Ca' Foscari
University
of Venice

**Department
of Economics**

Working Paper

Federico Etro

The Theory of Endogenous Market Structures: A Survey

ISSN: 1827-3580
No. 11/WP/2012

The Theory of Endogenous Market Structures: A Survey

Federico Etro

University of Venice at Ca' Foscari

This Draft: July 2012

Abstract

Most market structures are neither perfectly or monopolistically competitive: they are characterized by a small number of large firms engaged in strategic interactions in their production and investment decisions. Yet, most of our economic theories are still based on a simplified world where firms are either small price takers producing under constant returns to scale (perfect competition) or isolated price setters (monopolistic competition). The theory of EMSs analyzes markets in partial and general equilibrium where strategies affect entry and entry affects strategies, and only exogenous primitive conditions on technology and preferences affect the equilibrium outcome. Understanding market structures means to understand how many firms are active in a market, which strategies they adopt and how primitive conditions and policy shocks affect them in a static or dynamic perspective.

Keywords

Endogenous entry, oligopoly, sunk costs, general equilibrium

JEL Codes

L1, E20, E32, F12

Address for correspondence:

Federico Etro

Department of Economics
Ca' Foscari University of Venice
Cannaregio 873, Fondamenta S. Giobbe
30121 Venezia - Italy
Phone: (+39) 041 2349172
Fax: (+39) 041 2349176
e-mail: federico.etro@unive.it

This Working Paper is published under the auspices of the Department of Economics of the Ca' Foscari University of Venice. Opinions expressed herein are those of the authors and not those of the Department. The Working Paper series is designed to divulge preliminary or incomplete work, circulated to favour discussion and comments. Citation of this paper should consider its provisional character.

The Theory of Endogenous Market Structures: A Survey*

Most market structures are neither perfectly or monopolistically competitive: they are characterized by a small number of large firms engaged in strategic interactions in their production and investment decisions. Yet, most of our economic theories are still based on a simplified world where firms are either small price takers producing under constant returns to scale (perfect competition) or isolated price setters (monopolistic competition). The first case generates the indeterminacy of the market structure and therefore its irrelevance to explain economic phenomena: this is typical of neoclassical economics in trade (the Heckscher-Ohlin model), business cycles (the RBC model) and growth (the Solow model). The second case generates exogenous market structures in which a continuum of atomistic monopolists choose prices with constant mark ups independently from each other: this is typical of the Dixit and Stiglitz (1977) model applied to trade (Krugman, 1980), business cycles (Blanchard and Kiyotaki, 1987) or growth (Romer, 1990). All this comes at a loss in our understanding of what determines the structure of markets, i.e. how many firms are active, which pricing and production decisions they adopt, how much they invest in R&D, which other strategic investments they undertake, and how the market structure evolves over time. These factors are important to analyze markets in partial equilibrium, but also to understand how aggregate phenomena affect market structures in general equilibrium and how these structures affect trade, business cycle and growth. The analysis of endogenous market structures (EMSs) in industrial organization over the last decades has made it possible to expand our understanding of these aspects and new research is flourishing in multiple fields, even if communication between them is often limited. This survey is about this theoretical literature, tries to build a bridge between applications in different fields, and supports the idea that opening the black box of the market structure can be extremely useful to build a more realistic approach to theoretical analysis.

The EMSs approach analyzes jointly strategic interactions and endogenous entry decisions. Of course, each one of the two ingredients is well understood:

*I am extremely grateful to Olivier Blanchard, Andr de Palma, Jonathan Eaton, Jacques Thisse and Kresimir Zigic for insightful discussions.

most of industrial organization theory is based on models of imperfect competition such as Cournot, Bertrand and Stackelberg duopolies, and free entry typically characterizes models of perfect and monopolistic competition. However, the simultaneous consideration of both aspects is not common: the theory of EMSs analyzes markets in which strategies affect entry and entry affect strategies, and only exogenous primitive conditions on technology and preferences affect the equilibrium outcome. Understanding market structures means to understand how many firms are active in a market, which strategies they adopt and how primitive conditions and policy shocks affect them in a static or dynamic perspective.

Let us summarize the main insights of the literature on EMSs which hold independently from the form of competition (in quantities or prices) and under weak additional conditions. First of all, when the size of demand increases in a market, this market attracts new firms, which strengthens competition, reduces profit margins and, to break even, leads to an expansion of the production of each firm (competition effect). While this entry process creates social gains from the price reduction, it is typically excessive from a social point of view because it induces too much waste in start-up costs (Mankiw-Whinston business stealing effect). This opens up space for firms to gain a competitive advantage through different strategies: market leaders always tend to reduce their prices to expand their market share or even to deter entry (limit-pricing in contestable markets), but there are also gains from a strategic use of R&D investments or financial and contractual decisions that induce lower prices in the subsequent competition (aggressive leaders). When the same technology is the endogenous result of the R&D investment by all firms, fixed investments tend to increase with the size of the market, creating an upper bound on the number of firms that can be active (Sutton bounds approach). This justifies the same focus on highly concentrated markets even at the global level, and it is consistent with two general insights: larger markets tend to generate lower prices and higher production per firm. Natural consequences are that growth and opening up to trade create gains from price reductions of the traded goods (gains from competition). Globalization, however, induces some business destruction of domestic firms, which opens new roads for trade policy analysis. Moreover, in case of cost heterogeneity between firms, a selection effect emerges for which only the most productive firms find it profitable to export abroad (Melitz effect): the endogeneity of the number and productivity of the firms and of the exporters has important implications for our understanding of trade flows in general equilibrium. Finally, it is exactly in general equilibrium that the size of a certain market in a certain moment stops being exogenous: it depends on the level of competition of all the markets in different periods. The key consequence is that aggregate shocks create new intertemporal substitution mechanisms that amplify the propagation of the business cycle: booms attract entry, which strengthens competition, reduces prices and increases real wages, which in turn pushes consumption and labor supply. These channels are absent

under perfect or monopolistic competition, and they entirely depend on the endogeneity of entry, markups and individual production levels. Needless to say, the dynamic inefficiency of the EMSs asks for a reconsideration of normative principles in macroeconomic policy: this should promote entry in recessions and limit business creation in booms.

While the properties of a basic Cournot model with free entry represented a sort of folk theorem for a long time, the systematic analysis of EMSs has been developed through a number of key contributions, such as those of von Weizsäcker (1980) on U-shaped cost functions, Dasgupta and Stiglitz (1980) on endogenous sunk costs, Brander (1981) on trade implications, Mankiw and Whinston (1986) on welfare and Sutton (1991) on the empirical analysis of entry. More recently, a new theoretical wave of the literature has applied these tools to revisit standard theoretical results not only in partial equilibrium but also in general equilibrium. Some of the early dynamic general equilibrium models that stimulated this second generation of research include endogenous growth models with EMSs and endogenous sunk costs as in Peretto (1999), the pathbreaking analysis of trade with heterogeneous costs and endogenous entry by Melitz (2003) and Schumpeterian growth models with EMSs in the patent races as in Etro (2004). Most of the recent literature in the industrial organization tradition has been focused on the analysis of endogenous technology (Vives, 2008), limit pricing strategies (Etro, 2008; Kováč *et al.*, 2010) and various aspects of contract theory and antitrust theory in markets whose structure is endogenous. In the field of international trade recent important works have been focused on the analysis of the endogenous number and productivity of the exporting firms (Ghironi and Melitz, 2005; Eckel and Neary, 2010), on endogenous sunk costs in global markets (Sutton, 2007) and on foreign direct investments and strategic trade policy for markets with endogenous entry. In the field of macroeconomics recent attempts at augmenting RBC theory with oligopolistic competition (Etro and Colciago, 2010) and endogenous product variety (Bilbiie *et al.*, 2012) appear to open promising avenues for new research. Still at an early stage is the analysis of fiscal and monetary policy in environments where the number of competing firms is endogenous and time varying.

We are not aware of surveys on the wide theoretical literature analyzed here. All industrial organization textbooks deal with free entry equilibria but they peripherically touch on their implications in other fields, with the recent exception of Belleflamme and Peitz (2010, Chapter 3). Many surveys on the modern theories of international trade or macroeconomics focus on applications in their specific fields and, when they build bridges between fields as in Matsuyama (1996), they do it for the case of monopolistic rather than oligopolistic competition. On empirical aspects of entry and EMSs some good surveys are available (see Berry and Reiss, 2007), therefore we will not stress these aspects in what follows.

The survey is organized as follows. Section 1 analyzes EMSs in partial equilibrium, provides a microfounded analysis and derives some welfare results.

Section 2 extends the model to asymmetric interactions and endogenous technology. Section 3 introduces the analysis of imperfectly competitive markets in a general equilibrium framework and derives results for trade theory and policy. Section 4 moves to dynamic EMSs and applies them to analyze business cycle, trade and growth issues. Section 5 concludes.

1 Partial Equilibrium EMSs

In this section we briefly introduce a class of aggregative models of the market structure where the profit functions are exogenously given and the EMSs can be characterized in a rather general way.¹ Consider $n \geq 2$ firms choosing a strategic variable $x_i > 0$. All the strategies deliver for each firm i the following gross profits:

$$\Pi^i = \Pi(x_i, \beta_i) \quad (1)$$

where $\beta_i = \sum_{j=1, j \neq i}^n h(x_j)$ represents the effects induced by the strategies of all the other firms on the profits of firm i , for some increasing function $h(x)$. These effects exert a negative impact on profits ($\Pi_2 < 0$), and affect also the profit maximizing strategy, which is assumed to be unique. Such a framework nests standard models of competition with strategic substitutability ($\Pi_{12} < 0$) and complementarity ($\Pi_{12} > 0$), such as models of competition in quantities and prices and also patent races. Entry in the market requires a fixed cost F .

Under Nash competition between firms in the choice of their strategic variables, we can easily characterize the symmetric EMS with the pair (x, n) satisfying the profit maximizing and endogenous entry conditions:

$$\Pi_1[x, (n-1)h(x)] = 0 \quad \text{and} \quad \Pi[x, (n-1)h(x)] = F \quad (2)$$

where we imposed symmetry with $\beta = (n-1)h(x)$ and assumed unicity of the equilibrium. One can easily verify that the number of firms n is always decreasing in the size of the fixed cost of entry. Moreover, the strategy of each firm x is increasing with the fixed cost under strategic substitutability (i.e. the firm becomes more aggressive), and it is decreasing under strategic complementarity (i.e. the firm becomes more accommodating).

To exemplify this general model, consider a market with competition in quantities such that the strategy x_i represents the quantity produced by firm i . Any inverse demand $p_i = p(x_i, \beta_i)$ decreasing in both arguments generates profit functions nested in the specification (1) and a standard Cournot equilibrium. Ruffin (1971) provided the first example (we are aware of) for such an EMS and confirmed the Cournot conjecture, for which the equilibrium with homogenous

¹A growing literature is analyzing this kind of aggregative games (where profits depend on aggregate statistics of the strategies). See Acemoglu and Jensen (2011) for the case of an exogenous number of players and Anderson *et al.* (2012) for the case of an endogenous number, as here.

goods converges to the competitive one when the fixed cost tends to zero (an infinity of firms selling at a price equal to the marginal cost).

A classic example, that we will use repeatedly, is the one of a linear demand function for homogenous goods $p = a - X/E$, where X is total production and E can be interpreted as the size of the market, i.e. the endowment spent in the market or the number of consumers. If the marginal cost is constant (and exogenous), say c , the Cournot equilibrium generates production $x = E(a - c)/(n + 1)$ and gross profits $\Pi(n) = x^2/E$. With endogenous entry, the production of each firm is $x = \sqrt{FE}$, and the number of firms is:

$$n = \sqrt{s(a - c)^2} - 1 \quad (3)$$

where we defined $s \equiv E/F$ as the relative market size. The presence of a fixed entry cost limits the number of firms that can be active and generates an equilibrium price above the marginal cost $p = c + 1/\sqrt{s}$ which allows them to break even. When F increases, firms cover this fixed cost in part by expanding their production and in part by increasing their markups, which is possible thanks to a reduction of the number of active firms. On the other side, an increase in market size E attracts further entry, which strengthens competition so as to reduce markups and increase individual production in a less than proportional way. Von Weizsäcker (1980) has extended the example to the case of a U-shaped average cost function.²

Consider now models of competition in prices where p_i is the price of firm i . Any model with a direct demand $D^i = D \left[p_i, \sum_{j \neq i} g(p_j) \right]$ for substitute goods is nested in our general framework (1) under basic regularity conditions,³ and after setting $x_i \equiv 1/p_i$ and $h(x) = g(1/x)$. Examples include models of price competition with isoelastic demand functions, Logit demand, or any constant expenditure demand.⁴

1.1 Microfounded EMSs

In this section we endogenize market demand by utility maximizing consumers. The advantage is that a microfounded demand allows us to compare different

²Considering a quadratic cost function as $x_i^2/2$, von Weizsäcker (1980) derived the EMS with production $x = \sqrt{2FE/(2 + E)}$ for each firm and a number of firms $n = a\sqrt{s(2 + E)}/2 - 1 - E$, with similar properties to the constant marginal cost case.

³The assumptions needed are that $D_1 < 0$, $D_2 < 0$ and $g'(p) < 0$. This guarantees that goods are substitutes in a standard way. Notice that we exclude spatial models.

