiting distribution. To avoid this and identify the purely collusive limiting distribution the algorithm chooses a random starting state and lets the industry evolve until it reaches a state where collusion cannot be supported any more. Then, it draws a new state and repeats the simulation until 10^7 states have been evaluated.

A.V. Comments on Table III – The Number of Capacity States M

The assumption of an upper bound for capacity (M=6) impacts the economic interpretation of the results when the firms are fairly patient r<0.04 and engage in a preemption race strategy. As was explained in Section V, fairly symmetric firms may invest in capacities far beyond market demand until an asymmetry emerges that is sufficiently strong to establish one firm as dominant and the other as dominated. This strategy of overshooting investments is curtailed by the assumption of an upper bound for capacity (M=6). Hence, the firms invest in capacity until both hit the upper bound M. Evaluations of the model indicate that choosing a higher maximum number for capacity states M can solve this issue. However, this results in drastically increased computation times because of the curse of dimensionality.

To provide an example, with M=9 the effect of hitting the upper bound is less pronounced but still present. However, the computation time is approximately three times as high as in the case of M=6. This is especially relevant for parameter combinations with several equilibria where convergence may take several 10,000 iterations that – because of their consecutive nature – cannot be parallelized easily. Raising M to a level where the effect of the upper bound on M vanishes and repeating all calculations would require several months of CPU time. Additionally, the calculation would also be more burdensome for the simulation of the limiting distributions with 10^7 draws. Consequently, I continue with the assumption of M=6 and resort to two remedies. First, Table III reports no results for r=0.01 where the effect of hitting the upper bound of capacity is most pronounced. Second, I emphasize that the model may <u>over</u>state the stability of collusion when asymmetric capacity distributions result from preemption races. Especially when the discount rate is low (r<0.04) both firms may hit the upper bound of capacity, symmetric capacity states receive a high weight in the competitive limiting distribution. This symmetry stabilizes collusion. This effect vanishes for r≥0.04.

A.VI. Comments on Table III – Multiplicity of Equilibria

The numeric and, to some extent, stochastic nature of the algorithm causes some slight variation in the policy and value functions that are obtained in different runs of the algorithm. As a consequence, one requires a method for distinguishing this normal variation attributable to the numeric nature of the algorithm from the variation caused by the existence of multiple equilibria. This distinction is made using the sup norm distance measure (A25) that is calculated for every pair of policy functions (indexed by u and y) that are obtained in all runs of the algorithm for the same set of parameter values.

$$tolx_{u,y} = \max_{\mathbf{s},d} \left| \frac{x_u(\mathbf{s},d) - x_y(\mathbf{s},d)}{1 + |x_u(\mathbf{s},d)|} \right| \tag{A25}$$

An analogous measure $tolV_{u,y}$ is calculated for the value function.

Two solutions are considered (imperfect) representations of the same equilibrium if the values of the distance measures $tolx_{u,y}$ and $tolV_{u,y}$ are sufficiently small (i.e., differences in the values of the policy and value functions are attributed to the numeric nature of the search). When the distance measures are above the thresholds $tolx_{u,y}$≤5e-5 and $tolV_{u,y}$≤5e-6 the respective equilibria are defined to be distinct.

One unique equilibrium exists for each of the parameter combinations evaluated in the main text (i.e., δ=0.08 and r=0.02, δ=0.08 and r=0.08, as well as δ=0.02 and r=0.09). This is checked by restarting the algorithm 100 times for every combination of parameter values. The maximum tolerances between these solutions obtained in different restarts for the same set of parameters are as follows.

	max($tolx_{u,y}$)	max($tolV_{u,y}$)
δ=0.08, r=0.08, ρ=0.5	6e-7	1e-7
δ=0.02, r=0.09, ρ=0.5	6e-6	1e-6
δ=0.08, r=0.02, ρ=0.5	7e-7	9e-7

There is no evidence of a further collusive equilibrium than the one presented in the main text for δ=0.08, r=0.08, and ρ=0.5. This is also true for ρ=0.7 and ρ=0.9. Moreover, one does not find evidence of further collusive equilibria than the ones of type kA and kB presented for δ=0.02, r=0.09, and ρ=0.5. This is because the maximum distances of policy or value functions of the same type are very small. The two equilibria are distinct from each other because distances between equilibria of different types are no smaller than $tolV_{u,y}$=0.027 and $tolx_{u,y}$=0.24. The following distances are found between the collusive value or policy functions obtained in the 100 different restarts of the algorithm.

