NATIONAL RESEARCH UNIVERSITY
HIGHER SCHOOL OF ECONOMICS

Sergey Kokovin. Mathieu Parenti.
Jacques-François Thisse. Philip Ushchev

ON THE DILUTION OF MARKET POWER

BASIC RESEARCH PROGRAM
WORKING PAPERS

SERIES: ECONOMICS
WP BRP 176/EC/2017

On the Dilution of Market Power*

Sergey Kokovin† Mathieu Parenti‡ Jacques-François Thisse§ Philip Ushchev¶

October 18, 2017

Abstract

We show that a market involving a handful of large-scale firms and a myriad of small-scale firms may give rise to different types of market structure, ranging from monopoly or oligopoly to monopolistic competition through new types of market structure. In particular, we find conditions under which the free entry and exit of small firms incentivizes big firms to sell their varieties at the monopolistically competitive prices, behaving as if in monopolistic competition. We call this result the *dilution of market power*. The structure of preferences is the main driver for a specific market structure to emerge as an equilibrium outcome.

Keywords: dominant firms, monopolistically competitive fringe, monopolistic competition, oligopoly, contestability.

JEL Classification: D43, F12 and L13.

*We gratefully acknowledge comments by D. Spulber, E. Zhelobodko and participants of ESEM and EARIE conferences. The study has been funded by the Russian Academic Excellence Project '5-100' and grant RFBR 16-06-00101.

†Novosibirsk State University and NRU-Higher School of Economics (Russia). Email: skokov7@gmail.com

‡ECARES, Université Libre de Bruxelles and CEPR. Email: mathieu.parenti@gmail.com

§CORE-UCLouvain (Belgium), NRU-Higher School of Economics (Russia) and CEPR. Email: jacques.thisse@uclouvain.be

¶NRU-Higher School of Economics (Russia). Email: ph.ushchev@gmail.com

1 Introduction

According to Bruce D. Henderson, the founder of the Boston Consulting Group, "a stable competitive market never has more than three significant competitors." Using a sample of more than 160 U.S. industries, two base-time periods, and numerous performance measures, Uslay *et al.* (2010) find that most industries consist of three large generalists and numerous small and specialized producers, which succeed if they are able to operate in a niche market. However controversial the so-called "Rule of Three" may be, it seems unquestionable that many industries are dominated by a handful of big firms, which share the market with many small firms. Using a sample of $50,000$ U.S. firms, Hottman *et al.* (2016) observe that almost 90 percent of sales in a product group are produced by the 10 largest firms, while 98 percent of firms have market shares smaller than 2 percent. Similarly, the empirical trade literature documents the fact that a handful of firms account for a considerably large share of exports in many countries (Bernard *et al.*, 2012).

In this paper, we develop a simple, new theory which captures the interactions between big and small firms and fully characterizes the market outcome. The key idea is to combine two kinds of firms: a *discrete* number of atomic players, which represent big firms, and a *continuum* of nonatomic players, which represent small firms. Our setting thus blends oligopolistic and monopolistically competitive firms, which all produce different varieties of the same differentiated product. Formally, a big firm supplies a positive mass of varieties, whereas a small firm supplies a single variety. Owing to the difference in their product scope, firms adopt different attitudes toward competition: the big firms understand that they can strategically manipulate the market, whereas the small firms, which are each negligible to the market, treat the market conditions as given and choose their outputs (or prices) accordingly. Big and small firms differ here in *kind* because the latter are negligible to the market, while the former are not. This is to be contrasted with Melitz (2003) where the so-called big and small firms differ in *types* but not in kind because they are all negligible to the market (Neary, 2010).

In the wake of general-equilibrium models with imperfect competition (Hart, 1985) and the Stackelberg-Spence-Dixit literature on entry deterrence (Spence, 1977; Dixit, 1979), we assume that big firms are aware that their choices affect the size of the competitive fringe. More specifically, the market process is described as a two-stage game. At the first stage, a big firm chooses its outputs (or prices), anticipating the reactions of the small firms. At the second stage, the small businesses choose their outputs (or prices), treating the big firm's choice parametrically. In other words, the small firms enter or exit the market, whereas the big firms always stay in business. The same setting is used in the dominant firm model where one big firm chooses its sale price while anticipating the reactions of a large number of small firms which treat this price as a given (Markham, 1951). Note, however, the difference between this model, where the mass of small firms is exogenous, and ours, where it is endogenous. We discuss the staging of the game further in Section 2.

Our main finding is an effect new to the literature, which we christen the *dilution of market power*. The essence of this effect is that, despite being endowed with the ability to strategically manipulate the market, a large firm may find it rational to disregard this ability and mimic the behavior of small firms.[1] In other words, the large firms adopt the same aggressive pricing rule as the one followed by the firms belonging to the monopolistically competitive firms.[2] Therefore, *the market structure is observationally equivalent to monopolistic competition.*

We illustrate this effect by considering an economy where consumers have linear-quadratic preferences, while the supply side involves one large incumbent and a monopolistically competitive fringe. Whenever the equilibrium market structure is *hybrid*, i.e. it involves firms of both kinds, the dilution of the big firm's market power always occurs. But why is this so? Although the incumbent can manipulate the total output available on the market through its own output, it understands that the mass of small firms varies so that the equilibrium aggregate output is the same regardless of its own behavior. In other words, *the fringe acts as a buffer that stabilizes competition.* We will see that a necessary and sufficient condition for this to happen is that the big firm is not too big relative to the market. Otherwise, the incumbent chooses to either deter or blockade entry, as in Dixit (1979).

Next, we show that the dilution of market power keeps its relevance far beyond this simple example. The key-factor is that the demands faced by firms are *single-aggregate*, i.e. all the cross-effects in the demand system are captured by a scalar function whose value plays the role of a market aggregate (Pollak, 1972). This condition is satisfied for a broad class of demand systems, including those generated by additive preferences (Zhelobodko *et al.*, 2012), indirectly additive preferences (Bertoletti and Etro, 2017), and homothetic demand systems with a single aggregator (Matsuyama and Ushchev, 2017). We show that, when the market involves several big firms, all firms are heterogeneous and demands are single-aggregate, then any hybrid market outcome displays the dilution of market power. This is because the equilibrium market aggregate pinned down by free entry is independent of the big firms' choices. As a consequence,

[1] The dilution of market power has a strong contestability flavor. Indeed, the entry of small firms, which act non-strategically, suffices to discipline the big firm since this one chooses to sell its output as the small firms do. What is surprising (at least to us) is the fact that adopting such an aggressive behavior is rational on the part of the big firm. Note, however, the difference with Baumol (1980) and successors. The threat of entry per se is not sufficient here to make the market more competitive. It is the turnover of small firms that incentivizes the incumbent to behave as competitively as the small firms. Another major difference is that the key factor lies in the nature of preferences, rather than in cost considerations.

[2] By *pricing rule* we mean here an operator which maps the demand schedule faced by a firm into the profit-maximizing price as a function of the firm's marginal cost. Therefore, the same pricing rule does not mean that firms sell at the same price and share the same markups, as the big and small firms are likely to have different marginal costs.

both kinds of firms treat the market aggregate parametrically, but they do so for very different reasons: the small firms are by nature non-strategic, while the big firms accurately anticipate that the small firm's best response is flat in the domain of accommodated entry.

The dilution of market power also has a number of far-reaching implications. One such implication is the consequences of idiosyncratic technological shocks to large firms. Specifically, if a large firm is subject to, say, an exogenous productivity improvement, the equilibrium behavior of the other big firms remains the same. Another implication is that neither the emergence of new big firms nor the merger of a few of them give rise to major changes economy-wide. All these shocks only affect the size of the monopolistically competitive fringe, without causing any changes in the profit-maximizing strategies chosen by the firms in the fringe.

In contrast, the dilution of market power does not generally occur when the demand system involves two or several aggregates, as under quadratic preferences without an outside good (Demidova, 2017) and under homothetic preferences described by Kimball's (1995) flexible aggregator. In this case, big firms behave strategically. However, if we consider the possibility of multiple fringes (which may be interpreted, e.g., as populations of firms providing different quality levels), the dilution of market power is restored when the number of aggregates equals the number of fringes.

In short, *our setting allows for an endogenous determination of market structure.* Depending on the market size and the preference for diversity, there is either oligopolistic or monopolistic competition.

Related literature. The foregoing discussion shows that our paper is related to different strands of literature, including industrial organization, trade theory and general equilibrium under imperfect competition. In what follows, we discuss the most relevant contributions. It was shown in the 1970s that, when large traders are similar to each other, or when for each large trader there are small traders similar to it, the core of an exchange economy coincides with the set of competitive allocations. In other words, the market power of big traders is diluted (see Gabszewicz and Shitovitz, 1992, for a survey). Our results have a similar flavor. However, as suggested by Okuno *et al.* (1980), it is more natural to study such issues in a non-cooperative setting, which is what we accomplish in this paper. Industrial organization and trade models with multi-product firms include Allanson and Montagna (2005), Bernard *et al.* (2011), Dhingra (2013), Mayer *et al.* (2014). Although these authors use different approaches to model multi-product firms, they all assume that each firm is negligible to the market. We differ from all of them in that multi-product firms are able to manipulate the market. By showing that these firms may choose not to manipulate the market, we identify conditions for these various models to provide an accurate description of the functioning of markets.

Etro (2006, 2008) models the idea of big and small firms by assuming that a firm is big when it is the leader of a Stackelberg game and a firm is small when it is a follower. Hence, small firms are also able to manipulate the market outcome. Etro (2008) shows that the leaders are more aggressive than the followers when only big firms are free to enter the market in the second stage of the game. This difference in results is because, in our setting, the followers pursue the aggressive strategy of pricing at average cost. Therefore, the leaders cannot adopt a more aggressive behavior than the followers. Neary (2010) suggests a different approach in which firms choose to be big or small. Instead, we assume that firms are born big or small, but our results show that, under some conditions, the differences in kind is immaterial for the equilibrium outcome. Our paper is more directly linked to Shimomura and Thisse (2012) and Parenti (2017). Even though the dilution does not hold in their simultaneous game, these authors show that the presence of small firms incentivize the big firms to behave more aggressively than in a purely oligopolistic environment. This points to the same direction as the dilution of market power.

