Forward contracting and the endogenous activity of heterogeneous firms

Sébastien Mitraille* Henry Thille†

August 2, 2024

Abstract

Forward contracting in an N-firm quantity-setting oligopoly with heterogeneous costs introduces the possibility that relatively efficient firms deter the activity of inefficient rivals by reducing their margins. The equilibrium number of firms producing positive quantities can be any of $1, 2, ...N$ depending on the level of demand relative to firm-specific activity thresholds, with more firms active at higher demand levels. If only one firm is active, the Bertrand outcome is obtained. This potential reduction of the number of active firms may lessen the pro-competitive effect of forward sales, but does not eliminate it entirely. We explore the competition policy implications of the endogenous activity of firms, in particular for merger analysis.

JEL Classification: C72, D43, L13

Keywords: Forward contracting, oligopoly, cost heterogeneity, deterrence

Acknowledgement: We thank Kurt Annen, Simon Bélières, Chiara Canta, Zhiqi Chen, Debrah Meloso, and Geert van Moer for helpful comments, as well as participants in seminars at the Canadian Economics Association annual meeting, the Helsinki Graduate School of Economics, and the Toulouse Business School.

*TBS Business School, 20 Bd Lascrosses, 31000 Toulouse. E-mail: s.mitraille@tbs-education.fr

†Corresponding author. University of Guelph, Department of Economics and Finance, Guelph, Ontario, Canada.

1 Introduction

It is well established that the ability of firms to make sales though forward contracts in addition to selling spot may substantially affect the competitiveness of the market equilibrium. In the case of quantity competition, forward sales allow firms to preemptively capture a portion of future demand, which can lead to more aggressive competition. This pro-competitive effect of forward contracting has been analyzed for the most part with models of homogeneous firms, however, for many industries in which forward contracting is important firms operate with heterogeneous production costs, even if the products or services they sell are homogeneous. For example, non-ferrous metals markets have producers with heterogeneous costs (due to variations in the quality of ore deposits and countries in which the production costs are incurred) and centralized commodity exchanges, such as the London Metal Exchange, on which both forward and spot trading is conducted. Electricity generators are also characterized by heterogeneous costs (due to differing generation technologies) and both forward and spot sales opportunities (for example, day-ahead and hour-ahead contracts). Although the effect of forward sales has been extensively studied in the context of homogeneous firms, the more general case of heterogeneous firms has not seen much attention. By affecting the competitiveness of the equilibrium, forward sales with heterogeneous firms can affect the distribution of market shares and may also affect the number of firms that make positive sales in equilibrium. In particular, they raise the possibility that relatively efficient firms may use forward sales to squeeze the margins of less efficient rivals in order to reduce/eliminate their activity.[1]

We analyze the implications of firm heterogeneity on strategic forward contracting using a simple oligopoly model in which a set of firms that differ in their marginal cost can sell both in forward and spot markets. The number of active firms (those with positive sales) is endogenous as higher cost firms may find the equilibrium price at or below their marginal cost due to forward sales by their more efficient rivals. The possibility that forward trading affects the number of active firm leads to two potential types of equilibrium that differ depending on whether an inactive firm is "blockaded" or "deterred".[2] When all inactive firms are blockaded, the equilibrium is qualitatively similar to that in a homogeneous firm model with the same number of firms and average marginal cost, even though market shares are skewed towards relatively efficient firms. In a deterrence equilibrium, efficient firms non-cooperatively choose forward sales that result in price equal to the marginal cost of the deterred firm. It is possible that a deterrence equilibrium has only the lowest-cost firm active, in which case price is the same as in the heterogeneous firm Bertrand equilibrium.[3] In deterrence equilibria with more than one firm active there are a continuum of equilibria, however, they all have the same equilibrium price and aggregate sales, differing only with the allocation of the unique aggregate forward sales among the individual firms. We establish that for any distribution of marginal costs among firms, there exists a range of demand for which this results in a higher price than in a comparable homogeneous firm model. However, the equilibrium price remains below that in the Cournot equilibrium, so forward trading remains pro-competitive.

Our results have some bearing on competition policy. In particular, the presence of inactive firms can affect the analysis of mergers, and the redistribution of production across firms affects the relationships among concentration and production efficiency. Merger analysis is affected by

[1] Examples of this squeezing of margins include (Behar & Ritz (2017)), for OPEC against U.S. shale oil producers in 2014-2015, and Flamm & Reiss (1993), for Japanese semiconductor manufacturers against their U.S. rivals in the 1980s.

[2] We borrow the terms blockaded, deterred, and accommodated from the entry deterrence literature, however there is no entry in our model. Instead, we use these terms to refer to firms' "activity" in the market and the three terms coincide with price being respectively below, equal to, or above a firm's marginal cost.

[3] Allaz & Vila (1993) obtain equivalence to the Bertrand outcome in their model in the limit as the number of trading periods increases. In our model, it can be obtained with only one period of forward trading.

forward contracting as the profitability of a merger and its effect on both price and industry average cost depend significantly on the nature of both the pre- and post-merger equilibria. In the absence of synergies, and despite forward trading increasing competition, a merger between two firms always increases price. A merger may not result in a reduction of the number of active firms, as it may induce the substitution of a higher-cost inactive firm for a lower-cost merging firm. Regarding the relationship between concentration and production efficiency, forward trading reallocates market towards relatively efficient firms, so HHI is increased and industry average cost is decreased by forward trading.[4] Miller & Podwol (2020) find similar effects in their model with a fixed number of active firms. We present an example indicating that these effects are not significantly affected by the potential for forward trading to cause a reduction in the number of active firms.

Finally, we extend a duopoly version of our model to examine the impact of additional periods of forward trading. This exercise adds heterogeneous firms to the analysis of Allaz & Vila (1993), where, in the limit as the number of forward trading periods tends to infinity, the equilibrium approaches that of the Bertrand game with price equal to marginal cost. We find that the Bertrand outcome of price equal to the marginal cost of the higher cost firm is obtained with a finite number of trading periods. Therefore, the limiting result of Allaz & Vila (1993) is obtained with fewer periods of forward trading, although full efficiency is not obtained.[5]

There has been substantial interest in the strategic effects of forward sales since Allaz & Vila (1993) demonstrated that increasing the number of forward trading opportunities causes an increasingly competitive outcome. This pro-competitive effect of forward trading was questioned in a number of subsequent articles. Liski & Montero (2006) and Mahenc & Salanié (2004) demonstrate that competition can instead be reduced by forward trading, whereas Adilov (2012) shows that if firms choose their production capacity prior to the forward contracting decisions, the Cournot equilibrium re-emerges as the unique outcome. Breitmoser (2012) shows that if forward purchases of inputs are also possible, the equilibrium is contained in a range bounded by those in the Cournot, Stackelberg, and Allaz and Vila models. Mitraille & Thille (2020) show that the pro-competitive effects of forward trading may be reduced if demand is uncertain as firms may reduce forward sales in order to avoid over-committing in low demand states. Strategic contracting has been of particular interest in the analysis of electricity markets, where forward contracting plays a particularly significant role. Holmberg & Willems (2015) and Bushnell (2007) are examples of extensions of the Allaz & Vila (1993) framework to relevant aspects of trading in electricity markets.

Empirical evidence supporting the strategic role played by forward contracting has been found for a couple of energy industries. Wolfram (1999) examined market power following deregulation in the British electricity industry, finding that prices were above marginal cost, but not so much as suggested by the Cournot model. Forward contracting and the threat of entry were found to partially explain the low margins. The contracting in the British electricity market examined by Wolfram (1999) however was exogenously imposed by regulators, unlike the endogenous contracting in the theoretical literature. van Eijkel et al. (2016) examined the Dutch wholesale market for natural gas in which contracting was endogenous, finding evidence supportive of strategic contracting in addition to any risk-hedging effects.

Not much attention has been paid to role of firm heterogeneity in this literature, even though it is commonly a characteristic of markets for homogeneous commodities. Exceptions to this are de Frutos & Fabra (2012), Brown & Eckert (2017), and Miller & Podwol (2020), however the

[4] Salant & Shaffer (1999) show that increases in marginal cost heterogeneity among firms leads to an increased HHI and decreased total production cost under static Cournot competition. Our results demonstrate that forward sales exacerbate this effect.

[5] Forward trading cannot induce full efficiency in our model as that would require price dropping to the marginal cost of the lower cost firm.

nature of cost heterogeneity in their models does not allow the endogenous activity of higher cost firms as it does in ours.[6] Miller & Podwol (2020) is perhaps the closest of these to the current paper. They also explore the implications of firm heterogeneity on how forward contracting affects the relationship between concentration and market power as well as the welfare effects of mergers. We explore how endogenous firm activity affects some of these conclusions.

The endogenous activity of firms in our model is reminiscent of the literature examining entry deterrence with multiple incumbents (Gilbert & Vives (1986); Vives (1988)). Fixed costs play a central role in these models and generate discontinuous best-response functions which may lead to multiple equilibria of either accommodated or deterred entry. The latter involves a limit quantity chosen in equilibrium that deters the entrant by lowering its operating profit below its entry cost. In contrast, we have no fixed/entry costs in our model (and so have continuous best-responses) yet we also get a limit equilibrium in which price is maintained at a level that reduces the margin of a higher-cost firm to zero.[7] Another feature that distinguishes our model from an entry-deterrence model is that all firms may be active in both the forward and spot markets: there is no incumbent-entrant distinction, the only asymmetry among firms is in their marginal cost of production. A motivation for this distinction is seen in deregulated electricity markets. These generally have markets operating at many periods over the course of a day and given the substantial demand variation by time of day, not all generators operate at all times. Higher cost generators may only produce in peak hours, but they are available to potentially operate in non-peak hours as well. A high-cost generator may decide to operate in any given hour depending on the price relative to its marginal generation cost, as it already has the generation plant installed to produce electricity for this market.

We present our model in the next section and follow that with a simple duopoly version of the model that illustrates most of the forces at work. Section four presents the equilibrium for the full oligopoly game. The implications of heterogeneity for the competitive effects of forward trading are examined in section five and the implications for concentration, production efficiency, and market power are examined in section six. Some results regarding merger analysis are presented in section seven, and the duopoly example is extended to multiple forward trading periods in section eight. Section nine concludes.

2 Model

There are N producers each producing a quantity q_i at a cost of $C_i(q_i){=}c_iq_i$, $i{=}1, ..., N$.[8] Zero fixed costs are important as fixed costs would provide a reason for high-cost firms to choose zero production that is in addition to the effect on which we wish to focus: high-cost firms choosing zero production in response to a non-positive price-cost margin instead of a non-positive profit.[9] We index firms in increasing order of their marginal cost of production with firm one being the

[6] de Frutos & Fabra (2012) consider heterogeneous firms in an electricity market context, however forward commitments are exogenously chosen by a regulator in their model. Brown & Eckert (2017) and Miller & Podwol (2020) examine mergers among heterogeneous firms with forward contracting. However, firms differ only in the slopes of their increasing marginal cost functions, so price always exceeds firms' marginal cost and all firms are active in equilibrium.

[7] Hence, our "deterrence" of activity is reminiscent of the concern in competition policy with regard to a "price squeeze", or "margin squeeze", in which a vertically integrated firm that supplies an input to a downstream rival charges a high price in order to reduce the rival's margin by raising its costs (see Rey & Tirole (2007)). In contrast, forward sales squeeze the margin of a higher-cost rival by reducing the equilibrium price.

[8] This differs from Miller & Podwol (2020), who use $C_i(q_i){=}cq_i{+}q_i^2/2k_i$, where k_i is the level of capital stock for the firm. The key difference is that the linear component, c, is common across firms, which results in all firms being active in their model.

[9] In this sense our model differs from those in the entry deterrence literature, in which an entrant must pay a positive fixed/entry cost in order to enter, allowing incumbent firms to deter entry. See Gilbert & Vives (1986), Vives (1988), Etro (2006, 2008).

most efficient:[10] $0 \le c_1 < c_2 < ... < c_N$. Let $\bar{c}_N = \frac{1}{N} \sum_{i=1}^{N} c_i$ denote the average marginal cost across all firms, and $\bar{c}_k = \frac{1}{k} \sum_{i=1}^{k} c_i$ for $k < N$ denote the average cost of the k lowest-cost firms. The vector of marginal costs, $\boldsymbol{c} = (c_1, c_2, ..., c_N)$ is known to all firms.

Producers' sales on the forward and spot markets determine their total production. In the forward market each producer $i = 1, 2, ...N$ chooses a quantity x_i to sell to competitive agents at a price p_F. We define aggregate forward sales as $X = \sum_{i=1}^{N} x_i$, and the aggregate sales of i's rivals as $X_{-i} = \sum_{j \neq i} x_j$. The vector of forward sales, $(x_1, x_2, ..., x_N)$, and the forward price, p_F, are observable to all firms at the beginning of the second period.[11] In the second period spot market each producer may choose to produce an additional quantity y_i to sell at a price p_S, with $Y = \sum_{i=1}^{N} y_i$ and $Y_{-i} = \sum_{j \neq i} y_j$. A firm's total production is the sum of its forward and spot sales,[12] $q_i = x_i + y_i$. As marginal cost differs across firms, we must allow for the possibility that firms with high marginal cost may find it unprofitable to sell positive quantities in equilibrium. We say a firm is *active* in the spot market if $y_i > 0$ and *inactive* if $y_i = 0$. A firm is active on the forward market if $x_i > 0$ and inactive if $x_i = 0$.

A firm's profit consists of revenue from both forward and spot sales less total production cost:[13]

$$\pi_i = p_F x_i + p_S y_i - c_i (x_i + y_i). \tag{1}$$

Much of the strategic forward trading literature writes profit as a function of forward sales, x_i, and production, q_i, instead of x_i and y_i as we do in (1). We prefer to explicitly write profit in terms of the two types of sales firms make as a firm's activity in the spot market ($y_i > 0$ versus $y_i = 0$) is central to the analysis. The results would be identical if we instead expressed profit as the equivalent $\pi_i = p_F x_i + p_S (q_i - x_i) - c_i q_i$, as is done in much of the literature.

There is a single demand for the product by consumers represented by the linear inverse demand function $P(Q) = a - Q$. Given total forward sales determined in the first period, X, inverse demand in the second period spot market is then

$$p_S = P(X + Y) = a - X - Y. \tag{2}$$

Demand for forward sales is derived from the equilibrium spot price, p_S. The per-unit expected profit to a forward purchase is $p_S - p_F$, which is driven to zero by the competitive agents purchasing in the forward market.[14] Given perfect foresight among market participants, the inverse demand for forward sales, $p_F(X)$, will equal the equilibrium spot price that obtains when aggregate forward sales are X.

A strategy for a firm consists of its choice of forward sales, x_i, and spot sales, y_i in each sub-game. Sub-games differ with respect to the vector of forward sales for each firm. However, as we see from (1) and (2), the profit from spot sales depends on forward sales only through

[10] Although the analysis could be done allowing more than one firm to have the same marginal cost, the presentation of the results is substantially simplified by imposing that firms have distinct marginal costs.

[11] The assumption that firms observe the entire vector of forward sales is stronger than necessary. See Ferreira (2006) for an extended treatment. All that is required is that firm i knows x_i and X, and as X can be inferred from p_F, at a minimum firms only need to know their own forward sales and the forward price in our model.

[12] We assume that forward contracts imply delivery of the contracted quantities, which is a common approach taken in much of the strategic forward trading literature. An alternative interpretation is that the contracts are purely financial and are settled by a transfer in cash determined by the difference between the forward and spot prices (for example, Mahenc & Salanié (2004) and de Frutos & Fabra (2012) take this interpretation). Whether the contract settlement is through physical or cash settlement does not affect the results.

[13] This profit highlights the difference between forward sales and advance *production* (Saloner (1987), Pal (1991, 1996)), in which firms commit to a minimum level of sales. However, in those models committed sales are sold on the spot market at the spot price, so profit from committed production is not determined in the first period, unlike in the forward sales models.

[14] This is just a requirement that there be no arbitrage opportunities and is the standard condition in strategic forward trading models. See Ito & Reguant (2016) for an example in which arbitrage is imperfect.

their effect on the spot price. As the spot price depends only on aggregate forward sales, the distribution of forward sales across individual firms does not affect a firm's choice of spot sales. Consequently, we write $y_i(X)$ to denote firm i's spot sales strategy. The payoff for firm $i{=}1,...,N$ given strategies $\{x_j, y_j()\}$, $j{=}1,...,N$, is then

$$\pi_i = \left(a - X - \sum_{j=1}^{N} y_j(X)\right)(x_i + y_i(X)) - c_i\,(x_i + y_i(X))\,. \tag{3}$$

We analyze the sub-game perfect equilibrium of this game, in which we first find equilibrium spot sales, $y_i^*(X)$, for a given X, and then use this to determine equilibrium forward sales, x_i^*.

3 A duopoly example

Before presenting the equilibrium for the oligopoly model, we examine a duopoly version of it in order to illustrate the underlying forces driving the results. This represents an extension of Allaz & Vila (1993) (with one period of forward trading) to the case of heterogeneous firms. To simplify the exposition we set $c_1{=}0$ and $c_2{=}c{>}0$.

Spot market sub-games are differentiated by the forward sales, (x_1, x_2), and the forward price, p_F, determined in the first period. A firm's profit in a sub-game is $\pi_i{=}(p_F{-}c_i)x_i{+}(a{-}X{-}Y{-}c_i)y_i$, where the first term represents the profit from forward sales, which is sunk when a firm chooses its spot sales. As the firm only considers its spot market profit when choosing spot sales, sub-games mimic a Cournot game with demand reduced by aggregate forward sales, X. If both firms are active, the spot market equilibrium price and quantities are simply their Cournot values with the demand intercept given by $a{-}X$:

$$p_S(X) = \frac{a-X+c}{3}, \quad y_1(X) = \frac{a-X+c}{3}, \quad y_2(X) = \frac{a-X-2c}{3}. \tag{4}$$

However, this is only an equilibrium if $p_S(X){>}c$ (or, equivalently, $y_2(X){>}0$), which requires $X{<}a{-}2c$. For $X{\geq}a{-}2c$, firm two is not active and the monopoly outcome obtains: $p_S(X){=}y_1(X){=}(a{-}X)/2$. In summary, the equilibrium in a spot market sub-game with aggregate forward sales, X, has spot price and sales of

$$(p_S^*(X), y_1^*(X), y_2^*(X)) = \begin{cases} \left(\frac{a-X+c}{3}, \frac{a-X+c}{3}, \frac{a-X-2c}{3}\right), & \text{if } X{<}a{-}2c, \\ \left(\frac{a-X}{2}, \frac{a-X}{2}, 0\right), & \text{if } X{\geq}a{-}2c. \end{cases} \tag{5}$$

The spot price and spot sales are continuous functions of the aggregate forward position with a kink at $X{=}a{-}2c{\equiv}X_2^d$, where X_2^d is a threshold level of forwards sales for which firm two is active if $X{<}X_2^d$ and inactive if $X{\geq}X_2^d$. Clearly if $a{-}2c{<}0$ firm two is not active for any $X{\geq}0$, so we will assume $a{-}2c{>}0$ for the rest of this section.