⁴We can also mention a class of models used to study the market for innovations: patent races. Suppose that each firm invests x_i to create a new product, patent it and create a new business which replaces the existing one and gives right to a certain flow of profits. The chances that this happens are increasing in the investment of the firm but decreasing in the aggregate probability that any (other) firm wins the patent race. As long as this probability of innovation can be expressed as $\delta = \sum_{j=1}^n h(x_j)$ the expected profits of firm i can be written as in our general framework (1).

forms of competition in the same market and to evaluate EMSs from a welfare perspective. The disadvantage is that our assumptions on consumers' preferences will induce a loss of generality compared to the profit functions introduced in the previous section.⁵ We will impose two restrictions: the first is that utility is additively separable in the subutilities $u(C_j)$ from the consumption C_j of the goods produced by each firm j , and the second is that the subutilities $u(C_j)$ are identical and homothetic. The first condition is necessary for profits to be nested in (1), but the second one is restrictive, because it amounts to adopt the CES preferences introduced by Dixit and Stiglitz (1977):

$$U = \left[\sum_{j=1}^n C_j^{\frac{\theta-1}{\theta}} \right]^{\frac{\theta}{\theta-1}} \quad (4)$$

Here $\theta > 1$ is the degree of substitutability between goods: when $\theta \rightarrow \infty$ the goods become perfect substitutes, when $\theta \rightarrow 1$ they tend to complete independence. In the rest of the survey we will adopt this microfoundation but we will also mention when our results apply to the more general framework of the previous section.

Utility is maximized under a budget constraint where p_j is the price of good j , and E is the (exogenous) size of the market. The Dixit-Stiglitz preferences have been traditionally used to study monopolistic competition in which firms choose their strategies as isolated monopolists that do not take into account strategic interactions with other firms.⁶ In what follows, we will depart from this assumption and study competition between a small and endogenous number of interacting firms.

1.1.1 EMSs with Cournot competition

Utility maximization of the representative agent generates the inverse demand $p_i = x_i^{-1/\theta} E / \sum_j x_j^{(\theta-1)/\theta}$. Notice that in case of perfect substitutability, the inverse demand becomes hyperbolic in total consumption: $p = E/X$ for each firm. If we assume that firms produce at a constant (exogenous) marginal cost c , the gross profit function of firm i becomes:

$$\Pi^i = \frac{x_i^{\frac{\theta-1}{\theta}} E}{\sum_{j=1}^n x_j^{\frac{\theta-1}{\theta}}} - cx(i) \quad (5)$$

⁵For instance, one could derive the linear demand model for homogenous goods from a simple quadratic utility, but this would exclude general forms of imperfect substitutability between products. Extensions of the quadratic utility to account for product differentiation are available, but they do not generate models nested in (1) and they can hardly be introduced in a dynamic general equilibrium framework. Ideal for this purpose is the CES microfoundation.

⁶The first general characterization of monopolistic competition with endogenous entry with non-homothetic preferences has been developed by Zhelobodko *et al.* (2012).

which is nested in (1) with $h(x) = x^{(\theta-1)/\theta}$. For a given n the Cournot output of each firm can be derived as $x = (\theta - 1)(n - 1)E/\theta n^2 c$, which is associated with the equilibrium mark up (defined as the price-marginal cost ratio):

$$\mu(n) = \frac{\theta n}{(\theta - 1)(n - 1)} \quad (6)$$

The price is decreasing not only in the number of firms, but also in the degree of substitutability between goods, and only when the number of firms tends to infinity the mark up reaches the competitive limit. Gross profits can be now derived as $\Pi = (n + \theta - 1)E/\theta n^2$, and the endogenous entry condition leads to the following number of firms:

$$n = \frac{s}{2\theta} \left[1 + \sqrt{1 + \frac{4\theta(\theta - 1)}{s}} \right] \quad (7)$$

Larger markets attract more firms, which induces stronger competition and reduces the markup, as can be verified from (6). Moreover, more substitutability between goods strengthens competition reducing both the mark ups and the number of firms.

The results simplify drastically in case of homogenous goods, which provides a markup $\mu(n) = n/(n - 1)$. With endogenous entry, the number of firms is simply $n = \sqrt{s}$ implying $\mu = \sqrt{s}/(\sqrt{s} - 1)$. These relations show the simple link between the relative size of the market and the EMS. We can easily verify that increasing the size of a market the number of firms increases but less than proportionally, and the mark up decreases. More precisely, entry of at least n firms requires that the size of the market must be above the cut-off $n^2 F$. Therefore, if we want to double the number of active firms, we need a size of the market that is more than the double. Such a prediction can be tested in the presence of markets of different sizes, for instance professional or retail markets in different towns: a wide empirical literature (see Breshnan and Reiss, 1990, and Manuszak, 2002) has found encouraging support for such a view. Finally, when the size of the market increases, each firm has to produce at a larger scale - $x = (\sqrt{EF} - F)/c$ in case of homogenous goods. This happens because a larger market opens space for a larger number of firms, but this strengthens competition and reduces the mark ups, which requires a larger scale of production for each firm to cover the fixed costs.

1.1.2 EMSs with Bertrand Competition

The utility maximization problem can be used also to express the direct demand of each good i and the following gross profits:

$$\Pi^i = \frac{(p_i - c) p_i^{-\theta} E}{\sum_{j=1}^n p_j^{1-\theta}} \quad (8)$$

which is nested in (1) after setting $x(i) = 1/p_i$ and $h(x) = x^{\theta-1}$. The symmetric Bertrand equilibrium is characterized by a common price $p = \mu c$ with the following mark up:

$$\mu(n) = \frac{\theta n - \theta + 1}{(\theta - 1)(n - 1)} \quad (9)$$

and gross profits $\Pi = E/[1 + \theta(n - 1)]$. Given the number of firms, it is easy to verify that markup and profits are smaller than those under competition in quantities. The endogenous entry condition provides the following number of firms:

$$n = 1 + \frac{s - 1}{\theta} \quad (10)$$

which generates the equilibrium markup $\mu = \theta s/(\theta - 1)(s - 1)$. One can immediately verify that competition in prices implies a smaller number of firms compared to competition in quantities, but Bertrand competition is so much tougher that the final markups are still below those under Cournot competition. More importantly, the markups are always decreasing in the size of the market relative to the fixed cost. This happens because larger markets attract more competitors, which takes place now through a linear relationship rather than a concave one. Nevertheless, there are two implications that are common to both forms of competition: larger markets induce each firm to sell more and at a lower price. The positive relation between market size and average production per firm has been tested in a number of professional or retail markets in different towns by Campbell and Hopenhayn (2005) finding convincing evidence in its support.

1.2 Optimal market structures and the excess-entry result

Are EMSs efficient? Do they create insufficient or excessive entry? The answer to such a question depends on a number of factors. Consider first the case of models of competition in quantities when firms produce homogenous goods. It is trivial to realize that the first best would entail marginal cost pricing by a single (regulated) firm to minimize the waste in fixed costs of production. Less trivial is to compare the equilibrium with the second best emerging when the strategic interactions are taken as given but the number of firms can be chosen optimally (for instance with entry regulation). This problem has been solved independently by Mankiw and Whinston (1986) and Suzumura and Kiyono (1987), who have shown that there is a general tendency for too many firms to enter in the market.

To exemplify, consider the classic example with linear demand, which can be derived under a quadratic utility. In that case, consumer surplus is $X^2/2E$ and producer surplus is $n(\Pi(n) - F)$. Therefore, in Cournot equilibrium with

n firms, total welfare can be derived as:

$$W = \frac{(a-c)^2 n^2}{2E(n+1)^2} + n \left[\frac{(a-c)^2}{E(n+1)^2} - F \right]$$

This is maximized by the following number of firms:

$$n^* = \sqrt[3]{s(a-c)^2} - 1 \quad (11)$$

which is smaller than the equilibrium one in (3). This tendency toward excessive entry is due to the so-called “business stealing effect”. When deciding whether to enter in the market, each firm takes in consideration the consequences for its own profitability but not the reduction in profitability of the other firms: this externality leads to too many firms entering in the market from a social point of view. Each firm produces too little, but competition is strengthened to the point of reducing the price below the second best level and this increases consumer surplus. Nevertheless, the waste in fixed costs of production is large enough to reduce total welfare. It would be better to limit entry and allow each firm to produce more and exploit the scale effects associated with the presence of entry costs. The same outcome would emerge with any demand function for homogenous goods and any increasing marginal cost.⁷

When we depart from the extreme case of homogenous goods, things become more complex. Since consumers value product differentiation, even the first best requires multiple firms to be active, and it is not clear whether there are more or less firms in equilibrium compared to the second best. The reason is that a new externality emerges: when deciding whether to enter in the market, each firm takes in consideration the consequences for its own profitability, but not the social benefit of producing a new variety. Therefore, the market could either overprovide different varieties (for the business stealing effect) or may underprovide them. The original work of Dixit and Stiglitz (1977) has characterized the constrained-optimal market structure in case of utility given by (4). This is equivalent to one with a number of firms n^* and a common production level x^* for each firm that maximize welfare $W = n^{\theta/(\theta-1)}x$ subject to the resource constraint $px = cx + F$ with $p = E/nx$. This allows one to rewrite welfare as:

$$W = n^{\frac{\theta}{\theta-1}} \left(\frac{E}{cn} - \frac{F}{c} \right)$$

and derive the optimal number of firms:

$$n^* = \frac{s}{\theta} \quad (12)$$

One can easily verify that under both quantity and price competition, the EMS is characterized by excessive entry. Both forms of competition induce excessive

⁷For instance, in the model of von Weizsäcker (1980) with U shaped average costs, we have too many firms producing below the cost-minimizing level $x = \sqrt{2F}$.

mark ups that attract too many firms: under isoelastic preferences the business stealing effect prevails and each firm sells too little at an excessive high markup.⁸ This excess-entry result is also supported by some empirical evidence: in particular, Berry and Waldfogel (1999) have investigated EMSs in radio broadcasting, providing evidence that entry in this market appears to be systematically above the optimal level.

1.3 Beyond CES

One may wonder how general are our results when we depart from CES preferences. To understand this, it is easier to start from the recent contribution of Zhelobodko *et al.* (2012), who have provided the first characterization of the equilibrium with monopolistic competition in case of non-homothetic sub-utilities in consumption $u(x)$ and analyzed the comparative statics. First of all, they show that larger markets increase always the number of firms. Second, when the “relative love for variety” $r(x) = -xu''(x)/u'(x)$ is increasing there is an additional reason for larger markets to reduce prices and increase individual production: the elasticity of substitution is increased by the larger number of goods and competition gets tougher. However, the opposite situation is also possible: when $r(x)$ is decreasing entry induces a price increase. Welfare results are changed as well: Dixit and Stiglitz (1977) proved that when the elasticity of the sub-utility $\rho(x) = xu'(x)/u(x)$ is increasing (decreasing) the equilibrium is characterized by too few (many) goods. In the presence of strategic interactions, these additional effects are present but together with the basic competition effect, for which entry strengthens competition and tends to reduce prices: therefore, broadly speaking, we need rather extreme circumstances to overturn our key conclusions on the negative relation between market size and prices and the excess entry result.

Finally, we need to remark that the general inefficiency of the EMSs leaves space for two opportunities. The first is for the same firms, which may find alternative strategies to reorganize production in a more efficient way and turn this into additional profits: we will examine this possibility in the next section when studying different forms of commitment that firms can adopt to obtain a competitive advantage. The second opportunity is for the government, which may improve the allocation of resources with appropriate entry fees and production subsidies: it is easy to derive the optimal intervention in the examples above, but we will return to optimal policy issues in a more general environment.

⁸Bertrand competition typically generates higher welfare; however, in case of high substitutability and high fixed costs it may not be viable while Cournot competition could allow entry of at least two firms so as to generate higher welfare.

2 Inside the Black Box of the Market Structure

In this section we explore additional strategies adopted by firms to affect the competitive arena, and we extend the analysis in two natural directions. First, we reconsider the strategic process departing from the assumption of perfect symmetry between firms and we analyze the opportunity for firms to precommit to alternative strategies with the purpose of gaining a competitive advantage on the rivals. Second, we study the equilibrium in which all firms can precommit to a particular strategy, and we do it in the relevant case of firms investing to reduce their costs or improve the quality of their goods: this corresponds to extending the EMS approach to endogenize technology in the tradition of Sutton (1991, 1998).

2.1 Limit pricing and competitive advantage

Let us study asymmetric forms of competition in which some firms are able to pre-commit to their own strategies or to other investments before the rivals. Since many sectors are characterized by the presence of leaders which typically have larger market shares, taking them into account allows us to obtain a more realistic description of these sectors. For this purpose we study markets with a leader and endogenous entry of followers. The basic model of Stackelberg competition with endogenous entry has been characterized by Etro (2008) formalizing ideas from older literatures developed without game-theoretic foundations, including the dominant firm approach to the choice of leaders facing a competitive fringe (Markham, 1951), the *Bain-Sylos Labini-Modigliani* approach to limit pricing in oligopoly (Sylos-Labini, 1956, and Modigliani, 1958) and the contestable market theory (Baumol *et al.*, 1982).