Even closer to us is the approach developed by Anderson *et al.* (2015). The great merit of this paper is to link together results that are a priori disparate. Once oligopolistic firms have chosen to enter the market, their profits depend on their own action and an aggregate of all other firms' actions. Therefore, each firm wants to manipulate the market aggregate, so that the dilution of market power never holds. Rather, we are interested in studying the impact of a monopolistically competitive fringe on the market outcome. Furthermore, unlike Anderson *et al.* (2015), we do not impose restrictions on the functional form of the aggregate, which is given here by *any* function mapping firms' strategies into a scalar. Finally, our analysis is not restricted to the case of a single aggregate. We show that the dilution of market power holds true when the number of aggregates and fringes is the same. For all these reasons, the two papers are to be viewed as complements rather than substitutes.

The rest of the paper is organized as follows. The aim of Section 2 is to highlight the effect we call the dilution of market power. To achieve this, we use a simple dominant-firm-type model with linear-quadratic preferences, one big firm and a monopolistically competitive fringe. Section 3 provides a general setting with single-aggregate preferences, which clarifies how far we can go with this result. Section 4 discusses the cases of multiple aggregates and multiple fringes. Section 5 concludes.

2 The dominant-firm model revisited

The economy involves two goods - a horizontally differentiated good and a homogenous good - and one production factor - labor. Each consumer supplies inelastically one unit of labor. The labor market is perfectly competitive and labor is chosen as the numéraire.

As discussed in the introduction, our aim is to study a *hybrid* market structure involving big and small firms characterized by different market behaviors: a *big* firm can manipulate the market aggregate that affects profits, whereas a large number of *small* firms are unable to influence this aggregate. The small firms react to the big firm's behavior by choosing to enter/exit the market as well as their output volume once they are in business. This dictates the following modeling strategy: the supply side of the economy involves (i) a continuum of mass M of single-product firms and (ii) one multi-product firm that supplies a given range $[0, n]$ of varieties. We refer to $n > 0$ as the *scope* of the big firm and to the set of small firms as the *monopolistically competitive fringe*. While M is endogenous, we treat n as exogenous to insulate the pure impact of the big firm's scope on the equilibrium market structure. For example, we do not have to make any assumption about cannibalization effects. However, we will discuss at the end of this section what happens when n is endogenous.

Preferences. On the demand side, there is a unit mass of identical consumers. Consumers share the same linear-quadratic preferences nested into a linear upper-tier utility (Ottaviano *et al.*, 2002):

$$\mathcal{U} = Z + \mathbb{X} - \frac{\beta}{2}\left(\int_0^M x_i^2 \mathrm{d}i + \int_0^n X_k^2 \mathrm{d}k\right) - \frac{\mathbb{X}^2}{2}. \tag{1}$$

where

$$\mathbb{X} \equiv \int_0^M x_i \mathrm{d}i + \int_0^n X_k \mathrm{d}k \tag{2}$$

is the total consumption (or total output since there is a unit mass of consumers) of the differentiated good, Z the consumption of a homogeneous good which is chosen as the numéraire, x_i the consumption of the variety provided by the small firm $i \in [0, M]$, and X_k is the consumption of the variety $k \in [0, n]$ provided by the large firm. The interaction across varieties is captured by $\mathbb{X}^2$. To ease the burden of notation, the coefficient of $\mathbb{X}$ in (1) is normalized to 1 by factorizing this coefficient, while the coefficient of $\mathbb{X}^2$ is also normalized to 1 by choosing appropriately the unit of the differentiated good. Hence, a lower value of β means a weaker love for variety, a larger market size, or both.

Each consumer is endowed with one unit of labor and an equal ownership share of all firms. She maximizes her utility subject to the budget constraint:

$$Z + \int_0^M p_i x_i \mathrm{d}i + \int_0^n P_k X_k \mathrm{d}k \leq y,$$

where p_i denotes the price of variety i and P_k the price of variety k, while $y > 0$ is consumer's income.

Inverse demands and profits. First-order conditions for utility maximization yield the following inverse demand functions for each variety $i \in [0, M]$ and each variety $k \in [0, n]$:

$$p(x_i, \mathbb{X}) = 1 - \mathbb{X} - \beta x_i, \qquad p(X_k, \mathbb{X}) = 1 - \mathbb{X} - \beta X_k. \tag{3}$$

For simplicity, we assume in this section that all marginal costs to zero and denote by $f > 0$ a small firm's fixed cost. Therefore, firm i's profits are given by

$$\pi_i = (1 - \mathbb{X} - \beta x_i) x_i - f.$$

Note that we have normalized the market size to 1. If the market size were given by L, the fixed cost f would be replaced with f/L in the analysis developed below. Therefore, a larger market is equivalent to a lower fixed cost f, which facilitates the entry of small firms.

The big firm's profits are given by

$$\mathfrak{N}(\mathbf{X}) = \int_0^n (1 - \mathbb{X} - \beta X_k) X_k \mathrm{d}k,$$

where $\mathbf{X} \equiv (X_k)_{k \in [0,n]}$ is the big firm's production profile.

Since it is negligible to the market, small firm accurately treats the total output $\mathbb{X}$ as a parameter. By contrast, (2) shows that the big firm expects its action to affect the value of $\mathbb{X}$ through its total output $\mathbb{X}_{mp}$ defined as follows:

$$\mathbb{X}_{mp} \equiv \int_0^n X_k \mathrm{d}k. \tag{4}$$

The timing of the game is as follows. The incumbent moves first and the monopolistically competitive firms second. Besides the reasons discussed in the introduction, this staging may be justified on the following grounds. First, the difference in entry behavior fits a fairly robust empirical fact, i.e. the survival probability of a firm is positively correlated with its size. Second, the above staging captures the idea that the big firm is committed to the market due to the large investment this firm has to make to build its production capacity. By contrast, the assumption of free entry and exit reflects the high turnover characterizing small firms in many industries, probably because these firms invest little money to be in business. We seek a subgame perfect Nash equilibrium and solve the game by backward induction.

2.1 Stage 2: The small firms' equilibrium strategies

Assume that a hybrid market structure prevails in equilibrium. In this case, a small firm observes the choice made by the big firm through the total output $\mathbb{X}_{mp}$ and chooses its profit-maximizing output. This yields the equilibrium output, price and profits as a function of $\mathbb{X}$:

$$x^*(\mathbb{X}) = \frac{1-\mathbb{X}}{2\beta}, \qquad p^*(\mathbb{X}) = \frac{1-\mathbb{X}}{2}, \qquad \pi^*(\mathbb{X}) = \frac{1}{4\beta}(1-\mathbb{X})^2 - f.$$

It is readily verified that the zero-profit condition $\pi^*(\mathbb{X}) = 0$ implies that equilibrium total output is a constant equal to

$$\mathbb{X}^* = \mathbb{X}_D \equiv 1 - 2\sqrt{\beta f} < 1. \tag{5}$$

For $\mathbb{X}_D > 0$, we assume that $\beta < 1/(4f)$; otherwise, the small firms never enter because f is too high. In this case, the equilibrium total output is *independent* of the big firm's behavior: it depends only upon the parameter β, which captures the intensity of preference for diversity, and the small firms' fixed cost f, which determines the easiness of entry into the fringe. Note also that $\mathbb{X}^*$ is independent of the scope n of the big firm.

Substituting $\mathbb{X}^*$ into $x^*(\mathbb{X})$ and $p^*(\mathbb{X})$, we determine the small firm's equilibrium output and its equilibrium price:

$$x^* = \sqrt{\frac{f}{\beta}} > 0, \qquad p^* = \sqrt{\beta f} > 0. \tag{6}$$

Therefore, small firms' equilibrium price and output are independent of the big firm's choice, which implies that the big firm influences the monopolistically competitive fringe through the mass of small firms only. In addition, the zero-profit condition implies that the small firms price their varieties at their average cost: $p^* = f/x^*$.

Evaluating (2) at a symmetric outcome yields $\mathbb{X} = \mathbb{X}_{mp} + Mx$. Solving for M, we obtain the equilibrium mass of small firms conditional upon $\mathbb{X}_{mp}$:

$$M^*(\mathbb{X}_{mp}) = (1-\mathbb{X}_{mp})\sqrt{\frac{f}{\beta}} - 2f. \tag{7}$$

This expression shows that the fringe acts as a *buffer* that stabilizes the total output at the value $\mathbb{X}^*$ through a change in the equilibrium mass of small firms. When the big firm increases (decreases) its total output $\mathbb{X}_{mp}$, the size of the fringe shrinks (expands).

Observe that $\mathbb{X}_{mp} = \mathbb{X}_D$ is the unique solution to the equation $M^*(\mathbb{X}_{mp}) = 0$. Therefore, the monopolistically competitive fringe disappears when $\mathbb{X}_{mp}$ is sufficiently large:

$$\mathbb{X}_{mp} \geq \mathbb{X}_D,$$

which implies that $\mathbb{X}_D$ is the big firm's minimal total output that deters the entry of small firms. As a result, we have:

$$\mathbb{X}^*(\mathbb{X}_{mp}) = \max\{\mathbb{X}_{mp}, \mathbb{X}_D\}. \tag{8}$$

As shown by Figure 1, the behavior of the big firm affects the equilibrium value of the market aggregate if and only if $\mathbb{X}_{mp}$ exceeds $\mathbb{X}_D$. Otherwise, the big firm behaves as if it had no market power. Indeed, this firm finds it rational to forego manipulating the equilibrium value of the market aggregate $\mathbb{X}$ by changing its total output $\mathbb{X}_{mp}$.

Insert Figure 1 about here

Thus, although the big firm is non-negligible to the market, things work *as if* it were so. As a consequence, any hybrid outcome in which the equilibrium level $\mathbb{X}^*_{mp}$ of the big firm's output is such that $\mathbb{X}^*_{mp} < \mathbb{X}_D$, is observationally equivalent to a purely monopolistically competitive equilibrium obtained by replacing the big firm with a mass n of small firms. We call this effect *dilution of market power* and discuss it in more detail in sub-section 2.4.