Turning now to the choice of forward positions, using $p_F(X){=}p_S^*(X)$, firm one's profit reduces to

$$\pi_1(x_1, x_2) = p_S^*(X)(y_1^*(X) + x_1), \tag{6}$$

which is continuous given the continuity of $p_S^*(X)$ and $y_1^*(X)$, but is kinked at $X{=}a{-}2c$, or, from firm ones's point of view, at $x_1{=}a{-}2c{-}x_2$. Marginal profit for firm one is then given by

$$\frac{\partial \pi_1(x_1, x_2)}{\partial x_1} = \begin{cases} \frac{a-x_1-x_2+c}{9} - \frac{x_1}{3} & \text{if } x_1{<}a{-}2c{-}x_2 \\ \frac{-x_1}{2} & \text{if } x_1{>}a{-}2c{-}x_2. \end{cases} \tag{7}$$

Examining the limits to the left and right of $x_1{=}a{-}2c{-}x_2$, it is straightforward to show that[15]

$$\lim_{x_1\to(a-2c-x_2)^-} = -\frac{a-3c-x_2}{3} > -\frac{a-2c-x_2}{2} = \lim_{x_1\to(a-2c-x_2)^+}, \tag{8}$$

so marginal profit for firm one jumps downward when x_1 is increased past the threshold at which firm two becomes inactive. Notice that, under the assumption that $a{-}2c{>}0$, marginal profit for firm one is negative for $x_1{>}a{-}2c{-}x_2$, so the maximum forward sales that firm one would consider is the quantity that just deters the activity of firm two. Similarly, firm two's profits when choosing forward sales are $\pi_2(x_1,x_2){=}(p_S^*(X){-}c)(y_1^*(X){+}x_1)$ and marginal profit (for $x_2{<}a{-}2c{-}x_1$) is[16]

$$\frac{\partial\pi_2(x_1,x_2)}{\partial x_2} = \frac{a-x_1-x_2-2c}{9} - \frac{x_2}{3} \tag{9}$$

We can now describe the nature of the equilibrium choice of forward sales, determining the conditions under which firm two is active. For $a{\leq}2c$, firm two will be inactive for any choice of x_1 in which case firm one's optimal choice is the monopoly choice of $x_1{=}0$, i.e., $\partial\pi_1(0,0)/\partial x_1{=}\partial\pi_2(0,0)/\partial x_2{=}0$. Firm two's activity is blockaded in this case, with equilibrium forward sales of $x_1^*{=}x_2^*{=}0$ and price of $p_F^*{=}a/2$. For $a{>}2c$, firm two's marginal profit is positive when firm one chooses zero forward sales, $\partial\pi_2(0,0)/\partial x_2{=}(a{-}2c)/9{>}0$, so firm two will choose positive forward sales if firm one sets $x_1{=}0$. Firm one will choose the level of forward sales that deters the activity of firm two, $x_1{=}a{-}2c$, if its marginal profit jumps down from a positive value to a negative value at $x_1{=}a{-}2c$. This occurs if the lower limit in (8) is non-negative with $x_2{=}0$, or $a{\leq}3c$. The equilibrium forward sales in this case are $x_1^*{=}a{-}2c$ and $x_2^*{=}0$ with the equilibrium price $p_F{=}c$. Finally, if $a{>}3c$, firm one's marginal profit is negative at the level of forward sales that deters the activity of firm two, in which case firm one accommodates the activity of firm two. In this case, setting the first line in (7) and (9) equal to zero yields the equilibrium forward sales of $x_1^*{=}(a{+}2c)/5$ and $x_2^*{=}(a{-}3c)/5$ with the equilibrium price $p_F^*{=}(a{+}2c)/5$. We have demonstrated

Proposition 1. *In the duopoly forward trading game with $c_1{=}0$ and $c_2{=}c$, the activity of firm two is accommodated if $a{>}3c$, deterred if $2c{\leq}a{\leq}3c$, and blockaded if $a{<}2c$. The resulting forward price is*

$$p_F^* = \begin{cases} \frac{a}{2} & \text{if } a<2c, \\ c & \text{if } 2c\leq a\leq 3c, \\ \frac{a+2c}{5} & \text{if } a>3c. \end{cases} \tag{10}$$

Firm heterogeneity influences the incentives for firms to sell forward compared to those in Allaz & Vila (1993). For low levels of demand, firms do not make forward sales, which coincides with Allaz & Vila (1993) only for $N{=}1$ in their model. A monopolist does not make forward sales in the Allaz and Vila model as the only reason to do so is influence a rival firm as a Stackelberg leader does. We get the same outcome for low levels of demand where firm two is blockaded: firm one need not influence firm two's sales in this case. However, for intermediate demand levels, $2c{<}a{\leq}3c$, only firm one is active, but it must sell forward to deter the activity of firm two. So activity deterrence represents an additional motive for a monopolist to make forward sales. For high levels of demand, firm two is accommodated, and each firm sells forward for the same reason firms do in the Allaz & Vila (1993) duopoly. So both the "Stackelberg leadership" motive and the deterrence motive for selling forward are present in our model.

[15] The inequality in (8) holds as long as $a{>}x_2$, which is clearly true for any reasonably chosen x_2.

[16] Firm two's marginal profit also has a downward jump discontinuity at $x_2{=}a{-}2c{-}x_1$, but (9) is negative at this point, so this has no bearing on the equilibrium.

To compare the price level in our model to Allaz & Vila (1993), we need to determine an appropriate common marginal cost. For $c_1{=}c_2{=}c_H$, Allaz & Vila (1993) show that the equilibrium price is $p_F^H{=}(a{+}4c_H)/5$. This coincides with the price in an accommodation equilibrium if $c_H{=}c/2$, where each homogeneous firm has a marginal cost equal to the mean of those in our duopoly example ($\bar{c}_2{=}c/2$). For $a{\geq}3c$, cost heterogeneity does not affect price in the forward trading equilibrium, as long as the average marginal cost is held constant. However, $p_F^*{=}c{>}p_F^H$ for $a{\in}(2c,3c)$, so cost heterogeneity lessens the pro-competitive effect of forward trading.

Can this effect be so large as to eliminate the pro-competitive effect entirely? In the absence of forward trading, firm two is active in a Cournot equilibrium for $a{\geq}2c$, and since the Cournot price, $p^C{=}(a{+}c)/3$, is higher than firm two's marginal cost, forward trading remains pro-competitive when deterrence occurs: $p^C{>}c$ for $a{\in}(2c,3c)$. Firm one deters the activity of firm two by keeping the price at the latter's marginal cost, which is lower than the Cournot price. This establishes the following Corollary to Proposition 1:

Corollary 1. *Price is equal to that in the Allaz and Vila model for $a{\geq}3c$ and higher than that in the Allaz and Vila model for $2c{<}a{<}3c$. Forward trading remains pro-competitive as price is lower than that in the Cournot game, so $p^C{>}p_F^*{\geq}p_F^H$ for $a{>}2c$.*

Finally, it is notable that the deterrence equilibrium outcome is identical to that in a heterogeneous firm Bertrand duopoly, where the low cost firm prices at the marginal cost of the high cost firm. In Allaz & Vila (1993), the Bertrand outcome is obtained only in the limit as the number of forward trading periods becomes large. For some demand and cost conditions, forward trading obtains the Bertrand outcome with just one forward trading period when firms are heterogeneous.

4 Equilibrium in the oligopoly model

We now determine the equilibrium in the oligopoly game for $N{>}2$. As in the duopoly game of the previous section, the nature of the equilibrium depends on how the marginal payoff of relatively efficient firms behaves around the thresholds at which relatively inefficient firms become active/inactive. We begin with the spot market equilibria for the N-firm oligopoly.

4.1 Equilibrium in spot market sub-games

Just as in the duopoly model of Section 3, spot market sub-games are simply Cournot equilibria with inverse demand $p_S{=}(a{-}X){-}Y$. The set of firms that are active in the spot market depends on aggregate forward sales as higher-cost firms will not sell if X is sufficiently large. Consequently, the number of firms active on the spot market, which we denote $n(X){\leq}N$, will be determined by comparing the level of forward sales, X, to thresholds that determine whether a given firm is active in the spot market.

In each sub-game firm's spot sales are the solution to

$$\max_{y_i}\{(a{-}(X{+}Y){-}c_i)y_i{+}(p_F{-}c_i)x_i\} \qquad i{=}1,...,N, \tag{11}$$

with marginal profit

$$\frac{\partial\pi_i}{\partial y_i}{=}a{-}X{-}Y_{-i}{-}c_i{-}2y_i. \tag{12}$$

Suppose that all N firms are active, the equilibrium individual spot sales would be the same as in the heterogeneous-firm Cournot game with inverse demand intercept of $a{-}X$:

$$y_i^*{=}\frac{(a{-}X){+}\sum_{j=1}^{N-1}c_j{-}Nc_i}{N{+}1}{=}\frac{a{-}X{+}N\bar{c}_N}{N{+}1}{-}c_i \qquad i{=}1,2,...,N, \tag{13}$$

and

$$p_S^* = \frac{a - X + \sum_{j=1}^{N} c_j}{N+1} = \frac{a - X + N\bar{c}_N}{N+1}. \tag{14}$$

For all N firms to be active, price must exceed the marginal cost of the least efficient firm, which requires $a - X > (N+1)c_N - N\bar{c}_N$. For lower values of $a - X$, firm N is not active. Similarly, if a sub-game equilibrium has $k < N$ firms active, we have

$$y_i^* = \frac{a - X + \sum_{j=1}^{k} c_j - (k+1)c_i}{k+1} = \frac{a - X + k\bar{c}_k}{k+1} - c_i \qquad i = 1, 2, ..., k, \tag{15}$$

and

$$p_S^* = \frac{a - X + \sum_{j=1}^{k} c_j}{k+1} = \frac{a - X + k\bar{c}_k}{k+1}. \tag{16}$$

Price must exceed c_k for k to be active, or $a - X > (k+1)c_k - k\bar{c}_k$. However, for this to be an equilibrium, it must also be the case that the price is lower than firm $k+1$'s marginal cost, i.e., $a - X < (k+2)c_{k+1} - (k+1)\bar{c}_{k+1}$. Consequently, we can define threshold levels of aggregate forward sales that determine whether a firm is active or inactive on the spot market:

Definition 1. *Firm k is active on the spot market if $X < X_k^d$ and inactive on the spot market if $X \geq X_k^d$, where*

$$X_k^d = a + k\bar{c}_k - (k+1)c_k. \tag{17}$$

It is straightforward to show that the X_k^d form a decreasing sequence: $a - c_1 = X_1^d > X_2^d > X_3^d > ... > X_N^d$.[17,18] The number of firms active on the spot market when aggregate forward sales are X is then given by the number of firms for which $X < X_k^d$. In summary, we have

Proposition 2. *In a sub-game with aggregate forward sales of X, the number of active firms is given by*

$$n(X) = \sum_{j=1}^{N} \mathbb{1}_{[X < X_j^d]}, \tag{18}$$

equilibrium spot sales strategies are the continuous functions

$$y_i^*(X) = \max\left[\frac{a - X + n(X)\bar{c}_{n(X)}}{n(X)+1} - c_i, 0\right] \qquad \forall i, \tag{19}$$

and the spot market equilibrium price is the continuous function

$$p_S^*(X) = \frac{a - X + n(X)\bar{c}_{n(X)}}{n(X)+1}. \tag{20}$$

Proof. See Appendix A.

As the spot market sub-game with $X = 0$ is simply the Cournot game with heterogeneous firms, Proposition 2 also provides the Cournot equilibrium in this case, where a firm will be active in the Cournot equilibrium if $X_k^d > 0$. As each X_k^d varies with a and $\boldsymbol{c}$, the number of firms active in the Cournot equilibrium will also not necessarily equal N. We will be comparing the outcome of the forward sales game to that of the Cournot game below, so, for completeness, we state the Cournot equilibrium as a Corollary to Proposition 2:

Corollary 2. *The equilibrium in the Cournot game has n_C active firms, with $n_C = n(0) \leq N$, individual output $y_i^C = y_i^*(0)$, $i = 1, 2, ...N$, and price $p^C = p_S^*(0)$.*

[17] To see this, express $X_k^d > X_{k+1}^d$ as $c_{k+1} + (k+1)(c_{k+1} - c_k) > (k+1)\bar{c}_{k+1} - k\bar{c}_k$. As $(k+1)\bar{c}_{k+1} - k\bar{c}_k = c_{k+1}$, this reduces to $c_{k+1} > c_k$ which is true by our assumption regarding firms' marginal costs.

[18] Applying this definition to the duopoly model with $c_1 = 0$ and $c_2 = c$, we have $X_2^d = a + 2(0+c)/2 - 3c = a - 2c$, which is the same threshold as determined in Section 3.

4.2 Equilibrium forward sales

In equilibrium, the forward price is equal to the spot price, so

$$p_F^* = p_S^*(X) = a - \left(X + \sum_{i=1}^{n(X)} y_i^*(X) \right), \tag{21}$$

which represents the inverse demand faced by producers for their forward sales. Given p_F^* and the spot market equilibrium described in Proposition 2, the profit faced by a firm when choosing its forward sales in the first period is

$$\pi_i(x_i, X_{-i}) = (p_S^*(x_i + X_{-i}) - c_i)(y_i^*(x_i + X_{-i}) + x_i). \tag{22}$$

Both the spot sales strategies, $y_i^*(X)$, and the spot price, $p_S^*(X)$, are continuous, but kinked functions of aggregate forward sales. Consequently, the firm's profit is a continuous, kinked function of its forward sales, with marginal profit discontinuous where the level of forward sales induces a higher-cost firm to become active/inactive, i.e., where $x_i + X_{-i} = X_k^d$. Using Proposition 2 this profit is given by

$$\pi_i(x_i, X_{-i}) = \left(\frac{a - X + n(X)\bar{c}_{n(X)}}{n(X)+1} - c_i \right) \left(x_i + \max\left[0, \frac{a - X + n(X)\bar{c}_{n(X)}}{n(X)+1} - c_i \right] \right), \tag{23}$$

with $X = x_i + X_{-i}$. There are two sources of kinks in firm i's profit, and hence discontinuities in i's marginal profit: first, when firm i is indifferent between being active or inactive in the spot market, and, second, when firm $j > i$ is indifferent between being active or inactive in the spot market. The latter is due to the discontinuous nature of $n(X)$ as the number of firms changes when the thresholds, X_j^d, of less efficient firms are reached. The first term in (23) is simply $p_S^*(X) - c_i$, so if firm i is inactive on the spot market, ($p_S^*(X) < c_i$), it is also inactive on the forward market as profit is negative for any positive level of forward sales. This result means that we can ignore the kink at firm X_i^d (due to the max operator in (23)) when determining the firm's optimal forward sales since they will not be positive when the firm is not active on the spot market.

Marginal profit for active firms is then

$$\frac{\partial \pi_i(x_i, X_{-i})}{\partial x_i} = \left(\frac{a - X + n(X)\bar{c}_{n(X)}}{n(X)+1} - c_i \right) \frac{n(X)-1}{n(X)+1} - \frac{x_i}{n(X)+1}, \qquad \forall i \leq n(X), \tag{24}$$

which is discontinuous at $x_i = X_k^d - X_{-i}$ for each $k > i$ as there is a discrete change in $n(X)$ at these points.[19] Clearly, marginal profit for firm i is linear and decreasing apart from the points where $x_i + X_{-i} = X_k^d$, for $k > i$. From Proposition 2 we know that firm k will be active if $X < X_k^d$, or $x_i < X_k^d - X_{-i}$.

Lemma 1 in Appendix A establishes that marginal profit only jumps downward at these discontinuities, which implies that the firm's profit maximizing response to X_{-i} will be one of two qualitatively different choices: either i) it is optimal to choose x_i as a zero of (24), essentially ignoring the existence of firm $n(X)+1$, or ii) it is optimal to choose $x_i = X_{n(X)+1}^d - X_{-i}$ at a discontinuity in marginal profit, where the firm deters the activity of firm $n(X)+1$. For the former, setting (24) equal to zero yields

$$r_{i,n(X)}(X_{-i}) = \frac{n(X)-1}{2n(X)} \left(a + n(X)\bar{c}_n(X) - (n(X)+1)c_i - X_{-i} \right), \tag{25}$$

[19] Recall that $n(X)$ is piece-wise constant, which is why there are no $n'(X)$ terms in (24).

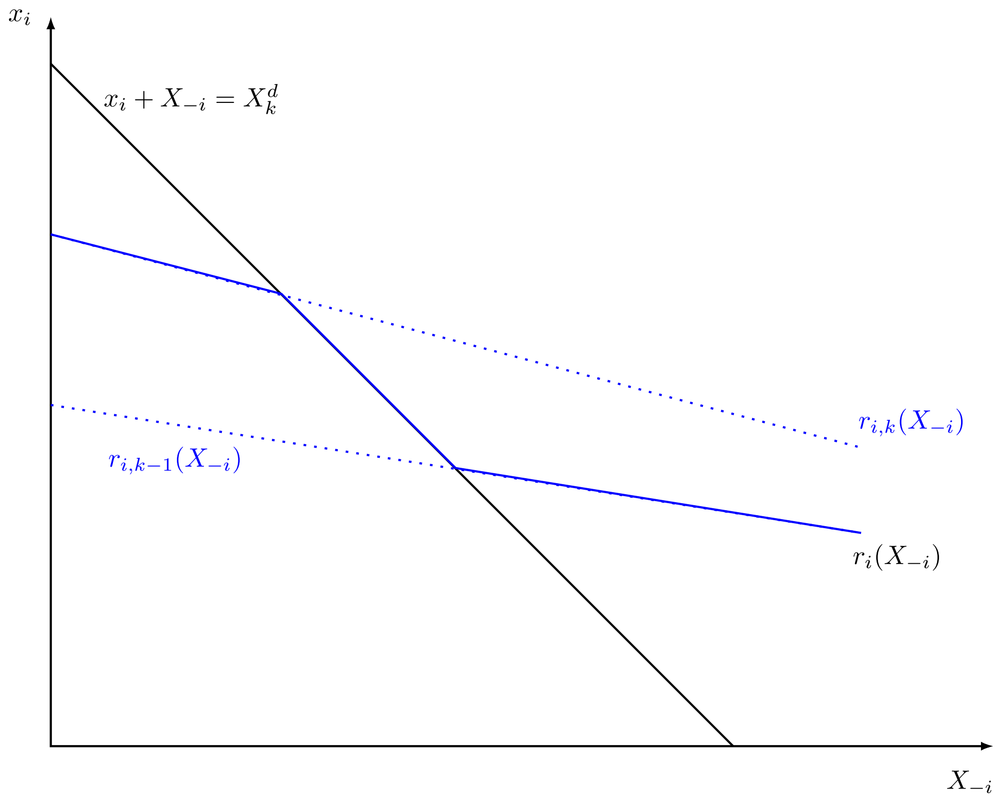

Figure 1: Best response $r_i(X_-i)$ with $k>i$. $r_i^k(X_-i)$ is the best-response when there are k firms active, and $x_{i,k}^d(X_{-i})$ is the threshold value for x_i that causes firm k to be inactive.

where the notation $r_{i,k}(X_{-i})$ for $k>i$, indicates the best-response for firm i when there are k active firms and the deterrence of firm $k+1$ is not optimal,[20] i.e. $r_{i,k}(X_{-i})+X_{-i} \geq X_{k+1}^d$. Consequently, a firm's best-response will either be of the form $r_{i,k}(X_{-i})$ or $X_k^d - X_{-i}$ depending on X_{-i}. Lemma 2 in Appendix A establishes the precise nature of the best-response functions, here we illustrate the best-response for firm $i<k$ in Figure 1. $r_{i,k}(X_{-i})$ is i's best-response when it results in $x_i < X_k^d - X_{-i}$ and $r_{i,k-1}(X_{-i})$ is i's best-response when it results in $x_i > X_k^d - X_{-i}$. For values of X_{-i} that result in $r_{i,k-1}(X_{-i}) \leq X_k^d - X_{-i} \leq r_{i,k}(X_{-i})$, the best-response is $X_k^d - X_{-i}$. The overall best-response, $r_i(X_{-i})$, is the solid line connecting these three segments. A similar pattern occurs for firm i's best-response function as it crosses each threshold, $X_k^d - X_{-i}$ for each $k=i+1, ..., N$.