To build on our earlier models, let us start from the case of a Stackelberg leader in a market with homogeneous goods and constant marginal costs. Notice that with a fixed number of entrants, a leader would typically expand production to reduce the production of each follower, so as to obtain larger profits than each follower (this is the usual outcome whenever firms' actions are strategic substitutes).⁹ With endogenous entry of firms, however, the equilibrium is rather different. A leader realizes that its own production perfectly crowds out the production of the followers without affecting total production. As a consequence, it is always optimal for the leader to expand production all the way to the point of forcing the exit of all the followers and remain alone in the market with a price that is low enough to avoid entry and still make positive profits. In the linear demand example the equilibrium price can be derived as $p = c + 2/\sqrt{s}$, while in the hyperbolic demand example, the equilibrium markup becomes $\mu = \sqrt{s}/(\sqrt{s} - 1)^2$: they are both higher than their Cournot counterparts, but they are still constrained by the entry threat (to the point that total

⁹For instance, under hyperbolic demand, with one leader and $n-1$ followers, the Stackelberg equilibrium mark up can be calculated as $\mu = n/(n - 1/2)$.

welfare increases). This form of limit pricing reproduces a classic outcome of the contestable market theory, but clarifies the determinants of the equilibrium in a game-theoretic framework. Even if the EMS is radically different from the case of Cournot competition, the endogeneity of entry leads to similar comparative statics. In particular, an increase in the market size forces the leader to produce more and reduce the markup, while profits go up: this is possible because production is reorganized in a more efficient way.¹⁰

The radical result of entry deterrence disappears when we introduce enough product differentiation under both forms of competition. Let us move to price competition, where the endogeneity of the market structure plays a rather surprising role. First of all, as well known, let us remember that a Stackelberg leadership in prices with a given number of followers leads to an accommodating behavior. A leader increases its price because this softens competition and induces each follower to increase its own price as well: the leader ends up with higher mark ups than the followers, and profits go up for the leader and even more for the followers. When entry is endogenous, however, an accommodating behavior is counterproductive because it simply attracts more competitors to the point of dissipating all profits. The optimal strategy for the leader is the opposite one, with a reduction of the price below the price of the followers: in our case of CES preferences, the leader reduces the mark up to the monopolistically competitive level $\mu = \theta/(\theta - 1)$, and the followers adopt the same markup as under symmetric Bertrand competition, with a reduction of the number of entrants. The interesting consequence is that the consumer surplus remains unchanged relative to the symmetric Bertrand model, but the leader has a larger market share with positive profits: the leadership is beneficial because it rationalizes production.

2.1.1 EMSs with Stackelberg competition

What we found above holds much more in general. Consider a market where firms have a profit function nested in the general formulation (1). A Stackelberg equilibrium with endogenous entry is characterized by a strategy for the leader x_L , a strategy for each follower x and a number of firms n that satisfy the profit maximizing and endogenous entry conditions for the followers:

$$\Pi_1 [x, h(x_L) + (n - 2)h(x)] = 0 \quad \text{and} \quad \Pi [x, h(x_L) + (n - 2)h(x)] = F \quad (13)$$

and a profit maximizing condition for the leader. Notice that (13) uniquely defines the strategy of the followers x and their aggregate statistic $\beta = h(x_L) +$

¹⁰Limit pricing requires marginal costs that are constant or not too much increasing. Otherwise the leader chooses a price equal to the marginal cost obtaining a larger production than each follower. For instance, in the model of von Weizsäcker (1980) with linear demand and quadratic costs the output of each follower x remains the same but the leader expands production to $x_L = x + x/E$ so as to reduce the number of followers without changing the price.

$(n - 2)h(x)$ independently from the strategy of the leader, which only affects the number of entrants. This allows one to derive easily the condition for the optimal strategy of the leader (Etro, 2008):¹¹

$$\Pi_1 [x_L, (n - 1)h(x)] = \Pi_2 [x_L, (n - 1)h(x)] h'(x_L) \quad (14)$$

Since the right hand side is negative, the leader is always aggressive compared to each follower (produces more, sells at a lower price), and each follower (when active) chooses the same strategy as in the Nash equilibrium with endogenous entry. Moreover, under our microfounded isoelastic preferences, one can show that consumer surplus is the same in both equilibria.

Whenever a firm is unable to pre-commit to its operative strategy, a competitive advantage can be always reached through strategic investments that affect the subsequent competition. Consider a market where the leader can undertake a commitment on a variable k which affects its profit function $\Pi(x_L, \beta_L, k)$. In case of an exogenous market structure, the analysis of the optimal strategies is quite complex, because it depends on the form of competition and on the impact of the investment on the subsequent strategy of the leader: Fudenberg and Tirole (1985) have shown that in markets with quantity competition leaders adopt strategies that lead to an aggressive behavior (expanding production), while in markets with price competition they adopt strategies aimed at implementing accommodating behavior (increasing prices). This result collapses in case of EMSs: leaders always aim at an aggressive behavior, because otherwise entry dissipates all the competitive advantage. The general principle is that, whenever the strategic investment makes the leader tough ($\Pi_{13} > 0$), there is a strategic incentive to overinvest and whenever the strategic investment makes the leader soft ($\Pi_{13} < 0$), there is a strategic incentive to underinvest, and all this is independent from the form of competition. However, as a consequence of the strategic commitment, the leader is always aggressive in the market, expanding its production or reducing its price to limit entry and conquer larger profits through a larger market share. Below we provide some applications of these general principles.¹²

2.1.2 Research & Development

To show how to apply the previous results, let us analyze investments which reduce the cost of production of a firm. Suppose that an investment k generates a marginal cost of production $c(k)$ which is decreasing in k . Then,

¹¹The optimality condition for the leader can be derived maximizing $\Pi[x_L, (n - 1)h(x)] = \Pi[x_L, \beta + h(x) - h(x_L)]$ where x and β have been proved to be independent of x_L . The three conditions hold in equilibrium if entry deterrence is not optimal for the leader, that is when the marginal costs are increasing enough or goods are differentiated enough.

¹²The general principles are presented in Etro (2006, 2008) and hold with multiple leaders and strategies. For extensions see Tesoriere (2008) on sequential entry, Cato and Oki (2012) on heterogeneity between leaders, Žigić (2012) on the impact of product differentiation on the equilibrium and Anderson *et al.* (2012) on further applications.

any form of quantity competition with profits for the leader $\Pi(x_L, \beta_L, k) = [p(x_L, \beta_L) - c(k)]x_L - k$ implies $\Pi_{13} = -c'(k) > 0$. Analogously, consider a model of price competition where the leader faces a direct demand D^i that takes a form nested in (1) under the change of variable $p_L = 1/x_L$ and has profits $\Pi(x_L, \beta_L, k) = [1/x_L - c(k)]D^i - k$. In this case, the cross derivative with respect to x_L and k is $\Pi_{13} = -c'(k)D_1/x_L^2 > 0$. This implies that under any form of competition, a firm has always a strategic incentive to overinvest in R&D to reduce its marginal cost of production:¹³ under competition in quantities this generates a subsequent commitment to expand production and therefore limit entry and gain market share over the rivals, under competition in prices this forces a subsequent reduction in the price which strengthens competition and allows the firm to conquer more profits through a larger demand. As well known, this differs from what emerges in simple duopolies: there we have a strategic incentive to reduce marginal costs to be aggressive in Cournot duopolies but not in Bertrand duopolies, where a firm would actually prefer higher costs to soften competition. It is the endogeneity of markets that induces firms to be aggressive. Since this holds always and for all the firms, in the next section we will study EMSs in which all firms do invest in R&D.

In some cases, aggressive leaders may even become predators and invest to the point of deterring entry. In a recent important work, Kováč *et al.* (2010) have provided a full characterization of the dynamic equilibrium in the continuous time, with a gradual cost reduction by the leader and endogenous entry of followers. They identify the conditions under which it is optimal for the leader to invest in R&D and accommodate entry or to undertake pre-emptive R&D investment (strategic predation) that eventually leads to the exit of all the followers. Their steady state outcomes correspond to those of the static game, but the dynamic path allows to examine the role of predatory strategies in a new way.

2.1.3 Corporate finance

The celebrated Modigliani-Miller theorem states the irrelevance of the financial structure for a firm with access to perfectly competitive markets. Brander and Lewis (1986) have shown that in the presence of strategic interactions and commitment power on the leverage choice, the Modigliani-Miller irrelevance breaks down and the optimal financial structure depends on the form of competition, whether in quantities or in prices, between a given number of firms. However, this is only a short term perspective: when a firm commits to a certain financial structure as a strategic device, it appears reasonable that this firm is taking into consideration not only the impact on the competition with given rivals, but also the impact on the entire equilibrium market structure expected in the

¹³This is not always the case for R&D to increase the quality of the goods and advertising: these investments increase demand and make it less elastic, with ambiguous impact on the cross derivative.

future. Etro (2010) has studied the optimal financial structure for firms active in markets whose structure is endogenous. Under all forms of competition, the optimal debt-equity ratio should be biased toward debt in case of relevant cost uncertainty: it is exactly in this case that debt induces the equity holders (or their managers) to choose their strategies focusing only on the positive (low-cost) scenarios and expand production or reduce prices. As noticed above, such an aggressive behavior is the only way to gain a competitive advantage in a market whose structure is endogenous. Moreover, one can show that the optimal debt is positively related to the fixed cost of entry, and converges to zero when the fixed cost tends to zero (the Modigliani-Miller theorem holds again when the market structure becomes competitive). Finally, the optimal debt decreases with the variance of the shock because greater uncertainty strengthens the debt commitment on the equity holders.

2.1.4 Incentive contracts

Contractual arrangements between different stakeholders of a firm are crucial for the efficiency of its production process and for its profitability. The literature on contract theory has widely studied these arrangements in the presence of informational asymmetries or contractual incompleteness within a principal-agent relation, but has rarely investigated the interaction between contracts within a firm and competition with other firms. Applying the principles outlined above, Etro (2011a) has derived a number of general results concerning the optimal contracts to be adopted by firms active in markets whose structure is endogenous. Moreover, these principles are independent from the form of competition taking place in the market.¹⁴ For instance, the profit-maximizing contract with the manager requires always a positive weight on sales to induce an aggressive behavior against the rivals. When there is also a problem of moral hazard of the manager engaged in cost-reducing tasks, the optimal contract should be augmented with high-powered incentives with a larger variable compensation compared to the competitors, always with the purpose of inducing an aggressive behavior. Analogously, in case of asymmetric information on the productivity of the same manager, the payment schedule should induce higher effort than the other firms in the market. Another implication is about franchising: vertical separation between an upstream producer and a downstream retailer should always entail wholesale prices below marginal costs for the downstream unit with the purpose of inducing an aggressive behavior of the retailer that generates larger profits to be extracted through a fixed fee.

¹⁴On a similar line of research, Anderson et al. (1997) have studied the gains from privatization of a public firm when this becomes a leader and faces endogenous entry, while Ino and Matsumura (2010) have provided further analysis on contractual and regulatory issues in markets with EMSs.

2.1.5 Implications for competition policy

The analysis of market leaders has wide implications for our understanding of abuse of dominance issues in antitrust policy (see Žigić and Maçi, 2011). What appear to be predatory strategies by market leaders can actually be welfare improving as long as they do not increase the costs of entry in the market: for instance, these are the cases analyzed by Creane and Konishi (2009) on unilateral technology transfers, and by Kováč *et al.* (2010) and Ishida *et al.* (2011) on cost reducing investments. Ino and Matsumura (2012) show that, contrary to common wisdom, a leadership induces always beneficial concentration in markets with homogenous goods and endogenous entry, in the sense that it increases both the Herfindahl-Hirschman Index of concentration and welfare.

Another interesting example is about tying. Whinston (1990) has shown that, when a monopolist in a primary market is also active in a secondary duopolistic market, tying strengthens competition and can be profitable only to deter entry and monopolize the secondary market, so as to hurt consumers. Contrary to this, when the secondary market is characterized by differentiated goods and endogenous entry, tying is a profitable device to reduce prices without fully deterring entry in the secondary market. This result rejects also the single-monopoly profit theorem of the Chicago school on antitrust, for which a monopolist in one market cannot use tying to leverage market power in another market where entry is free. In reality a monopolist can do that, because tying can create larger gains in the secondary market than losses in the primary one: once more, this is possible because of the inefficient pricing emerging under EMSs. However, contrary to the claims of the post-Chicago school, tying does not necessarily have a predatory purpose and it simply strengthens competition in the secondary market.¹⁵

Let us consider mergers finally. Any merger induces a more accommodating behavior of the merging units, which exerts an indirect effect on the competitors: in the traditional case of a fixed number of firms, the competitors become more aggressive under strategic substitutability, and more accommodating under strategic complementarity. For this reason, loosely speaking, mergers tend to be more profitable and anti-competitive under competition in prices. Once again, the situation changes when entry in the market is endogenous, as recently shown by Davidson and Mukherjee (2007), Erkal and Piccinin (2010a,b) and others. In such a case a hypothetical merger would generate the usual accommodating behavior of the merged firms, attracting new entry and reducing the profits after the merger: consequently, there is no strategic rationale for mergers when entry in the market is endogenous. The only way a merger can be profitable is by creating cost efficiencies: this conclusion exactly matches the informal insights of the Chicago school.¹⁶

¹⁵The result is in Etro (2011a). One can even show that under CES preferences tying does not affect consumer surplus in the secondary market, but total profits and welfare increase.

¹⁶Again, under CES preferences, the equilibrium after the merger implies the same consumer

2.2 Endogenous technology

A wide interest in the literature on EMSs has been dedicated to endogenize the technology used by the firms. One way to do this is to assume that firms invest in R&D to change their cost structure. In the previous section we noticed that firms can profit from preliminary R&D investments to expand their market share. But what happens if all firms can invest in cost reducing technologies? Both the number of firms active in the market and their strategies are then affected by the equilibrium investment, and it is crucial to understand how this happens.

The first theoretical analysis of EMSs with endogenous sunk costs is due to Dasgupta and Stiglitz (1980). They considered the case of isoelastic preferences with quantity competition and a specific R&D process, showing that a larger market size is associated with such a larger fixed investment in cost reductions that the endogenous number of firms remains invariant. Assume that the marginal cost depends on the fixed R&D investment k_i as in $c(k_i) = k_i^{-\varrho}$ for any firm i with $\varrho < 1$. Assume that investment and production strategies are taken simultaneously. It is relatively simple to calculate the equilibrium with endogenous entry, which implies the number of firms:

$$n = \frac{(1 + \varrho)(\theta - 1)}{\varrho(\theta - 1) - 1} \quad (15)$$

which is independent from the size of the economy. Nevertheless, since the marginal cost is now endogenous and decreasing in the investment, the equilibrium price is always decreasing in the size of the market. These outcomes justify the same focus of the EMS approach on markets with a small number of firms (even when the size of the market is global), and they are consistent with two robust insights of our models: larger markets generate lower prices and higher production (and R&D investment) per firm.