2.2 Stage 1: The large firm's equilibrium strategy

The big firm chooses its output anticipating the small firms' optimal responses. As a result, the big firm treats $\mathbb{X}^*$ as a given. Thus, the big firm's adjusted inverse demand for each its variety k, with or without a monopolistically competitive fringe, is defined by the following expression:

$$p(X_k, \mathbb{X}^*) = 1 - \mathbb{X}^*(\mathbb{X}_{mp}) - \beta X_k. \tag{9}$$

Combining (9) with (8), the big firm's profit function $\mathfrak{N}(\mathbf{X})$ may be written as follows:

$$\mathfrak{N}(\mathbf{X}) = \min\left\{\mathbb{X}_{mp} - \beta\int_0^n X_k^2 \mathrm{d}k - \mathbb{X}_{mp}^2,\ (1-\mathbb{X}_D)\mathbb{X}_{mp} - \beta\int_0^n X_k^2 \mathrm{d}k\right\}. \tag{10}$$

It follows immediately from (10) that the big firm's profit function is the minimum of two strictly quasi-concave functions, which implies that $\mathfrak{N}(\mathbf{X})$ is also strictly quasi-concave. Furthermore, as the profit function (10) is symmetric in $\mathbf{X}$, we can focus on symmetric production profiles only: $X_k = \mathbb{X}_{mp}/n$ for all $k \in [0, n]$. Indeed, for each asymmetric production profile $\mathbf{X}$ there exists a symmetric profile $\mathbf{X}'$ with the same level of aggregate output $\mathbb{X}_{mp}$ as $\mathbf{X}$ such that $\mathfrak{N}(\mathbf{X}) < \mathfrak{N}(\mathbf{X}')$.

Let $\Pi(\mathbb{X}_{mp})$ be the restriction of $\mathfrak{N}(\mathbf{X})$ to the diagonal, which depends only on the aggregate $\mathbb{X}_{mp}$:

$$\Pi(\mathbb{X}_{mp}) \equiv \begin{cases} \Pi_1(\mathbb{X}_{mp}) \equiv (1-\mathbb{X}_D)\mathbb{X}_{mp} - \frac{\beta}{n}\mathbb{X}_{mp}^2, & \text{if } \mathbb{X}_{mp} \leq \mathbb{X}_D, \\ \Pi_2(\mathbb{X}_{mp}) \equiv \mathbb{X}_{mp} - \left(1+\frac{\beta}{n}\right)\mathbb{X}_{mp}^2, & \text{if } \mathbb{X}_{mp} > \mathbb{X}_D. \end{cases}$$

The function Π is the lower envelope of Π_1 and Π_2, and the equilibrium value of $\mathbb{X}_{mp}$ is the maximizer of $\Pi(\mathbb{X}_{mp})$. This maximizer exists and is unique because $\Pi(\mathbb{X}_{mp})$ is continuous and strictly concave (see Figure 2).

2.3 The market equilibrium

We are now equipped to provide a full characterization of the equilibrium market structure.

Proposition 1. *Assume linear-quadratic preferences. (i) Entry is blockaded if and only if the scope of the incumbent is sufficiently broad:*

$$n > 2\beta\frac{\mathbb{X}_D}{1-2\mathbb{X}_D}. \tag{11}$$

(ii) Entry is deterred if and only if the scope of the incumbent is neither too broad nor too narrow:

$$2\beta\frac{\mathbb{X}_D}{1-\mathbb{X}_D} \leq n \leq 2\beta\frac{\mathbb{X}_D}{1-2\mathbb{X}_D}. \tag{12}$$

(iii) Entry is accommodated if and only if the scope of the incumbent is sufficiently narrow:

$$n < 2\beta\frac{\mathbb{X}_D}{1-\mathbb{X}_D}. \tag{13}$$

Hence, like in Dixit (1979), the following cases may arise: the entry of small firms is (i) blockaded, (ii) deterred, or (iii) accommodated by the incumbent. Dixit shows that the strategy chosen by the incumbent depends on the entry cost of the potential entrant relative to the market size. Proposition 1 shows under which conditions on the big firm's size each of these regimes emerges when entrants are small. The main distinctive feature of our approach lies in the endogenous entry and exit of small firms, a difference that has unsuspected (at least to us) implications.

The proof of Proposition 1 goes as follows. Observe that $\Pi(\mathbb{X}_{mp})$ is the lower envelope of two concave parabolas, while $\Pi_1(\mathbb{X}_{mp})$ and $\Pi_2(\mathbb{X}_{mp})$ satisfy the following properties: (i) $\Pi_1(0) = \Pi_2(0) = 0$, (ii) $\Pi_1(\mathbb{X}_D) = \Pi_2(\mathbb{X}_D) = \mathbb{X}_D - (1+\beta/n)\,\mathbb{X}_D^2$, and (iii) $\Pi_1'(\mathbb{X}_D) > \Pi_2'(\mathbb{X}_D)$. As illustrated by Figure 2, three cases may arise.

Insert Figure 2 about here

Blockaded entry. Assume that $\Pi_1'(\mathbb{X}_D) > \Pi_2'(\mathbb{X}_D) > 0$ (see Figure 2a). Since $\Pi_1(\mathbb{X}_{mp})$ and $\Pi_2(\mathbb{X}_{mp})$ are both increasing in the neighborhood of $\mathbb{X}_D$, the maximizer of $\Pi(\mathbb{X}_{mp})$ exceeds $\mathbb{X}_D$, which implies that the big firm is an unconstrained monopolist. It is readily verified that $\Pi_2'(\mathbb{X}_D) > 0$ holds if and only if the incumbent firm's scope is given by (11).

Since $\Pi_2(\mathbb{X}_{mp})$ is strictly concave, the big firm's profit-maximizing output and profits are obtained from the first-order-condition:

$$\mathbb{X}_{mp}^*(n) = \frac{n}{2(\beta+n)}, \qquad \Pi^*(n) = \frac{n}{4(\beta+n)}.$$

In short, if the incumbent is large enough, entry is sufficiently costly, or both, then the monopolist may accurately ignore the potential entry of small firms. Otherwise, entry can never be blockaded because the market size is too large for the big firm to ignore the small firms. In this case, does the incumbent deter or accommodate entry?

Entry deterrence. Assume now that $\Pi_1'(\mathbb{X}_D) > 0 > \Pi_2'(\mathbb{X}_D)$ (see Figure 2b). In this case, the maximizer $\mathbb{X}_{mp}^*$ of $\Pi(\mathbb{X}_{mp})$ is exactly the kink $\mathbb{X}_D$. Indeed, the incumbent chooses an output preventing the entry of small firms if and only if $\mathbb{X}_{mp}^* \geq \mathbb{X}_D$, which is equivalent to $M^*(\mathbb{X}_{mp}^*) = 0$. The inequalities $\Pi_1'(\mathbb{X}_D) > 0 > \Pi_2'(\mathbb{X}_D)$ hold if and only if the breadth of the big firm's scope is given by (12).

Since $\mathbb{X}_D$ is the smallest value of $\mathbb{X}$ that deters entry, it has the nature of a "limit output", which is the quantity counterpart of limit pricing. Under entry deterrence, the incumbent's total output and profits are equal to

$$\mathbb{X}_{mp}^*(n) = \mathbb{X}_D, \qquad \Pi^*(n) = \mathbb{X}_D\left[1 - \left(1 + \frac{\beta}{n}\right)\mathbb{X}_D\right].$$

Accommodating entry. Last, assume that $\Pi_2'(\mathbb{X}_D) < \Pi_1'(\mathbb{X}_D) < 0$ (see Figure 2c). Since $\Pi_1(\mathbb{X}_{mp})$ and $\Pi_2(\mathbb{X}_{mp})$ are both decre0asing in $\mathbb{X}_{mp}$ in the neighborhood of $\mathbb{X}_D$, so does $\Pi(\mathbb{X}_{mp})$. Therefore, the maximizer of $\Pi(\mathbb{X}_{mp})$ is smaller than $\mathbb{X}_D$. This occurs if and only if the big firm's scope n is small enough relative to the market size and given by (13). Maximizing $\Pi_1(\mathbb{X}_{mp})$ with respect to $\mathbb{X}_{mp}$ yields:

$$\mathbb{X}_{mp}^*(n) = \frac{n}{2\beta}(1 - \mathbb{X}_D) = n\sqrt{\frac{f}{\beta}}, \qquad \Pi^*(n) = \frac{n}{4\beta}(1 - \mathbb{X}_D)^2 = nf.$$

To sum up, as n steadily decreases, the incumbent's profits and markups (weakly) decrease, thus implying that the big firm's market power fades away. During this process, *the market outcome displays a gradual transition from pure monopoly to monopolistic competition* through entry deterrence. Note that the same holds when the market size steadily rises while n remains constant. Since accommodated entry leads to new and unsuspected results, we focus below on the case of hybrid markets.

2.4 The dilution of market power

When entry is accommodated, the incumbent faces an inverse demand for variety k that accounts for the mass $M^*(\mathbb{X}_{mp})$ of small firms that enter the market in the second stage. Although the incumbent is a priori able to manipulate $\mathbb{X}$ through $\mathbb{X}_{mp}$, it anticipates that the mass (7) of small firms will adjust in a way such that the equilibrium value of $\mathbb{X}$ is always equal to $\mathbb{X}^*$ regardless of the value taken by $\mathbb{X}_{mp}$. In other words, the incumbent accurately treats $\mathbb{X}^*$ as a parameter.

Plugging (5) into (9), it is straightforward to show that the profit-maximizing output and price of variety k under accommodating entry are given by

$$X^* = \sqrt{\frac{f}{\beta}}, \qquad P^* \equiv \sqrt{\beta f},$$

which are the same as those given by (6), that is, the equilibrium output and price of a small firm. Therefore, the big firm's equilibrium total output is $\mathbb{X}_{mp}^* = nx^*$, while its equilibrium profits are $\Pi^*(n) = nf$.

