2. The EMSs Approach to Macroeconomics

Sometimes the road less traveled is less traveled for a reason?
Jerry Seinfeld

In this chapter we start our review of an approach to macroeconomics that is in part alternative and in part complementary to the neoclassical one. We depart from the perfectly competitive environment, in the sense that firms do not take prices as given, but they do choose their strategies and they interact strategically. We focus not only on the choice of prices as the strategic variables, but also on the choice of output levels, and on the choice of entry to produce new or better goods. In most of the analysis of this book we adopt either symmetric Cournot competition or symmetric Bertrand competition as the main models of static strategic interactions, but we will occasionally introduce other forms of competition, as Stackelberg competition or models of imperfect collusion, and we propose a general approach that can be employed with more sophisticated competitive structures borrowed from research in the field of industrial organization. As a matter of fact, one of the main aims of this book is exactly to build a solid bridge between macroeconomics and industrial organization.

The new ingredients of the endogenous market structures (EMSs) approach will be on the supply side of the economy. The technological conditions will be characterized by positive fixed costs of entry so as to move beyond the constant returns to scale hypothesis. To a large extent, we will also depart from the neoclassical assumption that investment (of final goods) builds the physical capital that is used as factor of production together with labor supplied by the working class. That was a good assumption to describe production in the industrialization phase, characterized by the dominance of the secondary (manufacturing) sector and by the social conflict between capital and labor, but not such a good one to describe production in the modern age, dominated by the tertiary (service) sector and by the New Economy, where ideas, innovations, intellectual property rights and creativity are the main inputs needed to create new products, and where the value of start ups without any capital can be high because of these intangible inputs. For this reason, we will embrace a concept of investment (in terms of labor or consumption

F. Etro., *Endogenous Market Structures and the Macroeconomy*,
DOI: 10.1007/978-3-540-87427-0_2,

units) needed to enter in the market with new products (or with better products) produced through labor.[1] This will establish a two-way link between investment and market structure: profitability in the market attracts investment to create new products, and the creation of new products by means of investment enhances competition and reduces profitability in the market.

Finally, we will endogenize the entry decision of the firms as a rational profit maximizing decision. As we have seen at the end of the previous chapter, the New-Keynesian literature has taken into consideration the rational behavior of monopolistic firms in the choice of their profit maximizing prices, but it has typically neglected the rational behavior of the same firms in the choice of entering in the market if and only if positive profits can be expected. As a consequence, there was no link between profit opportunities and production or any other aggregate variable. Our analysis of the entry process leads to the final characterization of the EMSs.

In a sense, our approach can be seen as a natural evolution of the neoclassical approach, which has been guided by the attempt of introducing rational behavior in all the aspects of decision making. The rational theories of consumption and labor supply and the theory of rational expectations (as opposed to adaptive expectations) have been the building blocks of the neoclassical approach. However, a rational theory of entry in markets in which there are profit-maximizing strategic firms has not been introduced until recently.[2] This is the additional contribution of the EMSs approach.

An EMS is defined as *an equilibrium organization of a market where each firm chooses its own strategy to maximize profits taking as given the demand conditions and the strategies of the other firms, and where the number of firms is such that all of them make non-negative profits and further entry cannot provide positive profits.* We will often refer to a simplified situation with a symmetric equilibrium in which all firms choose the same strategy and they obtain the same profits, and we will approximate the exact equilibrium assuming that the number of firms is a natural number. In such a case an

[1] This does not mean that we will ignore the accumulation of stock variables, but only that they will play a different role: we will focus on the development of the stock market value and on the accumulation of innovative ideas.

[2] There is an old partial equilibrium literature which investigates the endogenous entry process on the basis of an adaptive mechanism rather than a rational one. Suppose that gross profits in a market with N_t firms at time t are $\Pi(N_t)$, and that entry of N_t^e new firms depends on the excess profits compared to a fixed cost F according to:

$$N_{t+1} = (1 - \delta_N)(N_t + N_t^e) \quad \text{with } N_t^e = \boldsymbol{\lambda}[\Pi(N_t) - F]$$

where $\delta_N \in (0, 1)$ is a rate of exit from the market and $\boldsymbol{\lambda} > 0$ parametrizes the speed of entry. The evolution of this system can exhibit monotonic or cyclical convergence to the steady state, but complex dynamics can emerge as well. The exogenous and adaptive nature of this process is its limit, which will be avoided by the EMSs approach, where the number of entrants N_t^e derives from an endogenous and rational process.

EMS is defined by *a pair* (x, N) *where* x *is the strategy adopted by each firm and* N *is the number of firms, and the equilibrium satisfies the conditions for profit maximization and endogenous entry.* Notice that the strategy can be given by the production level of the firms or by their prices in case of competition in the market respectively *à la* Cournot or *à la* Bertrand, or by the investment in R&D in case of competition for the market.