It is interesting to contrast this best-response function with those found in the literature on entry deterrence with multiple incumbents,[21] which exhibit a similar effect when a firm finds it optimal to set a limit output. For example, Gilbert & Vives (1986) show that best-response functions are discontinuous due to the presence of fixed entry costs that must be incurred by potential entrants. This leads to multiple equilibria with the coexistence of a non-deterrence equilibrium with a continuum of deterrence equilibria. Since we do not have fixed costs, best-response functions in our model are continuous, so we get a unique *type* of equilibrium, either involving non-deterrence or deterrence, but not both simultaneously.

The best-response functions in our model are continuous and bounded, which establishes the existence of a pure-strategy equilibrium, however, the equilibrium is not necessarily unique. To see this note that, if there is a deterrence equilibrium with, say, k active firms deterring firm $k+1$,

[20] Since the discontinuities in marginal profit coincide with those of $n(X)$, the number of firms is not affected by local variations in x_i in this case.

[21] Gilbert & Vives (1986); Vives (1988)

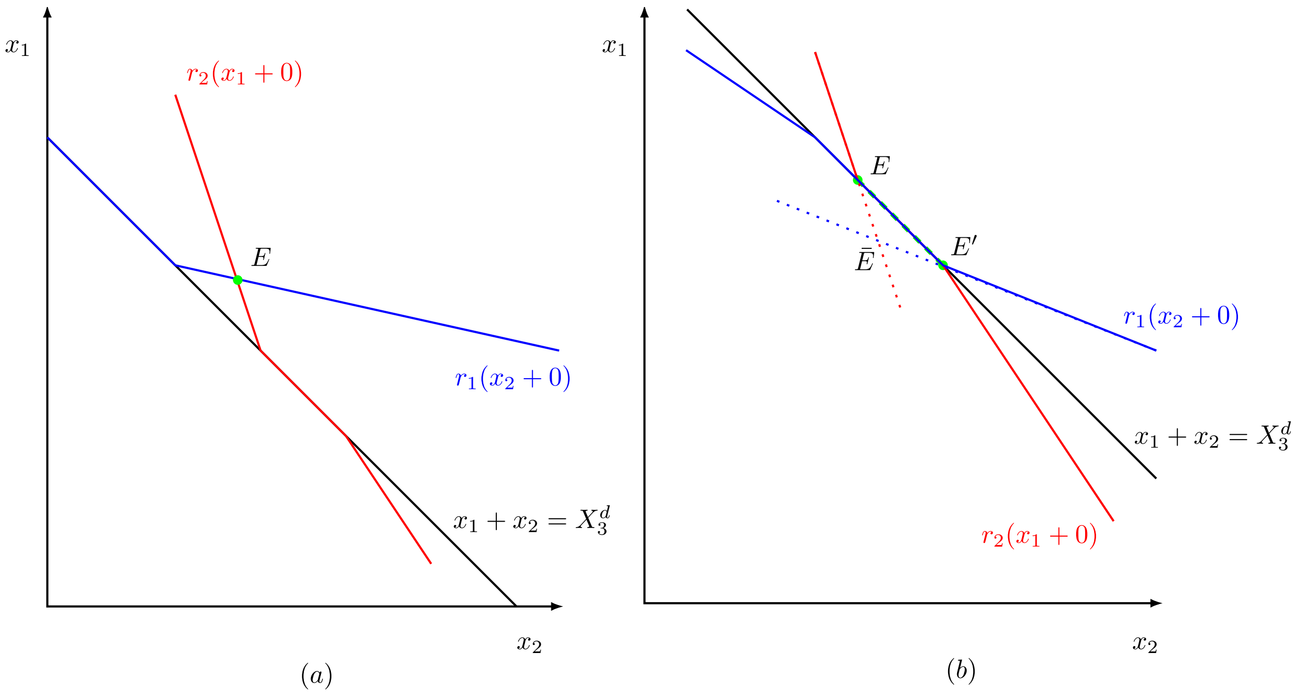

Figure 2: Best responses, three potential firms with x_3=0. (a) non-deterrence equilibrium, (b) deterrence equilibria.

each firm must be playing the best-response that deters firm k+1, i.e., $x_i(X_{-i})=X^d_{k+1}-X_{-i}$. So firms playing mutual best-responses results in the condition $\sum_{i=1}^{k} x_i^*=X^d_{k+1}$ from which we cannot determine a unique $(x_1^*,...,x_k^*)$. Of course aggregate sales are uniquely determined in such an equilibrium (at X^d_{k+1}), as is price (at c_{k+1}). The best-response functions do impose some limits on the individual forward sales since it is required that for each i=1, ..., k, X^*_{-i} is in the range of values for which the firm wishes to choose the deterrence level of sales, but this yields a continuum of possible equilibrium sales vectors, each resulting in the same aggregate forward sales and price. This range of possible deterrence equilibria does not appear in the duopoly model of Section 3 as there is only a single firm active in the deterrence equilibrium in that case.

We can illustrate the nature of the alternative types of equilibria graphically by focusing on the case with three potential firms and examine the types of situations that give rise to the duopoly outcome as an equilibrium, i.e. x_3^*=0. Figure 2 plots the best-response functions for firms one and for zero sales by firm three, i.e., $r_1(x_2+0)$ and $r_2(x_1+0)$. The two panels in the figure illustrate the two possible situations in which a duopoly emerges in equilibrium. Panel (a) depicts the situation in which the equilibrium, E, has the mutual best-response of firms one and two above X_3^d. In this case, a unique equilibrium occurs at point E with $x_1^*+x_2^*>X_3^d$. Panel (b) of Figure 2, illustrates the other possibility. The mutual best-responses occur on the line segment EE' where there are multiple equilibria, each with $x_1^*+x_2^*=X_3^d$. The point labelled $\bar{E}$ represents the equilibrium that would occur if firm three did not exist, with lower aggregate output than in any of the equilibria along EE'.

The proof of the following proposition formally establishes these results and provides the conditions on the model parameters under which each type of equilibrium obtains.

Proposition 3. *A pure-strategy, sub-game perfect Nash equilibrium exists with a unique forward price and aggregate forward sales. Furthermore, there is a sequence of thresholds,* $\alpha_1<\alpha_1^d<\alpha_2<\alpha_2^d<...<\alpha_{N-1}<\alpha_{N-1}^d<\alpha_N<\alpha_N^d\equiv\infty$, *for which the equilibrium number of active firms is* $n^*=\sum_{j=1}^{N}\mathbb{1}_{[a>\alpha_k]}$ *and the equilibrium is either* *i) a non-deterrence equilibrium for* $a \in (\alpha_{n^*}, \alpha^d_{n^*}]$ *with,*

$$X^* = \frac{n^*(n^*-1)}{n^{*2}+1}(a - \bar{c}_{n^*}) \quad \textit{and} \quad p^*_F = \frac{a + n^{*2}\bar{c}_{n^*}}{n^{*2}+1},$$

or

ii) a deterrence equilibrium for $a \in (\alpha^d_{n^*}, \alpha_{n^*+1}]$ *with,*

$$X^* = X^d_{n^*+1} \quad \textit{and} \quad p^*_F = c_{n^*+1}.$$

Proof: See Appendix A.

The demand thresholds[22] in Proposition 3 are of two types. First, the *activity thresholds*, $\alpha_1 < \alpha_2 < ... < \alpha_N$ determine whether a firm is active, so $a \in (\alpha_n, \alpha_{n+1}]$ means that there are n active firms[23] in equilibrium. The *deterrence thresholds* further subdivide these intervals, determining the nature of the equilibrium: there are n firms active who do not deter firm $n+1$ if $a \in (\alpha_n, \alpha^d_n]$ (firm $n+1$ is blockaded), and n firms active who deter the activity of firm $n+1$ if $a \in (\alpha^d_n, \alpha_{n+1}]$ (firm $n+1$ is deterred).

The activity and deterrence thresholds in Proposition 3 are derived in the proof from conditions on aggregate forward sales. However, we can provide an alternative, intuitive argument based on the equilibrium price. A firm will be active in equilibrium if the price exceeds its marginal cost so, given our ranking of marginal costs, it is clear that if firm $j > i$ is active, then so is firm i. Let $p^{nd}_F(n) = (a + n^2\bar{c}_n)/(n^2+1)$ denote the expression for price in a non-deterrence equilibrium with n active firms. For all N firms to be active it must be that the equilibrium price with N firms active, $p^{nd}_F(N)$, exceeds the marginal cost of all firms, in particular that of the least efficient firm, N. Firm N's activity threshold follows directly from $p^{nd}_F(N) > c_N$:

$$a > c_N + N^2(c_N - \bar{c}_N) \equiv \alpha_N. \tag{26}$$

Similarly, for an equilibrium to have $n < N$ active firms and $N-n$ inactive firms requires $p^{nd}_F(n) > c_n$ and $p^{nd}_F(n+1) \leq c_{n+1}$, i.e., the marginally active firm earns a positive margin, while the marginally inactive firm would not face a positive margin were it to be active. Again, we can express these two conditions in terms of the activity thresholds of firms n and $n+1$:

$$\alpha_n \equiv c_n + n^2(c_n - \bar{c}_n) < a \leq c_{n+1} + (n+1)^2(c_{n+1} - \bar{c}_{n+1}) \equiv \alpha_{n+1}. \tag{27}$$

In deriving the α_n's in this way, we implicitly assume that the equilibrium is of the non-deterrence type when a is slightly larger than α_n since we use $p^{nd}_F(n)$ at this point. To see this that this is valid, note that the limit of $p^{nd}_F(n)$ as a approaches α_n is c_n, which is strictly less than c_{n+1}, so the equilibrium is not a deterrence one. For $a > \alpha_n$ as long as $p^{nd}_F(n) < c_{n+1}$ the equilibrium is as described in part i) of Proposition 3. However, if a is such that $p^{nd}_F(n) = c_{n+1}$, firm $n+1$ is not indifferent between being active versus inactive, since $p^{nd}_F(n+1) < p^{nd}_F(n)$ as the activity of firm $n+1$ causes a discrete reduction in price. So for values of a slightly larger than that which results in $p^{nd}_F(n) = c_{n+1}$, price must remain at c_{n+1} as otherwise firm $n+1$ becomes active. The deterrence threshold, α^d_n follows immediately from $p^{nd}_F(n) = c_{n+1}$:

$$a = c_{n+1} + n^2(c_{n+1} - \bar{c}_n) \equiv \alpha^d_n. \tag{28}$$

Hence, the activity threshold for firm k can be derived as the value of a for which $p^{nd}_F(k) = c_k$ and the deterrence threshold where k firms deter the activity of firm $k+1$ as the value of a for which $p^{nd}_F(k) = c_{k+1}$.

[22]We present these as thresholds for the demand parameter a, but the α_k and α^d_k are themselves determined by the vector of marginal cost parameters, so these conditions represent subsets of the feasible parameters.

[23]We drop the * from the equilibrium number of firms when it is unambiguous in order to reduce notational clutter.

Proposition 3 establishes the unique equilibrium *aggregate* forward sales, but in the proof we note that individual forward sales are only uniquely determined in a non-deterrence equilibrium. In the deterrence case, there is a continuum of equilibrium individual forward sales allocations, however we can establish bounds on each firm's level of sales in such an equilibrium.[24] The following proposition characterizes the individual forward sales in the equilibrium:

Proposition 4. *In a forward sales equilibrium with n active firms, individual forward sales are determined as follows:*

i) If $a\in(\alpha_n, \alpha_n^d]$, individual sales of active firms are unique and given by

$$x_i^* = (n-1)\left(\frac{a+n^2\bar{c}_n}{n^2+1} - c_i\right), \quad \forall i \leq n.$$

ii) If $a\in(\alpha_n^d, \alpha_{n+1}]$, individual sales of active firms must satisfy the following conditions:

$$\sum_{i=1}^{n} x_i^* = X_n^d,$$

and

$$x_i^* \in [(n-1)(c_{n+1}-c_i), n(c_{n+1}-c_i)], \qquad \forall i \leq n.$$

Proof: See Appendix A.

Proposition 4 shows that the range of possible forward sales for a firm in a deterrence equilibrium is increasing in the efficiency of that firm. Indeed, the size of the interval of potential forward sales for firm i is $c_{n+1}-c_i$, so more efficient firms have a larger range of possible sales. In addition, both the maximum and minimum sales are higher for more efficient firms. As we wish to compute the output of individual firms in order to compare our results to alternative models of competition, we need to select particular deterrence equilibria to do so. We cannot use the Pareto criterion to identify a focal equilibrium as the aggregate forward sales and price in a deterrence equilibrium are the same for any of the possible combinations of individual forward sales. However, we can use the restrictions on individual firm sales in Proposition 4ii) to define two particular deterrence equilibria of interest: the most- and least-efficient equilibria. These represent the best- and worst-case scenarios for a deterrence equilibrium with n firms as they correspond to the lowest- and highest-cost allocations for producing the aggregate quantity X_n^d, and can be defined as follows:[25]

Definition 2. *The* **most-efficient deterrence equilibrium** *(MEDE) is that in which forward sales are concentrated in efficient firms as much a possible while still satisfying the conditions in Proposition 4ii), whereas the* **least-efficient deterrence equilibrium** *(LEDE) is that in which forward sales are concentrated in inefficient firms as much as possible.*[26]

We will use this definition later in the paper when we examine how market shares are affected by forward trading in our model.

[24] Figure 2(b) provides an illustration of these bounds as the point E is at the maximal sales for firm one and the minimal sales for firm two and the point E' is at the minimal sales for firm one and the maximal sales for firm two.

[25] A more formal definition is provided in Appendix B.2.

[26] In the case of firms one and two deterring firm three, the most- and least-efficient deterrence equilibria are illustrated in panel (b) of Figure 2 by points E and E' respectively.

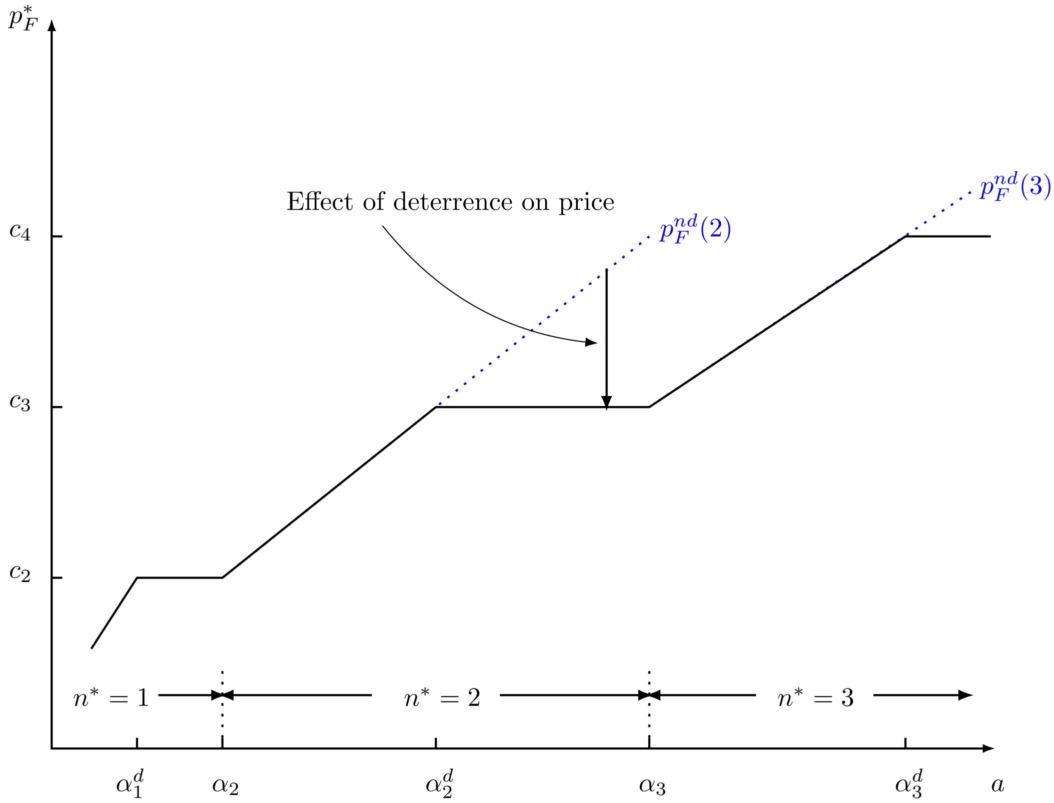

Figure 3: Equilibrium price versus a.

5 The competitive effects of forward trading

What are the implications of Proposition 3 for the equilibrium price, and hence, our understanding of the competitive effects of forward trading? We begin with an analysis of the influence that deterred firms have on price, and then turn to comparisons to homogeneous-firm and Cournot benchmark models to answer this question.

5.1 The influence of inactive firms on competition

To see how the endogenous activity of firms affects the equilibrium price, we illustrate how price varies with demand in Figure 3 for $N \geq 4$. Equilibrium price increases with a in a step-like pattern, which is due to the change in the type of equilibrium as demand varies (for fixed marginal costs). The equilibrium alternates between non-deterrence and deterrence as the activity of successively higher-cost firms is alternately blockaded, deterred, and then accommodated as a increases.