Stated differently, when the large firm accommodates the presence of small firms, the former chooses to sell the quantity x^* given by (6) for each of its varieties, which it prices at the same level as the varieties sold by the latter. In other words, if the incumbent is not "too" big relative to the market, *there is monopolistic competition.* In other words, the big firm's market power is *dissolved* in an ocean of small firms. This amounts to saying that the incumbent behaves like a *multidivisional firm* in which autonomous "profit centers" produce each a specific variety and maximize their own profits, while ignoring demand linkages within the firm's product range. While the divisionalization of a firm is often justified by the desire to prevent pyramiding management costs, our model provides a justification driven by the demand side only.

This shows that a firm able to manipulate the market may find it profit-maximizing to disregard its strategic power, an effect which we christen *the dilution of market power*. The above example and the analysis developed in the next section show that the dilution of market power owns nothing to the cost side; it is fully driven by the structure of the demand size. Note also that the empirical evidence suggests that the small firms may be significant in number, while their market share is very small. However, the size of the monopolistically competitive fringe is immaterial for the dilution of market power to hold. What matters is the existence of entry and exit flows in the fringe.

2.5 Endogenizing the big firm's product range

So far, we have assumed that the size n of the product range was given. In line with Spulber (1981), we assume that the established firm chooses its scope n before its output. At the stage 0 of the game, the big firm's profit function is as follows:

$$\Pi^*(n) = \begin{cases} \frac{n}{4(\beta+n)} & \text{if } n > 2\beta\frac{\mathbb{X}_D}{1-2\mathbb{X}_D}, \\ \mathbb{X}_D\left[1 - \left(1+\frac{\beta}{n}\right)\mathbb{X}_D\right] & \text{if } 2\beta\frac{\mathbb{X}_D}{1-\mathbb{X}_D} \leq n \leq 2\beta\frac{\mathbb{X}_D}{1-2\mathbb{X}_D}, \\ fn & \text{if } n < 2\beta\frac{\mathbb{X}_D}{1-\mathbb{X}_D}, \end{cases}$$

which can be shown to be strictly increasing, concave and once continuously differentiable (see Figure 3 for an illustration).

Insert Figure 3 here

Building on the spatial model of flexible manufacturing (Eaton and Schmitt, 1994) and following Mayer *et al.* (2014) who use linear-quadratic preferences, we assume that the big firm has a baseline variety at $k = 0$, which corresponds to its core competency. To supply another variety, the firm must incur an additional cost that increases with the "distance" to its baseline variety. For simplicity, we assume that the development cost of variety k is given by $c(k) = tk$ with $t > 0$. When the big firm's scope is n, the total cost is therefore equal to tn^2. However, our results remain qualitatively the same for more general specifications of the development cost $c(k)$.[3]

Maximizing $\Pi^*(n) - tn^2$ with respect to n and using Proposition 1 yields the following result.

[3] Our conclusions are unaffected when we account for scope economies by assuming that total costs are given by $F + tn^2$ where $F > 0$ is the cost of launching a R&D division.

Proposition 2. *Assume core competencies with $c(k) = tk$. (i) Entry is blockaded if and only if:*

$$t < \frac{(1-2\mathbb{X}_D)^3}{8\beta^2\mathbb{X}_D}.$$

(ii) Entry is deterred if and only if:

$$\frac{(1-2\mathbb{X}_D)^3}{8\beta^2\mathbb{X}_D} \leq t \leq \frac{f}{4\beta}\frac{1-\mathbb{X}_D}{\mathbb{X}_D}.$$

(iii) Entry is accommodated if and only if:

$$t > \frac{f}{4\beta}\frac{1-\mathbb{X}_D}{\mathbb{X}_D}. \tag{14}$$

Hence, we get the following intuitive conditions: the big firm accommodates entry, hence the dilution of market power holds, when moving away from core competency is costly, the market is large, and/or the preference for diversity is strong. In the remaining sections, we will focus on the regime of accommodated entry.

3 The dilution of market power under single-aggregate preferences

It is tempting to argue that the dilution of market power is an artefact of linear-quadratic preferences. In this section, we show that this property survives under a much more general condition, i.e. firms' inverse demand depends only on the firm's output and a *scalar* that aggregates the decisions made by all firms (Pollak, 1972). Consider $N \geq 1$ big firms, with firm $j = 1, .., N$ supplying each a mass $n_j > 0$ of varieties and producing at marginal cost $C_j > 0$. We discuss below the case where large firms can freely enter the market.

The small firms are heterogeneous in the sense of Melitz (2003). Prior to entry the small firms face uncertainty about their marginal cost but know the continuous distribution $\Gamma(c)$ from which the marginal cost c is drawn. To enter the market, the small firms must bear a sunk cost f_e. After entry, each firm observes its marginal cost c. In addition, an active c-type firm must incur a fixed production cost f, so that producing the quantity x_c involves a cost equal to $f + cx_c$.

Consider preferences such that firms face *single-aggregate inverse demands* $p(\cdot, \Lambda)$ where Λ is a scalar market aggregate which accounts for all the cross-effects within the demand system (examples are given below).[4] In this case, firms' profit functions may be written as follows:

$$\Pi_j = \int_0^{n_j} [p(X_{jk}, \Lambda) - C_j]\, X_{jk} \mathrm{d}k, \qquad j = 1, ..., N, \tag{15}$$

$$\pi(c) = [p(x_c, \Lambda) - c]\, x_c - f, \qquad c \in [0, \bar{c}], \tag{16}$$

where $\bar{c}$ is the cutoff cost.[5]

Assumption SA. *The inverse demands are single-aggregate and given by* $p(\cdot, \Lambda)$ *where the value of* Λ *is unaffected by the action of a single small firm. These demands are such that firms' profits* (15) *and* (16) *are continuous, strictly quasi-concave in their own strategy for all admissible* Λ*, and strictly monotone in the aggregate* Λ *for all admissible* X_{jk} *and* x_c*.*[6]

The strict quasi-concavity assumption is made for the best reply to be a well-defined function. That π_c strictly decreases (increases) with Λ means that the market aggregate Λ is a substitute (complement) of x_c.

Assumption SA is satisfied by the linear demand system (3) where $\Lambda = \mathbb{X}$. Other examples of demand systems that are widely used in industrial organization and international trade are given below.

1. Additive preferences. Consider the inverse demands derived from *additive preferences* (e.g. the CES and CARA):

$$\mathcal{U} = M_e \int_0^{\bar{c}} u(x_c) \mathrm{d}\Gamma(c) + \sum_{j=1}^{N} \int_0^{n_j} U_j(X_{jk}) \mathrm{d}k,$$

where u and U_j are strictly increasing and concave, with $u(0) = U_j(0) = 0$, while M_e is the mass of entrants. Denoting by λ the Lagrange multiplier of the budget constraint, the utility-maximizing conditions yield the following inverse demand functions:

$$P_{jk}(X_{jk}, \lambda) = \frac{U_j'(X_{jk})}{\lambda}, \qquad p(x_c, \lambda) = \frac{u'(x_i)}{\lambda}.$$

In this case, the market aggregate, which is the marginal utility of income, is pinned down by the budget constraint:

$$\Lambda \equiv \lambda = M_e \int_0^{\bar{c}} x_c u'(x_c) \mathrm{d}\Gamma(c) + \sum_{j=1}^{N} \int_0^{n_j} X_k U_j'(X_{jk}) \mathrm{d}k.$$

[4]Pollak (1972) shows that Marshallian demands are single-aggregate if and only if the inverse demands satisfy the same property.

[5]To avoid complex issues raised by income endogeneity which stems from redistribution of profits, we assume that firms are owned by absentee shareholders. Therefore, since labor is the numéraire, we have $y = 1$.

[6]Unlike Acemoglu and Jensen (2013) and Anderson *et al.* (2015), we do not assume the aggregator Λ to be additively separable in firms' strategies.

Clearly, Λ is unaffected by a change in x_c or in X_{jk}.[7]

2. Homothetic demand systems with a single aggregator. Another example is given by the *homothetic demand systems with a single aggregator* (HSA) studied by Matsuyama and Ushchev (2017). The HSA inverse demands are given by

$$p(x_c, \Lambda) = \frac{1}{\Lambda}\phi\left(\frac{x_c}{\Lambda}\right), \qquad p(X_{jk}, \Lambda) = \frac{1}{\Lambda}\Phi_j\left(\frac{X_{jk}}{\Lambda}\right), \tag{17}$$

where $\phi(\cdot)$ and $\Phi_j(\cdot)$ are decreasing functions whose elasticities do not exceed one in absolute value. The market aggregate Λ is implicitly defined by the solution to the following fixed-point condition:

$$\Lambda = M_e \int_0^{\bar{c}} x_c \phi\left(\frac{x_c}{\Lambda}\right) \mathrm{d}\Gamma(c) + \sum_{j=1}^{N} \int_0^{n_j} X_{jk}\Phi_j\left(\frac{X_{jk}}{\Lambda}\right) \mathrm{d}k,$$

which stems from combining (17) with the budget constraint.

The class of HSA preferences includes both the CES and the translog. The former is obtained by choosing power functions with the same exponent for $\phi(\cdot)$ and $\Phi_j(\cdot)$. The latter is a special case of (17) where $\phi(\cdot)$ and $\Phi_j(\cdot)$ are such that

$$\phi^{-1}(z) = \frac{1 - \gamma \ln z}{z}, \qquad \Phi_j^{-1}(z) = \frac{1 - \gamma_j \ln z}{z},$$

where γ and γ_j, $j = 1, ..., N$, are positive parameters.