In general, in the presence of multiple markets, each market k is characterized by an EMS with (x_k, N_k) and, in the presence of multiple periods, each period t is characterized by EMSs for each market (x_{kt}, N_{kt}) with associated dynamic paths for the equilibrium strategies and the equilibrium number of firms. These converge to steady state EMSs $(\tilde{x}_k, \tilde{N}_k)$ that depend on structural (technological, behavioral, strategic and policy) factors and can be interpreted as the long run EMSs. The crucial aspect of substituting perfect competition or exogenous market structures with EMSs in macroeconomics has to do with the link between demand and supply in general equilibrium. The demand functions perceived by the firms must be the result of the maximization of utility by rational consumers (or by a representative consumer), whose income includes both the remuneration of the factors of production and the eventual profits of the firms (that were zero in the neoclassical approach with perfect competition, or constant in models with an exogenous number of monopolistic firms). In a dynamic model, the discounted value of the firms'profits, represented by the stock market capitalization, reflects both the strategic interactions and the entry/exit process and it affects aggregate demand as well. Therefore, the EMSs approach creates a novel and complex channel that links competition, the stock market and the aggregate economy. In the book we will gradually introduce all these complex elements, but in this chapter we start sketching a simpler model with a single market and a single period to introduce the reader to the main aspects of the EMSs approach. Later on in the chapter we introduce a dynamic setup, and we provide preliminary discussions about the role of EMSs in explaining the determinants of the business cycle, the international trade between countries and the growth process.

In the analysis of industrial organization there are well developed studies on strategic interactions in the Cournotian tradition and on endogenous entry in the presence of fixed costs of production in the Marshallian tradition. The systematic adoption of both elements is more recent, but it is rapidly becoming the standard way to model market structures. One of the first characterizations of EMSs is due to Dasgupta and Stiglitz (1980), who studied competition in quantities and cost reducing strategies with homogenous goods and free entry. Only recently their results have been generalized to product differentiation and competition in prices by Vives (2008). However, the Dasgupta-Stiglitz model has largely inspired the investigations of Sutton (1991, 1998, 2008), who has analyzed markets with strategic interactions in the choice of production and quality, endogenous entry and endogenous sunk costs from both a theoretical and empirical point of view. The analysis of strategic investments and asymmetries in the presence of EMSs has been introduced only recently, with the first general characterization of Stackelberg equilibria with endogenous entry by Etro (2008,b).[3]

The modern empirical literature on EMSs started with the works of Bresnahan and Reiss (1987, 1990) and Berry (1992), which moved beyond the naive view for which lower mark ups are due to more competition associated with a larger number of firms. Such a mechanism definitely holds in the presence of exogenous market structures, but when entry is endogenous there is an opposite mechanism at work: lower mark ups attract a lower number of firms and higher mark ups attract a higher number of firms. In general, the empirical analysis of EMSs requires a different methodology. One possibility is an approach based on the effect that exogenous factors, as the size of demand or other technological conditions, have on the endogenous variables: mark ups, number of firms and production of these firms. Berry and Reiss (2007) review empirical studies within this approach, paying particular attention to equilibrium models that interpret cross-sectional variation in the number of firms or firm turnover rates, and to applications that analyze EMSs in airline, retail, professional, auction, lodging, and broadcasting markets. A more recent approach is based on the impact that entry conditions of different markets exert on the strategic behavior of some firms, and in particular on the leaders. When there are independent variables (or natural experiments) that can discriminate between markets with exogenous or endogenous entry, the predictions of the EMSs approach for the behavior of the leaders can be tested.[4]

The introduction of EMSs in macroeconomic analysis is very recent, even if the microeconomic tools have been available for a while. As we noticed in the previous chapter, the microfounded model of Dixit and Stiglitz (1977) has been widely used in the New-Keynesian macroeconomics assuming monopolistic behavior by an exogenous number of firms, therefore both strategic interactions and endogenous entry have been systematically neglected. The trade literature has mainly focused on one of the two aspects: endogenous entry of monopolistic firms in general equilibrium (Krugman, 1980) or strategic interactions between an exogenous number of firms in partial equilibrium (for instance Brander and Spencer, 1985). Growth theory has endogenized entry

[3] For a comprehensive survey on the industrial organization literature on EMSs see Etro (2007,a). On recent advances of the theory of EMSs in partial equilibrium see Erkal and Piccinin (2007), Ino and Matsumura (2007), Mukherjee (2008), Ishida, Matsumura and Matsushima (2008), Tesoriere (2008,a,b), Žigić (2008), Anderson and de Palma (2008) and Creane and Konishi (2009).

[4] For recent works within the first approach see Manuszak (2002), Mazzeo (2002), Campbell and Hopenhayn (2005), Czarnitzki and Etro (2009) and Chapter 4. For the second approach see Czarnitzki, Etro and Kraft (2008) and Chapter 5. See also Basker (2008) for an interesting analysis of entry in U.S. grocery distribution.

in the competition in the market neglecting strategic interactions (Romer, 1990) and has avoided any strategic consideration in the analysis of the competition for the market (Aghion and Howitt, 1992).

A recent class of models has augmented all these frameworks with the introduction of genuine EMSs, obtaining a number of new positive and normative predictions that we will examine in the next chapters. A few early works on the business cycle (summarized by Cooper, 1999) have introduced monopolistic behavior and endogenous entry in each period within otherwise standard neoclassical models. Other important works by Peretto (1996, 1999) have provided the first systematic attempt to introduce EMSs in the competition in the market in a dynamic general equilibrium model of endogenous growth, and to show the relevance of EMSs for the aggregate behavior of the economy. Etro (2004,a) has provided the first attempt to introduce EMSs in the competition for the market in a dynamic general equilibrium model of Schumpeterian growth. Ghironi and Melitz (2005) have nested trade models with monopolistic behavior and endogenous entry in a DSGE model of the open economy, and this important contribution has opened new research opportunities to study EMSs in macroeconomics. Etro (2007,b) has provided simple examples of the impact of EMSs on trade and business cycles: first, by analyzing strategic interactions and endogenous entry in trade theory and trade policy and second, by studying the impact of shocks on simple two periods models with EMSs. This chapter is based on some of the insights of that work.