In the case of a non-deterrence equilibrium with $n^* < N$, the active firms act as they would if the inactive firms did not exist. The inactive firms have no effect on the equilibrium price because their activity is blockaded. In Figure 3 this occurs for $a \in (\alpha_2, \alpha_2^d)$, where firms $2, 3, \ldots$ are blockaded, and for $a \in (\alpha_3, \alpha_3^d)$ where firms $3, 4, \ldots$ are blockaded. The dashed lines $p_F^{nd}(2)$ and $p_F^{nd}(3)$ illustrate the counterfactual price that would occur with only two or three potential firms. The effect of deterrence on price is shown by the distance from the dotted line to the flat part of the equilibrium price. For instance, for $a \in (\alpha_2^d, \alpha_3)$ the distance from the $N=2$ dashed line to c_3 represents the reduction in price caused by the deterrence behaviour of firms one and two, and measures the influence of the *inactive* firm three on the equilibrium price. This effect on price in a deterrence equilibrium can be computed in general by comparing the price which would obtain if the inactive firms did not exist, $p_F^{nd}(n^*)$, with the price in the deterrence equilibrium, c_{n^*+1}. We establish the following:

Proposition 5. *As a varies from α_n^d to α_{n+1}, the reduction in price attributable to the deterred firm varies from 0 to $n(c_{n+1}-\bar{c}_n)/(n^2+1)$.*

Proof. See Appendix A. □

Conditional on remaining in the deterrence region, the higher the marginal cost of the deterred firm, the larger the maximum price reduction due to deterrence. Furthermore, for a given cost disadvantage of the deterred firm, $c_{n+1}-\bar{c}_n$, this price difference is decreasing in n, so the effect of deterrence is more significant the fewer the number of active firms. We present some numeric computations of the size of this maximal price difference in Table 2 of Appendix B.1 where the effect on price is well over 10% for relatively concentrated market structures.

5.2 The effect of firm heterogeneity on the forward price

In order to examine the effects of heterogeneity on the equilibrium price, an appropriate homogeneous benchmark must be determined. A natural model to consider is one in which there are N firms, each with marginal cost $\bar{c}_N$. This is simply the N-firm extension of Allaz & Vila (1993) which can be found in Bushnell (2007), so we have

Definition 3 (Homogeneous-firm benchmark)**.** *The homogeneous-firm benchmark model has N firms each with marginal cost $\bar{c}_N$. The equilibrium price is given by $p_F^H = p_F^{nd}(N) = (a + N^2\bar{c}_N)/(N^2+1)$.*

As a non-deterrence equilibrium with N active firms obtains for $a > \alpha_N$, an immediate result following from Definition 3 is that prices in the heterogeneous- and homogeneous-firm models coincide when demand is high enough that all firms are active. Consider now a level of demand slightly lower than firm N's activity threshold, $a' = \alpha_N - \epsilon$. Equilibrium in the heterogeneous-firm model is now a deterrence one with price equal to c_N, whereas price in the homogeneous model declines to $(a' + N^2\bar{c}_N)/(N^2+1) < c_N$. This will remain true for any $a \in [\alpha_{N-1}^d, \alpha_N]$, so for levels of demand "near" that at which firm N is deterred, price is higher than that in the homogeneous firm benchmark. Firm heterogeneity lessens the pro-competitive effect of forward trading in this case. We summarize these results in *i)* and *ii)* of the following Proposition, which also establishes a necessary and sufficient condition for $p_F^* > p_F^H$ to be true for any level of demand.

Proposition 6.

i) Price coincides with that in the homogeneous-firm benchmark when all firms are active in equilibrium: $p_F^* = p_F^H \; \forall \, a \geq \alpha_N$.

ii) Price is higher than that in the N-firm homogeneous benchmark for a range of levels of demand below the activity threshold of firm N: there exists an $\hat{\alpha} < \alpha_{N-1}^d$ such that $p_F^* > p_F^H(N) \; \forall \, a \in (\hat{\alpha}, \alpha_N)$.

iii) If the distribution of marginal costs across firms is sufficiently skewed, price is higher than the N-firm homogeneous benchmark for all levels of demand less than the activity threshold of firm N: $p_F^* > p_F^H(N) \; \forall \, a < \alpha_N \iff c_2 > \bar{c}_N$.

Proof. See Appendix A. □

Part *ii)* of Proposition 6 establishes that the forward sales equilibrium with $n^*{<}N$ heterogeneous firms is less competitive than the homogeneous N-firm benchmark for some interval of the demand parameter a. It does not rule out the possibility that there are other intervals for a at which this occurs as well, however a full characterization of when price is above or below the homogeneous firm price is dependent on the distribution of marginal costs within $\boldsymbol{c}$, whereas Proposition 6 *ii)* holds for any $\boldsymbol{c}$. A simple demonstration of how the distribution of marginal costs affects the analysis is to examine the case in when firm one deters firm two, i.e., $a{\in}[\alpha_1^d, \alpha_2)$, which as was the case in the duopoly example above, mimics the Bertrand equilibrium with $p_F^*{=}c_2$. This has a lower price than that in the homogeneous firm case as long as $c_2{<}\bar{c}_N$, which is clearly feasible for $N{>}2$, but not so when $N{=}2$. So the reduction in the pro-competitive effects of forward trading that we found in the duopoly example of Section 3 is not a general result. We will illustrate these results further with an example below.

Part *iii)* of Proposition 6 provides a necessary and sufficient condition for firm heterogeneity to reduce the pro-competitive effects of forward trading for any level of demand. It requires that the distribution of marginal costs be sufficiently skewed, in that only firm one has a marginal cost below the mean.

5.3 The effect of forward trading on heterogeneous firm competition

Proposition 6 establishes that firm heterogeneity may lead to a reduction in the pro-competitive effect of forward trading due to fewer firms being active in equilibrium. A natural question that arises is whether this effect can be so strong as to result in a price higher than that in the heterogeneous-firm Cournot model. It is straightforward to see that this cannot be the case. Clearly, Proposition 2 and Corollary 2 imply that, if $n^*{=}n^C$, price is lower in the forward sales game due to $X^*{>}0$. When there are fewer firms active in the forward sales game, $n^*{<}n^C$, price in the forward sales game must be lower than firm n^C's marginal cost (as it is inactive in the forward sales game) while price in the Cournot game must be higher than firm n^C's marginal cost (as it is active in the Cournot game). Consequently, the reduction in the equilibrium number of active firms cannot completely offset the pro-competitive effects of forward sales and we have established

Proposition 7. *The reduction in active firms when heterogeneous firm engage in forward trading is not sufficient to eliminate the pro-competitive effects of forward trading:* $p_F^*{<}p^C$.

5.4 Illustration

We now illustrate the above results with an example in which we consider four potential producers with marginal costs given by $\boldsymbol{c}{=}(50, 55, 60, 65)$, and consider a varying from 55 to 200. These numbers are chosen simply because they generate figures in which the alternative equilibria are clearly discernible. The equilibrium price, p_F^*, is plotted in Figure 4, along with the corresponding Cournot price, p^C, and the homogeneous-firm benchmark price, p_F^H. Comparing the forward sales price with that from the homogeneous firm game, we see that the results of Proposition 6 hold. The equilibrium is less competitive with heterogeneous firms for a significant range of demand levels, $a{\in}[118, 185]$. The lower bound of this interval ($\hat{\alpha}$ in Proposition 6 *ii)*) is substantially below the deterrence threshold for firm four. Another region of $p_F^*{>}\bar{p}_F$ occurs for demand levels around the deterrence threshold for firm three. The equilibrium is more competitive with heterogeneous firms at low levels of demand in which firm one deters firm two, $a{\in}[60, 65]$, which is as predicted since $c_2{<}\bar{c}_4$ in this case.

Figure 4 also demonstrates the dramatic difference in the number of active firms between the forward sales model and the heterogeneous firm Cournot model. All firms are active in the

Cournot model at significantly lower levels of demand than in the forward sales model, n^C=4 for a>95 or so, and n^*=4 for a>185. In addition, there is no level of demand for which three firms are active in both models. So any evaluation of the effects of forward trading must consider this variation in the active number of firms when considering the effects of forward sales.

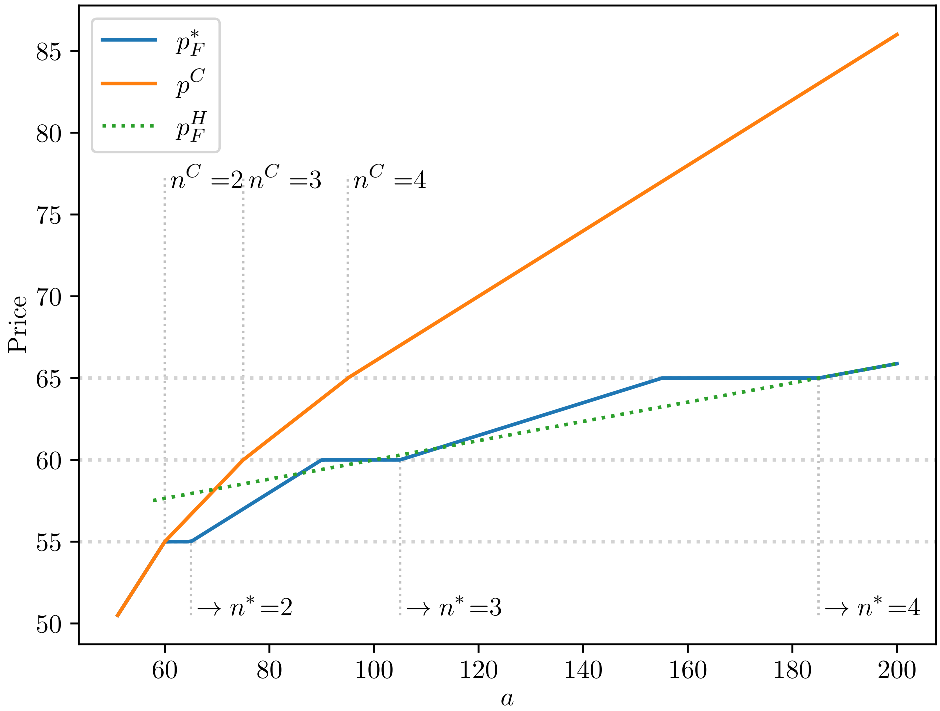

Figure 4: Equilibrium price, p^*_F, Cournot price, p^C, and homogeneous firm price, p^H_F versus a. Heterogeneous costs: $\boldsymbol{c}$=(50, 55, 60, 65); homogeneous cost c=57.5.

In order to illustrate part *iii)* of Proposition 6, consider a second example with a skewed distribution of marginal cost in which firm one has a substantial efficiency advantage: $\boldsymbol{c}$= (50, 59, 60, 61), which also has $\bar{c}_4$=57.5, so c_2>$\bar{c}_4$. The equilibrium prices are plotted in Figure 5, where we see the same general pattern of alternating deterrence and non-deterrence equilibria as a increase. However, now there is no level of demand for which p^*_F<p^H_F with more than one firm active. As shown in Proposition 6 *iii)*, c_2>$\bar{c}_N$ results in p^*_F>p^H_F even when firm one is deterring the activity of firm two. The lower limit $\hat{\alpha}$ of Proposition 6 *ii)* is lower than firm two's deterrence threshold leading to the wide range of demand levels for which price is higher in the heterogeneous firm case.

6 Implications for competition policy

In this section we explore how the endogenous activity of firms in our model affects some aspects of competition policy, in particular, merger analysis and the relationship between concentration and market power.

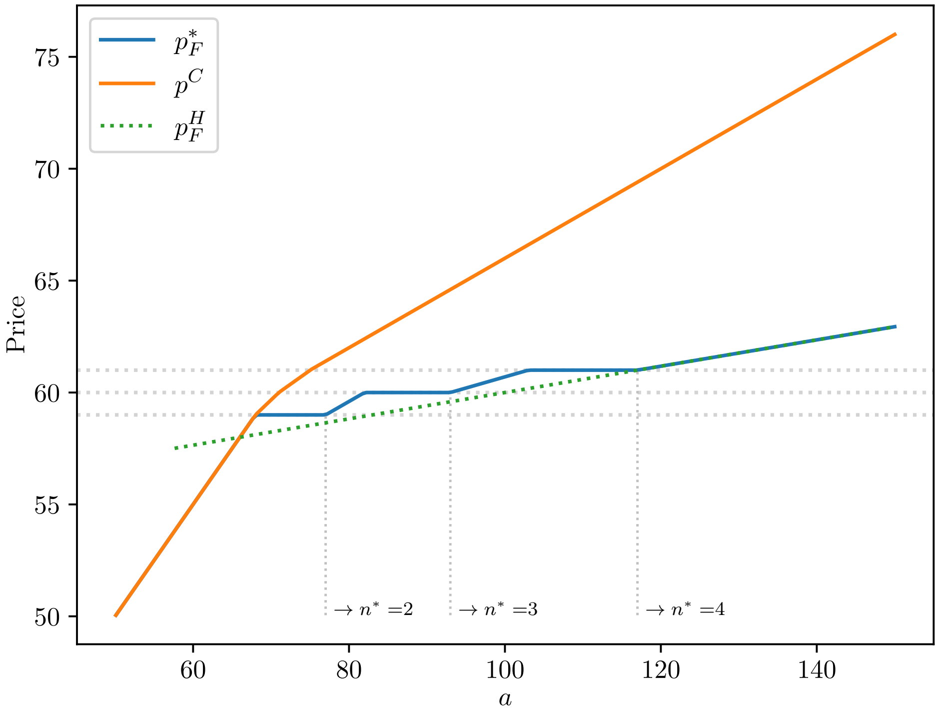

Figure 5: Equilibrium price, p_F^*, Cournot price, p^C, and homogeneous firm price, p_F^H versus a. Heterogeneous costs: $\boldsymbol{c}$=(50, 59, 60, 61); homogeneous cost c=57.5.

6.1 Mergers

The endogenous activity of firms in our model has implications for the analysis of mergers. We will assume that following a merger the merged firm operates at the marginal cost of the most efficient of the merging firms,[27] so a merger simply results in an elimination of the higher cost merging firm. Of course a more careful analysis of mergers would bring in capacity or capital in line with the analysis of Miller & Podwol (2020), but we simply wish to point to implications that stem from the different cost structure in our model that leads to the endogenous activity of firms.

The presence of inactive firms (either deterred or blockaded) may limit the degree to which price rises following a merger of active firms, whereas a merger between an active firm and a deterred inactive firm removes the ceiling on the price in a deterrence equilibrium, resulting in a price increase without a change in the set of active firms. We explore each of these possibilities in turn.

6.1.1 Mergers between active firms

Consider a pre-merger equilibrium with $n<N$ active firms. The effects of a merger between two of these firms depend on the nature of the pre- and post-merger equilibria, in particular whether either involves deterrence. Suppose that firm $n+1$ is blockaded in the pre-merger equilibrium so that the pre-merger price is $p_F^{Pre}=p_F^{nd}(n)$. If firm $n+1$ is also blockaded post-merger, then $p_F^{Post}=p_F^{nd}(n-1)>p_F^{nd}(n)=p_F^{Pre}$. Alternatively, if the post-merger equilibrium is one where firm $n+1$ is deterred, the increase in the post-merger price is limited by the marginal cost of firm $n+1$: $p_F^{Post}=c_{n+1}<p_F^{nd}(n-1)$. Price still increases due to the reduction in active firms, but this increase is limited by the potential activity of firm $n+1$. Finally, it is possible that firm $n+1$ is active in the post-merger equilibrium: $p_F^{Post}>c_{n+1}$. In this case, there is no change in the number of active firms, $n^{Post}=n^{Pre}$, yet price increases due to the substitution of the relatively inefficient firm $n+1$ for one of the lower cost firms that is removed by the merger. In each of these cases the merger results in a price increase, but the size of the price increase depends on the nature of the post-merger equilibrium.[28]

Now consider a merger of two active firms from a deterrence equilibrium. It is straightforward to establish that the post-merger equilibrium can no longer involve the deterrence of firm $n+1$, so that $p_F^{Post}>c_{n+1}$. This occurs because the non-deterrence price for the new set of active firms, $p_F^{nd'}(n'-1)$, is higher than $p_F^{nd}(n)$, which reduces the activity threshold of firm $n+1$ sufficiently for firm $n+1$ to be active with the new set of potential firms. Again, the merger does not decrease the number of active firms, rather it substitutes the relatively inefficient firm $n+1$ for the more efficient firm that disappears with the merger. As firm $n+1$ is active post-merger, price increases.

Summarizing, we have

Proposition 8. *A merger between two active firms:*

i) does not result in a reduction in the number active firms if the pre-merger equilibrium is a deterrence one, and may not result in a reduction in the number of active firms if the pre-merger equilibrium is a non-deterrence one.

[27] The merger does not generate synergies as defined in Farrell & Shapiro (1990), but represents the best outcome without synergies for the merging firms.

[28] There are other possibilities for the post-merger equilibrium, such as firm $n+1$ being active and firm $n+2$ being deterred or accommodated, however to keep the discussion simple we focus only on the potential activity of firm $n+1$. Note though, that accommodation of firm $n+2$ would result in an increase in the number of firms along with an increase in price due to the merger.

ii) increases price regardless of the nature of the pre-merger equilibrium.

Proof: See Appendix A.

Proposition 8 shows that a merger between two active firms is anticompetitive as it results in a price increase, regardless of whether it is profitable. This result is an analog to Spector (2003) who establishes that price increases for any profitable merger in a heterogeneous Cournot oligopoly with fixed costs. The possibility that an inactive firm becomes active following a merger means that the number of firms may not decline, which is a similar effect to that in Caradonna et al. (2024) where a merger may be followed by entry, but this still results in an increase in price. The model of Caradonna et al. (2024) is quite different from ours however, as they model differentiated products, price competition, and fixed entry costs.

6.1.2 Mergers with inactive firms

The fact that an inactive firm affects the equilibrium outcome in a deterrence equilibrium raises the possibility that an active firm might find it profitable to acquire the deterred firm.[29] Such a merger removes the ceiling on price due to the presence of the deterred firm and post-merger, Proposition 3 applies to the new set of potential firms: price will increase either to the non-deterrence price with n active firms, $p_F^{nd}(n)$, or to the marginal cost of the next higher cost firm, c_{n+2}. We have the following corollary to Proposition 3:

Corollary 3. *A merger of an active firm with a deterred firm results in an increase in price.*

As the deterred firm earns zero profit prior to the merger, the profitability of such a merger hinges only on how it affects the profit of the active merging firm. Clearly, for the duopoly example of Section 3, a merger from a deterrence equilibrium is profitable as it is a merger to monopoly. However, there is a monopoly prevailing pre-merger and the merger simply causes the elimination of an inactive firm who is not a direct competitor to the active firm in this market. This merger may not attract the attention of competition authorities even though it is welfare reducing.