Only recently, Vives (2008) has been able to generalize this analysis to general profit functions nested in a specification like (1) under quantity and price competition. He shows that an increase in the size of the market increases always individual investment and production but has ambiguous effects on the number of firms, which suggests that the isoelastic example above is rather peculiar. In the case of CES utility with price competition, Vives (2008) has derived the following equilibrium number of firms:

$$n = \frac{(1 - s\varrho)(\theta - 1) + s + \sqrt{[\theta - 1 + s - \varrho(\theta - 1)s]^2 + 4\varrho s\theta(\theta - 1)}}{2\theta} \quad (16)$$

surplus as before. Erkal and Piccinin (2010a) extend this analysis to more complex demand functions: under competition in prices with a demand system derived from a quadratic utility, a merger increases the prices of the merged firms and reduces the prices of the other firms while increasing entry (nevertheless, in the absence of cost efficiencies, the impact on consumer surplus is typically negative). Erkal and Piccinin (2010,b) extend the analysis to R&D cooperation.

which is now increasing in the relative market size $s = E/F$. Finally, notice that the spirit of the other earlier results holds: there is a tendency toward excess entry and market leaders tend to be aggressive through strategic precommitments.¹⁷

2.2.1 The Sutton approach

Sutton (1991) has forcefully proposed the necessity of endogenizing the costs of entry as a preliminary sunk investment for the active firms. In his perspective, the initial investment affects profitability by expanding the demand of the firm through increased quality (or even advertising). The general hypothesis advanced by Sutton (1991), known as the *bounds approach*, is that an increase of the size of the market tends to increase the number of firms only initially, and can then lead to a stable maximum number of firms, that is to a lower bound on the level of concentration in the market.

The theoretical framework proposed by Sutton (1991) requires firms that invest in a first stage $z_i = k_i^\beta$ to obtain quality k_i , and in a second stage they compete in quantities. The parameter $\beta > 0$ reflects the productivity of the investment in quality. Since there is perfect substitutability between quality-adjusted production levels, all firms must provide the same price-quality ratio $p_j/k_j = \lambda$. Therefore industry revenues satisfy $\sum_j p_j x_j = \lambda \sum_j k_j x_j = E$ and firms profits can be derived as:

$$\Pi^i = \frac{k_i x_i E}{\sum_{j=1}^n k_j x_j} - cx(i)$$

It is standard to derive the Cournot equilibrium for given quality levels:

$$x_i = \frac{E(n-1)}{ck_i \sum_{j=1}^n k_j^{-1}} \left(1 - \frac{n-1}{k_i \sum_{j=1}^n k_j^{-1}} \right)$$

which generalizes our earlier result with constant quality. Given this, the preliminary stage of quality choice has one symmetric equilibrium with $k = [2E(n-1)^2/\beta n^3]^{1/\beta}$ which generates net profits decreasing in n and the following endogenous number of firms:

$$n = 1 + \frac{\beta + \sqrt{\beta(8+\beta)}}{4} \quad (17)$$

This is again independent from the size of the market. The introduction of an exogenous fixed cost would lead to a slightly more general result: an increase

¹⁷The model can be extended to uncertainty and asymmetric information on the productivity of the managers in the cost-reduction activity to study equilibrium principal-agent contracts. Etro and Cella (2012) find out a positive relation between competition and high-powered incentives and show that larger markets amplify the effort differentials between good and bad managers.

in the size of the market would increase the number of firms with an upper bound given by (17). Sutton (1991) has also emphasized the possibility of a non-monotone relation between size and entry, but only due to a shift of regime from a traditional one with zero endogenous investment for all firms (where a larger size increases the number of firms) to an “escalation” phase in which market expansion leads to higher investments and higher concentration (a larger size reduces the number of firms). A wide empirical literature, has tried to verify the hypothesis of such an inverted-U relation between market size and number of firms. Sutton (1991, 1998) has provided evidence on markets with large investments in advertising and in R&D. More recently, Ellickson (2007) has looked at the supermarket industry in the U.S. and shown that the minimum level of concentration is highest in the smallest markets, decreases for a range, and then hits a lower limit beyond which it does not fall and might even increase.

2.2.2 A new example

The models of Dasgupta and Stiglitz (1980) and Sutton (1991) are more specific than what may appear at first sight.¹⁸ As shown by Vives (2008), the relation between market size and number of firms is ambiguous in general. Here we show that a very simple setup delivers the non-monotone relation which was conjectured by Sutton (1991). Consider the standard Cournot model with linear demand, marginal cost $c_i = c - \sqrt{k_i}$ and a positive exogenous entry cost $F > 0$. For simplicity, consider the Nash equilibrium in the simultaneous choice of output and R&D investment. With endogenous entry, this implies the investment $k = FE/(4 - E)$ and the production $x = 2\sqrt{k}$. Therefore, in equilibrium, the number of firms is:

$$n = \sqrt{s(a - c)^2(1 - E/4) + E/2} - 1 \quad (18)$$

which depends on the size of the market - compare it with the fixed technology case of (3). First of all, notice that when the positive exogenous sunk cost F goes to zero the endogenous investment z goes to zero as well, and the number of firms goes to infinity: in spite of the endogenous sunk costs, the lower bound on concentration is always zero in this case. The reason is that entry reduces individual sales, profits and incentives to invest, and all of them must vanish (in the long run) when there are not exogenous obstacles to entry. This should not be surprising because the model does not satisfy the conditions found by Sutton (1991) and Vives (2008) for an upper bound on the number of firms. When the exogenous fixed cost is positive, however, the function of the number of firms is U-shaped in the size of the market. To understand the rationale

¹⁸Tandon (1984) has extended the standard Cournot model with linear demand to endogenous entry and sunk costs to confirm that the endogenous number of firms is independent from the size of the market. However, he focused on another very specific example, which we generalize below to obtain new results.

for this non-monotone relation notice that the R&D investment increases more than proportionally with the size of the market. In small markets with few firms and low investment, the mark ups are high enough that an increase in demand is able to expand production and profits as much as needed to attract a larger number of firms. However, in larger markets with many firms, high investments and low mark ups, an increase in demand cannot expand production enough to increase profits and attract more firms. For a given high number of firms, an expansion of the market is going to increase the entry costs more than the gross profits, which in equilibrium requires a smaller number of active firms. Once again the price is decreasing in the size of the market and the production of each firm is increasing in it, as in all the earlier models.¹⁹

3 General Equilibrium EMSs

Until now the size of the market has been taken as exogenous in a partial equilibrium context. More realistically, the structure of a market affects the size of the demand of this market and of other markets in general equilibrium: for instance, a more competitive market may attract more consumption because of lower prices relative to a less competitive market, or, as we will see in the next section, a temporary shock which temporarily reduces markups may temporarily push consumption. In this section we extend the microfounded model of EMSs to a static general equilibrium setup with two factors of production and two sectors.

Consider an economy with total endowments of capital and labor K and L . One sector is perfectly competitive and produces a homogenous good, and the other one is characterized by product differentiation and EMSs. Preferences take the Cobb-Douglas form $V = U^\gamma Y^{1-\gamma}$ over the homogeneous good Y and the differentiated good with the usual consumption index U as in (4). Technology is characterized by a linearly homogenous production function for the competitive sector $Y = F(K_Y, L_Y)$, which is associated with the constant marginal cost $c_Y(w, r)$, and by increasing returns for the differentiated good. In particular, let us assume that the production of each variety requires a fixed cost of η units of labor and takes place according to a linearly homogenous function $x = g(k_X, l_X)$ where k_X and l_X are the inputs used by a representative firm. This function is associated with a constant unitary cost of production $c_X(w, r)$. Therefore, the profit function of each firm can be expressed as (5) under Cournot competition or (8) under Bertrand competition, with total spending corresponding to a fraction γ of total income, $E = \gamma(wL + rK)$ and a marginal cost $c = c_X(w, r)$, while the fixed cost is $F = \eta w$.

¹⁹The inverted-U relation shows up also when the strategies are taken sequentially, investment first and production after (as assumed by Sutton, 1991). In such a case, the equilibrium number of firms is $n = (a - c)^2 s - \sqrt{(a - c)^4 s^2 - 1/2} - 1/2$, which is again an inverted-U curve in market size.

Let us assume that the homogenous good is the *numeraire* and p is the equilibrium relative price of the varieties of the differentiated good. The equilibrium market structure is characterized by the free entry condition:

$$c_Y(w, r) = 1 \quad (19)$$

in the competitive sector, and the endogenous entry condition:

$$x[p - c_X(w, r)] = \eta w \quad (20)$$

in the oligopolistic sector. All the endogenous prices p are given by a markup on the marginal cost:

$$p = \mu(n)c_X(w, r) \quad (21)$$

where the markup $\mu(n)$ is given by (6) under competition in quantities and by (9) under competition in prices. The market clearing conditions for the factor markets employ the *Shepherd Lemma* for which the labor (capital) requirement for a unitary production in sector $j = X, Y$ is given by $a_{Lj} = \partial c(w, r)/\partial w$ ($a_{Kj} = \partial c(w, r)/\partial r$). Therefore, the equilibrium conditions for the labor and capital markets are:

$$a_{LY}Y + a_{LX}nx + \eta n = L \quad (22)$$

$$a_{KY}Y + a_{KX}nx = K \quad (23)$$

Finally, the market clearing conditions for the the market for goods can be combined in a single one by *Walras' Law*:

$$(1 - \gamma)np = \gamma Y \quad (24)$$

The system of six equations (19)-(24) can be solved for the six unknowns w , r , p , x , n and Y .

As an example, consider the case of Cobb-Douglas production functions $Y = K^{\alpha_Y} L^{1-\alpha_Y}$ and $x = k^{\alpha_X} l^{1-\alpha_X}$. This generates marginal cost functions $c_j(w, r) = w^{\alpha_j} r^{1-\alpha_j} / \alpha_j^{\alpha_j} (1 - \alpha_j)^{1-\alpha_j}$ for $j = Y, X$, which allow one to solve explicitly for the unique equilibrium with closed form solutions depending on the form of competition. Focusing on the simplest case, let us assume competition in quantities with homogenous goods ($\theta \rightarrow \infty$) and $\gamma = 1/2$. The equilibrium number of firms can be derived as:

$$n = \sqrt{\frac{L}{\eta} + \frac{\alpha_X^2}{4(2 - \alpha_Y - \alpha_X)^2}} - \frac{\alpha_X}{2(2 - \alpha_Y - \alpha_X)} \quad (25)$$

which provides residually all the other equilibrium values. The total number of firms increases less than proportionally with the endowment of labor, which is the general equilibrium counterpart of the market size effect obtained before. Notice that the number of firms is independent from the capital endowment because we assumed that the entry cost is a labor cost. The general cases of

imperfect substitutability (finite θ) with quantity or price competition are more cumbersome but lead to similar qualitative results. An important task for future research remains to investigate the conditions for existence, unicity and stability of general equilibrium models with EMS (as well known, the problem is much more complex than under perfect competition) and to examine their welfare properties.

3.1 International trade

The above general equilibrium model can be extended to a two-country setup with frictionless trade in goods (and not in inputs).²⁰ It is relatively easy to check that a trading equilibrium with production of both countries in both sectors induces factor price equalization and equal output x and markups $\mu(n^W)$ for all firms (under both price and quantity competition), with n^W total number of domestic and foreign firms. The equilibrium system is augmented with the market clearing conditions for the foreign country and provides a total number of firms which increases less than proportionally with global population. Therefore, markups are reduced when the country opens up and firms' production levels are increased, inducing gains from competition. Assuming imperfect competition in the capital intensive sector and perfect competition in the labor intensive sector, the trading equilibrium induces intraindustry trade in the capital intensive sector with positive net exports of the capital abundant country.²¹

Depending on the general markup functions, the degree of substitutability between goods, and the size of the fixed costs relative to the size of the global market, the generalized $2x2x2$ model would converge to traditional models:

- when substitutability between goods θ approaches infinity for a given fixed cost η (as in the example above) and a given world population, the model converges to the Brander (1981) model under competition in quantities as extended to free entry and general equilibrium by Lahiri and Ono (1995) and Shimomura (1998): this suggests why the general model inherits some of the basic properties of the imperfectly competitive models in which trade increases competition.

- when the fixed cost shrinks (η is reduced) and substitutability between goods is increased (θ increases indefinitely) for a given global population, the model converges asymptotically to the Heckscher-Ohlin model with an increasing number of firms in the capital intensive sector and $\mu \rightarrow 1$, so that both sectors are perfectly competitive in the limit: this suggests why the general model inherits some of the basic neoclassical properties on interindustry trade;

- when the population expands for a given fixed cost η and a given degree of substitutability θ , the model converges to the Krugman model extended to

²⁰Here we assume identical technologies between countries and firms. Cost heterogeneity as a source of trade in the Ricardian tradition has been introduced by Melitz (2003) and will be considered later.

²¹The number of domestic firms is now increasing in the ratio between domestic and foreign capital stock.

general equilibrium as in Helpman and Krugman (1985), with monopolistic competition between an increasing number of firms in the capital intensive sector and $\mu \rightarrow \theta/(\theta - 1)$: this suggests why the general model inherits some of the basic properties of the monopolistic competition models on intraindustry trade;

The general case presents three sources of gains from trade: gains from competition related to the reduction of markups, gains from comparative advantage related to factor endowments and gains from variety related to the increase in product differentiation. Notice that in the absence of labor market imperfections, business destruction is inconsequential for the agents: they switch jobs reallocating their work between a smaller number of firms that have a larger share of the global market. Nevertheless, job destruction due to globalization can have dramatic effects in the presence of labor market rigidities, both in terms of unemployment and income inequalities.²² Sutton (2007) has introduced his model with endogenous sunk costs in a trade context, and he has confirmed the beneficial effect of globalization on international prices. He emphasizes how global markets are characterized by a lower bound to quality, below which firms cannot sell however low their local wages are in general equilibrium. The range of quality levels of the goods produced shifts upwards after opening up to trade. The key question is then how wage adjustments can compensate for low levels of productivity and quality, and under the assumption that material inputs are independently tradable, the model shows that productivity differences can be fully offset by wage differences, but differences in quality cannot. This implies that the initial impact of globalization may be associated with a welfare reduction in countries with intermediate levels of capability. Nevertheless, these countries may be the most important gainers as capabilities are transferred in subsequent phases.