The chapter is organized as follows. Section 2.1 introduces strategic interactions and endogenous entry in general models of competition in and for the market. Section 2.2 restricts the attention to microfounded profit functions in partial equilibrium focusing on competition in quantities and in prices with endogenous entry. Section 2.3 studies the particular case of isoelastic utility which will be adopted in multiple applications in the following chapters. Section 2.4 applies the EMSs approach to the simplest dynamic model, that is a two periods exchange economy with endogenous entry in each period. Section 2.5 extends the simple analysis to a general equilibrium context. Section 2.6 develops the first full fledged dynamic model with endogenous entry in the long run and characterizes the equilibrium and steady state EMSs. The analysis keeps savings and labor supply as exogenous, postponing their endogenous characterization to the next chapter. Nevertheless, this simple model allows us to derive in Sections 2.7-2.9 preliminary implications for the three main topics of the book. Section 2.10 concludes.

2.1 EMSs in Partial Equilibrium

When we want to analyze the endogenous structure of a market the first step is to characterize the profit functions of the firms active in this market and to understand how these firms interact strategically. The second step is to understand which firms are endogenously going to be active in this market and to study how demand and supply conditions affect entry and the strategies of the firms. The third step is to understand how the aggregate demand conditions have determined the profit functions of the firms under consideration, which allows us to introduce the market under investigation in a microfounded framework. The fourth step is to introduce this framework in a general equilibrium context.

In this section we focus on the first two steps and we briefly introduce a general class of models of the market structure (used already in the partial equilibrium analysis of Etro, 2007,a) where the profit functions are exogenously given and the EMSs can be characterized in a general way. In the next section, we will restrict our attention to a subset of this class of models where the profit functions are endogenously derived from the utility maximizing behavior of the consumers.

Consider N firms choosing a strategic variable $x(i) > 0$ with $i = 1, 2, ..., N$. These strategies deliver for each firm i the gross profit function:

$$\Pi(i) = \Pi\left[x(i), \beta_i\right] \tag{2.1}$$

where $x(i)$ is the strategy of firm i and we assume that gross profits have always a unique maximum in $x(i)$: $\Pi_1 \gtreqless 0$ for any $x \lesseqgtr \hat{x}$ for some profit maximizing strategy $\hat{x}$. The second argument represents the effects (or spillovers) induced by the strategies of the other firms on firm i's profits, summarized by $\beta_i = \sum_{j=1,j\neq i}^{N} h(x(j))$ for some function $h(x)$ which is assumed positive, differentiable and increasing; these spillovers exert a negative effect on profits, $\Pi_2 < 0$, and of course they affect the profit maximizing strategy. This general framework nests models of competition with strategic substitutability ($\Pi_{12} < 0$), and with strategic complementarity ($\Pi_{12} > 0$). In the former case, typical of Cournot competition, there may be multiple asymmetric equilibria (with firms choosing different strategies), and in the latter case, typical of Bertrand competition, there may be multiple symmetric equilibria. Cooper and John (1988) have emphasized the Keynesian implications of models with multiple equilibria derived from strategic complementarities,[5] but in this book we will not stress this issue, and we will focus on unique symmetric equilibria.

Finally, we assume that entry requires a fixed sunk cost F, so that the net profits of firm i are:

$$\pi_i = \Pi\left[x(i), \beta_i\right] - F$$

Given these profit functions, under the standard assumption of Nash competition between the firms, we can easily characterize the symmetric EMS with the pair (x, N) satisfying the profit maximizing condition:

[5] See also Diamond (1982), Hart (1982) and Murphy, Shleifer and Vishny (1989) for related Keynesian models.

$$\Pi_1\left[x,(N-1)h(x)\right]=0 \tag{2.2}$$

and the endogenous entry condition:

$$\Pi\left[x,(N-1)h(x)\right]=F \tag{2.3}$$

where we used the equilibrium condition $\beta=(N-1)h(x)$.

Such an EMS satisfies a number of properties that are widely discussed in Etro (2007,a). The main properties are the following. First, the number of firms N is always decreasing in the size of the fixed cost of entry (relative to the size of the market).[6] Second, the strategy of each firm x is increasing with the fixed cost of entry (relative to the size of the market) under strategic substitutability, i.e. the firm becomes more aggressive, and it is decreasing under strategic complementarity, i.e. the firm becomes more accommodating. Third, any firm would gain by committing, before entry occurs, to a more aggressive strategy than x, which would reduce the endogenous number of firms N. Fourth, any firm would also gain by committing to strategic investments that lead to a more aggressive behavior than x, which would reduce the endogenous number of firms N.