3. Indirectly additive preferences. A further example of single-aggregate demand systems may be obtained from *indirectly additive preferences*, that is, the indirect utility is as follows (recall that $y = 1$):

$$\mathcal{V} = M_e \int_0^{\bar{c}} v\left(p_c\right) \mathrm{d}\Gamma(c) + \sum_{j=1}^{N} \int_0^{n_j} V_j\left(P_{jk}\right) \mathrm{d}k,$$

where $v(\cdot)$ and $V_j(\cdot)$ are decreasing, convex, and twice differentiable. The corresponding inverse demand system is given by

$$p_c = \left(v'\right)^{-1}(\Lambda x_c), \qquad P_{jk} = \left(V_j'\right)^{-1}(\Lambda X_{jk}),$$

where the market aggregate Λ is given by the solution to the budget constraint:

[7] Since Λ is the only variable that accounts for income, the properties derived below hold true when the individual income is made endogeneous through the redistribution of the big firms' profits.

$$M_e \int_0^{\bar{c}} x_c \left(v'\right)^{-1} (\Lambda x_c) \mathrm{d}\Gamma(c) + \sum_{j=1}^{N} \int_0^{n_j} X_{jk} \left(V_j'\right)^{-1} (\Lambda X_{jk}) \mathrm{d}k = 1.$$

In sum, the Assumption SA is satisfied for a broad class of preferences.

Consider the second stage of the sequential game in which the big firms move first and the small ones second. Being negligible to the market, each small firm accurately treats the market aggregate Λ as a parameter and determines its best reply function $x^*(\Lambda)$. This function is well defined under Assumption SA. This assumption and the envelope theorem imply that the optimal profit function $\pi_c^*(\Lambda)$ is monotone in Λ. Therefore, for any given Λ, the cutoff condition

$$\pi_c^*(\Lambda) \equiv \pi_c(x_c^*(\Lambda), \Lambda) = 0.$$

has at most one solution $\bar{c}(\Lambda)$, which is also monotone in Λ. More specifically, $\bar{c}(\Lambda)$ increases (decreases) with Λ if and only if $\pi_c^*(\Lambda)$ increases (decreases) with Λ. Hence, the zero-profit condition

$$\int_0^{\bar{c}(\Lambda)} \pi_c^*(\Lambda) \mathrm{d}\Gamma(c) - f_e = 0 \tag{18}$$

has a unique solution Λ^*. In other words, as long as entry is accommodated, the equilibrium value Λ^* of the market aggregate is uniquely determined and independent of the actions chosen by the big firms.

In the first stage of the game, the big firms anticipate the equilibrium value Λ^*. Hence, each big firm j chooses the quantity X_{jk}^* of its variety k that maximizes its profits $[p(X_{jk}, \Lambda^*) - C_j]\, X_{jk}$, which depend only upon X_{jk}.

Consequently, we have the following proposition.

Proposition 3. *Assume a hybrid market structure with big and small firms. Under Assumption SA, the big firms behave like the small firms.*

This result has several important implications.

Being non-strategic is a rational strategy. A hybrid market functions "as if" all firms were to operate under monopolistic competition. Stated differently, the small firms incentivize the big firms to refrain from reducing their output, which renders these firms more aggressive. Furthermore, since the big and small firms are heterogeneous, they do not sell at the same prices. However, under Assumption SA, they adopt *the same pricing rule*. In short, Proposition 2 shows that the dilution of market power holds true for a much broader class of preferences than the linear-quadratic utility.

The effects of idiosyncratic shocks. Consider the consequences of a small shock on the marginal cost C_j of a large firm j. We have seen that the equilibrium value Λ^*, the cutoff cost $\overline{c}^*$, and the the small firms' pricing behavior are independent of C_j. Furthermore, it is readily verified that, if entry is accommodated in equilibrium, then the behavior of the incumbents (other than j) does not depend on C_j. Indeed, the first-order condition of a large firm $i \neq j$ is given by

$$\frac{\partial p(X_{ik}, \Lambda^*)}{\partial X_{ik}} X_{ik} + p(X_{ik}, \Lambda^*) = C_i, \tag{19}$$

for all $k \in [0, n_i]$. Since the profit function is strictly quasi-concave in every firm's own actions (Assumption AS), the profit-maximizing output X_{ik}^* of firm i's kth variety is uniquely determined by (19). Moreover, X_{ik}^* depends solely on the equilibrium market aggregate Λ^* and firm i's own marginal cost C_i. Therefore, the equilibrium outputs X_{ik}^* and prices $P_{ik}^* \equiv p(X_{ik}^*, \Lambda^*)$ of any large firm $i \neq j$ are independent on C_j.

In sum, when the market structure is hybrid, *idiosyncratic shocks to large firms do not trigger aggregate shocks on these firms*; they affect the mass M_e^* of entrants that shrinks or expands in response to the shock. However, a change in the cost distribution across small firms does generate aggregate shocks. Indeed, as seen from (18), a shift in $\Gamma(\cdot)$ changes the value of Λ^*, which in turn affects the cutoff $\overline{c}(\Lambda^*)$ and the pricing strategies of all firms.

Entry or merger is neutral. Assume that an additional big firm enters the market. If the market is still hybrid after entry, the big firms do not react because the value of Λ^* is unaffected. However, the fringe shrinks for Λ^* to remain constant. In the same vein, when two or several big firms choose to merge, the other big firms do not react. In the limit, the number N^* of large firms can be pinned down through different mechanisms, e.g. free entry. When N^* is determined, the argument developed above still applies when the number of big firms is N^*. In particular, the free entry of large firms need not deter the entry of small firms, and thus the dilution of market power may also hold under free entry of big firms.[8]

If a shock is not too strong, the market is stabilized by the entry and exit of small firms. By contrast, if the shock is sufficiently strong to trigger the disappearance of the fringe, the big firms adopt a strategic behavior and react to exogenous cost shocks, entry or merger.

The above neutrality effects concur with Redding and Weinstein (2017) who show empirically that supply-side distributional assumptions play a very limited role for understanding revealed comparative advantage across countries and sectors, while the demand-side is key.

[8] See Norman and Thisse (1999) for an example in a spatial model of product differentiation.

4 The dilution of market power under two aggregates

So far we have discussed the case in which the inverse demand functions are single-aggregate. In the baseline model of Section 2, the aggregate is the total output, while under additive preferences, the aggregate is the marginal utility of income. These two aggregates capture different types of market interactions. Total output stems directly from demand linkages across varieties, but it disregards the impact of income. In contrast, the marginal utility of income captures the substitution effects channeled through the budget constraint, but it ignores how total consumption affects the utility derived from consuming a specific variety. Therefore, it seems natural to consider the more general case where both effects are at work, that is, demands are *double-aggregate*. In other words, firms' inverse demands involve two independent market aggregates. To achieve our goal, we go back to the quadratic utility of Section 2 but without an outside good.

4.1 Two aggregates and one fringe

Consider the basic setting of Section 2 where the utility is given by

$$\mathcal{U} = \mathbb{X} - \frac{\beta}{2}\left(\int_0^M x_i^2 \mathrm{d}i + \int_0^n X_k^2 \mathrm{d}k\right) - \frac{\mathbb{X}^2}{2}. \tag{20}$$

Note the difference with (1) that includes two goods. Because the differentiated product is now the only consumption good, the marginal utility of income is endogenous and affects the demand for each variety.

The utility-maximizing conditions yield the following inverse demands for variety $i \in [0, M]$ and variety $k \in [0, n]$:

$$p(x_i, \mathbb{X}, \lambda) = \frac{1 - \mathbb{X} - \beta x_i}{\lambda}, \qquad p(X_k, \mathbb{X}, \lambda) = \frac{1 - \mathbb{X} - \beta X_k}{\lambda},$$

which depend on the two aggregates $\mathbb{X}$ and λ. Note that λ and $\mathbb{X}$ have a different impact on demands: a lower λ rotates clockwise the demand schedules around the saturation point, whereas a lower $\mathbb{X}$ shifts the demand schedules upward by increasing the sole intercept. Thus, when both $\mathbb{X}$ and λ fall (rise), each firm faces a higher (lower) demand. Furthermore, λ is constant when preferences are quasi-linear, e.g. linear-quadratic. The same holds in partial equilibrium models with an outside good. What renders λ variable here is that the budget constraint is binding.

The small firm i's profit is given by

$$\pi_i = p(x_i, \mathbb{X}, \lambda)x_i - f.$$

Since the small firms treat the aggregates $\mathbb{X}$ and λ parametrically, their profit functions are strictly concave in x_i. Of course, any equilibrium of the fringe features symmetry among small firms: $x_i = x$ all $i \in [0, M]$.

In the second stage, the large firm's total output $\mathbb{X}_{mp}$ defined by (4) is treated parametrically by the small firms. A small firm's first-order condition yields:

$$x^*(\mathbb{X}) = \frac{1-\mathbb{X}}{2\beta}, \qquad p^*(\mathbb{X}) = \frac{1-\mathbb{X}}{2\lambda^*(\mathbb{X})}, \tag{21}$$

where $\lambda^*(\mathbb{X})$ is pinned down by the zero-profit condition:

$$\lambda^*(\mathbb{X}) = \frac{(1-\mathbb{X})^2}{4\beta f}. \tag{22}$$

This expression shows that the equilibrium value of the aggregate λ varies with the aggregate $\mathbb{X}$: the marginal utility of income decreases as the total consumption rises. Therefore, we must know $\mathbb{X}$ to determine λ.

Since marginal costs are zero, the labor market balance implies that the equilibrium mass of entrants is given by $M^* = 1/f > 0$. Combining this with $\mathbb{X} = \mathbb{X}_{mp} + M^* x^*(\mathbb{X})$ and (21), we find that the impact of the big firm on the market aggregate is described as follows:

$$\mathbb{X}^*(\mathbb{X}_{mp}) = \begin{cases} \frac{1}{1+2\beta f} + \frac{2\beta f}{1+2\beta f}\mathbb{X}_{mp}, & \mathbb{X}_{mp} \leq 1, \\ \mathbb{X}_{mp}, & \mathbb{X}_{mp} > 1. \end{cases} \tag{23}$$

Figure 4 provides an illustration. Comparing Figures 1 and 4 shows why the big firm now affects the equilibrium value of $\mathbb{X}$: there is no flat spot in Figure 4. As a result, the equilibrium of the second stage depends on the choice $\mathbb{X}_{mp}$ made by the big firm in the first stage. To put it differently, *the large firm always exploits its strategic power to manipulate the market outcome.* This is to be contrasted with Proposition 1 where $\mathbb{X}^*$ is independent of $\mathbb{X}_{mp}$ when n is not too large.