3.2 Implications for Trade policy

Static trade models can provide interesting insights on the role of strategic trade policy, which is usually driven by a profit shifting rationale. Investigations on import tariffs for a domestic market with endogenous entry have been limited to very particular cases until now: those of Cournot competition, with domestic firms producing identical goods that are imperfectly substitutable for those produced by the foreign firms, either in a segmented market (Venables, 1985) or in an integrated market (Horstmann and Markusen, 1986). Since both the domestic and foreign firms end up with zero profits (a reasonable assumption only in the very long run), any profit shifting rationale for trade policy disappears,

²²One could think of sectors producing highly differentiated goods with competition in prices as sectors where innovation and design are the fruit of skilled labor: in these sectors business destruction due to globalization is limited. On the other side, sectors with homogeneous goods and competition in quantities can be seen as sectors characterized by a standard production process employing low-skilled workers: for these sectors, business destruction due to globalization is radical. Further investigations in this direction may enlighten the analysis of the impact of globalization on unemployment and wage inequality.

and the optimal policy is only aimed at improving the terms of trade, exactly as in the neoclassical context. It is crucial to extend these analysis to more general EMSs in the domestic market and also to the case of a fixed number of domestic producers, which is more realistic when the international firms are already active elsewhere and their endogenous entry in the domestic market is faster than the creation of new domestic firms.²³

Trade policy for the domestic exporters has been first analyzed by Horstmann and Markusen (1986) in the case of endogenous entry of both domestic and foreign firms. Again any profit shifting rationale for trade policy disappears in this case (because all domestic and foreign firms end up with zero profits), and the optimal policy is only aimed at improving the terms of trade, typically with an export tax. As shown in Etro (2011b), things are radically different when a domestic exporter is competing with an endogenous number of international competitors for a third market. Export subsidies become always the optimal unilateral policy because they are the only way to provide a strategic advantage to the domestic firm. This happens not only in markets characterized by Cournot competition, but under any form of competition: in particular, export subsidies are optimal also under Bertrand competition, which is in contrast with the celebrated result on the optimality of export taxes found by Eaton and Grossman (1986) for price duopolies. This is another consequence of the principle for which any commitment to an aggressive strategy is beneficial in front of endogenous entry. To verify this, express the profit of the domestic firm in the international market as $\Pi(x_H, \beta_H, k)$, where k is now an export subsidy. Under any form of competition, one can verify that a subsidy implies $\Pi_{13} > 0$, therefore there is always a strategic advantage from the subsidy, which turns the domestic firm into a more aggressive competitor against the international rivals. Since this increases domestic profits more than the direct social cost of the subsidy, the optimal unilateral trade policy requires always an export subsidy.

The above result has also implications for exchange rate policy and R&D policy for international markets: devaluations and R&D subsidies can be used to provide a strategic advantage to the domestic firms engaged in the international competition. For other trade applications see De Santis and Stähler (2004) and Markusen and Stähler (2011) who examine the impact of FDIs in foreign markets with EMSs.

4 Intertemporal EMSs

In this section we study how EMSs evolve over time. To do this, we build on the macroeconomic literature on endogenous entry (Peretto, 1999, Ghironi and Melitz, 2005, Etro and Colciago, 2010, and Bilbiie *et al.*, 2012) and develop

²³In this case, the optimal tariff depends on the elasticity of demand: with isoelastic preferences it can be shown to be negative (i.e. an import subsidy).

a dynamic stochastic general equilibrium model where the market for goods is characterized by imperfect competition and endogenous business creation. The number of firms, the markups and the production of each firm, therefore, change over time. The model will be useful to study macroeconomic issues.

Let us start from the symplest model of an economy with an infinite horizon inspired to Etro and Colciago (2010) and Bilbiie *et al.* (2012). In each period firms can pay a fixed entry cost to start producing from the following period. They produce a homogenous good and compete *à la* Cournot in each period. Production requires a single input, labor, which is inelastically provided by the representative consumer. Investment in business creation is driven by savings. We adopt the simplifying assumption for which savings are a constant fraction s of income. However, income includes both labor income and profits. This allows us to obtain a dynamic model in which it is not investment in physical capital that generates the accumulation of the reproducible input over time, as in the neoclassical framework, but it is entry of new firms that generates the gradual creation of new productive business. Since entry strengthens competition, it also induces a sort of decreasing marginal productivity of business creation, just like capital accumulation reduces the marginal productivity of capital in the Solow model. Here, entry strengthens competition and reduces the marginal profitability of subsequent entry. Therefore, the model generates a gradual convergence toward a steady state which is similar to the one of the Solow model.²⁴

Let us sketch the model, which is fully solved in the Appendix. Consider a representative market for a homogenous good with n_t firms active in each period t . Each firm i produces x_{it} according to the linear production function $x_{it} = A_t l_{it}$, where A_t is the productivity of labor, common to all firms, and l_{it} is the labor input used by firm i . Given the nominal wage W_t , the constant marginal cost of production is $c_t = W_t/A_t$. Cournot competition generates the equilibrium price $p_t = \mu_t c_t$ at time t , where the time varying markup is:

$$\mu_t = \frac{n_t}{n_t - 1} \quad (26)$$

In every period a fraction $\delta \in (0,1)$ of the firms exits from the market for exogenous reasons and n_t^e new firms enter in the market. Therefore, the number of firms follows the equation of motion:

$$n_{t+1} = (1 - \delta) (n_t + n_t^e) \quad (27)$$

²⁴The toy model is a simplified version of Colciago and Etro (2010). The general version with endogenous savings and labour supply with differentiated goods and capital accumulation has been developed in works by Jaimovich and Floetotto (2008), Etro and Colciago (2010) and Bilbiie *et al.* (2012) under different market conditions, but the present toy model allows us to introduce most of the basic applications to macroeconomics. An early model emphasizing imperfect competition in the RBC framework was developed by Rotemberg and Woodford (1992), but its focus was on collusion.

introduced in macroeconomic analysis by Ghironi and Melitz (2005). Entry occurs until the real gross value of a new firm, calculated as the present discounted value of its future expected profits, equates the fixed cost of entry,²⁵ assumed equal to η_t/A_t units of labor. In the Appendix, we show that the equality between savings (out of labor income and profits) and investment (in the creation of new firms) and the market clearing condition for the labor market can be used to rewrite (27) as follows:

$$n_{t+1} = n_t (1 - \delta) + \frac{s(1 - \delta)A_t L_t}{\eta_t \left(1 - \frac{1-s}{n_t}\right)} \quad (28)$$

This equation fully describes the evolution of the number of firms, and therefore of the markups, of the individual production and of individual output over time.

Assume that the labor force is constant at the level L at each point in time, and that the aggregate productivity is fixed with $A_t = A$ and $\eta_t = \eta$. Let us consider the stationary situation to characterize the long run EMSs. Since the right hand side of (28) is increasing in the current number of firms but with a declining slope (smaller than one for a number of firms large enough), we can conclude that the dynamic path of the economy is stable around its unique steady state. When the initial number of firms is low, savings contribute to create new firms, but new firms strengthen competition reducing the profits and the incentives to enter. The steady state number of firms can be derived as:

$$n = 1 + s \left[\frac{(1 - \delta) AL - \eta \delta}{\eta \delta} \right] \quad (29)$$

which is increasing in the savings rate s , in the productivity level A and in the labor force L , and decreasing in the exit rate δ and in the relative size of the fixed costs η . The equilibrium endogenously generates imperfect competition between a positive but limited number of firms producing the homogenous good, with a steady state mark up which is characterized by the opposite comparative statics of the number of firms.²⁶ In its simplicity, this model can be used for multiple purposes, and in the next paragraphs we will provide a short overview of those that will be at the core of the EMSs approach to macroeconomics.

²⁵Devereux *et al.* (1996), Cooper (1999) and Jaimovich and Floetotto (2008) have studied economies where firms pay a fixed cost in each period and entry dissipates all profits in each period. This leads to a less realistic dynamic path. More recently Chang and Lai (2012) have extended the endogenized the fixed costs in a simple macroeconomic model of endogenous entry.

²⁶Of course, the dynamic path of output and consumption (and of the real wage and the interest rate) can be determined residually from the evolution of the number of firms. When the latter increases toward its steady state value, output increases as well toward its steady state value $\tilde{Y} = AL - \eta \delta / (1 - \delta)$. This does not depend on the savings rate: a larger propensity to save increases entry and the number of firms, which enhances competition and wages, but decreases consumption which reduces the profits, and the two effects balance each other.

4.1 Business cycle

Our characterization of the dynamic market structures and of the incentives to create new firms gives rise to a new mechanism of propagation and amplification of shocks that has nothing to do with capital accumulation, intertemporal substitution of labor and price rigidities, all elements that are absent until now. The new mechanism is entirely driven by the relation between profits, firm's value, entry and mark ups.

To see the mechanisms at work in the simple model of the previous section, let us re-introduce a variable aggregate productivity A_t to study the reaction of the EMSs and of the aggregate variables to exogenous shocks and verify the business cycle properties of the model. Consider a temporary positive shock to A_t : this would suddenly increase the productivity and the profits of the existing firms, which in turn would increase their stock market value and attract entry. The temporary increase in the number of firms would strengthen competition so as to reduce the markup, enhance production and increase the real wages (while dampening the impact on the profits). The proportional allocation of output between consumption and savings, which are invested in business creation, contributes to spread gradually the effects of the shock over time. Notice that the impact of the shock on the aggregate variables operates through the stock market, which reflects the value of the firms, the incentives to enter in the market and the impact on competition and on the mark ups. The dynamics of the stock market are due to the presence of imperfect competition between the firms, which generates large operative profits whose expected discounted value is affected by the shocks and affects the entry process. Under perfect competition (for $\eta \rightarrow 0$) any additional propagation mechanism would disappear.

In case of a temporary but persistent technology shock, the effects are stronger. The impulse response of the number of firms becomes hump shaped when the autocorrelation of the shock is high enough, savings are high enough and the exit rate is low enough - this can be seen log-linearizing (28). In this case, the shock induces a gradual increase of the stock market value of the firms and of their number, associated with a gradual reduction of the mark ups: only after a few periods these variables start returning toward their initial levels. Therefore, the EMS approach generates entry-driven cycles characterized by procyclical entry and profits with countercyclical markups. All these three aspects are supported by empirical evidence: for instance, Rotemberg and Woodford (1992) have provided wide evidence of countercyclical mark ups and procyclical profits and Bilbiie *et al.* (2012) have documented the procyclicality of entry.

4.1.1 RBC with oligopolistic competition

The model above has been augmented by Colciago and Etro (2010) with endogenous savings, capital accumulation and labor supply: savings introduce a new

propagation mechanism based on the positive effect of competition on demand, and labor supply strengthens the propagation of the shock through a standard mechanism of intertemporal substitution due to the impact of shocks on real wages through a general equilibrium effect (more competition increases the real wages). The structure of the model is the same as above, except for the fact that the representative household maximizes the following utility:

$$U = E_t \sum_{t=0}^{\infty} \beta^t \left[\log C_t - (1 - u_t) \frac{L_t^{1+1/\varphi}}{1 + 1/\varphi} \right] \quad (30)$$

where β is the discount factor, φ is the Frish elasticity of labor supply and $C_t = \sum_{j=1}^{n_t} x_{jt}$ is the total consumption of a representative homogenous good produced by n_t firms at time t . For now, let us consider unemployment u_t equal to zero in the absence of labor market frictions: all fired workers are immediately employed by a new firm. The representative sector is characterized by firms with a Cobb-Douglas production function $x_{it} = A_t k_{it}^\alpha l_{it}^{1-\alpha}$ that compete in quantities, which leads to the markups (26). Aggregate capital $K_t = \sum_{j=1}^{n_t} k_{jt}$ accumulates in a traditional way and its marginal productivity determines the interest rate. A Euler condition for the optimal savings and an endogenous entry condition close the model. The equilibrium system can be reduced to three equations that fully determine the dynamics of C_t , K_t , and n_t along a stable equilibrium path. In steady state with exogenous labor supply, the equilibrium number of firms can be derived as:

$$n = \frac{\Phi + (1 - \alpha)(1 - \beta) + \alpha\beta\delta}{\alpha\beta\delta + 2(1 - \alpha)(1 - \beta + \beta\delta)} \quad (31)$$

where:

$$\Phi = \sqrt{4\beta(1 - \alpha)(1 - \delta)[1 - \beta(1 - \delta)] \frac{A}{\eta} + [(1 - \alpha)(1 - \beta) + \alpha\beta\delta]^2}$$

The number of firms is increasing in the discount factor β , decreasing in the exit rate δ and increasing less than proportionally with the ratio between total factor productivity A and entry cost parameter η . Moreover, an increase of α increases the steady state number of firms: a higher elasticity of output to capital reduces the wage and therefore the cost of entry, which promotes business creation and competition while reducing the markups.

The model nests the standard RBC framework when $\eta \rightarrow 0$ and $\theta \rightarrow \infty$, but provides a much richer performance in general. In reaction to a standard supply shock, the model is able to generate a larger impact on output, consumption and labor compared to a standard RBC model, together with procyclical responses of entry and profits and a countercyclical response of mark ups. Under standard calibration, the performance of the model with Cournot competition and homogenous goods improves substantially for what concerns the second moments of both output and hours worked (output volatility is 1.69, hours volatility 1.30,

close to those obtained from U.S. data). Moreover, the EMS approach delivers a contemporaneous negative correlation between output and markups similar to that found in the data. Profits' volatility is also positive at 1.57, but the low variability of both mark-ups and profits compared to the data remains an issue to be addressed in future research.