Most of the common models of competition in the market, that is in the choice of production or pricing for given products, are nested in our general specification. For instance, consider a market with competition in quantities such that the strategy $x(i)$ represents the quantity produced by firm i. The corresponding inverse demand for firm i is $p(i)=p\left[x(i),\sum_{j\neq i}x(j)\right]$ which is decreasing in both arguments (if goods are substitutes). With a generic cost function $c(x(i))$ with $c'(\cdot)>0$, it follows that the gross profits for firm i are:

$$\Pi(i)=x(i)p\left[x(i),\beta_i\right]-c(x(i)) \tag{2.4}$$

with $\beta_i=\sum_{j\neq i}x(j)$. Examples include the case of linear demand $p(i)=a-\sum_{j=1}^{N}x(j)$ for any i, the class of isoelastic demand functions, and other common cases.

Consider now models of competition in prices where $p(i)$ is the price of firm i. Any model with a direct demand $D_i=D\left[p(i),\sum_{j\neq i}g(p(j))\right]$ such that $D_1<0$, $D_2<0$ and $g'(p)<0$ is nested in our general framework after setting $x_i\equiv 1/p_i$ and $h(x(i))=g(1/x(i))$. This specification guarantees that goods are substitutes in a standard way since $\partial D_i/\partial p(j)=D_2 g'(p(j))>0$. Examples include models of price competition with isoelastic demand, Logit demand, or any constant expenditure demand. Adopting, just for simplicity, a constant marginal cost c, we obtain the gross profits for firm i:

[6] Notice that the size of the fixed cost must be compared to the size of the market, which determines the profit opportunities, therefore we can think of F as a the fixed cost relative to the market size. In other words, if gross profits were $\Pi(i)=E\Pi\left[x(i),\beta_i\right]$ with E as a size parameter, the comparative statics of F would be the same as that of F/E.

$$\Pi(i)=\left(\frac{1}{x(i)}-c\right)D\left(\frac{1}{x(i)},\beta_i\right)=(p(i)-c)\,D\left[p(i),\beta_i\right] \tag{2.5}$$

with $\beta_i=\sum_{j\neq i} g(1/x(j))$. This model is nested in our general framework as well.

Notice that in a dynamic framework where entry costs F are born once and the firm remains active over time, the gross value of the firm can be seen as the discounted sum of its profits, something that should reflect the stock market capitalization of the same firm. If r is the constant interest rate, this corresponds to:

$$V(i)=\frac{\Pi(i)}{r} \tag{2.6}$$

and the endogenous entry condition equates this to the fixed cost of entry, so that:

$$V(i)=F \quad \Longleftrightarrow \quad \Pi(i)=rF \tag{2.7}$$

This dynamic framework can be easily extended with an exogenous probability of exit from the market, for instance due to the introduction of a new and better product.[7]

Models of competition for the market focus exactly on the competition to innovate and associate the exit of the incumbent firm with the introduction of a new and better product. These models are also known as patent races because they represent contests to obtain profits from intellectual property rights associated with innovations (which typically provide a temporary monopolistic power). Assume that firms invest a flow of resources in the continuous time to obtain an innovation of exogenous value V^M according to a stochastic process *à la* Poisson. If $x(i)$ is the flow of investment of firm i determining an instantaneous probability of innovation $h(x(i))$, which is assumed to be positive, increasing and strictly concave, the gross value of firm i can be derived as:

$$V(i)=\frac{h(x(i))V^M-x(i)}{r+\delta_N} \tag{2.8}$$

[7] In a discrete environment where $\delta_N\in[0,1)$ is the exit rate in the presence of N firms, the gross value from being in the market at time is:

$$V(i)=\left(\frac{1-\delta_N}{1+r}\right)\Pi(i)+\left(\frac{1-\delta_N}{1+r}\right)^2\Pi(i)+...=\frac{(1-\delta_N)\Pi(i)}{r+\delta_N}$$

In the continuous time, with the instantaneous rate of exit $\delta_N\in[0,\infty)$, we have:

$$V(i)=\int_{t=0}^{\infty}\Pi(i)e^{-(r+\delta_N)t}dt=\frac{\Pi(i)}{r+\delta_N}$$

where $\delta_N = \sum_{j=1}^{N} h(x(j))$ is the instantaneous probability of innovation, which corresponds to the rate of exit of the incumbent firm (now endogenous). It is easy to verify that this case is nested in our general framework after decomposing the exit rate as $\delta_N = h(x(i)) + \beta_i$. Assuming again that the entry cost F is born once and the firms keep doing research until an innovation emerges, endogenous entry must satisfy:

$$V(i) = F \quad \Longleftrightarrow \quad rV(i) = h(x(i))V^M - \delta_N V(i) - x(i) \tag{2.9}$$

whose second expression equates the return on the value of the firm $rV(i) = rF$ with the expected net return from the R&D investment. This takes into account the expected net gain from innovation $h(x(i))\left[V^M - V(i)\right] - x(i)$, and the expected loss in case others innovate $\beta_i V(i)$.

In all these models, we can derive the EMSs and characterize the equilibrium pair (x, N) as a function of the exogenous variables, which is a starting point for comparative static analysis and for the study of the strategic behavior of firms in a realistic market environment. This class of models has proved to be quite useful to investigate a number of positive and normative issues at the microeconomic level. Etro (2007,a) reviews the applications of the EMSs approach to strategic investments in R&D, advertising, quality choices, product differentiation, debt financing and other financial decisions, dynamic forms of competition, issues related to network effects, bundling, vertical restraints, price discrimination, mergers, collusion and liberalizations, and discusses the main implications for antitrust policy.