	max($tolx_{u,y}$)	max($tolV_{u,y}$)	
δ=0.08, r=0.08, ρ=0.5	3e-7	1e-6	
δ=0.02, r=0.09, ρ=0.5	2e-6	5e-7	eq. of same type kA or kB

A.VII . Robustness Checks – The Persistence Parameter ρ

The article presents results for a persistence of demand ρ=0.5. In the following, I show that the level of demand persistence has a small effect on the strategic patterns in the present model. The most visible effect of more persistent demand is the emergence of a greater number of equilibria.

This is shown based on a calculation of the competitive and collusive policy and value functions of the firms for all parameter combinations $r \in \{0.01, 0.02, ..., 0.1\}$, $\delta \in \{0.01, 0.02, ..., 0.1\}$, and $\rho \in \{0.5, 0.7, 0.9\}$ given a=4, b=0.1, θ=0.9, D=9, M=6, and η=10. The sample parameter combinations r=0.02 and δ=0.08, r=0.08 and δ=0.08, and r=0.09 and δ=0.02 that were presented and discussed in the main text are also evaluated for $\rho \in \{0.1, 0.3, 0.35, 0.4, 0.45\}$.

Tables IA and IIA show that the results presented in the article are robust to changes in ρ. The tables show the values of the asymmetry measure Δs_{c9} across the parameter space (spanned by δ and r) when assuming ρ=0.7 or ρ=0.9. The existence of multiple equilibria deteriorates the convergence properties of the algorithm and may result in lengthy computation times. Therefore, for ρ=0.7 and ρ=0.9, I do not analyze parameter combinations with δ<0.05 and r<0.06, where multiple equilibria are most likely. To detect multiple equilibria, the algorithm is run 10 times on every combination of parameters. Tables IA and IIA suggest the following findings:

1. Changes in the persistence of demand only have a small impact on the distribution of equilibrium types across the parameter space spanned by δ and r. One merely observes a faint effect that somewhat asymmetric capacity distributions (Δs_{c9}=1) spread out into the regions where quite asymmetric (Δs_{c9}>1) or symmetric (Δs_{c9}=0) capacities had been found previously.
2. A higher persistence of demand ρ results in a greater number of equilibria (typically with the same value of Δs_{c9}) for some combinations of parameter values. For example, for δ=0.08, r=0.08, one finds one competitive equilibrium for ρ=0.5 and one for ρ=0.7. However, one finds

two competitive equilibria for ρ=0.9 as is indicated by an asterisk. The longer time spent in every demand state allows the firms to choose from a greater variety of similar policies. There are two equilibria for ρ=0.9. The competitive limiting distributions of these equilibria are shown in Figure 1A.

3. Figure 1A shows that all four competitive equilibria are very similar. A higher persistence of demand only somewhat alters the characteristics of the equilibrium. The longer time that a firm expects to stay in any demand state makes the firms adjust their capacities more closely to the state of demand. However, these differences are minor. The collusive equilibria that are found for different values of ρ share an even greater degree of similarity than the competitive equilibria. Therefore, the incremental value of collusion is very similar for the evaluated values of ρ, as can be seen from Figure 2A.

These results also apply for $\rho \in \{0.1, 0.3, 0.35, 0.4, 0.45\}$. With these values of ρ it is more likely that demand moves in a different state d in period t+1 than to remain in the same state as in period t. Assuming such volatile demand does not affect the main characteristics of the competitive equilibria much. However, more volatile demand implies that capacities cannot be adjusted to current demand conditions well. Therefore, one finds somewhat higher capacities in demand state d=1 for ρ=0.1 than for ρ=0.5. Capacities are targeted to a greater degree at average demand. This also implies that in case of fairly asymmetric capacities the number of capacity states $\boldsymbol{s}$ increases where a drop in demand raises the incremental value of collusion. This is because competition in states of low demand is particularly intense when capacities are high.