Etro and Colciago (2010) have introduced imperfect substitutability in the model, so that C_t is now the Dixit-Stiglitz consumption index and the markups are now as in (6) under competition in quantities and as (9) under competition in prices for a number of firms n_t in each period. Product differentiation improves the performance of the model with physical capital: when $\theta = 6$, the variability of output and labor reach 1.74 and 1.54 without significant losses on consumption variability. Other forms of imperfect competition, such as asymmetric competition with leaders and followers or imperfect collusion (through conjectural variations models), improve even more the performance of the model (Colciago and Etro, 2010), suggesting that additional work to integrate the industrial organization of markets into macroeconomic models could be quite fruitful.

Bilbiie *et al.* (2012) have augmented a similar model with monopolistic competition to the case of a time-varying substitutability between goods. This depends on the number of firms because of translog preferences, introduced by Feenstra (2003), which generate the elasticity of substitution $\theta_t = 1 + \sigma n_t$: goods become closer substitutes as the number of varieties increases. This implies that the benefits to consumers of additional varieties and the profit incentive for firms to develop new products decrease during booms and increase in recessions because of a change in preferences. For this reason we can refer to this approach as a demand-side explanation for markup countercyclicalities as opposed to the supply-side explanation advanced by Etro and Colciago (2010). Under equivalent calibration the performance of the two models is similar: the Etro-Colciago model displays a higher and more persistent variation in the number of firms and profits, while the model of Bilbiie *et al.* (2012) shows a higher elasticity of the markup function with respect to the number of producers. It is worth emphasizing that the supply side and the demand side explanations are designed to capture different effects. The Bilbiie *et al.* (2012) framework is better suited to describe an economy, or a sector, where entry is associated with product innovation: it identifies the effects on markups due to an increased number of varieties. In Etro and Colciago (2010), instead, entry increases competition and affects markups independently of the degree of substitutability which characterizes goods provided by new entrants. The entry process in many traditional sectors and especially in the service sectors, which are a large part of developed economies, are mostly associated with the creation of new businesses in existing markets (as opposed to new markets). None of these traditional businesses will affect the extensive margin as defined in the statistics in terms of new consumer products, but they will be relevant nevertheless.

4.1.2 Labor and credit market imperfections

Colciago and Rossi (2012) have extended the model to account for labor market frictions which reduce the variability of wages and increase the variability of employment, showing that the EMS approach can contribute to explain also stylized facts concerning the process of job creation and destruction. Since a job is lost when a firm goes bankrupt, which occurs with probability δ , unemployment u_t in (30) must follow a process as $u_t = (1 - m_t)u_{t-1} + \delta(1 - u_{t-1})$, where m_t is the rate of matching between vacancies and unemployed workers (which should be increasing in labor demand and decreasing in the unemployment force). Employment is chosen to maximize the joint surplus of a job match and wages are the result of Nash bargaining between firms and workers (as in the static model with endogenous markups of Blanchard and Giavazzi, 2003). In steady state, the unemployment rate must be increasing in the rate of business destruction and decreasing in all the factors that improve the matching technology. Endogenizing the matching function, Colciago and Rossi (2012) analyze a new channel through which structural parameters affect the labor force and, through that, the EMSs. In line with U.S. data, they find that new firms account for a relatively small share of overall employment, but they create a relevant fraction of new jobs. Moreover, in response to a technology shock the labor share decreases on impact and overshoots its long run level, exactly what we observe in the data. Finally, the propagation on labor market variables is much stronger than in the standard search models because booms increase competition and through that wages and hiring activity.

Gil-Moltó and Varvarigos (2012) have introduced oligopolistic competition as above in a OLG model where agents can choose whether to provide labour or to become entrepreneurs and compete in the production of intermediate goods. The idea that entry is determined through occupational choice has major implications for the industry's dynamics. They find that the industry's convergence to the steady state equilibrium occurs through cyclical fluctuations, despite the lack of any type of exogenous shocks. Furthermore, the path of convergence is not uniquely determined, implying that differences in economic performance may not necessarily reflect differences in either structural characteristics or initial conditions.

Also credit market imperfections could be introduced in the basic framework to generate frictions in the matching of savings and investments. As long as these frictions delay the process of business creation, the impact of a temporary shock can be fundamentally affected, and if they alter the cost of entry in the long run, the result could be steady state EMSs with a lower number of firms and therefore higher mark ups. When access to credit for business creation is positively correlated with the current profitability or with the stock market value because of endogenous constraints on loans (that require collaterals in terms of current profits or stocks), mechanisms of propagations of the shocks could be magnified.

4.1.3 Fiscal policy

Dynamic EMSs generate new implications for fiscal policy, most of which remain to be investigated. From a positive point of view, the impact of fiscal policy is different compared to standard RBC models. A temporary increase in government spending contributes to generate new entry of firms and higher labor demand, which induces a competition effect analogous to the one emerging after a technology shock: as a consequence, private consumption may overshoot and wages increase on impact.

From a normative point of view, the optimal fiscal policy should adjust the inefficiencies of the equilibrium market structure, which generates a new case for countercyclical taxation. First of all, notice that in steady state, there are always too many firms producing too little and selling at an excessive price, just as in the static environment. However, this inefficiency can be even more radical in an intertemporal environment, because excessive investment in business creation can reduce consumption possibilities both in the short and long run (Etro and Colciago, 2010). The optimal allocation of resources can be re-established with a countercyclical production subsidy. Moreover, if the labour supply is endogenous, the optimal policy requires also a countercyclical wage subsidy. Lewis (2008) confirms these results in a related model with endogenous entry of firms active for a single period and with monopsonistic competition in the labor market: the presence of monopoly power in the supply of labor requires an even higher wage subsidy to restore optimality. All this supports the traditional arguments in favor of countercyclical taxation, but notice that this has not the traditional role of minimizing tax distortions (it is optimal also when lump sum taxes are available): countercyclical taxation should promote entry in recessions and refrain it in booms.

A complete characterization of the Ramsey-optimal fiscal policy for a related environment with monopolistic competition is due to Chugh and Gironi (2011). In the absence of production subsidies, dividend payments should be taxed in the long run in a countercyclical way. This policy balances monopoly incentives for product creation with consumers' welfare benefit of product variety, and is in radical contrast to the neoclassical result for which the optimal taxation of dividends is zero in the long run. Moreover, the optimal fiscal policy induces dramatically smaller, but efficient, fluctuations of both capital and labor markets than in a calibrated exogenous policy.

4.1.4 Monetary policy

Recent research has examined monetary policy issues in models augmented with costly price adjustments *à la* Rotemberg. Faia (2012) has focused on Bertrand competition, while Bilbiie *et al.* (2008) and Bilbiie *et al.* (2011) have considered basic monopolistic competition with CES preferences and time-varying elasticity of substitution *à la* Feenstra (2003). They generate a modified New-Keynesian

Phillips curve that ties inflation dynamics to the relative price of investment in new firms (for given expected inflation the current inflation rate is positively related to equity prices) and to the number of firms, which introduces a new degree of endogenous inflation persistence. Lewis (2009) has also augmented the model in a number of directions, showing that sticky wages, costs of entry with congestion externalities and an endogenous exit rate can improve substantially the performance of the model in matching the empirical evidence on the impact of monetary shocks. This class of models can be useful to study new channels of monetary policy: during recessions monetary policy could stimulate business creation through interest rate reductions aimed at increasing the (stock market) value of firms and promoting investments in business creation, and during booms it could limit excessive investments with a tight monetary policy.

None of the above models, however, takes in consideration sticky prices and strategic interactions at the same time. Therefore, a key task for future research remains the introduction of genuine Bertrand competition under Calvo pricing. Real rigidities associated with strategic effects may generate interesting interactions with the nominal rigidities. We can provide a simple insight on this point. Let us reconsider competition in prices between firms producing differentiated goods, and let us assume that a fraction λ of the n_t firms cannot adjust the nominal price, maintaining the pre-determined price level $\bar{p}_{t-1}$, while the fraction $1-\lambda$ can reoptimize. It is easy to show that the Bertrand equilibrium is characterized by new nominal prices reset with the following generalization of the markup formula (9):

$$\mu_t = \frac{\theta n_t \Gamma_t - \theta + 1}{(\theta - 1)(n_t \Gamma_t - 1)} \quad \text{with } \Gamma_t = \lambda \left(\frac{p_t}{\bar{p}_{t-1}} \right)^{\theta-1} + 1 - \lambda$$

which shows that the presence of firms that do not adjust their prices leads also the optimizing firms to adjust less their own prices. This is a consequence of the strategic complementarity between price setters (at the microeconomic level), which is usually ignored in the New-Keynesian analysis, even in the literature focusing on strategic complementarities at the aggregate level (that emphasizes mainly equilibrium multiplicity). In concentrated markets, when firms increase their prices other firms are induced to increase their prices as well, but when only some firms increase their prices, the other firms increase their prices *by less*. As long as the price level is increasing, in the sense that $p_t > \bar{p}_{t-1}$, prices are reset at a new level which is decreasing in the fraction of firms that do not adjust. This suggests that small nominal rigidities can have larger real effects in the presence of strategic interactions. In a dynamic context, where firms adjust their nominal prices to maximize the discounted value of future profits, these forms of strategic interactions would affect the New Keynesian Phillips curve as well.

4.2 International finance

Dynamic models with endogenous entry of firms can be extended to an open economy framework to study trade and international finance. The key papers in the area are those by Melitz (2003) and Ghironi and Melitz (2005), which have augmented standard open economy models with heterogeneity between firms to study issues related to trade and exchange rate fluctuations. Both models abstract from strategic interactions, but their novel analysis of dynamic entry opens space for new investigations of EMSs in the open economy.

4.2.1 The Melitz model

The pathbreaking work of Melitz (2003) has extended the standard Krugman (1980) model to a dynamic setup similar to the one adopted above. Consider a domestic country with a representative agent endowed with a standard intertemporal utility as (30). Each firm i produces $x_{it} = \omega(i)A_t l_{it}$ where $\omega(i)$ is a firm specific factor which is known by each firm only after entry. Exporting abroad requires an extra fixed cost in units of effective labor in every period and transport costs, in form of iceberg costs - selling one unit of good abroad requires shipping $1/d > 1$ units. In each period, entry of domestic firms occurs until the discounted value of expected profits equals the fixed costs of entry, and given the exogenous exit rate δ , the number of firms follows the equation of motion (27).

The foreign country is characterized in the same exact way, possibly with different average productivity, entry cost and export cost. Given the nominal exchange rate E_t (units of domestic currency per units of foreign currency) and the price indexes of the two countries P_t and P_t^* , the consumption-based real exchange rate (units of domestic consumption for units of foreign consumption) is defined as $Q_t = E_t P_t^* / P_t$. In a neoclassical world where all goods are traded (as here in the absence of trade frictions) this should be always unitary, so as to satisfy the purchasing power parity. However, here the endogeneity of the number of goods produced, traded and consumed in each country implies further deviations from the PPP.

Each domestic firm chooses a price for domestic demand in home currency and, in case it is productive enough (relative to the export costs), a price for its exports in foreign currency: under basic monopolistic behavior, the optimal prices are given by a mark up $\mu = \theta/(\theta - 1)$ over the marginal cost, which is simply $W_t/A_t\omega(i)$ for domestic production, and $(W_t/A_t\omega(i))/dQ_t$ for exports. Total profits are the sum of the profits from domestic and foreign sales. Notice that more productive firms set lower prices but earn higher profits, therefore, given the fixed cost of exports in every period, only the firms with productivity above a certain cut-off (depending on the fixed cost of export) do actually export their goods. This “selection effect” implies that there are gains from trade derived from the greater efficiency of foreign exporters: beyond the traditional

gains from variety we have the gains from selection. The strength of the Melitz model, however, is mainly in its capacity to generate a number of new predictions that are widely supported by empirical research: exporters are few highly productive and larger firms that are active in selected foreign countries in line with adjusted gravity equations (see for instance Helpman *et al.*, 2008).

The selection effect emerges also with strategic interactions and cost heterogeneity. Early insights in this sense have been developed by Bernard *et al.* (2003), but the first fully fledged extensions of the Melitz (2003) model to EMSs with strategic interactions are due to Eckel and Neary (2010) and Eaton *et al.* (2012). Eckel and Neary (2010) consider a small number of multiproduct firms engaged in static Cournot competition over a continuum of goods and endogenize the range of goods produced by each firm and the same number of firms. Globalization affects the scale and scope of firms through a competition effect and a demand effect. Confirming recent empirical evidence, their results suggest that multiproduct firms in conjunction with flexible manufacturing play an important role in the impact of international trade liberalization. In particular, the model highlights a source of gains from trade (productivity increases as firms become “leaner” and concentrate on their core competence), but also a source of losses from trade (product variety may fall). Eaton *et al.* (2012) have simulated a model with Bertrand competition and endogenous entry of heterogeneous firms, reproducing a number of stylized facts, including a concave relation between market size and number of firms which is due to the competition effect. Additional theoretical work to study the endogenous number of firms and exporters in models with cost heterogeneity and strategic interactions appears to be quite promising.