In this book, however, we want to introduce EMSs in a macroeconomic framework, therefore all of the above mentioned exogenous variables are going to be endogenized sooner or later. For instance, in Chapter 3 we study competition in the market within dynamic stochastic general equilibrium models of the aggregate economy, therefore the demand functions derive from endogenous choices of utility maximizing agents (and policymakers as well), and the cost functions depend on the technology but also on the equilibrium in the market for inputs. In Chapter 4 we study open economies in which decisions taken by firms and consumers in the foreign markets (and by policymakers as well) affect the profit functions of the domestic firms. In Chapter 5 we study models of competition for the market where the value of innovations and the interest rate depend on the general equilibrium of the economy (and by the action of policymakers once again), and they affect accordingly the expected profits of the firms investing in R&D.

Of course, a preliminary investigation of the EMSs in partial equilibrium must be our next step, and we now proceed in this direction focusing on a restricted class of static models of competition in the market whose demand structure can be easily derived from consumers' behavior.

2.2 Microfounded EMSs

In this section we follow the industrial organization literature and analyze a single static market with multiple products characterized by a set of demand functions that are directly derived from the optimal choices of a representative agent with an exogenous endowment. All the firms face common technological conditions. Given these elements, we derive the EMSs in the case of competition in quantities and in prices in partial equilibrium. This framework will be introduced in dynamic and general equilibrium macroeconomic models in later sections.

Consider a representative agent with the following utility depending on the consumption of N goods:[8]

$$U = U\left[\sum_{j=1}^{N} u\left(C(j)\right)\right] \tag{2.10}$$

where $C(j)$ is consumption of good j, $u(C) > 0$, $u'(C) > 0$ with $u''(C) \leq 0$, and $U(\cdot)$ is a positive and increasing function.[9] Notice that these preferences exhibit "love for variety", in the sense that spreading consumption through a larger number of goods increases utility: this reflects complementarities in consumption. The above utility is maximized under the budget constraint:

$$\sum_{j=1}^{N} p(j)C(j) = E \tag{2.11}$$

where $p(j)$ is the price of good j and E is the exogenous endowment of the representative agent. In partial equilibrium this endowment is taken as given. Utility maximization provides the demand for each good and allows us to analyze competition in quantities or in prices. Here we analyze the general case, but at a first reading, one may want to skip the rest of this section and move to the traditional case of isoelastic sub-utilities considered in Section 2.3.

2.2.1 EMSs with competition in quantities

Let us derive the inverse demand functions for the different goods. Utility maximization for each good i implies $u'(C(i)) = \lambda p(i)$ with λ La-

[8] As well known, this specification is due to Dixit and Stiglitz (1977), whose results are commented below. However, the original Dixit-Stiglitz model did not take into account strategic interactions.

[9] Moreover, we assume the regularity condition $u'(C) + Cu''(C) > 0$.

grange multiplier of the budget constraint.[10] Multiplying each side by $C(i)$, summing up over all goods and using the budget constraint, one obtains $\lambda = \sum_j C(j)u'(C(j))/E$. Therefore, if we define with $x(i)$ the production of good i, its inverse demand can be written as:

$$p(i) = \frac{u'(x(i))E}{\sum_{j=1}^{N} x(j)u'(x(j))} \tag{2.12}$$

which is increasing in the endowment and decreasing in the production of each good.[11] Notice that with linear sub-utilities ($u''(C) \to 0$) we would obtain the particular case of homogenous goods as a limiting outcome; namely, the inverse demand would become hyperbolic:

$$p = \frac{E}{\sum_{j=1}^{N} x(j)}$$

for every firm.

If each firm produces at a constant marginal cost c, the gross profit function is:

$$\Pi(i) = \frac{x(i)u'(x(i))E}{\sum_{j=1}^{N} x(j)u'(x(j))} - cx(i) \tag{2.13}$$

which is nested in our general formulation (2.1) with $\beta_i = \sum_{j \neq i} x(j)u'(x(j))$.

In case of Cournot competition between N firms, each one chooses its own output $x(i)$ to maximize profits given the strategies of the other firms, and in the symmetric equilibrium one can derive the following output per firm:

$$x = \frac{(N-1)\left[u'(x) + xu''(x)\right]E}{N^2 u'(x)c}$$

which generates the following price:

$$p = \mu^Q(N,x)c \ \text{ with } \mu^Q(N,x) = \frac{Nu'(x)}{(N-1)\left[u'(x) + xu''(x)\right]} \tag{2.14}$$

where the index Q stands for competition in quantities. The mark up rule is decreasing in the number of firms, but in general it also depends on the

[10] The problem of maximization of (2.10) s.v. (2.11) is equivalent to the problem of maximization of $\sum_{j=1}^{N} u(C(j))$ under the same constraint, which generates the Lagrangian:

$$\mathcal{L} = \sum_{j=1}^{N} u(C(j)) + \lambda \left[E - \sum_{j=1}^{N} p(j)C(j)\right]$$

Its maximization with respect to the consumption of all goods and the Lagrange multiplier λ provides the optimal consumption plan.

[11] This holds under our restrictions on the sub-utilities.

individual production x, and we assume that it is non-decreasing in x.[12] Budget balance requires $x = E/Np$, or $\mu^Q(N) = E/cNx$, which together with (2.14) uniquely defines the mark up as a decreasing function of the number of firms $\mu^Q(N)$. Moreover, for a given number of firms, the mark up is non-increasing in the marginal cost c, and non-decreasing in the endowment E.