$\delta - r$	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
0.01						1 0	1 0	1 0	0	0
0.02						1 0	1 0	1 0	0	0
0.03						1	1 0	1 0	1 0	0
0.04						1	1	1	1 0	0
0.05	>1	>1	>1	>1*	1	1	1	1	1	1 0
0.06	>1	>1	>1	>1	>1*	1	1	1	1	1
0.07	>1	>1	>1	>1*	>1* 1	1	1	1	1	1
0.08	>1	>1	>1	>1	>1	>1	1	1	1	1
0.09	>1	>1	>1	>1	>1	>1	1	1	1	1
0.1	>1	>1	>1	>1	>1	>1	1	1	1	1

Table IA: Values of $\Delta s_{c9}(r,\delta)$ *for* $\rho=0.7$

$\delta - r$	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
0.01						1 0	1 0	1 0	1 0	0*
0.02						1 0	1 0	1 0	1 0	0*
0.03						1	1	1 0	1 0	1 0
0.04						1*	1	1	1 0	1 0
0.05	>1*	>1	>1	>1	1*	1	1	1	1	1
0.06	>1	>1	>1	>1	>1	1	1*	1	1	1
0.07	>1	>1	>1	>1	>1	1	1*	1	1	1
0.08	>1	>1	>1	>1	>1	>1 1	1	1*	1	1
0.09	>1*	>1	>1	>1	>1	>1 1	1	1*	1	1
0.1	>1*	>1	>1	>1	>1	>1	>1 1	1	1	1

Table IIA: Values of $\Delta s_{c9}(r,\delta)$ *for* $\rho=0.9$

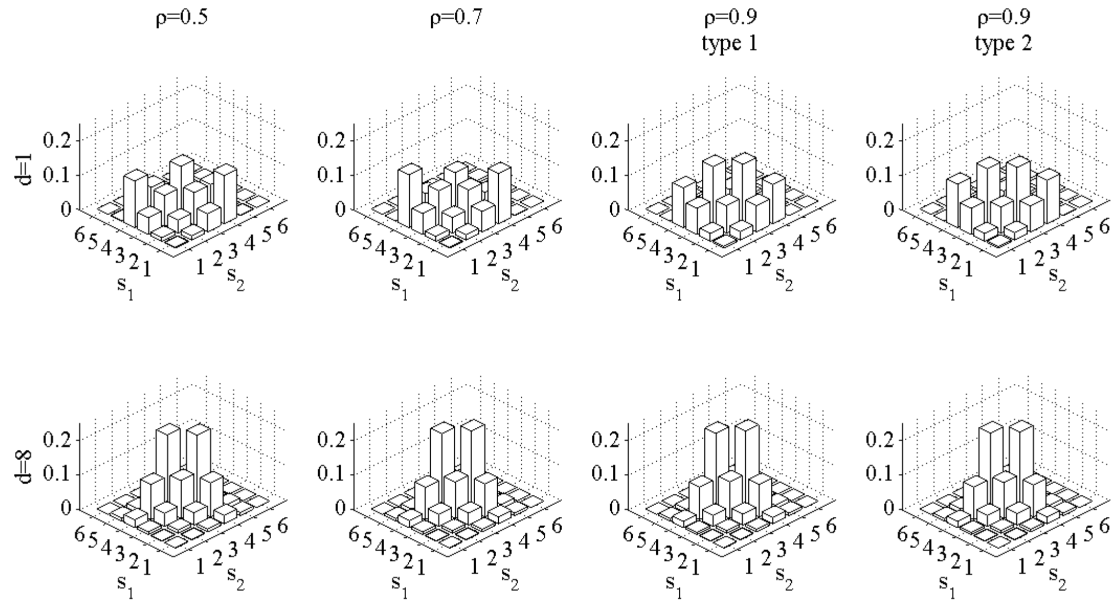

Figure 1A: Competitive limiting distributions with δ=0.08, r=0.08, and η=10

Notes: There are two equilibria for ρ=0.9 labeled type 1 and type 2.