For the oligopoly case, consider the profitability of an active firm merging with firm $n+1$ from a deterrence equilibrium. A sufficient condition for the active firm in the merger to increase its profit with the elimination of firm $n+1$ is that its total sales, $x_i+y_i^*(X)$ increase. Spot sales, $y_i^*(X)$, increase for all firms due to the reduction in aggregate forward sales caused by the merger, so an even weaker sufficient condition for profit to increase is that the firm's forward sales increase. In a deterrence equilibrium, a firm is most likely to increase forward sales following a merger if it finds itself selling the minimum quantity possible in that equilibrium. The following corollary to Proposition 4 establishes that this is indeed the case.

Corollary 4. *Any firm selling its minimum quantity, $x_i^*=n(c_{n+1}-c_i)$, in a deterrence equilibrium increases its profit by merging with firm $n+1$.*

Proof: See Appendix A.

Corollary 4 suggests that firms may have an incentive to break out of a deterrence equilibrium by way of merging with the deterred firm. The merger results in a reduction in consumer surplus due to the increase in price, but the effects on average production cost depend on the pre-merger

[29]Such a merger could be seen as reminiscent of Burns (1986), who studies the incentives of competitors to merge with a trust at distressed asset values, applying the model to the American Tobacco Company's acquisition of inactive rivals.

distribution of market shares and which firm undertakes the merger, so the net welfare effects of the merger depend on model parameters.[30]

For a concrete example, consider the illustration underlying Figure 4 described in subsection 5.4 where the equilibrium has the three most efficient firms deterring firm four, which occurs for $a \in [155, 185]$ in this example. Using the mid-point of this interval, $a=170$, the effects of a merger are presented in Table 1. For the pre-merger equilibrium, we consider each of the deterrence equilibria from Definition 2, MEDE and LEDE, along with a "median" scenario in which a firm's share of aggregate forward sales matches its market share in the non-deterrence equilibrium that would obtain in the absence of any higher cost inactive firms.[31] We label this scenario for the pre-merger equilibrium Non-Deterrence Market Share (NDMS). From the LEDE, firms one and three each see an increase in profit from merging with firm four, from the MEDE, firms two and three see an increase in profit, and from the NDMS equilibrium all firms see an increase in profit. So in each case, there are firms that gain and so would be willing to merge with firm four. The effect on price, and hence, consumer surplus is of course the same in each case, consumer welfare is harmed by the elimination of an inactive firm. The effect on total profit is positive in each case, and declining in the efficiency of the pre-merger production allocation. The increase in profit is large enough in a LEDE pre-merger that total surplus rises slightly. Industry average cost also declines in this case. If the pre-merger equilibrium is the most-efficient one, then the merger reduces the efficiency of the allocation of production across firms, resulting in a smaller increase in profit and, hence, a reduction in total surplus due to the merger. Although this is just a particular example, the results indicate that mergers with inactive firms are unlikely to be neutral with respect to welfare, and likely negative with respect to consumer surplus.

	Pre-merger equilibrium		
	LEDE	NDMS	MEDE
Price	2.3%	2.3%	2.3%
Profit			
Firm 1	21.0%	7.2%	-9.2%
Firm 2	-0.8%	13.4%	32.3%
Firm 3	27.0%	32.5%	69.3%
Total	14.0%	11.0%	5.1%
Industry average cost	-0.4%	0.1%	1.3%
Consumer surplus	-2.8%	-2.8%	-2.8%
Total surplus	0.1%	-0.4%	-1.4%

Table 1: Percentage effects of merger from a three-firm deterrence equilibrium for $a=170$ and $\boldsymbol{c}=(50, 55, 60, 65)$. LEDE = least-efficient deterrence equilibrium; MEDE = most efficient deterrence equilibrium; NDMS = non-deterrence market share equilibrium.

An important caveat to the above is that, under the interpretation of inactive firms as

[30] For jurisdictions in which a consumer surplus standard is applied to merger policy, such mergers would potentially be opposed, even though the merger does not remove an active competitor. In other jurisdictions that apply a total surplus standard the potential efficiency associated with a reduction in industry average cost may be relevant. See Duhamel (2003) for a discussion of these issues.

[31] We do not include this selection in Definition 2 as it is not guaranteed to satisfy the bounds on x_i in Proposition 4. It does satisfy these bounds for this particular example.

being active in some other related market (such as off-peak versus peak electricity markets), a merger with an inactive firm in one market needs to be addressed in the broader context of all markets in which these firms *may* compete, not just those in which they are both active. The same caution applies to mergers between active firms: if one of the merging firms is deterred in another market, the welfare effects of the merger needs to consider the possibility that price also rises in this other market, even though the number of competitors (active firms) is not affected in that market.

6.2 Concentration, production efficiency, and market power

With heterogeneous firms, forward trading also results in a reallocation of production across firms, so we now explore the implications of this reallocation for market concentration, average cost of production, and market power. In the case of a non-deterrence equilibrium, Proposition 4i) implies that lower-cost firms sell more forward, and, as more lower-cost firms also sell more spot, market share is reallocated towards relatively efficient firms. Indeed, it is straightforward to use the non-deterrence equilibrium quantities, x_i^* and $y_i^*(X^*)$, to obtain firm i's market share:

$$s_i \equiv \frac{x_i^* + y_i^*(X^*)}{X^* + Y^*} = \frac{1}{n} + \frac{(\bar{c}_n - c_i)(n^2+1)}{n(a - \bar{c}_n)}. \tag{29}$$

We can use Corollary 2 to determine market shares in the Cournot game with N firms, s_i^C and comparing this to s_i for $n{=}N$ yields

$$s_i - s_i^C = \frac{\bar{c}_N - c_i}{(a - \bar{c}_N)}(N-1). \tag{30}$$

Forward trading results in higher market shares for relatively efficient firms (those with $c_i < \bar{c}_N$) and lower market shares for relatively inefficient firms (those with $c_i > \bar{c}_N$). Hence, forward sales result in a reallocation of market share towards relatively efficient firms compared to the Cournot equilibrium, resulting in an increase in the Herfindahl-Hirschmann Index (HHI). Furthermore, this reallocation of market share to relatively efficient firms implies that industry average cost, $AC \equiv \sum_{i=1}^{n} s_i c_i$, declines with forward trading in this case. Notice that we get the same result for any situation in which $n_C = n^* < N$, however, there may not be a feasible set of parameters that generates this outcome. We have established:

Proposition 9. *In a non-deterrence equilibrium, if $n^* = n^C$, forward trading results in an increase in the HHI and a reduction in AC.*

Miller & Podwol (2020), in their Corollary 1, find a similar result in that firms with sufficiently low capital stocks see their output decrease with forward trading, concluding that forward trading may increase concentration in their model. Although the increase in concentration identified in Proposition 9 is valid for any distribution of marginal costs among firms, it is limited to a non-deterrence equilibrium with $n^* = n^C$, which applies to only a subset of feasible demand and cost parameters of our model.

To see how Proposition 9 generalizes to situations in which $n^* \neq n^C$, in the remainder of this subsection we use the illustration underlying Figure 4 to compare concentration, average cost, and market power to that in the Cournot model as well as in the homogeneous firm benchmark of Definition 3. In order to do so we use the most- and least-efficient deterrence equilibria (MEDE and LEDE) from Definition 2 to illustrate the range of possibilities in the case of deterrence equilibria. We plot the HHI versus the level of demand in Figure 6, which shows that the endogenous number of active firms and the possibility of deterrence equilibria do not affect the conclusion that forward trading results in an increase in concentration. Both the MEDE and LEDE deterrence equilibria result in a non-monotonic relationship between the level of demand

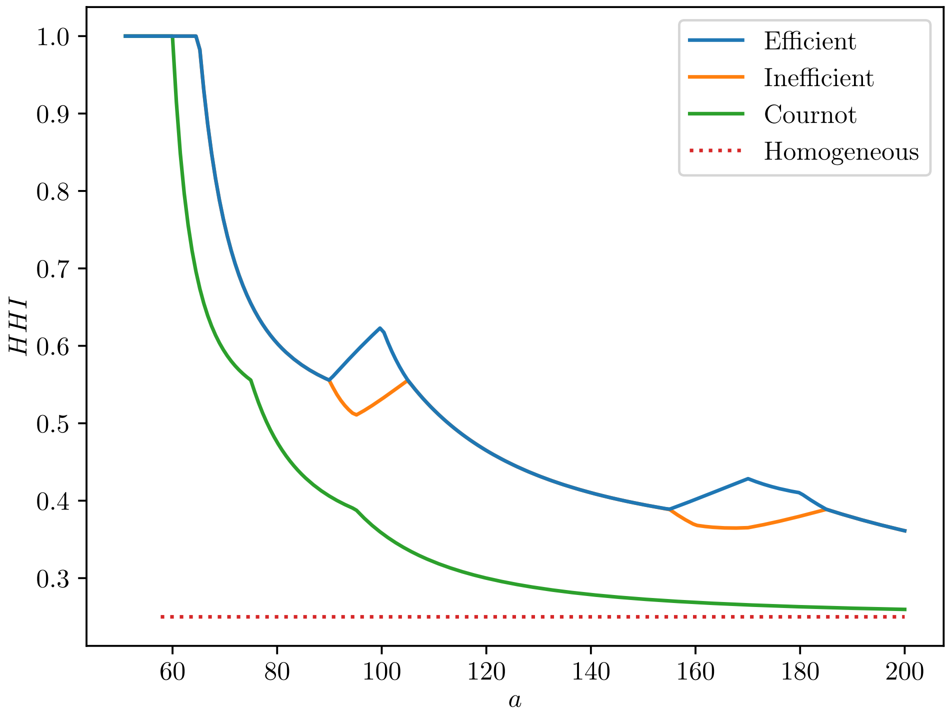

Figure 6: Hirschmann-Herfindahl Index versus a for $\boldsymbol{c}$=(50, 55, 60, 65).

and concentration, however the HHI in deterrence equilibria remain substantially larger than that in the Cournot equilibrium.

Industry average cost is plotted in Figure 7. Even in the least-efficient deterrence equilibria, average cost in this example is well below that in the Cournot model. The magnitude of the difference is not insignificant in proportional terms: the maximal percentage difference between the Cournot case and the efficient equilibrium is 6.2% in Figure 7.

We also plot the HHI and AC for the homogeneous-firm benchmark in Figures 6 and 7. As these are the same as in the homogeneous-firm Cournot model, we can identify to what extent the differences are driven by cost heterogeneity versus forward trading. The effect of heterogeneity under Cournot competition can be seen by the difference between the Cournot and Homogeneous lines in Figures 6 and 7: The HHI is increased and AC is decreased by cost heterogeneity. The addition of forward trading amplifies this effect substantially.

Finally, regarding market power, consider the effect of forward trading on the Lerner Index, $L = \sum_{i=1}^{n} s_i(p_F - c_i)/p_F$. The reallocation of market shares to relative efficient firms suggests L should increase with forward trading as market share is increased for firms with higher margins. However, the pro-competitive effect of forward trading suggests a decrease in L as price decreases. We plot the Lerner Index in Figure 8. Compared to the Cournot equilibrium, clearly the effect of lower price dominates that of the increasing market shares of high margin firms in this case as the Lerner Index is substantially below that in the Cournot equilibrium, even though the HHI is higher. Although this effect was also a feature of Miller & Podwol (2020), the contribution of Figure 8 is to suggest that it survives the endogenous activity of firms present in our model. The Lerner Index is also substantially higher than that in the N-firm homogeneous firm forward sales game due to the fact that more efficient firms have both higher market shares and higher price-cost margins than in the homogeneous firm case.

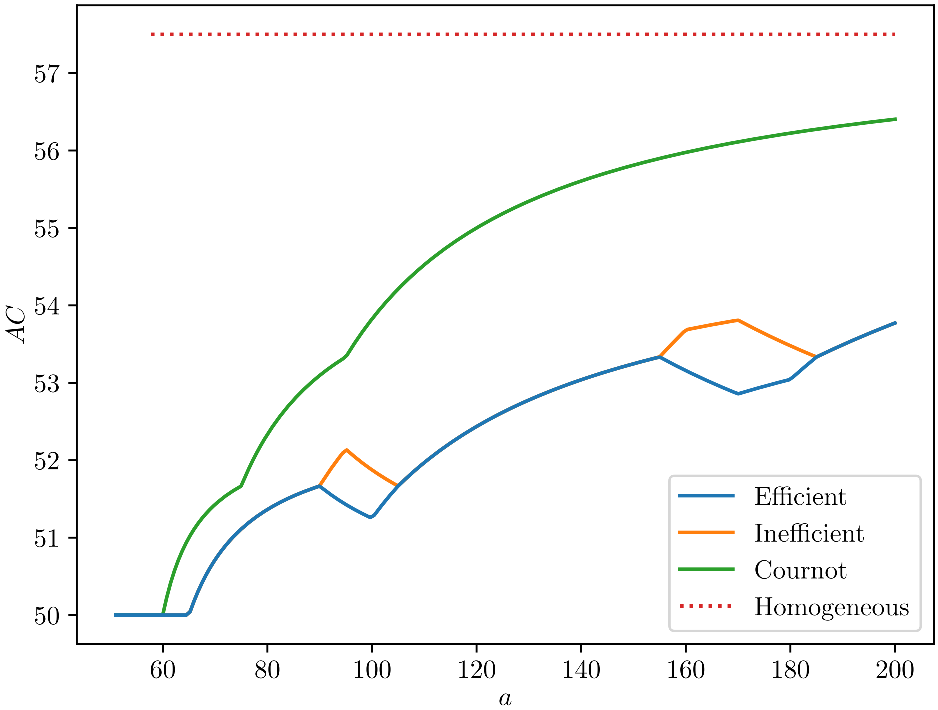

Figure 7: Industry average cost versus a for $\boldsymbol{c}$=(50, 55, 60, 65).

However, this higher Lerner Index does not necessarily correspond to a welfare loss as price levels can be higher, lower, or the same in the two models, as we see from Figure 4.

7 Multiple periods of forward trading

For a homogeneous firm duopoly, Allaz & Vila (1993) show that the competitive effect of forward trading strengthens as the number of forward trading opportunities increases. In the limit, as the number of trading periods tends to infinity, price approaches marginal cost. In this section, we apply our duopoly model from Section 3 to explore the effect of additional trading periods in the heterogeneous firm model.

Proposition 6 establishes that in a non-deterrence equilibrium with all N firms active the price is the same as in the homogenous-firm with each firm having a marginal cost of $\bar{c}_N$. In the duopoly example with c_1=0 and c_2=c and one period of forward trading, recall that a non-deterrence equilibrium occurs for a>3c for which the price p_F^*=(a+2c)/5. This is the same price as obtains from the Allaz and Vila model with each firm having a marginal cost of $c/2$. Conjecturing that this price equivalence between the two models holds as we add trading periods,[32] we can use the result established in Allaz & Vila (1993)'s Proposition 3.1, which showed that with T periods of forward trading prior to the spot market, the equilibrium price will be $p_F^{AV}(T)=c/2+(a-c/2)/(3+2T)$. This price tends to $c/2$ as $T\to\infty$ in the case of homogeneous firms. However, in the heterogeneous firm case, this is the equilibrium price only if it exceeds the marginal cost of firm two: $p_F^{AV}(T)>c$. It is easily seen that this requires $a>(T+2)c\equiv\alpha_1^d(T)$, which is the deterrence threshold below which firm one deters the activity of

[32]We confirm that this conjecture holds in the proof of the following proposition.

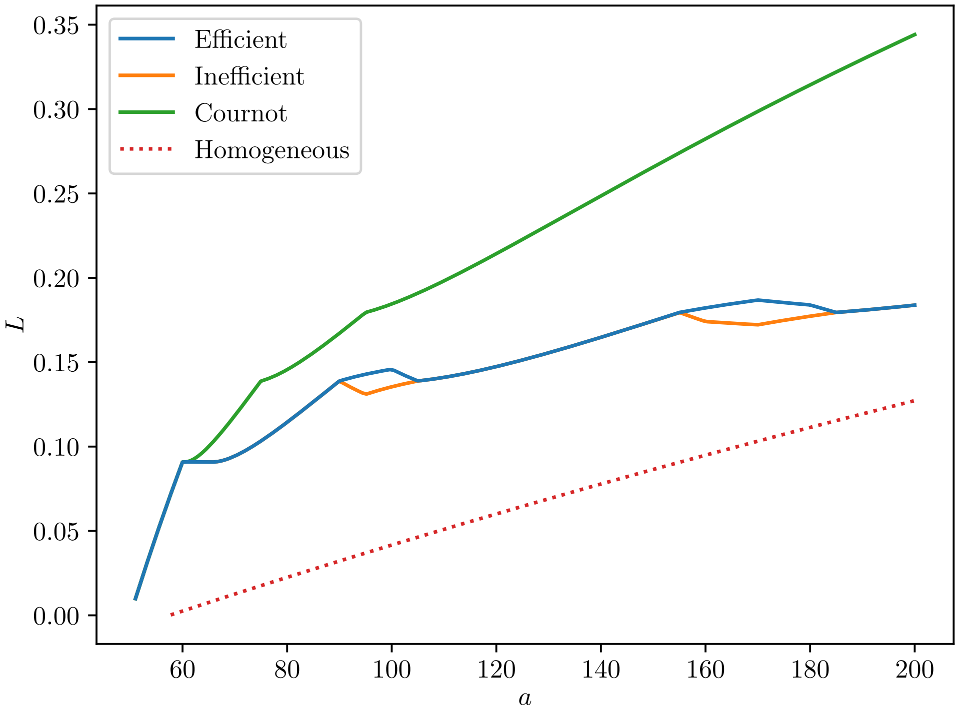

Figure 8: Lerner Index versus a for $\boldsymbol{c}$=(50, 55, 60, 65).

firm two. By lowering the price in a non-deterrence equilibrium, additional periods of forward trading have the effect of reducing the set of demand levels for which firm two is active.

Proposition 10. *In the duopoly game with c_1=0 and c_2=c,*

i) the set of demand levels in which deterrence is the outcome, $a \in [2c, (T+2)c]$, increases with T, and

ii) for any demand level $a > c$, the Bertrand outcome is obtained with a finite number of forward trading periods.

Proof: See Appendix A.

Proposition 10 demonstrates another difference to Allaz & Vila (1993), where, in the limit as the number of trading periods increases, the efficient outcome is obtained. As the efficient outcome in their model is the homogeneous-firm Bertrand outcome, we have a similar result in *ii)*, however the Bertrand outcome is not efficient in the heterogeneous firm case. Instead, the outcome is that firm one is active with price equal to the marginal cost of firm two. Furthermore, with heterogeneous firms, the Bertrand outcome is obtained with a finite number of trading periods. This was of course demonstrated in Section 3, where we obtain the Bertrand outcome for $a \in [2c, 3c]$ with just one period of forward trading. As the upper bound of this interval increases with the number of forward trading periods, the Bertrand outcome will occur after a finite number of trading periods.