Insert Figure 4 about here

In the foregoing analysis, we implicitly assume that the monopolistically competitive fringe exists ($x^* > 0$), which holds if and only if $\mathbb{X}_{mp} < 1$. Combining equation (23) with (21) and (22) yields the second-stage equilibrium: $x^*(\mathbb{X}_{mp})$, $p^*(\mathbb{X}_{mp})$ and $\lambda^*(\mathbb{X}_{mp})$. By studying the first stage, it is possible to determine the condition on the size n of the big firm for this inequality to be satisfied.

4.2 Two aggregates and two fringes

We now assume there are two monopolistically competitive fringes to capture the idea that the two fringes supply varieties having different qualities, as in Fajgelbaum *et al.* (2011). In our setting, this can be reformulated by assuming that firms in fringe 1 produce at a lower marginal cost but bear a higher fixed cost than firms belonging to fringe 2: $c_1 < c_2$ and $f_1 > f_2$.

Preferences are still given by (20), while the inverse demands for the small firms are now defined by

$$p(x_{s,i}, \mathbb{X}, \lambda) = \frac{1 - \mathbb{X} - \beta x_{s,i}}{\lambda},$$

where s is the fringe index, $s = 1, 2$. Therefore, we have:

$$\mathbb{X} \equiv \int_0^{M_1} x_{1,i} \mathrm{d}i + \int_0^{M_2} x_{2,i} \mathrm{d}i + \mathbb{X}_{mp}, \tag{24}$$

where $\mathbb{X}_{mp}$ is again the large firm's total output defined by (4).

A small firm in fringe s maximizes its profits:[9]

$$\pi_s(x_s) = \left(\frac{1 - \mathbb{X} - \beta x_s}{\lambda} - c_s \right) x_s - f_s, \qquad s = 1, 2.$$

The profit-maximizing outputs and prices of small firms are given, respectively, by

$$x_s^*(\mathbb{X}, \lambda) = \frac{1 - \mathbb{X} - \lambda c_s}{2\beta}, \qquad \pi_s^*(\mathbb{X}, \lambda) = \frac{(1 - \mathbb{X} - \lambda c_s)^2}{4\beta\lambda} - f_s, \qquad s = 1, 2. \tag{25}$$

Combining $\pi_s^*(\mathbb{X}, \lambda)$ with the free entry conditions $\pi_1^* = \pi_2^* = 0$, we find that the equilibrium values λ^* and $\mathbb{X}^*$ must satisfy the following relationships:

$$c_1 + 2\sqrt{\frac{f_1 \beta}{\lambda}} = \frac{1 - \mathbb{X}}{\lambda} = c_2 + 2\sqrt{\frac{f_2 \beta}{\lambda}}. \tag{26}$$

Since $c_1 < c_2$ and $f_1 > f_2$, the equation $c_1 + 2\sqrt{f_1\beta/\lambda} = c_2 + 2\sqrt{f_2\beta/\lambda}$ has a unique positive solution λ^*, while $\mathbb{X}^*$ is pinned down by plugging λ^* into (26):

$$\lambda^* = 4\beta \left(\frac{\sqrt{f_1} - \sqrt{f_2}}{c_2 - c_1} \right)^2, \qquad \mathbb{X}^* = 1 - 4\beta \frac{\left(\sqrt{f_1} - \sqrt{f_2}\right)\left(c_2\sqrt{f_1} - c_1\sqrt{f_2}\right)}{(c_2 - c_1)^2}. \tag{27}$$

As implied by (27), the equilibrium values λ^* and $\mathbb{X}^*$ of the two aggregates are *independent* of the big

[9] In what follows, we disregard the variety index i because of symmetry.

firm's behavior.[10] Put differently, when both monopolistically competitive fringes are active, *the market equilibrium exhibits the dilution of market power*, even though preferences are no longer single aggregate.

But does a market outcome with two active fringes exist? We show in Appendix that, if β is not too high, while the values of the supply-side parameters satisfy

$$c_1 < \frac{2 - \left(c_2\sqrt{f_1} - c_1\sqrt{f_2}\right)}{\sqrt{f_1} - \sqrt{f_2}} < c_2,$$

then there exists a threshold value $\overline{n} > 0$ such that both fringes are active if and only if $n < \overline{n}$. This result agrees with what we have seen in Proposition 1: *the dilution of market power holds true when the big firm is not too big.*

4.3 Two aggregates: a summary

We summarize the main findings of this section in the following proposition.

Proposition 4. *Assume a quadratic utility without an outside good. (i) If the market equilibrium involves one monopolistically competitive fringe, then the dilution of market power does not hold. (ii) If the market involves two monopolistically competitive fringes, then the big firm behaves like the small firms.*

In the foregoing, we have chosen to work with quadratic preferences because this allows us to highlight the role played by two market aggregates. However, it should be clear that the dilution of market power keeps its relevance under other preferences generating double-aggregate demand systems such as the Kimball's (1995) flexible aggregator (Matsuyama and Ushchev, 2017). The key factor is the relationship between the number of aggregates and the number of monopolistically competitive fringes.

5 Concluding remarks

In this paper, we distinguish between two kinds of firms, i.e. multi-product firms and single-product firms. The former can a priori manipulate the market whereas the latter cannot because they are negligible. Our results are best illustrated by our new version of the dominant firm model. Depending upon its scope, a dominant firm can either accommodate a monopolistically competitive fringe, or deter entry, or behave like an unconstrained monopolist. The novel case arises when entry is accommodated while firms' profits depends on a single aggregate such as the total output or the marginal utility of income. In this case, we

[10] To guarantee that $\mathbb{X}^* > 0$, we must assume that β is too large:

$$\beta < \frac{(c_2 - c_1)^2}{4\left(\sqrt{f_1} - \sqrt{f_2}\right)\left(c_2\sqrt{f_1} - c_1\sqrt{f_2}\right)}.$$

have shown that this aggregate is determined by the sole entry and exit of small firms. Therefore, even when the dominant firm has a relatively large market share, this firm finds it profit-maximizing to disregard its ability to manipulate the market, practicing instead a “divisionalization strategy” akin to monopolistic competition. More generally, we have seen that the presence of a monopolistically competitive fringe may vastly affect the behavior of large firms in that the former disciplines the latter.

Our results suggest that consumers' preferences, more than producers' costs, determine the market structure, implying that the on-going emphasis put on cost heterogeneity could well be exaggerated. This is also in line with the recent empirical findings of Hottman *et al.* (2016) and Redding and Weinstein (2017). In short, our framework allows the market structure to be endogenized by determining conditions for oligopolistic competition, monopolistic competition, or hybrid forms of competition to emerge as an equilibrium outcome.

References

[1] Acemoglu, D., Jensen, M.K. 2013. Aggregate comparative statics. *Games and Economic Behavior* 81: 27 – 49.

[2] Allanson, P., Montagna, C. 2005. Multiproduct firms and market structure: An explorative application to the product life cycle. *International Journal of Industrial Organization* 23: 587 – 97.

[3] Anderson, S.P., Erkal, N., Piccinin, D. 2015. Aggregative oligopoly games with entry. University of Virginia, memo.

[4] Baumol, W.J. 1980. Contestable markets: An uprising in the theory of industry structure. *American Economic Review* 70: 1 – 15.

[5] Bernard, A.B., Jensen, J.B., Redding, S.J., Schott, P.K. 2012. The empirics of firm heterogeneity and international trade. *Annual Review of Economics* 4: 283 – 313.

[6] Bernard, A.B., Redding, S.J., Schott, P.K. 2011. Multi-product firms and trade liberalization. *Quarterly Journal of Economics* 126: 1271 – 318.

[7] Bertoletti, P., Etro, F. 2017. Monopolistic competition when income matters. *Economic Journal* 127: 1217 – 43.

[8] d'Aspremont, C., Dos Santos Ferreira, R. Gérard-Varet, L.-A. 1996. On the Dixit-Stiglitz model of monopolistic competition. *American Economic Review* 86: 623 – 9.

[9] Demidova, S. 2017. Trade policies, firm heterogeneity, and variable markups. *Journal of International Economics* 108: 260 – 73.

[10] Dhingra, S. 2013. Trading away wide brands for cheap brands. *American Economic Review* 103: 2554 – 84.

[11] Dixit, A.K. 1979. A model of duopoly suggesting a theory of entry barriers. *Bell Journal of Economics* 10: 20 – 32.

[12] Dixit, A.K., Stiglitz, J.E. 1977. Monopolistic competition and optimum product diversity. *American Economic Review* 67: 297 – 308.

[13] Eaton, B.C., Schmitt, N. 1994. Flexible manufacturing and market structure. *American Economic Review* 84: 875 – 88.

[14] Eckel, C., Neary, J.P. 2010. Multi-product firms and flexible manufacturing in the global economy. *Review of Economic Studies* 77: 188 – 217.

[15] Etro, F. 2006. Aggressive leaders. *RAND Journal of Economics* 37: 146 – 54.

[16] Etro, F. 2008. Stackelberg competition with endogenous entry. *Economic Journal* 118: 1670 – 97.

[17] Fajgelbaum, P., Grossman, G.M., Helpman, E. 2011. Income distribution, product quality, and international trade. *Journal of Political Economy* 119: 721 – 65.

[18] Gabszewicz, J.J., Shitovitz, B. 1992. The core in imperfectly competitive economies. In: R. Aumann and S. Hart (eds.), *Handbook of Game Theory with Economic Applications. Volume 1.* Amsterdam: North-Holland, 459 – 83.

[19] Hart, O. 1985. Imperfect competition in general equilibrium: An overview of recent work. In K. Arrow and S. Honkapohja (eds.), *Frontiers in Economics.* Oxford: Basil Blackwell, 100 – 49.