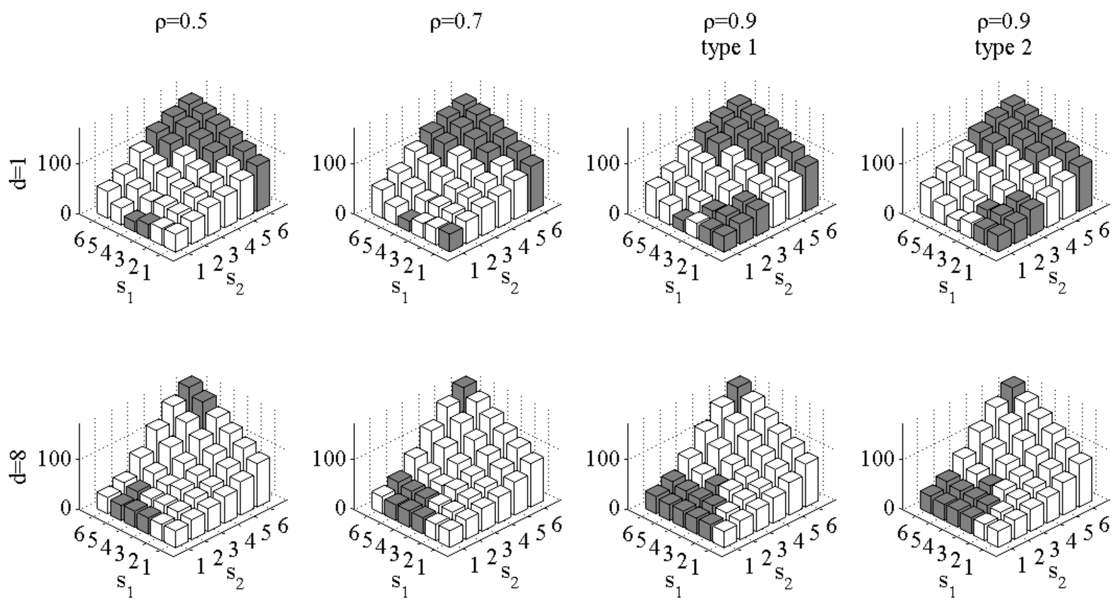

Figure 2A: Incremental value of collusion with δ=0.08, r=0.08, and η=10

A.VIII . Robustness Checks – The Investment Parameter α

The results of the model are only somewhat sensitive to changes in the value of the investment parameter α as can be seen when evaluating the model for $\alpha \in \{0.09375, 0.125, 0.25. 0.375, 0.5, 0.625, 6.25\}$. Given the parameterization of the model (and average investment expenditures per period of about x_j=5 monetary units) α=0.0625 implies an average time to build of about four periods per unit of capacity. With α=6.25 average investment expenditure of 5 monetary units results in an average construction period of about one period per unit of capacity. Therefore, an increase in α lowers the price of one unit of capacity. These lower investment costs induce the firms to choose higher and, thus, somewhat more symmetric capacities both in the competitive and in the collusive equilibria. However, the different competitive equilibrium types can still be distinguished. The choice of higher capacities is especially pronounced in collusion. Everything else equal, this lowers the incremental value of collusion. This is particularly true for a small firm while the incremental value of collusion of a dominant firm may even rise. This is because of the somewhat greater symmetry in competition, i.e., the low investment costs would make it difficult for this firm to retain its dominant position even in competition.

A.IX . Robustness Checks – The Product Differentiation Parameter θ

Besanko and Doraszelski [2004] and Besanko et al. [2010.1] show that greater product differentiation (i.e., a lower value of θ) makes firms more symmetric, i.e., $\Delta s_{c9}(\theta)<\Delta s_{c9}(\theta')$ if $\theta<\theta'$. The greater symmetry emerges because the softening of competition reduces the importance of capacities as a competitive advantage and, thus, eliminates preemption races. The observation of more symmetric capacities can be inferred from Figure 3A that shows competitive limiting distributions for θ=0.75 and compares to Figure 3 in the main text with θ=0.9. The greater symmetry of capacities also makes the incremental value of collusion of a small and a large firm more symmetric and prevents situations with

Ω_j<0. Therefore, by making the firms more symmetric greater product differentiation removes obstacles to collusion. Insofar this model also contributes to the theoretical literature challenging the common notion that collusion can most frequently be observed in markets for fairly homogeneous goods (e.g., Ross [1992]).