Although we do not demonstrate this, we conjecture that a similar result would obtain in the more general case of an oligopoly with N potential firms. As additional forward trading periods reduces price in any non-deterrence equilibrium the number of active firms will decline as T increases. For any firm other than the most efficient, as the number of forward trading periods increases, the level of demand required for those firms to be active also increases.

8 Concluding remarks

The results we have obtained can be applied more broadly to other situations in which firms compete in advance for sales. A similar force is present when the good is storable, either by consumers[33] or by speculators[34]. Consumption that is met though inventories, whether by consumers themselves or by speculators' resale, reduce the demand faced by producers in the same manner that forward sales do. These models differ from ours in that there is not a single spot market, but rather a sequence of markets with a separate demand in each. The closest of these models to ours is Anton & Das Varma (2005), who examine a homogeneous duopoly competing over two periods. The equilibrium may be one of two types: either prices in the two periods do not induce storage and the equilibrium is simply the repetition of the Cournot equilibrium, or storage occurs and prices are linked as when inventories are carried price the price difference cannot exceed the marginal cost of storage. It is the latter type of equilibrium that is similar to the forward trading equilibrium: the excess of sales over first period consumption are the equivalent of forward sales and the first period price is the equivalent of the forward price. The possibility to increase sales in the first period by induce storage through the lower price leads to more aggressive competition and lower prices in the same way that forward sales do. Anton & Das Varma (2005) show that in the absence of storage costs and with demand that does not vary over time, the only equilibrium is one with storage. So the results of this study apply directly to homogeneous product markets where consumer storage is possible.

Our results were derived using a simple homogeneous good model of quantity competition with linear demand and constant marginal cost. We conclude with a discussion of the potential robustness of these results to the underlying modelling assumptions. An alternative price competition model with heterogeneous costs would be a substantial departure and would likely reverse the pro-competitive effects as it does in Mahenc & Salanié (2004). However, the discrete change in the number of active firms when price reaches the marginal cost of higher-cost firms would likely still generate discontinuities in marginal profit as it does in our quantity choice game.

The linear demand assumption is important for generating the closed-form solutions for equilibrium variables. However, the key feature driving much of the analysis is the discontinuous nature of marginal profit when price reaches a level at which a higher-cost firm becomes active. This feature is likely present under more general demand specifications due to the discrete change in the number of active firms that occurs at these points.

The assumption of constant marginal cost can also be relaxed,[35] but the deterrence equilibria rely on firms' marginal cost at zero production be different, not that marginal cost is constant for all levels of output. For purely quadratic costs, $C_i(q_i)=c_iq_i^2$, for example, the deterrence equilibria will not exist as it is not possible to deter the activity of any firm with a non-zero price. However, as long as $C_i'(0)$ differs among firms, deterrence equilibria may be possible, and would represent an extension of the Miller & Podwol (2020) model to allow for the endogenous activity of firms. Another interesting extension to the model would be to relax the zero fixed cost assumption and allow for an earlier period in which firms choose to enter by paying an entry cost.

[33] Hendel & Nevo (2006), Dudine et al. (2006), Antoniou & Fiocco (2019), Anton & Das Varma (2005), and Guo & Villas-Boas (2007)

[34] Mitraille & Thille (2009) and Mitraille & Thille (2014)

[35] Breitmoser (2013) shows that increasing marginal cost reduces the pro-competitive effect of forward sales for homogeneous firms.

As competition is more intense with forward trading than it is in the Cournot model, it is likely that *entry* deterrence would be easier with forward trading than when Cournot competition obtains in the post-entry game.

Finally, the model analyzed here was deterministic. It is natural to question how the introduction of uncertainty affects the results, especially since forward trading is often thought to arise as a response to uncertainty. Mitraille & Thille (2020) have shown that the nature of equilibrium depends on the degree of uncertainty. If uncertainty is "minor" then certainty equivalence holds and the results are qualitatively the same as in the deterministic case. We explore whether similar results translate to the heterogeneous firm case in a companion paper.

References

Adilov, N. (2012). Strategic use of forward contracts and capacity constraints. *International Journal of Industrial Organization*, 30(2), 164–173.

Allaz, B. & Vila, J.-L. (1993). Cournot competition, forward markets and efficiency. *Journal of Economic Theory*, 59(1), 1–16.

Anderson, S. P., Erkal, N., & Piccinin, D. (2020). Aggregative games and oligopoly theory: Short-run and long-run analysis. *The RAND Journal of Economics*, 51(2), 470–495.

Anton, J. J. & Das Varma, G. (2005). Storability, market structure, and demand-shift incentives. *The RAND Journal of Economics*, 36(3), 520–543.

Antoniou, F. & Fiocco, R. (2019). Strategic inventories under limited commitment. *The RAND Journal of Economics*, 50(3), 695–729.

Behar, A. & Ritz, R. A. (2017). OPEC vs US shale: Analyzing the shift to a market-share strategy. *Energy Economics*, 63, 185–198.

Breitmoser, Y. (2012). On the endogeneity of Cournot, Bertrand, and Stackelberg competition in oligopolies. *International Journal of Industrial Organization*, 30(1), 16–29.

Breitmoser, Y. (2013). Increasing marginal costs are strategically beneficial in forward trading. *Economics Letters*, 119(2), 109–112.

Brown, D. P. & Eckert, A. (2017). Electricity market mergers with endogenous forward contracting. *Journal of Regulatory Economics*, 51(3), 269–310.

Burns, M. R. (1986). Predatory Pricing and the Acquisition Cost of Competitors. *Journal of Political Economy*, 94(2), 266–296.

Bushnell, J. (2007). Oligopoly equilibria in electricity contract markets. *Journal of Regulatory Economics*, 32(3), 225–245.

Caradonna, P., Miller, N., & Sheu, G. (2024). Mergers, Entry, and Consumer Welfare.

de Frutos, M.-A. & Fabra, N. (2012). How to allocate forward contracts: The case of electricity markets. *European Economic Review*, 56(3), 451–469.

Dudine, P., Hendel, I., & Lizzeri, A. (2006). Storable good monopoly: The role of commitment. *American Economic Review*, 96(5), 1706–1719.

Duhamel, M. (2003). On the Social Welfare Objectives of Canada's Antitrust Statute. *Canadian Public Policy / Analyse de Politiques*, 29(3), 301–318.

Etro, F. (2006). Aggressive leaders. *The RAND Journal of Economics*, 37(1), 146–154.

Etro, F. (2008). Stackelberg Competition with Endogenous Entry. *The Economic Journal*, 118(532), 1670–1697.

Farrell, J. & Shapiro, C. (1990). Horizontal Mergers: An Equilibrium Analysis. *The American Economic Review*, 80(1), 107–126.

Ferreira, J. L. (2006). The Role of Observability in Futures Markets. *Topics in Theoretical Economics*, 6(1).

Flamm, K. & Reiss, P. C. (1993). Semiconductor Dependency and Strategic Trade Policy. *Brookings Papers on Economic Activity. Microeconomics*, 1993(1), 249–333.

Gilbert, R. & Vives, X. (1986). Entry Deterrence and the Free Rider Problem. *The Review of Economic Studies*, 53(1), 71–83.

Guo, L. & Villas-Boas, J. M. (2007). Consumer stockpiling and price competition in differentiated markets. *Journal of Economics & Management Strategy*, 16(4), 827–858.

Hendel, I. & Nevo, A. (2006). Sales and consumer inventory. *The RAND Journal of Economics*, 37(3), 543–561.

Holmberg, P. & Willems, B. (2015). Relaxing competition through speculation: Committing to a negative supply slope. *Journal of Economic Theory*, 159, 236–266.

Ito, K. & Reguant, M. (2016). Sequential Markets, Market Power, and Arbitrage. *American Economic Review*, 106(7), 1921–1957.

Liski, M. & Montero, J.-P. (2006). Forward trading and collusion in oligopoly. *Journal of Economic Theory*, 131(1), 212–230.

Mahenc, P. & Salanié, F. (2004). Softening competition through forward trading. *Journal of Economic Theory*, 116(2), 282–293.

Miller, N. H. & Podwol, J. U. (2020). Forward Contracts, Market Structure and the Welfare Effects of Mergers. *The Journal of Industrial Economics*, 68(2), 364–407.

Mitraille, S. & Thille, H. (2009). Monopoly behaviour with speculative storage. *Journal of Economic Dynamics and Control*, 33(7), 1451–1468.

Mitraille, S. & Thille, H. (2014). Speculative storage in imperfectly competitive markets. *International Journal of Industrial Organization*, 35, 44–59.

Mitraille, S. & Thille, H. (2020). Strategic advance sales, demand uncertainty and overcommitment. *Economic Theory*, 69(3), 789–828.

Nocke, V. & Schutz, N. (2018). Multiproduct-Firm Oligopoly: An Aggregative Games Approach. *Econometrica*, 86(2), 523–557.

Novshek, W. (1984). Finding All n-Firm Cournot Equilibria. *International Economic Review*, 25(1), 61–70.

Pal, D. (1991). Cournot duopoly with two production periods and cost differentials. *Journal of Economic Theory*, 55(2), 441–448.

Pal, D. (1996). Endogenous Stackelberg equilibria with identical firms. *Games and Economic Behavior*, 12(1), 81–94.

Rey, P. & Tirole, J. (2007). Chapter 33 A Primer on Foreclosure. In M. Armstrong & R. Porter (Eds.), *Handbook of Industrial Organization*, volume 3 (pp. 2145–2220). Elsevier.

Salant, S. W. & Shaffer, G. (1999). Unequal Treatment of Identical Agents in Cournot Equilibrium. *American Economic Review*, 89(3), 585–604.

Saloner, G. (1987). Cournot duopoly with two production periods. *Journal of Economic Theory*, 42(1), 183–187.

Spector, D. (2003). Horizontal mergers, entry, and efficiency defences. *International Journal of Industrial Organization*, 21(10), 1591–1600.

van Eijkel, R., Kuper, G. H., & Moraga-González, J. L. (2016). Do firms sell forward for strategic reasons? An application to the wholesale market for natural gas. *International Journal of Industrial Organization*, 49, 1–35.

Vives, X. (1988). Sequential entry, industry structure and welfare. *European Economic Review*, 32(8), 1671–1687.

Wolfram, C. D. (1999). Measuring Duopoly Power in the British Electricity Spot Market. *The American Economic Review*, 89(4), 805–826.

A Appendix

Proof of Proposition 2.

To find the equilibrium in our heterogeneous firm game, we apply the technique to solving aggregative games established in Novshek (1984), Nocke & Schutz (2018), and Anderson et al. (2020).

The first order condition for the maximization of (11) directly gives the continuous best response

$$y_i(Y_{-i},X)=\begin{cases} \frac{a-X-c_i}{2}-\frac{1}{2}Y_{-i} & \text{if } Y_{-i}<a-X-c_i \\ 0 & \text{if } Y_{-i}\geq a-X-c_i \end{cases} \tag{31}$$

Adding Y_{-i} to both sides of this expression gives the total spot sales compatible with the optimization behaviour of firm i, $Y(Y_{-i},X)$

$$Y(Y_{-i},X)=\begin{cases} \frac{a-X-c_i}{2}+\frac{1}{2}Y_{-i} & \text{if } Y_{-i}<a-X-c_i \\ Y_{-i} & \text{if } Y_{-i}\geq a-X-c_i \end{cases} \tag{32}$$

As this function is strictly increasing and continuous, it can be inverted to obtain the aggregate spot sales of firm i's competitors, $Y_{-i}(Y,X)$, compatible with the optimization behaviour of firm i, as a response to Y:

$$Y_{-i}(Y,X)=\begin{cases} 2Y-(a-X-c_i) & \text{if } Y<a-X-c_i \\ Y & \text{if } Y\geq a-X-c_i \end{cases} \tag{33}$$

Using $Y_{-i}(Y,X)=Y-y_i$, solve for y_i, which yields the best response of firm i to the industry spot sales Y. We denote this expression $y_i(Y,X)$:

$$y_i(Y,X)=\begin{cases} a-X-c_i-Y & \text{if } Y<a-X-c_i \\ 0 & \text{if } Y\geq a-X-c_i \end{cases} \tag{34}$$

As marginal costs are ordered, thresholds at which firms are active on the spot market are ordered: $a-X-c_N<...<a-X-c_i<...<...a-X-c_1$ so that it is possible to find the equilibrium aggregate spot sales by solving $Y=\sum_{i=1}^{N} y_i(Y,X)$ for each possible value of X. However, the number of active firms must be determined in order to do this. Let k be the number of active firms, so firms $1,...,k$ have positive spot sales and firms $k+1,...,N$ have zero spot sales. Using (34) total spot sales, denoted $Y_k^*(X)$, solves $Y_k^*(X)=\sum_{i=1}^{k} y_i(Y_k^*(X),X)$, giving

$$Y_k^*(X)=\frac{k(a-X)-k\bar{c}_k}{k+1}. \tag{35}$$

For this to be an equilibrium it must be the case that $Y_k^*(X)<a-X-c_k$, so that firms $1,...,k$ are active, and that $Y_k^*(X)\geq a-X-c_{k+1}$, so that firms $k+1,...,N$ are inactive. So, for active firms, this requires $X\leq a+k\bar{c}_k-(k+1)c_k=X_k^d$, whereas for inactive firms it requires $X\geq a+k\bar{c}_k-(k+1)c_{k+1}=X_{k+1}^d$. Hence, firm $i=1,2,...,N$ is active in equilibrium if, and only if, $X\leq X_i^d= a+i\bar{c}_i-(i+1)c_{i+1}$, where $X_1^d>X_2^d>...>X_N^d$.

In the sub-game defined by X, firm i is active if $X<X_i^d$ and, since the X_i^d are decreasing in i, the number of active firms, $n(X)$, is simply the number of firms for which $X<X_j^d$, or

$$n(X)=\sum_{j=1}^{N} \mathbb{1}_{[X<X_j^d]}. \tag{36}$$

For active firms we have

$$\begin{aligned} y_i^*(X) &= y_i(Y^*_{n(X)}(X), X) \\ &= a - X - c_i - \frac{n(X)}{n(X)+1}(a - n(X)\bar{c}_{n(X)} - X) \\ &= \frac{a - X + n(X)\bar{c}_{n(X)}}{n(X)+1} - c_i \end{aligned} \tag{37}$$

and equilibrium price is

$$\begin{aligned} p_S^*(X) &= a - X - Y^*_{n(X)}(X) \\ &= a - X - \frac{n(X)}{n(X)+1}(a - X - n(X)\bar{c}_{n_X})) \\ &= \frac{a - X + n(X)\bar{c}_{n(X)}}{n(X)+1}. \end{aligned} \tag{38}$$

To establish continuity of spot sales at the threshold X_n^d, where the number of firms changes, we have

$$\begin{aligned} \lim_{X \to X_n^{d-}} y_i^*(X) &= \frac{a - X_n^d + n\bar{c}_n}{n+1} - c_i \\ &= \frac{a - (a + n\bar{c}_n - (n+1)c_n) + n\bar{c}_n}{n+1} - c_i \\ &= c_n - c_i, \end{aligned} \tag{39}$$

and

$$\begin{aligned} \lim_{X \to X_n^{d+}} y_i^*(X) &= \frac{a - X_n^d + (n-1)\bar{c}_{n-1}}{n} - c_i \\ &= \frac{a - (a + n\bar{c}_n - (n+1)c_n) + (n-1)\bar{c}_{n-1}}{n} - c_i \\ &= \frac{(n+1)c_n - n\bar{c}_n + (n-1)\bar{c}_{n-1}}{n} - c_i \\ &= \frac{nc_n + c_n - \sum_{j=1}^{n} c_j + \sum_{j=1}^{n-1} c_j}{n} - c_i \\ &= c_n - c_i. \end{aligned} \tag{40}$$

Therefore, spot sales are continuous at the threshold X_n^d, and, consequently, so is price. □

Lemma 1. *Given X_{-i}, marginal profit for firm i is strictly decreasing in x_i with downward jump discontinuities at values $x_i = X_k^d - X_{-i}$ for $k > i$ and $X_{-i} < X_k^d$.*

Proof of Lemma 1. At the points of discontinuity in firm i's marginal profit, (24), $X_k^d - X_{-i}$ for $k > i$, we have $p_S^*(X_k^d) = c_k$, where there are k active firms for x_i slightly below $X_k^d - X_{-i}$ and $k-1$ active firms for x_i slightly above $X_k^d - X_{-i}$. Examining the limit of (24) as x_i approaches $X_k^d - X_{-i}$ from below and above we have

$$\lim_{x_i \to X_k^d - X_{-i}^{\;-}} \frac{\partial \pi_i(x_i, X_{-i})}{\partial x_i} = (c_k - c_i)\frac{k-1}{k+1} - \frac{X_k^d - X_{-i}}{k+1} \tag{41}$$

and

$$\lim_{x_i \to X_k^d - X_{-i}^{\;+}} \frac{\partial \pi_i(x_i, X_{-i})}{\partial x_i} = (c_k - c_i)\frac{k-2}{k} - \frac{X_k^d - X_{-i}}{k}. \tag{42}$$

Since $\frac{k-2}{k}<\frac{k-1}{k+1}$, $\frac{-1}{k}<\frac{-1}{k+1}$, and $c_k-c_i>0$, we have

$$\lim_{x_i \to X_k^d - X_{-i}^{\ -}} \frac{\partial \pi_i(x_i, X_{-i})}{\partial x_i} > \lim_{x_i \to X_k^d - X_{-i}^{\ +}} \frac{\partial \pi_i(x_i, X_{-i})}{\partial x_i}, \tag{43}$$

so marginal profit jumps down at $x_i = X_k^d - X_{-i}$. □

Lemma 2. *An active firm i's best response to X_{-i}, $r_i(X_{-i})$, is a decreasing and continuous function with the following properties:*

1. *$r_i(X_{-i})$ alternates between $r_{i,k}(X_{-i})$ and $X_k^d - X_{-i}$ for a decreasing sequence of $k>i$ as X_{-i} increases from 0 to X_i^d,*
2. *$r_i(X_{-i})=0$ for $X_{-i} \geq X_i^d$.*

Proof of Lemma 2.