The equilibrium gross profits become the following decreasing function of the number of firms:

$$\Pi\left(N\right) = \frac{\left[u'(x) - (N-1)xu''(x)\right]E}{N^2 u'(x)} \tag{2.15}$$

Now, let us use the fact that there is a fixed cost of entry in the market F. Then, when entry is endogenous, the number of firms must be such that these profits are zero. One can solve for the equilibrium number as:

$$N^Q = \sqrt{\frac{E}{F}\left(1 + \frac{xu''(x)}{u'(x)}\right) - \frac{xu''(x)E}{2u'(x)F}} - \frac{xu''(x)E}{2u'(x)F} \tag{2.16}$$

The profit maximizing condition (2.14) and the endogenous entry condition (2.16) together provide the equilibrium value for the pair (x^Q, N^Q), and therefore fully characterize the EMS in partial equilibrium. In this general case the analysis of the comparative statics is complex, but a special case can help us to derive a few basic results.

Consider the case of homogenous goods, corresponding to the limiting case of linear sub-utilities ($u''(x) = 0$). Now the mark up boils down to $\mu^Q(N) = N/(N-1)$, which is decreasing in the number of firms (and independent from marginal costs and endowment), but always larger than one. This allow us to consider the effect of strategic interactions in an otherwise standard setup with perfectly substitutable goods (which has been traditionally studied only under perfect competition in the neoclassical tradition of macroeconomics). Under endogenous entry the number of firms becomes simply:

$$N^Q = \sqrt{\frac{E}{F}} \tag{2.17}$$

which sets the equilibrium mark up at:

$$\mu^Q = \frac{\sqrt{E}}{\sqrt{E} - \sqrt{F}} \tag{2.18}$$

This relations show the simple link between the endowment of the representative agent and the cost of entry on one side, and the EMS on the other side.[13] We can easily verify that increasing the size of a market the number

[12] This turns out to be true under weak conditions on the preferences.

[13] Only when the fixed costs of production tend to zero, the market structure approximates the perfectly competitive one, with infinite firms producing an infinitesimal amount of the uniform good at a price equal to the marginal cost.

of firms increases but less than proportionally, and the mark up decreases. More precisely, the entry of at least N firms requires an endowment above the minimum level N^2F: in other words, if we want to double the number of active firms, we need an endowment that is more than the double.[14]

Finally, when the endowment increases, each firm has to produce at a larger scale, according to:

$$x^Q = \frac{\sqrt{EF} - F}{c}$$

This happens because a larger expenditure 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.[15] In conclusion, notice that one could study alternative models of competition in quantities, as the Stackelberg model, in which one firm is the leader and has a first mover advantage in the choice of its production level. We will analyze this case later on.

2.2.2 EMSs with competition in prices

The utility maximization problem can be used also to express the direct demand functions. This allows us to analyze the case of competition in prices. In particular, inverting the utility maximizing condition $u'(C(i)) = \lambda p(i)$ we have $C(i) = u'^{-1}[\lambda p(i)]$, which must be decreasing in $\lambda p(i)$ by the concavity of the subutility function. Using our expression for the Lagrange multiplier $\lambda = \sum_j C(j)u'(C(j))/E$, we obtain the following function for the direct demand of good i:

$$C(i) = u'^{-1}\left[\frac{p(i)}{E}\sum_{j=1}^{N}\left\{u'^{-1}[p(j)]\right\}\right] \tag{2.19}$$

which is increasing in the endowment, decreasing in $p(i)$ and increasing in the other prices. Gross profits become:

$$\Pi(i) = [p(i) - c]\, u'^{-1}\left[\frac{p(i)}{E}\sum_{j=1}^{N}\left\{u'^{-1}[p(j)]\right\}\right] \tag{2.20}$$

---

[14] Such a prediction can be generalized to models of competition in quantities with imperfect competition, but not to models of competition in prices. It can be tested in the presence of markets of different sizes, for instance professional or retail markets in different towns. A wide empirical literature (Breshnan and Reiss, 1987; Manuszak, 2002) has found encouraging support for this view (see Chapter 4).

[15] This prediction holds in more general models of competition in quantities and prices as well. Campbell and Hopenhayn (2005) provide convincing empirical evidence in its support (see Chapter 4).

which are nested in our general formulation (2.1) with $\beta_i = \sum_{j \neq i} u'^{-1}[p(j)]$.

In a symmetric Bertrand equilibrium, all firms choose a profit maximizing price:

$$p = \mu^P(N)c \;\text{ with } \mu^{P\prime}(N) < 0 \tag{2.21}$$

where $\mu^P(N)$ is an implicit expression for the mark up (which may depend on marginal cost and endowment). Notice that budget balance requires a demand equal to $E/Np = E/N\mu^P(N)c$ for each firm, therefore the equilibrium gross profits must be:

$$\Pi(N) = \frac{\left[\mu^P(N) - 1\right]E}{\mu^P(N)N} \tag{2.22}$$

Under endogenous entry, the number of firms must be such that these profits are zero. An implicit expression for the equilibrium number of firms can be derived as follows:

$$N^P = \frac{\left[\mu^P(N^P) - 1\right]E}{F\mu^P(N^P)} \tag{2.23}$$

The profit maximizing and endogenous entry conditions (2.21) and (2.23) provide together the equilibrium values for the pair (p, N^P), and therefore fully describe the EMS under symmetric competition in prices. Also in this case, generality does not allow us to obtain simple comparative statics results, but the example of the next section will clarify the relation between exogenous and endogenous variables. Finally, notice that also in case of competition in prices one could study the role of a leader within a Stackelberg model, as we will do in the example of the next section. However, before focusing on this example, we need to derive the optimal market structure in this general model.