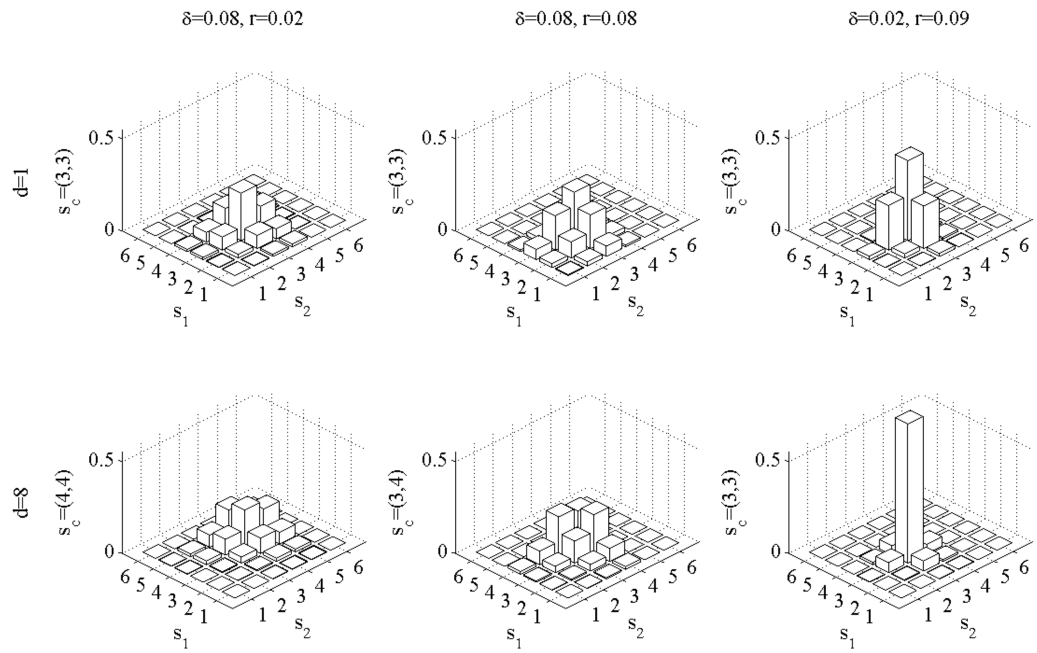

Figure 3A: Competitive, unconditional limiting distributions (ρ=0.5, η=10, θ=0.75)

A.X. Robustness Checks – The Hardness of Capacity Constraints η

The properties of the model were established under the assumption of close to hard capacity constraints (η=10). The same properties apply when assuming even harder capacity constraints (e.g., η=40). This is different when capacity constraints are fairly soft (e.g., η=2.5). This alters Property 8: When higher output can be produced without causing a (considerable) increase of marginal costs the firms decide to keep capacity more stable across demand states than would be the case with harder capacity constraints, i.e., $\Delta\bar{q}_{c,j,avg}(\eta)<\Delta\bar{q}_{c,j,avg}(\eta')$ if $\eta<\eta'$. At the same time, the intensification of competition in response to a decline in demand is relatively moderate, i.e., $\Delta\pi_{c,j,avg}(\eta)<\Delta\pi_{c,j,avg}(\eta')$ if $\eta<\eta'$. This changes Property 5: Without capacity constraints the additional profits from collusion fall along with demand in all capacity states (i.e., $\pi_{k,j}(\boldsymbol{s},d)-\pi_{c,j}(\boldsymbol{s},d)<\pi_{k,j}(\boldsymbol{s},d+1)-\pi_{c,j}(\boldsymbol{s},d+1)\ \forall\ \boldsymbol{s}$) as has been shown by Fabra [2006]. With very soft capacity constraints (e.g., η=2.5) one finds an increase only for very low capacities.

A similar finding applies for the incremental value of collusion Ω_j. This can be seen in Figure 4A showing the incremental value of collusion of firm 1 in the sample industry with δ=0.08, r=0.08, and ρ=0.5 for $\eta \in \{2.5, 10, 40\}$. With fairly soft capacity-constraints (η=2.5) the incremental value of collusion moves counter-cyclically only in few capacity states (see the white bars). This makes cartel formation in response to a decline in demand less likely and underlines the importance of capacity constraints for the observable patterns of cartel formation.