We start with the best response for firm N, $r_N(X_{-N})$, which is the simplest because there are no higher-cost rival firms to deter, so the marginal profit for firm N is continuous at any $x_N \geq 0$. Using (24) and $n(X)=N$, equating its marginal profit to zero provides a candidate best-response

$$x_N = \frac{N-1}{2N}\left(a + N\bar{c}_N - (N+1)c_N - X_{-N}\right), \tag{44}$$

which is strictly positive if

$$X_{-N} < a + N\bar{c}_N - (N+1)c_N = X_N^d. \tag{45}$$

If this condition on X_{-N} does not hold, firm N's profit is negative for any positive sales. Therefore, when $X_{-N} \geq X_N^d$, firm N is better off not selling forward, $x_N=0$. To summarize,

$$r_N(X_{-i}) = \begin{cases} \frac{N-1}{2N}\left(a + N\bar{c}_N - (N+1)c_N - X_{-N}\right) & \text{if } X_{-N} < X_N^d \\ 0 & \text{if } X_{-N} \geq X_N^d. \end{cases} \tag{46}$$

A similar approach is used to find the best response functions of the other firms, however, these are complicated by the discontinuity in marginal profit at the activity thresholds of higher cost rivals. Suppose that X_{-i} is such that firm i is considering a level of forward sales consistent with $k>i$ active firms, i.e. $n(X)=k$. One of two possibilities must be true: either i) firm i's best response occurs on the downward sloping part of its marginal profit, between $X_{k+1}^d - X_{-i}$ and $X_k^d - X_{-i}$, or ii) firm i's best response occurs at $X_{k+1}^d - X_{-i}$, where i's marginal profit jumps from positive to negative. We will examine each case in turn, deriving the bounds on X_{-i} under which they obtain.

Case i) $r_i(X_{-i}) \in (X_{k+1}^d - X_{-i}, X_k^d - X_{-i})$: In this case, i's best-response is

$$r_i(X_{-i}) = r_{i,k}(X_{-i}) = \frac{k-1}{2k}\left(a + k\bar{c}_k - (k+1)c_i - X_{-i}\right), \tag{47}$$

which is in the interval $(X_{k+1}^d - X_{-i}, X_k^d - X_{-i})$ if marginal profit is positive at the lower limit of the interval and negative at the upper limit:

$$\lim_{x_i \to X_{k+1}^d - X_{-i}^{\ +}} \frac{\partial \pi_i(x_i, X_{-i})}{\partial x_i} > 0 \quad \text{and} \quad \lim_{x_i \to X_k^d - X_{-i}^{\ -}} \frac{\partial \pi_i(x_i, X_{-i})}{\partial x_i} < 0. \tag{48}$$

This requires

$$\lim_{x_i \to X_{k+1}^d - X_{-i}^{\ +}} \frac{\partial \pi_i(x_i, X_{-i})}{\partial x_i} = (c_{k+1} - c_i)\frac{k-1}{k+1} - \frac{X_k^d - X_{-i}}{k+1} > 0, \tag{49}$$

which reduces to

$$X_{-i}>X^d_{k+1}-(k-1)(c_{k+1}-c_i). \tag{50}$$

For the upper bound,

$$\lim_{x_i\to X^d_k-X_{-i}{}^-}\frac{\partial\pi_i(x_i,X_{-i})}{\partial x_i}=(c_k-c_i)\frac{k-1}{k+1}-\frac{X^d_k-X_{-i}}{k+1}<0, \tag{51}$$

which reduces to

$$X_{-i}<X^d_k-(k-1)(c_k-c_i). \tag{52}$$

Consequently, (47) is firm i's best response to X_{-i} if

$$X^d_{k+1}-(k-1)(c_{k+1}-c_i)<X_{-i}<X^d_k-(k-1)(c_k-c_i). \tag{53}$$

This range is clearly not empty since $X^d_{k+1}<X^d_k$ and $c_{k+1}>c_k$.

Case ii) $r_i(X_{-i})=X^d_{k+1}-X_{-i}$: This requires that i's marginal profit jumps downward from positive to negative at $x_i=X^d_{k+1}-X_{-i}$, or

$$\lim_{x_i\to X^d_{k+1}-X_{-i}{}^+}\frac{\partial\pi_i(x_i,X_{-i})}{\partial x_i}<0 \qquad \text{and} \qquad \lim_{x_i\to X^d_{k+1}-X_{-i}{}^-}\frac{\partial\pi_i(x_i,X_{-i})}{\partial x_i}>0. \tag{54}$$

The first condition, immediate from (50), is

$$X_{-i}<X^d_{k+1}-(k-1)(c_{k+1}-c_i), \tag{55}$$

while the second condition requires

$$\lim_{x_i\to X^d_{k+1}-X_{-i}{}^-}\frac{\partial\pi_i(x_i,X_{-i})}{\partial x_i}=(c_{k+1}-c_i)\frac{k}{k+2}-\frac{X^d_{k+1}-X_{-i}}{k+2}>0, \tag{56}$$

which reduces to

$$X_{-i}>X^d_{k+1}-k(c_{k+1}-c_i). \tag{57}$$

So $X^d_{k+1}-X_{-i}$ is firm i's best response to X_{-i} if

$$X^d_{k+1}-k(c_{k+1}-c_i)<X_{-i}<X^d_{k+1}-(k-1)(c_{k+1}-c_i), \tag{58}$$

which is a non-empty interval for $c_i<c_{k+1}$.

The general form of the best-response function for firms $i=2,...,N$ is then

$$r_i(X_{-i})=\begin{cases} r_{i,N}(X_{-i}), & \text{if } X_{-i}\leq X^d_N-(N-1)(c_N-c_i),\\ X^d_N-X_{-i} & \text{if } X^d_N-(N-1)(c_N-c_i)<X_{-i}\leq X^d_N-(N-2)(c_N-c_i),\\ \dots & \\ r_{i,k}(X_{-i}) & \text{if } X^d_{k+1}-(k-1)(c_{k+1}-c_i)<X_{-i}\leq X^d_k-(k-1)(c_k-c_i),\\ X^d_k-X_{-i} & \text{if } X^d_k-(k-1)(c_k-c_i)<X_{-i}\leq X^d_k-(k-2)(c_k-c_i),\\ \dots & \\ r_{i,i}(X_{-i}) & \text{if } X^d_{i+1}-(i-2)(c_{i+1}-c_i)<X_{-i}<X^d_i\\ 0 & \text{if } X_{-i}\geq X^d_i. \end{cases} \tag{59}$$

Each component of the best-response in (59) is clearly downward-sloping, and since $r_{i,k}(X^d_k-(k-1)(c_k-c_i))=(k-1)(c_k-c_i)$ and $r_{i,k-1}(X^d_k-(k-2)(c_k-c_i))=(k-2)(c_k-c_i)$, the best-response functions are also continuous. □

Proof of Proposition 3.

We apply the same method used in the proof of Proposition 2, add X_{-i} to firm i's best-response, (59), determined in the proof of Lemma 2. However, the derivation of the inclusive best-response, $\tilde{r}_i(X)$, from $r_i(X_{-i})$ is complicated in this case due to the possibility of the equilibrium being a deterrence one. We show below that this possibility results in $\tilde{r}_i(X)$ being a correspondence rather than a function.

We first consider possible equilibria with at least two active firms. The possibility of a monopoly for firm one is considered separately below. Add X_{-i} to the individual best response we obtain the industry sales X compatible with firm i's optimization, which we denote $X^{(i)}(X_{-i})$:

$$X^{(i)}(X_{-i})=\begin{cases} \frac{N-1}{2N}\left(a+N\bar{c}_N-(N+1)c_i\right)+\frac{N+1}{2N}X_{-i}, & \text{if } X_{-i}\leq X_N^d-(N-1)(c_N-c_i),\\ X_N^d & \text{if } X_N^d-(N-1)(c_N-c_i)<X_{-i}\leq X_N^d-(N-2)(c_N-c_i),\\ \ldots & \\ \frac{k-1}{2k}\left(a+k\bar{c}_k-(k+1)c_i\right)+\frac{k+1}{2k}X_{-i}, & \text{if } X_{k+1}^d-(k-1)(c_{k+1}-c_i)<X_{-i}\leq X_k^d-(k-1)(c_k-c_i),\\ X_k^d & \text{if } X_k^d-(k-1)(c_k-c_i)<X_{-i}\leq X_k^d-(k-2)(c_k-c_i),\\ \ldots & \\ \frac{i-1}{2i}\left(a+i\bar{c}_i-(i+1)c_i\right)+\frac{i+1}{2i}X_{-i}, & \text{if } X_{i+1}^d-(i-2)(c_{i+1}-c_i)<X_{-i}<X_i^d\\ X_{-i} & \text{if } X_{-i}\geq X_i^d. \end{cases} \tag{60}$$

As each $r_i(X_{-i})$ is continuous, the function $X^{(i)}(X_{-i})$ is also continuous. However, it is only weakly increasing, since it is constant for $X_k^d-(k-1)(c_k-c_i)<X_{-i}\leq X_k^d-(k-2)(c_k-c_i)$.

In order to determine the inclusive best reply, we need to invert $X^{(i)}(X_{-i})$ to obtain the forward sales of i's rivals that generate aggregate sales of X when firm i plays its best-response. Since $X^{(i)}(X_{-i})$ is equal to a constant X_k^d for $k=i+1,\ldots,N-1$, when deterrence occurs, there is a range of values for X_{-i} that result in $X=X_k^d$. In particular, for $X_k^d-(k-1)(c_k-c_i)<X_{-i}<X_k^d-(k-2)(c_k-c_i)$ the inverse of the function $X^{(i)}(X_{-i})$ is a correspondence. We let $X_{-i}^{(i)}(X)$ denote this correspondence:

$$X_{-i}^{(i)}(X)=\begin{cases} \frac{2N}{N+1}X-\frac{N-1}{N+1}\left(a+N\bar{c}_N-(N+1)c_i\right), & \text{if } X<X_N^d,\\ \in\left(X_N^d-(N-1)(c_N-c_i),X_N^d-(N-2)(c_N-c_i)\right), & \text{if } X=X_N^d,\\ \ldots & \\ \frac{2k}{k+1}X-\frac{k-1}{k+1}\left(a+k\bar{c}_k-(k+1)c_i\right), & \text{if } X_{k+1}^d<X<X_k^d,\\ \in\left(X_k^d-(k-1)(c_k-c_i),X_k^d-(k-2)(c_k-c_i)\right), & \text{if } X=X_k^d,\\ \ldots & \\ \frac{2i}{i+1}X-\frac{i-1}{i+1}\left(a+i\bar{c}_i-(i+1)c_i\right), & \text{if } X_{i+1}^d<X<X_i^d\\ X_{-i} & \text{if } X\geq X_i^d. \end{cases} \tag{61}$$

We now can solve for x_i in $X-x_i=X_{-i}^{(i)}(X)$ to obtain the inclusive best-response, $\tilde{r}_i(X)$:

$$\tilde{r}_i(X)=\begin{cases} \frac{N-1}{N+1}(a+N\bar{c}_N-(N+1)c_i)-\frac{N-1}{N+1}X, & \text{if } X<X_N^d, \\ \in((N-2)(c_N-c_i),(N-1)(c_N-c_i)), & \text{if } X=X_N^d, \\ \dots & \\ \frac{k-1}{k+1}(a+k\bar{c}_k-(k+1)c_i)-\frac{k-1}{k+1}X, & \text{if } X_{k+1}^d<X<X_k^d, \\ \in((k-2)(c_k-c_i),(k-1)(c_k-c_i)), & \text{if } X=X_k^d, \\ \dots & \\ \frac{i-1}{i+1}(a+i\bar{c}_i-(i+1)c_i)-\frac{i-1}{i+1}X, & \text{if } X_{i+1}^d<X<X_i^d \\ 0 & \text{if } X\geq X_i^d. \end{cases} \tag{62}$$

The general procedure for finding the equilibrium to aggregative games would now sum the individual inclusive best-responses and solve for the equilibrium X. It is not quite so straightforward in this case due to the inclusive best-responses being correspondences. However, as the condition for individual firms to choose a deterrence level of forward sales is not firm-specific ($X=X_k^d$), we can aggregate the inclusive best-response correspondences to form an aggregate inclusive best-response function. Either the aggregate inclusive best-response is corresponds to a non-deterrence level of forward sales, $\sum_{i=1}^k \tilde{r}_i(X)\in(X_{k+1}^d,X_k^d)$, or it corresponds to a deterrence level of forward sales, $\sum_{i=1}^k \tilde{r}_i(X)=X_{k+1}^d$. We examine these two possibilities in turn, deriving the conditions on model parameters under which they obtain.

If $X^*\in(X_{k+1}^d,X_k^d)$, aggregating (62) yields

$$X^*=\frac{k(k-1)}{k+1}(a-\bar{c}_k-X^*) \quad\Longrightarrow\quad X^*=\frac{k(k-1)}{k^2+1}(a-\bar{c}_k). \tag{63}$$

For this to be an equilibrium it must lie in the interval (X_{k+1}^d,X_k^d). For $X^*<X_k^d$ we require

$$\frac{k(k-1)}{k^2+1}(a-\bar{c}_k)<a+k\bar{c}_k-(k+1)c_k, \tag{64}$$

which can be expressed as

$$a>c_k+k^2(c_k-\bar{c}_k)\equiv\alpha_k. \tag{65}$$

For $X^*>X_{k+1}^d$ we require

$$\frac{k(k-1)}{k^2+1}(a-\bar{c}_k)>a-(k+2)c_{k+1}+(k+1), \tag{66}$$

which reduces to

$$a<c_{k+1}+k^2(c_{k+1}-\bar{c}_k)\equiv\alpha_k^d. \tag{67}$$

Hence, the equilibrium is of the non-deterrence type with k active firms if $\alpha_k<a<\alpha_k^d$. To allow this statement to be applied for $k=N$, we define $\alpha_N^d=\infty$, since if N firms are active, there is no other firm to deter.

Notice that the activity threshold level of demand for firm $k+1$ is strictly higher than the deterrence threshold for firm $k+1$: $\alpha_{k+1}=c_{k+1}+(k+1)^2(c_{k+1}-\bar{c}_{k+1})>c_{k+1}+k^2(c_{k+1}-\bar{c}_k)=\alpha_k^d$. For $a\in[\alpha_k^d,\alpha_{k+1}]$, the equilibrium is a deterrence one, with k active firms producing X_{k+1}^d in aggregate. Although aggregate forward sales are uniquely determined in this case, individual forward sales are not. There are a continuum of equilibrium forward sales vectors that must satisfy $x_i^*\in((k-2)(c_k-c_i),(k-1)(c_k-c_i))$ and $X^*=X_{k+1}^d=\sum_{i=1}^k x_i^*$.

Finally, consider the possibility that the equilibrium has only one active firm, firm one. From the inclusive best reply, it should be noted that when $X{>}X_2^d$, $x_1(X){=}0$.[36] Since $X{=}0$ in this case, it can only be an equilibrium if $X_2^d{<}0$, or $a{<}2c_2{-}c_1{=}\alpha_1^d$, in which case firm two's activity is blockaded. We define $\alpha_1{=}c_1$, so this case occurs for $a{\in}(\alpha_1,\alpha_1^d)$, firm one is active on the spot market even though it is not on the forward sales market. For $a{\in}[\alpha_1^d,\alpha_2]$, firm one chooses $x_1^*{=}X_2^d$, deterring the activity of firm two.

In summary, given $(c_1,c_2,...,c_N)$, there are threshold levels of demand, $\alpha_1{<}\alpha_1^d{<}...{<}\alpha_{N-1}^d{<}\alpha_N{<}\alpha_N^d{=}\infty$, for which the number of active firms in equilibrium is given by $n^*{=}\max\{k|a{>}\alpha_k\}$. The equilibrium is a non-deterrence one if $\alpha_{n^*}{\le}a{<}\alpha_{n^*}^d$ and a deterrence one if $\alpha_{n^*}^d{\le}a{<}\alpha_{n^*+1})$. In each case, the aggregate level of forward sales, X^*, is unique, and, consequently, so is the equilibrium price $p_F^*{=}p_S^*(X^*)$.

□

Proof of Proposition 4.

a) For $a{\in}(\alpha_n,\alpha_n^d]$, Proposition 3 establishes that aggregate forward sales are $X^{nd}(n){=}\frac{n(n-1)}{n^2+1}(a{-}\bar{c}_n)$. Using the inclusive best-response, (62), we have $x_i^{nd}(n){=}\tilde{r}_i(X^{nd}(n))$ we have

$$\begin{aligned}
x_i^{nd}(n)&=\frac{n-1}{n+1}\left(a+n\bar{c}-(n+1)c_i-\frac{n(n-1)}{n^2+1}(a-\bar{c})\right)\\
&=\frac{n-1}{n+1}\left(\frac{n+1}{n^2+1}a+\frac{n^2(n+1)}{n^2+1}\bar{c}-(n+1)c_i\right)\\
&=(n-1)\left(\frac{a+n^2\bar{c}_n}{n^2+1}-c_i\right) \qquad (68)
\end{aligned}$$

and price $p_F^{nd}(n){=}p_S^*(X^{nd}(n))$ or

$$\begin{aligned}
p_F^{nd}(n)&=\frac{a+n\bar{c}_n-X^{nd}(n)}{n+1},\\
&=\frac{a+n\bar{c}_n-\frac{n(n-1)}{n^2+1}(a-\bar{c}_n)}{n+1},\\
&=\frac{1}{n+1}\left(\frac{n+1}{n^2+1}a+\frac{n^2(n+1)}{n^2+1}\bar{c}_n\right),\\
&=\frac{a+n^2\bar{c}_n}{n^2+1}. \qquad (69)
\end{aligned}$$

b) For $a{\in}(\alpha_n^d,\alpha_{n+1}]$, Proposition (3) establishes that there is a deterrence equilibrium with aggregate forward sales of X_{n+1}^d, and, consequently, a forward price of $p_F^*{=}c_{n+1}$. From the best response functions given in (59) in the proof of Lemma 2, firm $i{\le}n$ will choose the level of sales that deters firm $n{+}1$ if

$$X_{n+1}^d-(n)(c_{n+1}-c_i)<X_{-i}<X_{n+1}^d-(n-1)(c_{n+1}-c_i). \qquad (70)$$

The maximum sales that firm i can have in a deterrence equilibrium occurs when X_{-i} is at the lower bound of this range:

$$x_i^{max}=X_{n+1}^d-(X_{n+1}^d-(n)(c_{n+1}-c_i))=n(c_{n+1}-c_i). \qquad (71)$$

[36]Note that the efficient firm is not indifferent among all combinations of forward and spot sales adding up to the monopoly output in this case since any $x_1{>}0$ results in a price lower than the monopoly price as the firm cannot commit to not make additional spot sales at a lower price. This is essentially the same problem that is faced by a durable goods monopolist.