4.2.2 The Ghironi-Melitz model

Ghironi and Melitz (2005) have analyzed further macroeconomic implications of the Melitz model. The number of domestic firms n_t , the number of domestic exporters n_{Xt} , and also the average nominal price of domestic firms at home p_t and abroad p_{Xt} are all endogenous variables. This implies that the domestic price index is:

$$P_t = [n_t p_t^{1-\theta} + n_{Xt}^* p_{Xt}^{*1-\theta}]^{1/(1-\theta)} = (n_t + n_{Xt}^*)^{1/(1-\theta)} \tilde{P}_t \quad (32)$$

where we defined the average nominal price level through $\tilde{P}_t^{\theta-1} = r_t p_t^{1-\theta} + (1 - r_t) p_{Xt}^{*1-\theta}$, with $r_t = n_t / (n_t + n_{Xt}^*)$ fraction of domestic firms on the total number of firms selling in the domestic country. Of course, similar formulas apply in the foreign country with $P_t^* = (n_t^* + n_{Xt})^{1/(1-\theta)} \tilde{P}_t^*$. Domestic exports and imports can be expressed as:

$$EXP_t = Q_t n_{Xt} \left(\frac{p_{Xt}}{P_t^*} \right)^{1-\theta} C_t^* \quad \text{and} \quad IMP_t = n_{Xt}^* \left(\frac{p_{Xt}^*}{P_t} \right)^{1-\theta} C_t$$

The model can be closed 1) under financial autarchy with a condition for balanced trade equating domestic exports and imports, from which we can solve for the real exchange rate Q_t ,²⁷ and 2) under free capital flows with a dynamic path for the current account.

Imagine now that the domestic economy is hit by a temporary increase in aggregate productivity A_t . In a standard model with an exogenous number of firms, this would have the typical effect of increasing domestic demand for both national and foreign goods, exerting an upward pressure on wages. Since foreign productivity is not changed, this would lead to an increase in the relative prices abroad which is in contradiction with a wide empirical evidence. However, in the Ghironi-Melitz model where entry is endogenous, the opposite result occurs. The main reason is that the positive shock increases domestic consumption and profits, which attracts entry of new firms in the domestic economy. This increases labor demand at home compared to abroad, and exerts upward pressure on the domestic wages, which causes the relative price of non-traded goods at home to increase relative to foreign. Meanwhile, the positive impact on the profits of the foreign firms exerts an indirect expansionary effect abroad, which contributes to propagate the boom across countries.²⁸

Similar mechanisms would take place in case of international trade in bonds issued by the two countries. In such a set-up the domestic current account reflects the changes in aggregate holdings of domestic and foreign bonds, which must match the opposite changes of the foreign country. Bond trading allows the domestic economy hit by a positive productivity shock to finance entry of new firms by running a current account deficit. Home households borrow from abroad to finance higher initial investments in new firms, while foreign households share the benefits of higher domestic productivity by lending. However,

²⁷Notice that the theoretical counterpart to the empirical real exchange rate (which is usually calculated as the ratio between average nominal price abroad in domestic currency and average nominal price at home) can be derived as:

$$\tilde{Q}_t = \frac{E_t \tilde{P}_t^*}{\tilde{P}_t} = Q_t \left(\frac{n_t^* + n_{Xt}}{n_t + n_{Xt}^*} \right)^{\frac{1}{\theta-1}}$$

which is more interesting for practical analysis than Q .

²⁸Another phenomenon emphasized by Ghironi and Melitz (2005) has to do with the endogenous number of non-tradable goods. The higher relative cost of labor at home reduces the fraction of domestic exporters and increases the fraction of foreign ones. Since only the most efficient firms export their goods, the average price of domestic imports increases (because of the new and more costly imported goods) and the average price of foreign imports decreases (because only less costly imported goods remain). Finally, the relative increase in the number of domestic goods induces the domestic consumers to switch toward domestic production. As long as this is relatively more expensive than the imported production (which arrives from the most efficient foreign firms), there is another effect in favor of the appreciation of the real exchange rate. Ghironi and Melitz (2005) calibrate and simulate the model confirming these results, and showing that a positive technology shock at home increases the average domestic price level compared to the foreign one so as to appreciate the empirical real exchange rate ($\tilde{Q}_t$ decreases).

in case of a temporary shock, the domestic economy runs a deficit initially, but a surplus afterward.

Taking strategic interactions into account in the Ghironi-Melitz model would augment this mechanism with a competition effect analogous to the one emerging in the closed economy. Entry in the domestic country would strengthen competition, reduce the mark ups and increase the real wages even more, dampening to some extent the increase in the number of new firms - once again, in the presence of full fledged EMSs, globalization brings more benefits through a tendency toward mark up reduction rather than through increased number of goods. The entry process would increase labor demand at home, putting additional pressure for an increase in the domestic wage. Moreover, effects of intertemporal substitution would strengthen the domestic boom and the associated entry mechanisms. An interesting work by Atkeson and Burnstein (2008) has introduced strategic interactions in a related quantitative model: when this is parameterized to match salient features of the data on international trade and market structure in the U.S., it can reproduce deviations from relative PPP similar to those observed in the data because firms choose to price-to-market. More research appears to be fruitful to understand the role of the competition effect in open economy models and to examine how pricing-to-market depends on international trade costs and various aspects of the market structures. However, simpler models than the Ghironi-Melitz model are probably needed to make substantial progress in this area.

4.3 Endogenous growth

Finally, we can briefly turn our attention to the process of business creation as a source of growth. As we have seen, our simple model with a process of entry following (28) confirms that the growth rate should be declining toward its steady state level because of the decreasing marginal incentives to create new businesses. As suggested by the theory of endogenous growth (Romer, 1990), perpetual income growth can emerge as the result of externalities in the accumulation of knowledge (i.e.: A_t increases with the number of firms active in the market because each one brings new knowledge and experience to the production process with spillovers on the whole sector) or because innovators build on the shoulders of giants (i.e.: new products are easier to be created). Suppose that the fixed cost of creating new products is decreasing with the amount of goods invented, as with $\eta_t = \bar{\eta}/n_t$. Then, the equation of accumulation of the number of firms (28) becomes:

$$n_{t+1} = (1 - \delta) n_t + \frac{s(1 - \delta) A L n_t^2}{\bar{\eta} (n_t - 1 + s)} \quad (33)$$

which implies a process of perpetual growth for n_t converging to the long run growth rate:

$$\tilde{g} = \frac{s(1 - \delta)AL}{\bar{\eta}} - \delta \quad (34)$$

This is positive as long as the savings rate is high enough or the rate of exit is low enough; income grows at the same rate. Notice that, as in the Romer model, also here long run growth is increasing in the size of the labor force (which implies that opening up to trade would lead to larger growth rates, rather than larger output levels). Moreover, the growth rate is increasing in the savings rate and decreasing in the rate of business destruction and in the size of the costs of entry. However, notice that, contrary to the traditional result of the endogenous growth theory *à la* Romer (1990), the endogeneity of the market structure generates a gradual convergence of the growth rate to its long run level, which is also empirically more plausible. At the beginning of the growth process the incentives to create new firms are high and the rate of increase in the number of firms is high. While firms enter and competition becomes more intense, the rate of entry decreases and the growth rate of production decreases with it. In the long run, the growth rate remains constant because the increase in productivity associated with business creation maintains high the incentives to create new firms. A similar growth process characterized by EMSs in the competition in the market emerges in the model of Peretto (1996, 1999), in which the fixed entry costs are endogenized following the Sutton (1991) approach.

A deeper way in which the creation of new products induces growth is related to vertical (rather than horizontal) innovations. As suggested by the modern revival of the Schumpeterian tradition, this takes place when firms invest not just to create new products, as we assumed until now, but to create them at a lower cost. Of course, this allows us to increase total production through innovations, which is the essence of growth driven by endogenous technological progress. This mechanism relies on the market structure of the innovative activity (rather than the market structure of the productive activity, on which we focused until now). The analysis of patent races allows us to endogenize creative destruction: the rate of exit from the market δ is not exogenous anymore, but depends on the rate of creation of better and cheaper products replacing the old ones, which in turn derives from how many firms invest in R&D and how much each firm invests.

Also here, most of the literature has been based on perfect competition or monopolistic behavior in the market for innovation: in the first case a no-arbitrage condition pins down aggregate R&D investment without identifying how many firms invest in R&D, which ones (incumbents or outsiders) and how much (as in the initial work by Aghion and Howitt, 1992), in the second case one firm or an exogenous number of firms are engaged in R&D investment (as in most models by Aghion and Howitt, 2009). Recent models such as those

of Peretto (2003), Etro (2004), Denicolò and Zanchettin (2010) or Chu, Cozzi and Galli (2012) have contributed to endogenize the structure of the market for innovation. These models allow one to determine how many firms invest in R&D and how much, and under which conditions incumbents and outsiders invest and win the patent races driving growth.

5 Conclusion

In this survey we put together a wide and dispersed literature that analyzes the endogenous structure of markets in different contexts. Many results, as those about the impact of market size on entry and prices, the inefficiency of the decentralized equilibria, the role of strategic investments and the role of policy are common to many models, both in static partial equilibrium or in dynamic general equilibrium. Much more work is needed to improve the communication between different fields adopting the same tools, but the impression is that the literature is slowly developing a potentially new framework to study multiple issues: this framework is not based on perfect or monopolistic competition, but on more realistic forms of imperfect competition between an endogenous number of competitors.

References

- Acemoglu, Daron and Martin Jensen, 2011, Aggregate Comparative Statics, mimeo, M.I.T.
- Aghion, Philippe and Peter Howitt, 1992, A Model of Growth Through Creative Destruction, *Econometrica*, 60, 2, 323-51
- Aghion, Philippe and Peter Howitt, 2009, *The Economics of Growth*, MIT Press
- Anderson, Simon, Andrè de Palma and Jacques Thisse, 1997, Privatization and Efficiency in a Differentiated Industry, *European Economic Review*, 41, 1635-54
- Anderson, Simon, Nisvan Erkal and Daniel Piccinin, 2012, Aggregative Games and Entry, mimeo, University of Virginia
- Atkeson, Andrew and Ariel Burnstein, 2008, Pricing to Market, Trade Costs, and International Relative Prices, *The American Economic Review*, 98, 5, 1998-2031
- Baumol, William, John Panzar and Robert Willig, 1982, *Contestable Markets and the Theory of Industry Structure*, San Diego, Harcourt Brace Jovanovich
- Belleflamme, Paul and Martin Peitz, 2010, *Industrial Organization. Markets and Strategies*, Cambridge, Cambridge University Press
- Benassy, Jean Pascal, 1996, Monopolistic Competition, Increasing Returns to Specialization and Output Persistence, *Economics Letters*, 52, 187-91

- Bernard, Andrew, Jonathan Eaton, Bradford Jensen and Samuel Kortum, 2003, Plants and Productivity in International Trade, *The American Economic Review*, 93, 4, 1268-90
- Berry, Steven and Joel Waldfogel, 1999, Free Entry and Social Inefficiency in Radio Broadcasting, *The RAND Journal of Economics*, 30, 3, 397-420
- Berry, Steven and Peter Reiss, 2007, Empirical Models of Entry and Market Structure, Ch. 29 in *Handbook of Industrial Organization*, M. Armstrong and R. Porter Eds., Elsevier, 3, 1845-86
- Bilbiie, Florin, Fabio Ghironi and Marc Melitz, 2008, Monetary Policy and Business Cycles with Endogenous Entry and Product Variety, in *NBER Macroeconomic Annual 2007*, D. Acemoglu, K. Rogoff and M. Woodford Eds., 299-353
- Bilbiie, Florin, Fabio Ghironi and Marc Melitz, 2012, Endogenous Entry, Product Variety, and Business Cycles, *Journal of Political Economy*, 120, 2, 304-45
- Bilbiie, Florin, Ippei Fujiwara and Fabio Ghironi, 2011, Optimal Monetary Policy with Endogenous Entry and Product Variety, mimeo, Boston University
- Blanchard, Olivier and Francesco Giavazzi, 2003, Macroeconomic Effects of Regulation and Deregulation in Goods and Labor Markets, *Quarterly Journal of Economics*, 3, 879-907
- Brander, James, 1981, Intra-industry Trade in Identical Commodities, *Journal of International Economics*, 11, 1, 1-14
- Brander, James and Tracy Lewis, 1986, Oligopoly and Financial Structure: The Limited Liability Effect, *The American Economic Review* 76, 5, 956-70
- Bresnahan, Timothy and Peter Reiss, 1990, Entry in Monopoly Markets, *Review of Economic Studies*, 57, 531-53
- Campbell, Jeffrey and Hugo Hopenhayn, 2005, Market Size Matters, *Journal of Industrial Economics*, 53, 1, 1-25
- Cato, Susumu and Ryoko Oki, 2012, Leaders and Competitors, *Journal of Economics*, in press
- Chang, Cheng-wei and Ching-chong Lai, 2012, Markups and the Number of Firms in a Simple Model of Imperfect Competition, *Economics Letters*, 116, 3, 277-80
- Chugh, Sanjay and Fabio Ghironi, 2011, Optimal Fiscal Policy with Endogenous Product Variety, NBER Working Paper 17319
- Chu, Angus, Guido Cozzi and Silvia Galli, 2012, Does intellectual monopoly stimulate or stifle innovation?, *European Economic Review*, 56, 4, 727-46
- Colciago, Andrea and Federico Etro, 2008, Real Business Cycles with Cournot Competition and Endogenous Entry, *Journal of Macroeconomics*, Vol. 32, 4, 1101-17
- Colciago, Andrea and Lorenza Rossi, 2012, Endogenous Market Structures and the Labor Market, mimeo, University of Milan, Bicocca
- Cooper, Russell, 1999, *Coordination Games*, Cambridge University Press
- Creane, Anthony and Hideo Konishi, 2009, The unilateral incentives for technology transfers: Predation (and deterrence) by proxy, *International Journal of Industrial Organization*, 27, 3, 379-389