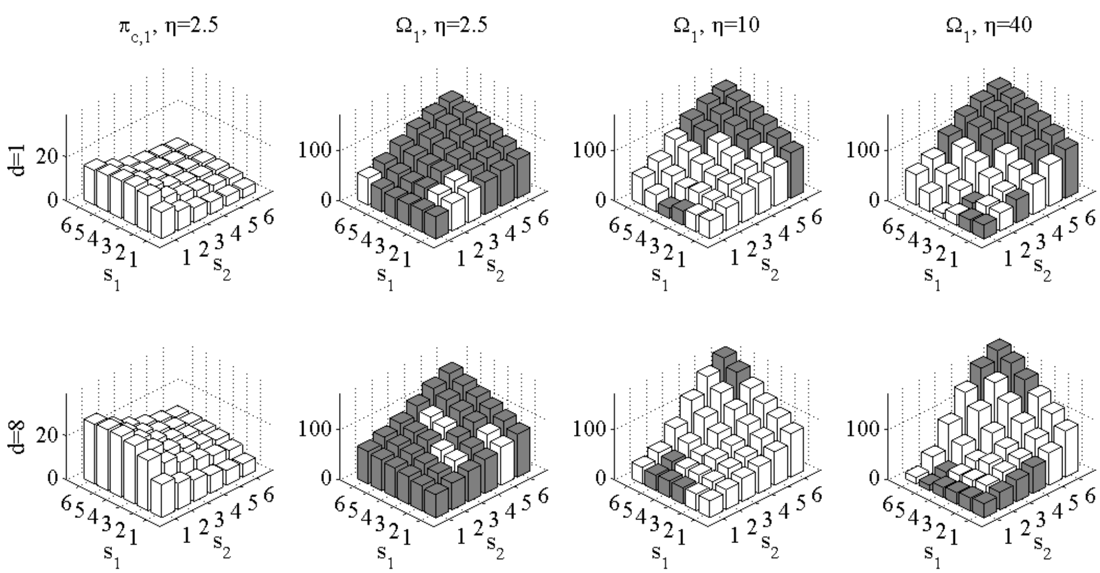

Figure 4A: Incremental value of collusion of firm 1 (Ω_1) with δ=0.08, r=0.08, and ρ=0.5

Notes: Gray bars indicate the capacity states where $\Omega_1(\boldsymbol{s},d)\leq\Omega_1(\boldsymbol{s},d+1)$ applies.

A.XI . Case Evidence

Both the structure and the results of the model are in line with case evidence. Case studies (e.g., Grout and Sonderegger [2005], Herold and Paha [2015]) suggest the hypothesis that changes in demand were causal for cartel formation in more than 50% of the cases prosecuted by the European Commission between 2001 and 2010. However, this evidence alone does not suggest a conclusive pattern. Some cartels were formed in times of rising demand while others were established in times of falling demand. Some demand shocks preceding cartel formation were quite persistent, others supposedly temporary. They occurred abruptly or slowly.

The inconclusive effect of demand on cartel formation suggests that demand affects the establishment of collusion through its interaction with other factors. The European Commission attributes a role to excess capacities in 13 of the 41 cases analyzed by Herold and Paha [2015] and reports an increase in the intensity of competition prior to the establishment of 28 of the 41 conspiracies. These are the factors that also raise the incremental value of collusion in the present model that helps to explain the puzzling effect of demand on cartel formation.

The model predicts that falling demand raises firms' incremental value of collusion if the lower demand softens capacity constraints (Property 5), which is particularly likely when capacities are durable (Property 12) and/or the decline in demand was sudden and perceptible (Property 13). This may explain certain features of, for example, the Norwegian cement cartel that was established in 1923 after a recession had contributed to the emergence of excess capacities (Steen and Sørgard [1999], Röller and Steen [2006]). Similar conditions were observed prior to the establishment of the more recent European steel beams conspiracy (DG Comp Case No COMP/F/38.907).

The predictions of the model depend on the assumptions of semi-collusion (i.e., coordination on prices but not on investments) and Nash bargaining. Semi-collusion has been reported, for example, for the Norwegian cement cartel (Steen and Sørgard [1999]), the steel beams conspiracy, as well as the conspiracies in nitrogenous fertilizer, synthetic fibers, plastics and aluminum (Davidson and Deneckere [1990]). The assumption of Nash [1950, 1953] bargaining allows avoiding ambiguity by concentrating on a scheme without side payments. Profits are distributed according to the *earnings follows output* principle (Bain [1948]), i.e., "each firm receives revenue only from the output it produces and sells itself." The numerical evaluations of the model show that firm j's share of collusive outputs under Nash bargaining roughly equals its capacity-share (i.e., $q_{k,j}/(q_{k,j}+q_{k,-j})\approx s_j/(s_j+s_{-j})$). One case where "the cartel [decided] to reward domestic market shares based on the members' share of total capacity" was the Norwegian cement cartel (Röller and Steen [2006]). They (*ibid.*) also suggest that this form of output-allocation caused the capacity-expansion observable in the Norwegian cement industry in the period 1956-1967, which is also predicted by the present model. Output allocation in the form of quotas could also be observed in the steel beams conspiracy.