2.2.3 Optimal EMSs

In their pathbreaking work on monopolistic pricing with product differentiation and free entry, Dixit and Stiglitz (1977) characterized the (constrained) optimal market structure.[16] This is given by a common production level x^O for each firm and by the number of firms N^O that maximize utility $U = U[Nu(x)]$ subject to the zero profit constraint $x(p-c) = F$ and to the resource constraint $pxN = E$, that is:

$$\max_x \frac{Eu(x)}{F + cx}$$

[16] The constraint refers to the zero profits of the firms. The unconstrained first best would adopt marginal cost pricing to maximize utility under the resource constraint.

The optimality condition can be written as:

$$x^O = \frac{F\rho(x^O)}{[1-\rho(x^O)]\,c}$$

where $\rho(x) \equiv u'(x)x/u(x)$ is the elasticity of the subutilities (notice that, contrary to the case of EMSs, the optimal production per firm is independent from the total endowment). Of course this is only an implicit expression unless the sub-utility is isoelastic. The corresponding optimal number of firms can be implicitly written as:

$$N^O = \frac{\left[1-\rho(x^O)\right]E}{F} \tag{2.24}$$

which is linear in the endowment. Comparing this with (2.23), one can notice that the optimal market structure is compatible with a mark up $\mu^O = 1/\rho(x^O)$. Therefore, the EMS under competition in prices is efficient if and only if the equilibrium mark up happens to coincide with the inverse of the elasticity of the utility function.

In the next section, we will explore a particular case of our model in which sub-utilities are isoelastic. In this case the equilibrium mark up in the short run (i.e. with exogenous entry) is higher than the optimal one under both forms of competition, and depends on the number of firms and on the degree of substitutability between goods, but not on the marginal cost and the endowment. This simplifies things at the cost of loosing (in the short run, but not in the long run) the impact of supply and demand conditions on mark ups. Future research should try to take into account more general preferences that deliver richer short-run interactions between supply and demand conditions and mark ups.[17]

2.3 EMSs with Isoelastic Sub-utility

Let us simplify our analysis by introducing isoelastic subutilities, which will be used in a large part of this book. Assume that preferences depend on the consumption of the N goods according to the following index:

[17] More complex utility functions have been usefully analyzed. Feenstra (2003) has introduced translog preferences which microfound an elasticity of substitution increasing in the number of goods. Bertoletti, Fumagalli and Poletti (2008) have introduced a new class of "increasing elasticity of substitution" preferences finding that, even under constant returns to scale, a rise in the number of firms can be price-increasing under both monopolistic and Cournot competition (notice that, despite the price increase, consumers benefit from a rise in the number of monopolistic competitors because of higher product diversity, therefore higher prices are associated with higher consumer welfare).

$$U = \left[\sum_{j=1}^{N} C(j)^{\frac{\theta-1}{\theta}} \right]^{\frac{\theta}{\theta-1}} \tag{2.25}$$

where $\theta > 1$ is the degree of substitutability between goods. When $\theta \to \infty$ the goods become perfect substitutes and generate a hyperbolic demand, when $\theta \to 1$ they tend to complete independence. Of course, intermediate values of θ are associated with imperfect substitutability.

Notice that the elasticity of the sub-utility is $\rho(x) = u'(x)x/u(x) = (\theta - 1)/\theta$, which is constant. Using the results of the previous section, this allows us to determine the (constrained) optimal market structure as characterized by a number of firms:

$$N^O = \frac{E}{\theta F} \tag{2.26}$$

The optimal number of firms can be obtained if the firms adopt a mark up $\mu^O = \theta/(\theta - 1)$ and produce

$$x^O = \frac{F(\theta - 1)}{\theta c} \tag{2.27}$$

which are both independent from the endowment. Incidentally, this is exactly what would emerge if firms were behaving as monopolistic price setters (as in Dixit and Stiglitz, 1977), ignoring the impact of their choices on the price index.[18] However, our interest here, is not on the monopolistic behavior of an infinity of firms, but on strategic interactions between a limited number of firms active in the market.

To microfound the profit function, notice that the representative consumer allocates its endowment E across the available goods with prices $p(i)$ according to the direct demand function:

$$C(i) = C \left(\frac{p(i)}{P} \right)^{-\theta} = \frac{p_t(i)^{-\theta}}{P^{1-\theta}} CP = \frac{p(i)^{-\theta} E}{P^{1-\theta}} \qquad i = 1, 2, ..., N \tag{2.28}$$

where P is a price index defined as:

$$P = \left[\sum_{j=1}^{N} p(j)^{1-\theta} \right]^{\frac{1}{1-\theta}} \tag{2.29}$$

such that total expenditure satisfies $E = \sum_{j=1}^{N} p(j)C(j) = CP$.