The minimum sales that firm i can have in a deterrence equilibrium occurs when X_{-i} is at the upper bound of this range:

$$x_i^{min}{=}X_{n+1}^d{-}(X_{n+1}^d{-}(n{-}1)(c_{n+1}{-}c_i)){=}(n{-}1)(c_{n+1}{-}c_i). \tag{72}$$

□

Proof of Proposition 5. Clearly for $a{=}\alpha_n^d$, $p_F^{nd}(n){-}c_{n+1}{=}0$ from the definition of α_n^d. For $a{=}\alpha_{n+1}$, $p_F^{nd}(n){=}\frac{c_{n+1}+(n+1)^2(c_{n+1}-\bar{c}_{n+1})+\bar{c}_n}{n^2+1}$, so

$$p_F^{nd}(n){-}c_{n+1}{=}\frac{c_{n+1}{+}(n^2{+}2n{+}1)c_{n+1}{-}(n{+}1)(\sum_{j=1}^n c_j{+}c_{n+1}){+}n\sum_{j=1}^n c_j{-}(n^2{+}1)c_{n+1}}{n^2{+}1}, \tag{73}$$

$$={\frac{n(c_{n+1}{-}\sum_{j=1}^n c_j)}{n^2{+}1}}, \tag{74}$$

$$={\frac{n(c_{n+1}{-}\bar{c}_n)}{n^2{+}1}}. \tag{75}$$

□

Proof of Proposition 6. i) follows directly from Proposition 3 and Definition 3.

To establish *ii)*, we first establish that $p_F^*{>}p_F^H$ for $a{\in}(\alpha_{N-1}^d,\alpha_N)$. Since $p_F^H{=}\frac{a+N^2\bar{c}_N}{N^2+1}$, at $a{=}\alpha_N$ we have $p_F^H{=}p_F^*{=}c_N$ by the definition of α_N. For $a{\in}[\alpha_{N-1}^d,\alpha_N)$, p_F^* remains equal to c_N, but $p_F^H{<}c_N$ since it is increasing in a and $a{<}\alpha_N$ on $[\alpha_{N-1}^d,\alpha_N)$. Continuity of p_F^* with respect to a implies that $p_F^*{>}p_F^H$ for $a{=}\alpha_{N-1}^d{-}\epsilon$ for some $\epsilon{>}0$. Consequently, there exists an $\hat{\alpha}{<}\alpha_{N-1}^d$ such that $p_F^*{>}p_F^H$ for $a{\in}(\hat{\alpha},\alpha_N)$.

For *iii)*, we wish to find conditions under which $\hat{\alpha}{<}\alpha_2$. First consider whether $\hat{\alpha}{<}\alpha_{N-1}$. As $\partial p_F^*/\partial a{=}\frac{1}{(N-1)^2+1}{>}\frac{1}{N^2+1}{=}\partial p_F^H/\partial a$, it follows that $p_F^*{>}p_F^H\ \forall a{\in}[\alpha_{N-1},\alpha_N]$ if, and only if, $p_F^H(\alpha_{N-1}){<}c_{N-1}$, where $p_F^H(\alpha_{N-1})$ refers to the value of p_F^H when $a{=}\alpha_{N-1}$. Conditional on $p_F^H(\alpha_{N-1}){<}c_{N-1}$, the same logic implies that $p_F^*{>}p_F^H\ \forall a{\in}[\alpha_{N-2},\alpha_{N-1}]$ iff $p_F^H(\alpha_{N-2}){<}c_{N-2}$. Applying this argument recursively establishes that $p_F^*{>}p_F^H\ \forall a{\in}[\alpha_2,\alpha_N)$ iff $p_F^H(\alpha_i){<}c_i\ \forall i{=}2,...,N{-}1$.

Using the definitions of p_F^H and α_n we have $p_F^H(\alpha_n){=}\frac{c_n+n^2(c_n-\bar{c}_n)+N^2\bar{c}_N}{N^2+1}{<}c_n$, which simplifies to $\bar{c}_N{<}c_n$, $n{=}2,...,N{-}1$. As $c_2{<}c_n$, for $n{=}3,...,N{-}1$, we have $p_F^*{>}p_F^H$, $\forall a{\in}[\alpha_2,\alpha_N)$ iff $c_2{>}\bar{c}_N$. □

Proof of Proposition 8. As the pre-merger equilibrium has n firms deterring firm $n{+}1$, we must have $\alpha_n^d{<}a{<}\alpha_{n+1}$. Consider a merger between firms i and j with $c_i{<}c_j$, so the merger essentially removes firm j. This merger has the effect of changing the deterred firms activity threshold from α_{n+1} to α'_{n+1}. The proposition is proved by showing that $\alpha'_{n+1}{<}\alpha_n^d$, so for the fixed level of demand a, firm $n{+}1$ will be active.

To derive α'_{n+1}, let $n'{=}n{-}1$ be the number of firms active pre-merger that remain post-merger. Firm $n{+}1$ will be active post-merger if $p_F^{nd'}(n'{+}1){>}c_{n+1}$, where $p_F^{nd'}(n'{+}1)$ is the price in a non-deterrence equilibrium with the n' surviving pre-merger active firms joined by firm $n{+}1$. This non-deterrence price is given by

$$p_F^{nd'}(n'{+}1){=}\frac{a{+}(n'{+}1)^2\bar{c}'_{n'+1}}{(n'{+}1)^2{+}1}$$

$$={\frac{a{+}n(\sum_{k\neq j}^n c_k{+}c_{n+1})}{n^2{+}1}}, \tag{76}$$

where we have used $n'{=}n{-}1$ and $\bar{c}'_{n'+1}{=}(\sum_{k\neq j}^{n} c_k{+}c_{n+1})/n$ which is the average marginal cost after replacing firm j with firm $n{+}1$. Setting $p_F^{nd'}(n'{+}1){=}c_{n+1}$ determines the new activity threshold for firm $n{+}1$:

$$\alpha'_{n+1}{=}(n^2{+}1)c_{n+1}{-}n(\sum_{k\neq j}^{n} c_k{+}c_{n+1}). \tag{77}$$

Using α_n^d from (28), we have $\alpha'_{n+1}{<}\alpha_n^d$ if

$$(n^2{+}1)c_{n+1}{-}n(\sum_{k\neq j}^{n} c_k{+}c_{n+1}){<}(n^2{+}1)c_{n+1}{-}n\sum_{k=1}^{n} c_k \tag{78}$$

$$n(\sum_{k=1}^{n} c_k{-}\sum_{k\neq j}^{n} c_k){-}nc_{n+1}{<}0 \tag{79}$$

$$n(c_j{-}c_{n+1}){<}0, \tag{80}$$

which is clearly true as $c_j{<}c_{n+1}$. As $a{>}\alpha_n^d$, it must be the case that post-merger $a{>}\alpha'_{n+1}$ and firm $n{+}1$ is active, so the post-merger number of active firms does not decline, establishing *i)*. The activity of firm $n{+}1$ implies that the merger must cause an increase in price, which, along with the increase in price when a merger occurs between active firms in a non-deterrence equilibrium, establishes *ii)*. □

Proof of Corollary 4. Profit to firm i pre-merger when selling $x_i^*{=}n(c_{n+1}{-}c_i)$ is $\pi_i^{pre}{=}n(c_{n+1}{-}c_i)^2$. If post-merger there is a non-deterrence equilibrium with n firms, profit is $\pi_i^{post}{=}n(\frac{a+n^{*2}\bar{c}_{n^*}}{n^{*2}+1}{-}c_i)^2{=}n(p_F^{nd}(n){-}c_i)^2$. Alternatively, if post-merger there is a deterrence equilibrium with firm $n{+}2$ deterred, profit is $\pi_i^{post}{=}n(c_{n+2}{-}c_i)^2$. Both $p_F^{nd}(n){>}c_{n+1}$ and $c_{n+2}{>}c_{n+1}$, so post-merger profit is higher in either cases. □

Proof of Proposition 10. We demonstrate that with T periods of trading prior to the spot market (which meets in period $T{+}1$) the equilibrium forward price in the duopoly example when firm two is active is equal to that in Allaz & Vila (1993) when each firm has marginal cost $c/2$, $p_F^{AV}(T){=}c/2{+}(a{-}c/2)/(3{+}2T)$, which we simplify here to $p_F^{AV}(T){=}\frac{a+(T+1)c}{3+2T}$. The price for $T{+}1$ given in Proposition 1 satisfies this. Let $\mathcal{X}_t$ denote the cumulative aggregate forward sales at the start of period t, and x_{it} the forward sales by firm i in period t. For the duopoly example we have $x_{11}{=}(a{+}2c)/5$, $x_{21}{=}(a{-}3c)/5$, $\mathcal{X}_2{=}(2a{-}c)/5$, and $\mathcal{X}_1{=}0$.

Working backwards, in period $T{+}1$ the spot market equilibrium price is simply that in Proposition 1 but with $\mathcal{X}_{T+1}$ replacing X. Let $V_{it}(\mathcal{X}_t)$ represent the value function for firm i with $\mathcal{X}_t$ the cumulative forward sales prior to t. For the spot market period we have $V_{iT+1}(\mathcal{X}_{T+1}){=}(p_S^*(\mathcal{X}_{T+1}){-}c_i)y_i^*(\mathcal{X}_{T+1})$, which, using the results from the duopoly section yields

$$V_{1T+1}(\mathcal{X}_{T+1}){=}\left(\frac{a{-}\mathcal{X}_{T+1}{+}c}{3}\right)^2 \tag{81}$$

and

$$V_{2T+1}(\mathcal{X}_{T+1}){=}\left(\frac{a{-}\mathcal{X}_{T+1}{-}2c}{3}\right)^2. \tag{82}$$

In period T, firm one solves

$$\max_{x_{1T}} p_{FT}x_{1T}{+}V_{1T+1}(\mathcal{X}_{T+1}), \tag{83}$$

which, using $p_{FT}{=}p_S^*(\mathcal{X}_{T+1})$ and $\mathcal{X}_{T+1}{=}\mathcal{X}_T{+}x_{1T}{+}x_{2T}$, can be written as

$$\max_{x_{1T}} p_S^*(\mathcal{X}_T{+}x_{1T}{+}x_{2T})x_{1T}{+}V_{1T+1}(\mathcal{X}_T{+}x_{1T}{+}x_{2T}). \tag{84}$$

Assuming a non-deterrence equilibrium, the necessary condition for this maximization problem is

$$p_S^*(\mathcal{X}_T{+}x_{1T}{+}x_{2T}){+}x_{1T}p_S^{*'}{+}V_{1T+1}'{=}0, \tag{85}$$

or

$$\frac{1}{3}\left(a{-}\mathcal{X}_T{-}x_{1T}{-}x_{2T}{+}c{-}x_{1T}{-}\frac{2}{3}(a{-}\mathcal{X}_T{-}x_{1T}{-}x_{2T}{+}c)\right){=}0, \tag{86}$$

from which we have firm one's inclusive best response

$$x_{1T}{=}\frac{a{-}\mathcal{X}_T{-}X_T{+}c}{3}. \tag{87}$$

A similar procedure applied to firm two yields from which we have firm one's inclusive best response

$$x_{2T}{=}\frac{a{-}\mathcal{X}_T{-}X_T{-}2c}{3}. \tag{88}$$

Adding together these inclusive best responses and solving for X_T results in

$$X_T{=}\frac{2(a{-}\mathcal{X}_T){-}c}{5} \tag{89}$$

and using this in $p_{FT}^*{=}p_S^*(\mathcal{X}_T{+}X_T)$ gives

$$p_{FT}^*(\mathcal{X}_T){=}\frac{a{-}\mathcal{X}_T{+}2c}{5}. \tag{90}$$

Notice that for $\mathcal{X}_T{=}0$, this is the solution found in the duopoly example section, and as $\mathcal{X}_T{=}0$ for $T{=}1$, we have $p_{F1}^*(0){=}\frac{a+2c}{5}{=}p_F^{AV}(1)$, and using $p_F^{AV}(1){=}c$ we get firm two's activity threshold $\alpha_2(1){=}3c$. Using this solution in the value function we have

$$\begin{aligned}V_{1T}(\mathcal{X}_T)&{=}p_{FT}^*(\mathcal{X}_T)x_{1T}{+}V_{1T+1}(\mathcal{X}_t{+}X_T)\\&{=}2\left(\frac{a{-}\mathcal{X}_T{+}2c}{5}\right)^2\end{aligned} \tag{91}$$

and

$$\begin{aligned}V_{2T}(\mathcal{X}_T)&{=}(p_{FT}^*(\mathcal{X}_T){-}c)x_{2T}{+}V_{2T+1}(\mathcal{X}_t{+}X_T)\\&{=}2\left(\frac{a{-}\mathcal{X}_T{-}3c}{5}\right)^2.\end{aligned} \tag{92}$$

Applying the same approach for period $T{-}1$, firms one and two solve

$$\max_{x_{1T-1}}\ p_S^*(\mathcal{X}_{T-1}{+}x_{1T-1}{+}x_{2T-1})x_{1T-1}{+}V_{1T}(\mathcal{X}_{T-1}{+}x_{1T-1}{+}x_{2T-1}) \tag{93}$$

and

$$\max_{x_{2T-1}}\ (p_S^*(\mathcal{X}_{T-1}{+}x_{1T-1}{+}x_{2T-1}){-}c)x_{2T-1}{+}V_{2T}(\mathcal{X}_{T-1}{+}x_{1T-1}{+}x_{2T-1}). \tag{94}$$

The same solution procedure for a non-deterrence equilibrium results in

$$X_{T-1}{=}\frac{2(a{-}\mathcal{X}_{T-1}){-}c}{7},\ \text{and} \qquad p_{FT-1}{=}\frac{a{-}\mathcal{X}_{T-1}{+}3c}{7} \tag{95}$$

so

$$V_{1T-1}(\mathcal{X}_{T-1}){=}3\left(\frac{a{-}\mathcal{X}_{T-1}{+}3c}{7}\right)^2,\ \text{and} \qquad V_{2T-1}(\mathcal{X}_{T-1}){=}3\left(\frac{a{-}\mathcal{X}_{T-1}{-}4c}{7}\right)^2. \tag{96}$$

If $T{=}2$, we use $\mathcal{X}_1{=}0$ to establish $p^*_{F1}(0){=}\frac{a+3c}{7}{=}p^{AV}_F(2)$ and so $\alpha_2(2){=}4c$.

Applying another iteration the same procedure results in the period $T-2$ non-deterrence equilibrium price of

$$p_{FT-2}{=}\frac{a-\mathcal{X}_{T-2}+4c}{9}. \tag{97}$$

For $T{=}3$, $\mathcal{X}_1{=}0$ and we have $p^*_{F1}(0){=}\frac{a+4c}{9}{=}p^{AV}_F(3)$ and $\alpha_2(3){=}5c$.

Tabulating the equilibrium price and firm two's activity threshold for $T{=}1,2,3$ gives

T	1	2	3	...
p^*_{F1}	$\frac{a+2c}{5}$	$\frac{a+3c}{7}$	$\frac{a+4c}{9}$	...
$\alpha_2(T)$	$3c$	$4c$	$5c$	...

The pattern followed by the forward price as T increases is the same as that in Allaz & Vila (1993): the equilibrium forward price in the first period of T trading periods is $p^*_{F1}(0){=}\frac{a+(T+1)c}{3+2T}{=}p^{AV}_F(T)$ which results in an activity threshold for firm two of $\alpha_2(T){=}(T{+}2)c$, establishing *i)* of the proposition. Given a level of demand a, $a{<}\alpha_2(T)$ occurs for $T{>}a/c-2$, which is finite. So the Bertrand outcome, price equal to firm two's marginal cost, occurs with a finite number of forward trading periods, which establishes *ii)* of the proposition. □

B Supplementary material

B.1 The effect of deterrence on price

The maximum price reduction that is attributable to deterrence can be expressed as a proportion of the deterrence price:

$$\frac{p_F^{nd}(n)-c_{n+1}}{c_{n+1}}=\frac{n}{n^2+1}\left(1-\frac{\bar{c}_n}{c_{n+1}}\right). \tag{98}$$

This depends only on the number of active firms, and the marginal cost of the deterred firm relative to the average marginal cost of active firms. We tabulate the percentage price effect for a variety of values of n and $\bar{c}_n/c_{n+1}$ in Table 2.

n vs $\bar{c}_n/c_{n+1}$	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%
1	25%	23%	20%	18%	15%	13%	10%	8%	5%	3%
2	20%	18%	16%	14%	12%	10%	8%	6%	4%	2%
3	15%	14%	12%	11%	9%	8%	6%	5%	3%	2%
4	12%	11%	9%	8%	7%	6%	5%	4%	2%	1%
5	10%	9%	8%	7%	6%	5%	4%	3%	2%	1%
6	8%	7%	6%	6%	5%	4%	3%	2%	2%	1%
7	7%	6%	6%	5%	4%	4%	3%	2%	1%	1%
8	6%	6%	5%	4%	4%	3%	2%	2%	1%	1%
9	5%	5%	4%	4%	3%	3%	2%	2%	1%	1%
10	5%	4%	4%	3%	3%	2%	2%	1%	1%	0%
11	5%	4%	4%	3%	3%	2%	2%	1%	1%	0%
12	4%	4%	3%	3%	2%	2%	2%	1%	1%	0%
13	4%	3%	3%	3%	2%	2%	2%	1%	1%	0%
14	4%	3%	3%	2%	2%	2%	1%	1%	1%	0%
15	3%	3%	3%	2%	2%	2%	1%	1%	1%	0%
16	3%	3%	2%	2%	2%	2%	1%	1%	1%	0%
17	3%	3%	2%	2%	2%	1%	1%	1%	1%	0%
18	3%	2%	2%	2%	2%	1%	1%	1%	1%	0%
19	3%	2%	2%	2%	2%	1%	1%	1%	1%	0%
20	2%	2%	2%	2%	1%	1%	1%	1%	0%	0%

Table 2: Maximal price variation for various n and $\bar{c}_n/c_{n+1}$.

B.2 The most- and least-efficient deterrence equilibria

In a deterrence equilibrium with n active firms, the **most-efficient deterrence equilibrium** has individual advance sales concentrated as much as possible in efficient firms. This means that there is a "marginal" firm, firm m, for which sales of all firms $i<m$ are at their maximum defined in Proposition 4b), and sales of all firms $j>m$ are at their minimum defined in Proposition 4b). The marginal firm then sells the residual amount necessary for aggregate sales to equal the required deterrence level, X_n^d. Formally,

$$x_i^* = \begin{cases} n(c_{n+1}-c_i), & \text{for } i<m, \\ X_n^d - \sum_{j\neq m} x_j^*, & \text{for } i=m, \\ (n-1)(c_{n+1}-c_i), & \text{for } i>m. \end{cases} \tag{99}$$

The **least-efficient deterrence equilibrium** is determined similarly, except that the marginal firm, m' (not necessarily equal to m above), is defined by all firms more efficient than m' sell at their minimum and all firms less efficient than m' sell at their maximum:

$$x_i^* = \begin{cases} (n-1)(c_{n+1}-c_i), & \text{for } i<m', \\ X_n^d - \sum_{j\neq m'} x_j^*, & \text{for } i=m', \\ n(c_{n+1}-c_i), & \text{for } i>m'. \end{cases} \tag{100}$$

In each case, the marginal firm, m or m', is uniquely determined by X_n^d, and, hence, depends on the level of $a\in(\alpha_n^d, \alpha^n]$.