[20] Hottman, C., Redding, S.J., Weinstein, D.E. 2016. Quantifying the sources of firm heterogeneity. *Quarterly Journal of Economics* 131: 1291 – 364.

[21] Kimball, M. 1995. The quantitative analytics of the basic neomonetarist model. *Journal of Money, Credit and Banking* 27: 1241 – 77.

[22] Markham, J.W. 1951. The nature and significance of price leadership. *American Economic Review* 41: 891 – 905.

[23] Matsuyama, K., Ushchev, P. 2017. Beyond CES: Three alternative classes of flexible homothetic demand systems. CEPR Discussion Paper N° 12210.

[24] Mayer, T., Melitz, M., Ottaviano, G.I.P. 2014. Market size, competition, and the product mix of exporters. *American Economic Review* 104: 495 – 536.

[25] Melitz, M. 2003. The impact of trade on intraindustry reallocations and aggregate industry productivity. *Econometrica* 71: 1695 – 725.

[26] Neary, J.P. 2010. Two and a half theories of trade. *The World Economy* 33: 1 – 19.

[27] Norman, G., Thisse, J.-F. 1999. Technology choice and market structure: Strategic aspects of flexible manufacturing, *Journal of Industrial Economics* 47: 345 – 72.

[28] Okuno, M., Postlewaite, A., Roberts, J. 1980. Oligopoly and competition in large markets. *American Economic Review* 70: 22 – 31.

[29] Ottaviano, G.I.P., Tabuchi, T., Thisse, J.-F. 2002. Agglomeration and trade revisited. *International Economic Review* 43: 409 – 36.

[30] Parenti, M. 2017. International trade with David and Goliath. *Journal of International Economics*, forthcoming.

[31] Pollak, R. A. 1972. Generalized separability. *Econometrica* 40: 431 – 53.

[32] Redding, S.J., Weinstein, D.E. 2017. Accounting for micro and macro patterns of trade. Columbia University, memo.

[33] Shimomura, K.-I., Thisse, J.-F. 2012. Competition among the big and the small. *RAND Journal of Economics* 43: 329 – 47.

[34] Spence, A.M. 1977. Entry, capacity, investment and oligopolistic pricing. *Bell Journal of Economics* 8: 534 – 44.

[35] Spulber, D.F. 1981. Capacity, output, and sequential entry. *American Economic Review* 71: 503 – 14.

[36] Uslay, C., Altintig, Z. Ayca, Winsor, R.D. 2010. An empirical examination of the "Rule of Three": Strategy implications for top management, marketers, and investors. *Journal of Marketing* 74: 20 – 39.

[37] Zhelobodko, E., Kokovin, S., Parenti, M., Thisse, J.-F. 2012. Monopolistic competition in general equilibrium: Beyond the constant elasticity of substitution. *Econometrica* 80: 2765 – 84.

Appendix

We know that

$$M_1 x_1^* + M_2 x_2^* = \mathbb{X}^* - nX^*, \tag{A.1}$$

while the budget constraint implies

$$M_1 p_1^* x_1^* + M_2 p_2^* x_2^* = 1. \tag{A.2}$$

As p_s^* and x_s^*, $s = 1, 2$, are pinned down by plugging (27) into (25), (A.1) – (A.2) is a system of two linear equations with two unknowns: M_1 and M_2.

Lemma 1. *The system* (A.1) – (A.2) *has a unique solution.*

Proof. Because $c_1 < c_2$, and because the pass-through under linear demands equals 1/2, we have:

$$p_2^* - p_1^* = \frac{c_2 - c_1}{2} > 0. \tag{A.3}$$

Hence, the determinant of the system (A.1) – (A.2) is always strictly positive:

$$\beta(c_2 - c_1) x_1^* x_2^* > 0.$$

Hence, the coefficient matrix is non-degenerate. Q.E.D.

For both fringes to be active, (A.1) – (A.2) must have a *positive* solution.

Lemma 2. *The system* (A.1) – (A.2) *has a positive solution if and only if the following inequalities hold*:

$$\frac{n}{2\beta + n} + \frac{2\beta}{2\beta + n} \frac{1}{p_2^*} < \mathbb{X}^* < \frac{n}{2\beta + n} + \frac{2\beta}{2\beta + n} \frac{1}{p_1^*}. \tag{A.4}$$

Proof. Observe that the slopes of the (A.1)- and (A.2)-loci in the (M_1, M_2)-plane are equal, respectively, to x_1^*/x_2^* and to $p_1^* x_1^*/p_2^* x_2^*$. Combining this with $p_2^* > p_1^*$, which is implied by (A.3), we find that the (A.1)-locus is steeper than the (A.2)-locus. Hence, the system (A.1) – (A.2) has a positive solution if and only if (i) the intercept of the (A.1)-locus with the M_2-axis is higher than that of the (A.2)-locus, and (ii) the intercept of the (A.2)-locus with the M_1-axis is further to the right than that of the (A.1)-locus. It is readily verified that (i) and (ii) are equivalent to (A.4). Q.E.D.

We are now equipped to prove that the dilution of market power occurs when the big firm is not too big.

Proposition A. *Assume that the following inequalities hold*:

$$c_1 < \frac{2 - \left(c_2\sqrt{f_1} - c_1\sqrt{f_2}\right)}{\sqrt{f_1} - \sqrt{f_2}} < c_2. \tag{A.5}$$

Then, $\overline{\beta} > 0$ *and* $\overline{n} > 0$ *exist such that both fringes are active if and only if* $\beta < \overline{\beta}$ *and* $n < \overline{n}$.

Proof. The inequalities (A.4) can be rewritten as follows:

$$K_1 < n + 2\beta < K_2, \tag{A.6}$$

where K_1 and K_2 are constants defined by

$$K_s \equiv \frac{(c_2 - c_1)^2}{2\left(\sqrt{f_1} - \sqrt{f_2}\right)\left(c_2\sqrt{f_1} - c_1\sqrt{f_2}\right)}\left[1 - \frac{2}{c_s\left(\sqrt{f_1} - \sqrt{f_2}\right) + c_2\sqrt{f_1} - c_1\sqrt{f_2}}\right]$$

for $s = 1, 2$. Furthermore, it is readily verified that (A.5) is equivalent to $K_1 < 0 < K_2$. Thus, it remains to show that $n + 2\beta < K_2$ for the proposition hold. Set

$$\overline{\beta} \equiv \frac{K_2}{2} = \frac{(c_2 - c_1)^2}{4\left(\sqrt{f_1} - \sqrt{f_2}\right)\left(c_2\sqrt{f_1} - c_1\sqrt{f_2}\right)}\left[1 - \frac{2}{c_2\left(\sqrt{f_1} - \sqrt{f_2}\right) + c_2\sqrt{f_1} - c_1\sqrt{f_2}}\right],$$

and assume $\beta \leq \overline{\beta}$. Setting $\overline{n} \equiv K_2/2\beta$ and using (A.6) completes the proof. Q.E.D.