The model also predicts that declining demand lowers firms' incremental value of collusion when the effect of capacity on the intensity of competition is small, for example, when the firms had already produced at excess capacity and thus had competed fiercely prior to the demand shock (Property 5). Then, the firms would rather establish collusion after a positive demand shock. These results may possibly explain why five firms in the animal-feed phosphates industry (DG Comp Case No COMP/38866), who had already produced at excess capacity, sat together in 1969 to share an *increase* in demand.

The model is in line with case evidence suggesting that cartels are often observed when firms had similar market shares in competition (Properties 10 and 11). One would conclude that symmetric firms have a higher tendency to self-select into collusion. Given that, first, symmetric capacities (i.e., symmetric market shares in competition) facilitate cartel formation and that, second, capacities in collusion remain symmetric, the model predicts that market shares will remain fairly stable over time as has been observed in many cartels (Harrington [2006]).

REFERENCES

Allen, B., Deneckere, R., Faith, T., and Kovenock, D., 2000, 'Capacity precommitment as a barrier to entry: A Bertrand-Edgeworth approach,' *Economic Theory*, 15, No. 3, pp. 501-530

Bain, J., 1948, 'Output Quotas in Imperfect Cartels,' *The Quarterly Journal of Economics*, 62, pp. 617-622

Besanko, D. and Doraszelski, U., 2004, 'Capacity Dynamics and Endogenous Asymmetries in Firm Size,' *The RAND Journal of Economics*, 35, pp. 23-49.

Online Appendix available at: http://goo.gl/z1x6CA

Besanko, D., Doraszelski, U., Lu, L. and Satterthwaite, M., 2010.2, 'Lumpy Capacity Investment and Disinvestment Dynamics,' *Operations Research*, 58, pp. 1178-1193

Deneckere, R. and Kovenock, D., 1996, 'Bertrand-Edgeworth duopoly with unit cost asymmetry,' *Economic Theory*, 8, pp. 1-25

Fershtman, C. and Pakes, A., 2000, 'A Dynamic Oligopoly with Collusion and Price Wars,' *The RAND Journal of Economics*, 31, pp. 207-236

Grout, P. and Sonderegger, S., 2005, 'Predicting Cartels. A Report Prepared for the Office of Fair Trading,' *OFT Economic Discussion Paper*, 773, http://goo.gl/xGjTPf.

Harrington, J., 2006, 'How do Cartels Operate?,' *Foundations and Trends in Microeconomics* (Now Publishers, Hanover)

Herold, D. and Paha, J., 2015, 'Predicting Cartel Formation,' *Work in progress*

Kreps, D. and Scheinkman, J., 1983, 'Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes,' *The Bell Journal of Economics*, 14, pp. 326-337

Maggi, G., 1996, 'Strategic Trade Policies with Endogenous Mode of Competition,' *The American Economic Review*, 86, pp. 237-258

Nash, J., 1950, 'The Bargaining Problem,' *Econometrica*, 18, pp.155-162

Nash, J., 1953, 'Two-Person Cooperative Games,' *Econometrica*, 21, pp. 128-140

Pakes, A. and McGuire, P., 1994, 'Computing Markov-Perfect Nash Equilibria: Numerical Implications of a Dynamic Differentiated Product Model,' *The RAND Journal of Economics*, 25, pp. 555-589

Ross, T., 1992, 'Cartel stability and product differentiation,' *International Journal of Industrial Organization*, 10, pp. 1-13

Röller, L.-H. and Steen, F., 2006, 'On the Workings of a Cartel: Evidence from the Norwegian Cement Industry,' *The American Economic Review*, 96, pp. 321-338

Steen, F. and Sørgard, L., 1999, 'Semicollusion in the Norwegian Cement Market,' *European Economic Review*, 43, pp. 1775-1796