- Dasgupta, Partha and Joseph Stiglitz, 1980, Industrial Structure and the Nature of Innovative Activity, *The Economic Journal*, 90, 266-93
- Davidson, Carl and Arijit, Mukherjee, 2007, Horizontal Mergers with Free Entry, *International Journal of Industrial Organization*, 25, 1, 157-72
- Denicolò, Vincenzo and Piercarlo Zanchettin, 2010, Competition, Market Selection and Growth, *The Economic Journal*, 120, 761-85
- De Santis, Roberto and Frank Stahler, 2004, Endogenous Market Structures and the Gains from Foreign Direct Investment, *Journal of International Economics*, 64, 545-65
- Devereux, Michael, Allen Head and Beverly Lapham, 1996, Aggregate Fluctuations with Increasing Returns to Specialization and Scale, *Journal of Economic Dynamics and Control*, 20, 627-56
- Dixit, Avinash and Joseph Stiglitz, 1977, Monopolistic Competition and Optimum Product Diversity, *The American Economic Review*, 67, 297-308
- Eaton, Jonathan and Gene Grossman, 1986, Optimal Trade and Industrial Policy under Oligopoly, *Quarterly Journal of Economics*, 101, 386-406
- Eaton, Jonathan, Samuel S. Kortum and Sebastian Sotelo, 2012, International Trade: Linking Micro and Macro, NBER WP 17864
- Eckel, Carsten and Peter Neary, 2010, Multi-Product Firms and Flexible Manufacturing in the Global Economy, *Review of Economic Studies*, 77, 1, 188-217
- Erkal, Nisvan and Daniel Piccinin, 2010a, Welfare-Reducing Mergers in Differentiated Oligopolies with Free Entry, *Economic Record*, 86, 273, 178-84
- Erkal, Nisvan and Daniel Piccinin, 2010b, Cooperative R&D under Uncertainty with Free Entry, *International Journal of Industrial Organization*, 28, 1, 74-85
- Etro, Federico, 2004, Innovation by Leaders, *The Economic Journal*, 114, 4, 495, 281-303
- Etro, Federico, 2006, Aggressive Leaders, *The RAND Journal of Economics*, 37, 1, 146-54
- Etro, Federico, 2008, Stackelberg Competition with Endogenous Entry, *The Economic Journal*, 118, 1670-97
- Etro, Federico, 2010, Endogenous Market Structures and the Optimal Financial Structure, *Canadian Journal of Economics*, 43, 4, 1333-1352
- Etro, Federico, 2011a, Endogenous Market Structures and Contract Theory: Delegation, Principal-Agent Contracts, Screening, Franchising and Tying, *European Economic Review*, 55, 4, 463-79
- Etro, Federico, 2011b, Endogenous Market Structures and Strategic Trade Policy, *International Economic Review*, 52, 1, 63-84
- Etro, Federico and Andrea Colciago, 2010, Endogenous Market Structure and the Business Cycle, *The Economic Journal*, 120, 549, 1201-34
- Etro, Federico and Michela Cella, 2012, Equilibrium Principal-Agent Contracts. Competition and R&D Incentives, *Journal of Economic & Management Strategy*, in press

- Faia, Ester, 2012, Oligopolistic Competition and Optimal Monetary Policy, *Journal of Economic Dynamics & Control*, forthcoming
- Feenstra, Ronald, 2003, A Homothetic Utility Function for Monopolistic Competition Models, without Constant Price Elasticity, *Economics Letters*, 78, 79-86
- Fudenberg, Drew and Jean Tirole, 1984, The Fat Cat Effect, the Puppy Dog Ploy and the Lean and Hungry Look, *The American Economic Review, Papers and Proceedings* 74, 361-8
- Ghironi, Fabio and Marc Melitz, 2005, International Trade and Macroeconomic Dynamics with Heterogenous Firms, *Quarterly Journal of Economics*, 120, 865-
- Gil-Moltó, Maria José and Dimitrios Varvarigos, 2012, Industry Dynamics and Indeterminacy in an OLG Economy with Endogenous Occupational Choice, WP 12/09, University of Leicester
- Helpman, Elhanan and Paul Krugman, 1985, *Market Structure and Foreign Trade. Increasing Returns, Imperfect Competition and the International Economy*, MIT Press
- Helpman, Elhanan, Marc Melitz and Yona Rubinstein, 2008, Estimating Trade Flows: Trading Partners and Trading Volumes, *Quarterly Journal of Economics*, 123, 2, 441-87
- Horstmann, Ignatius and James Markusen, 1986, Up the Average Cost Curve: Inefficient Entry and the New Protectionism, *Journal of International Economics*, 20, 225-47
- Ino, Hiroaki and Toshihiro Matsumura, 2010, What role should Public Enterprises play in Free-entry Markets?, *Journal of Economics*, 101, 3, 213-30
- Ino, Hiroaki and Toshihiro Matsumura, 2012, How Many Firms Should Be Leaders? Beneficial Concentration Revisited, *International Economic Review*, in press
- Ishida, Junichiro, Roshihiro Matsumura and Noriaki Matsushima, 2011, Market Competition, R&D and Firm Profits in Asymmetric Oligopoly, *Journal of Industrial Economics*, 59, 484-505
- Jaimovich, Nir and Max Floetotto, 2008, Firm Dynamics, Markup Variations, and the Business Cycle, *Journal of Monetary Economics*, 55, 7, 1238-52
- Krugman, Paul, 1980, Scale Economies, Product Differentiation, and the Pattern of Trade, *The American Economic Review*, 70, 950-9
- Kováč, Eugen, Viatcheslav Vinogradov and Krešimir Žigić, 2010, Technological Leadership and the Persistence of Monopoly under Endogenous Entry, *Journal of Economic Dynamic & Control*, 34, 8, 1421-41
- Lahiri, Sajal and Yoshiyasu Onu, 1995, The Role of Free Entry in an Oligopolistic Heckscher-Ohlin Model, *International Economic Review*, 36, 3, 609-24
- Lewis, Vivien, 2009, Business Cycle Evidence on Firm Entry, *Macroeconomic Dynamics*, 13(5), 605-24
- Mankiw, Gregory and Michael Whinston, 1986, Free Entry and Social Inefficiency, *The RAND Journal of Economics*, 17, 1, 48-58

- Manuszak, Mark, 2002, Endogenous Market Structure and Competition in the 19th Century American Brewing Industry, *International Journal of Industrial Organization*, 20, 673-932
- Markham, Jessy, 1951, The Nature and Significance of Price Leadership, *The American Economic Review*, 41, 891-905
- Markusen, James and Frank Stähler, 2011, Endogenous Market Structure and Foreign Market Entry, *Review of World Economics*, 147, 2, 195-215
- Matsuyama, Kiminori, 1995, Complementarities and Cumulative Processes in Models of Monopolistic Competition, *Journal of Economic Literature*, 33, 2, 701-29
- Melitz, Marc, 2003, The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity, *Econometrica*, 71, 6, 1695-725
- Modigliani, Franco, 1958, New Developments on the Oligopoly Front, *Journal of Political Economy*, 66, 3, 215-32
- Peretto, Pietro, 1996, Sunk Costs, Market Structure and Growth, *International Economic Review*, 37, 895-923
- Peretto, Pietro, 1999, Cost Reduction, Entry, and the Interdependence of Market Structure and Economic Growth, *Journal of Monetary Economics*, 43, 173-96
- Peretto, Pietro, 2003, Endogenous Market Structure and the Growth and Welfare Effects of Economic Integration, *Journal of International Economics*, 60, 177-201
- Romer, Paul, 1990, Endogenous Technological Change, *Journal of Political Economy*, 98, 5, S71-102
- Rotemberg, Julio and Michael Woodford, 1992, Oligopolistic Pricing and the Effects of Aggregate Demand on Economic Activity, *Journal of Political Economy*, 100, 6, 1153-207
- Ruffin, Roy, 1971, Cournot Oligopoly and Competitive Behaviour, *Review of Economic Studies*, 38, 4, 493-502
- Shimomura, Koji, 1998, Factor Income Function and an Oligopolistic Heckscher-Ohlin Model of International Trade, *Economics Letters*, 61, 91-100
- Suzumura, Kotaro and Kazuharu Kiyono, 1987, Entry Barriers and Economic Welfare, *Review of Economic Studies*, 54, 1, 157-67
- Sutton, John, 1991, *Sunk Costs and Market Structure*, MIT Press
- Sutton, John, 1998, *Technology and Market Structure*, MIT Press
- Sutton, John, 2007, Quality, Trade and the Moving Window: the Globalisation Process, *The Economic Journal*, 117, F469-98
- Sylos-Labini, Paolo, 1956, *Oligopolio e Progresso Tecnico*, Giuffrè, Milano
- Tandon, Pankaj, 1984, Innovation, Market Structure, and Welfare, *The American Economic Review*, 74, 3, 394-403
- Tesoriere, Antonio, 2008, Endogenous Timing with Infinitely Many Firms, *International Journal of Industrial Organization*, 26, 1381-88
- Venables, Anthony, 1985, Trade and Trade Policy with Imperfect Competition:

- the Case of Identical Products and Free Entry, *Journal of International Economics*, 19, 1-19
- Vives, Xavier, 2008, Innovation and Competitive Pressure, *Journal of Industrial Economics*, 56, 3, 419-69
- von Weizsäcker, C.C., 1980, A Welfare Analysis of Barriers to Entry, *Bell Journal of Economics*, 11, 2, 399-420
- Zhelobodko, Evgeny, Sergey Kokovin, Mathieu Parenti and Jacques-François Thisse, 2012, Monopolistic Competition in General Equilibrium: Beyond the CES, *Econometrica*, forthcoming
- Žigić, Krešimir and Ilir Maçi, 2011, Competition Policy and Market Leaders, *Economic Modelling*, 28, 1042-49
- Žigić, Krešimir, 2012, Stackelberg Leadership with Product Differentiation and Endogenous Entry: Some Comparative Static and Limiting Results, *Journal of Economics*, 99, 3, 251-66

Appendix: Solving for the dynamic EMSs

In the dynamic model of EMSs, every period is characterized by total expenditure in the market $E_t = p_t C_t = p_t \sum_{j=1}^{n_t} x_{jt}$, where p_t is the equilibrium price equating consumption demand C_t , for the moment taken as given, and supply by all the firms in period t . Nominal profits for firm i are:

$$\Pi_t^i = x_{it} (p_t - c_t) = \frac{x_{it} E_t}{\sum_{j=1}^{n_t} x_{jt}} - \frac{W_t x_{it}}{A_t} \quad (35)$$

We assume that firms cannot credibly commit to future production strategies, therefore they play Cournot competition in each period. If at time t firm i chooses its production x_{it} to maximize its profits taking as given the production of the other firms, the equilibrium generates the following equilibrium price at time t :

$$p_t = \frac{n_t}{n_t - 1} \left(\frac{W_t}{A_t} \right) \quad (36)$$

which provides individual profits $\Pi_t = E_t/n_t^2$ in nominal terms. Since the equilibrium price of the consumption good is p_t , in general equilibrium it is convenient to express all the variables in units of consumption, that is in real terms (alternatively one can use the consumption good as the numeraire). Then, the real profits $\pi_t(n_t) \equiv \Pi_t/p_t$ become:

$$\pi_t(n_t) = \frac{C_t}{n_t^2}$$

and the real wage $w_t = W_t/p_t$ can be derived from the equilibrium pricing relation as $w_t = (n_t - 1) A_t/n_t$. Therefore each firm produces $x_t = C_t/n_t$.

The real gross value of a new firm V_t can be calculated as the present discounted value of its future expected profits. In each period entry occurs until this expected

value equates the fixed cost of entry. Since all firms produce the same homogenous good, it is reasonable to assume that entry of a new firm requires only an extra labor activity to prepare production (rather than a specific monetary investment in R&D to create a new or better product), therefore we assume that the fixed cost of entry F_t is equal to η/A_t units of labor, where $\eta > 0$. Given the wage, the endogenous entry condition amounts to:

$$V_t = F_t = \eta \frac{(n_t - 1)}{n_t}$$

Investment is destined to the creation of new firms. Given the fixed costs of entry F_t and the number of entrants n_t^e , total investment is:

$$I_t = n_t^e F_t = \frac{\eta(n_t - 1)n_t^e}{n_t}$$

where we used the endogenous entry condition.

Assume that the number of workers is given by L_t and each one supplies a unit of labor in each period. Real income in each period must be the sum of profits and labor income, $Y_t = n_t \pi_t(n_t) + w_t L_t = C_t/n_t + w_t L_t$. This income must be allocated between consumption C_t and savings S_t in each period. The market clearing condition that equates savings and investments in every period links the equilibrium number of active firms to the equilibrium interest rate in each period. Therefore, the interest rate depends on the stock market evaluation of the return on the investment in business creation, which depends on the strategic interactions between firms and on the entry/exit process. Finally, total labor demand equates the exogenous labor supply in each period.

To close the model we need to introduce a consumption function. Following the Solow approach, we assume that savings are an exogenous fraction s of income, and consumption is $C_t = (1 - s)Y_t$. From the aggregate resource constraint derived above, this implies:

$$Y_t = \frac{(1 - s)Y_t}{n_t} + \frac{(n_t - 1)}{n_t} A_t L_t$$

where we used the expressions for consumption and wage rate. Solving for income we obtain:

$$Y_t = \frac{(n_t - 1)A_t L_t}{n_t - (1 - s)}$$

which is an increasing function of productivity and labor force, but also of the number of firms and of the propensity to consume $(1 - s)$. The last effects have a Keynesian flavor), even if they operate on the supply side of the economy (rather than on the demand side). Given the number of active firms, a stronger propensity to consume increases aggregate demand and total profits, which in turn increases total output (on multiplier effects in these models see Matsuyama, 1996). Of course, an increase in the number of firms strengthens competition and reduces the profits while increasing labor income. However, as long as part of income is saved and not consumed, the reduction

in total profits is more than compensated by the increase in labor income, so that more firms lead to higher output.

Applying the equality of savings $S_t = sY_t$ with investments I_t as define dabove, we can solve for the equilibrium number of new firms:

$$n_t^e = \frac{sn_t A_t L_t}{\eta(n_t - 1 + s)}$$

Plugging the above expression in (27) we have our final result for the evolution of the number of firms (28).