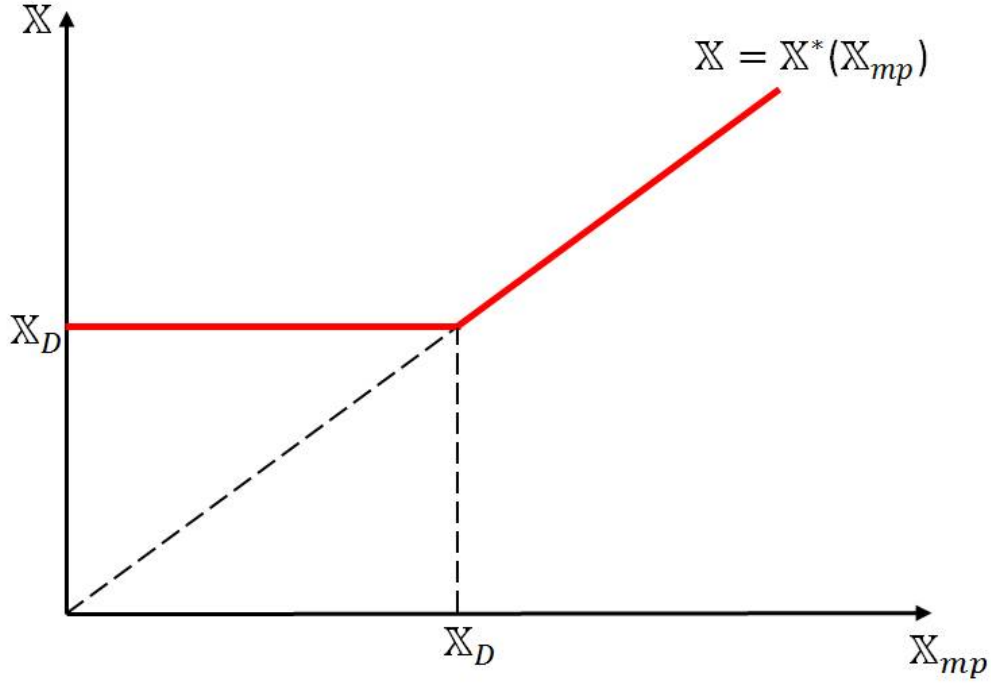

Figure 1

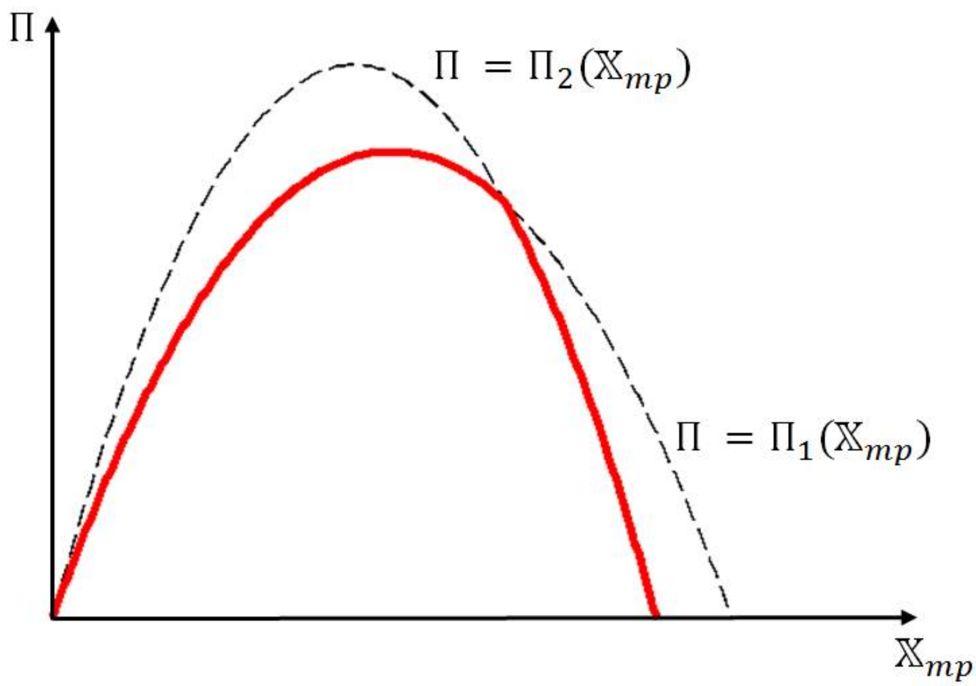

Figure 2a

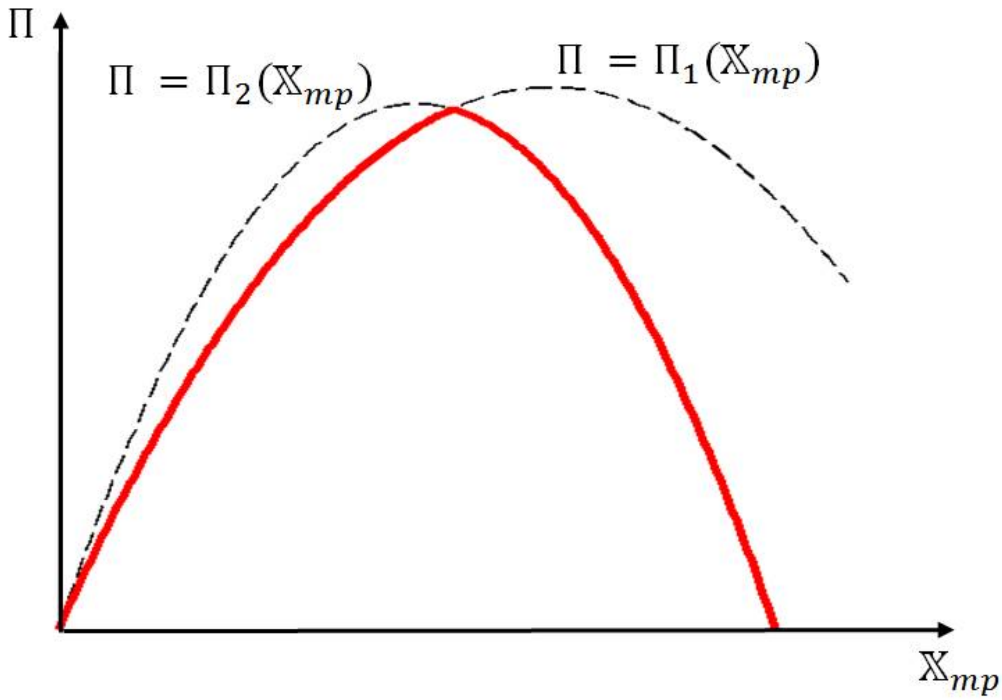

Figure 2b

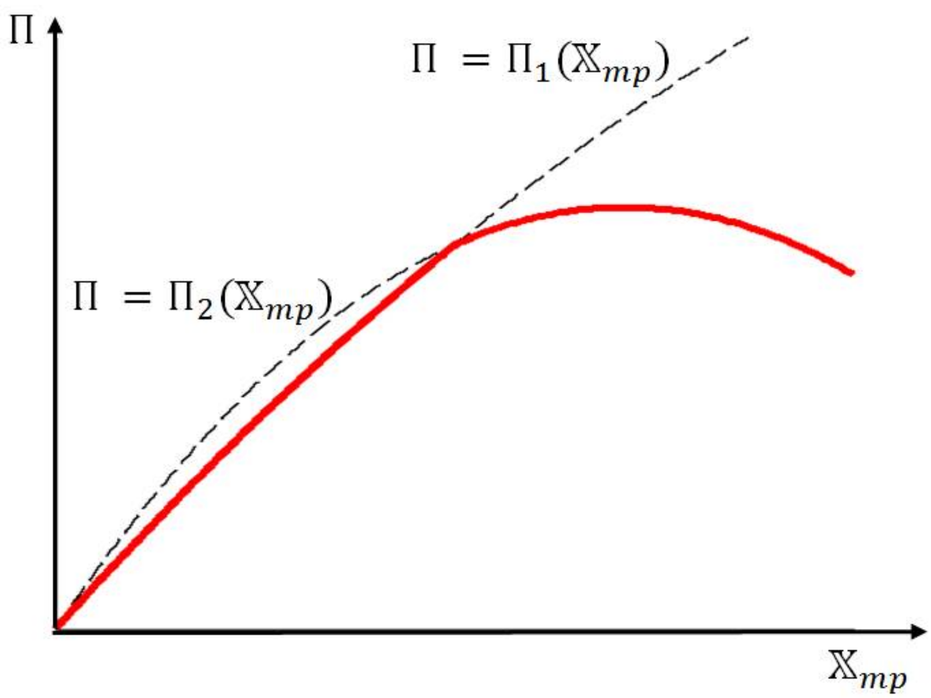

Figure 2c

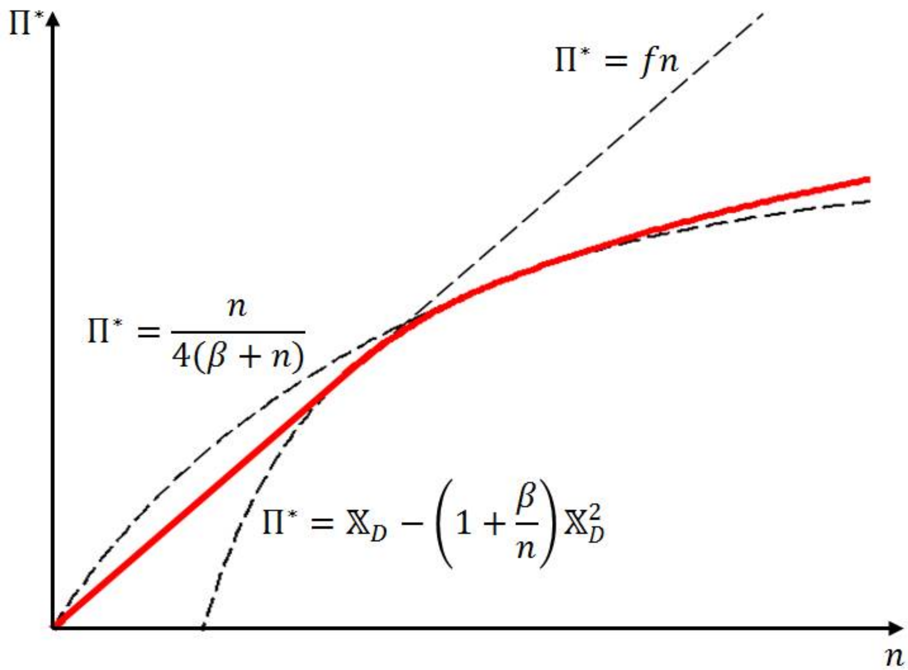

Figure 3

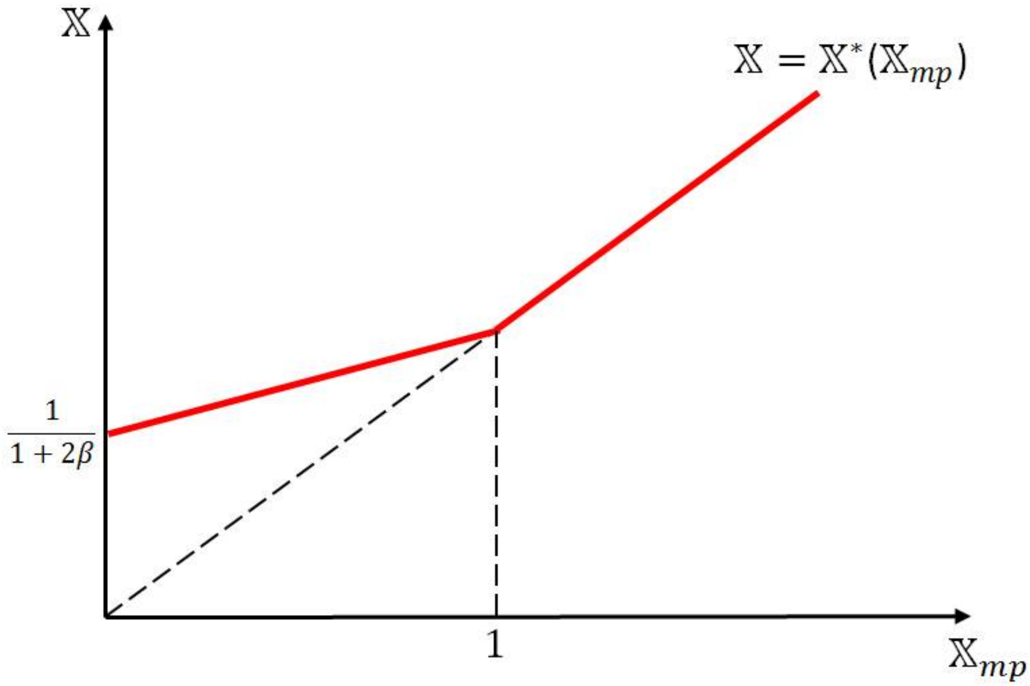

Figure 4

Contact details:

Philip Ushchev

National Research University Higher School of Economics (Moscow, Russia).

Center for Market Studies and Spatial Economics, leading research fellow.

E-mail: fuschev@hse.ru