[18] Of course, this would be a reasonable assumption in the presence of an infinity (or a very high number) of firms selling different varieties of the same good - by the way, this is a situation at odds with the same concept of monopolistic behavior.

Inverting the direct demand functions, we can also derive the system of inverse demand functions:

$$p(i) = \frac{x(i)^{-\frac{1}{\theta}} E}{\sum_{j=1}^{N} x(j)^{\frac{\theta-1}{\theta}}} \qquad i = 1, 2, ..., N \tag{2.30}$$

where $x(i)$ is the consumption of good i.

In the following sections we analyze different forms of competition that can take place between the firms and derive the associated EMSs.

2.3.1 Cournot competition

First, let us consider competition in quantities. Using the inverse demand function, we can express the profit function of a firm i as a function of its output $x(i)$ and the output of all the other firms:

$$\begin{aligned} \Pi(i) &= [p(i) - c]\, x(i) = \\ &= \frac{x(i)^{\frac{\theta-1}{\theta}} E}{\sum_{j=1}^{N} x(j)^{\frac{\theta-1}{\theta}}} - cx(i) \end{aligned} \tag{2.31}$$

Assume now that each firm chooses its production $x(i)$ taking as given the production of the other firms. The first order conditions:

$$\left(\frac{\theta-1}{\theta}\right) \frac{x(i)^{-\frac{1}{\theta}} E}{\sum_j x(j)^{\frac{\theta-1}{\theta}}} - \left(\frac{\theta-1}{\theta}\right) \frac{x(i)^{\frac{\theta-2}{\theta}} E}{\left[\sum_j x(j)^{\frac{\theta-1}{\theta}}\right]^2} = c$$

for all firms $i = 1, 2, ..., N$ can be simplified imposing symmetry of the Cournot equilibrium. This generates the individual output:

$$x = \frac{(\theta-1)(N-1)E}{\theta N^2 c} \tag{2.32}$$

Substituting into the inverse price, one obtains the equilibrium price $p = c\theta N/(\theta-1)(N-1)$, which is associated with the equilibrium mark up:

$$\mu^Q(\theta, N) = \frac{\theta N}{(\theta-1)(N-1)} \tag{2.33}$$

which is a particular case of (2.14). Notice that the mark up is decreasing in the degree of substitutability between products θ, with an elasticity $\epsilon_\theta^Q = 1/(\theta-1)$. As long as the number of firms is finite, the markup remains positive for any degree of substitutability. Finally, the mark up is decreasing and convex in the number of firms and it tends to $\theta/(\theta-1) > 1$ for $N \to \infty$.

Its elasticity is $\epsilon_N^Q = 1/(N-1)$, which is decreasing in the number of firms (the mark up decreases with entry at an increasing rate) and independent from the degree of substitutability between goods.

Gross profits can be expressed as:

$$\Pi^Q(\theta, N) = \frac{(N+\theta-1)E}{\theta N^2} \tag{2.34}$$

If the fixed cost of entry is F, entry will take place and will reduce the individual profits as long as the gross profits are higher than this fixed cost. In equilibrium, the zero profit condition leads to the following number of firms:

$$N^Q = \frac{E}{2\theta F}\left[1+\sqrt{1+\frac{4\theta(\theta-1)F}{E}}\right] \tag{2.35}$$

which is larger than the optimal number (2.26). This excessive entry result generalizes to a wider context (Mankiw and Whinston, 1986) and has also found some empirical evidence.[19] Moreover, the equilibrium number of firms increases in a less than proportional way with the size of the market (E/F), contrary to what happens in the case of monopolistic behavior of each firm (or in the optimal market structure). Larger markets induce stronger competition, as can be verified from the equilibrium markup:

$$\mu^Q(\theta, N^Q) = \frac{\theta}{(\theta-1)}\left(1-\frac{2\theta F}{E+\sqrt{E^2+4\theta(\theta-1)FE}}\right)^{-1} \tag{2.36}$$

which is decreasing in ratio E/F. This implies that the size of a market has to more than double to allow the entry of a double number of firms. Nevertheless, comparing this EMS with the (constrained) optimal market structure we can conclude that competition in quantities leads to an excessive mark up and to an excessive number of firms.

Finally, we can calculate the production of each firm as:

$$x^Q = \frac{F(\theta-1)}{\theta c}\Phi\left(1-\Phi\right) \quad \text{with } \Phi = \frac{2\theta F}{E+\sqrt{E^2+4\theta(\theta-1)FE}} \tag{2.37}$$

which is decreasing in the marginal cost of production and increasing and concave in the endowment. The former result shows that positive cost shocks induce a larger production by each firm. The latter shows that positive demand shocks (increasing the endowment of the consumers) increases the production of each firm as well: this happens because each firm has to produce more to cover the same fixed costs at a lower mark up.

[19] Berry and Waldfogel (1999) have investigated EMSs in radio broadcasting, providing evidence that entry is systematically above the optimal level.

2.3.2 Bertrand competition

Let us now consider competition in prices. In each period, the gross profits of firm i can be expressed as:

$$\Pi(i) = \frac{[p(i) - c]\, p(i)^{-\theta} E}{\left[\sum_{j=1}^{N} p(j)^{-(\theta-1)}\right]} \tag{2.38}$$

Firms compete by choosing their prices. Contrary to the traditional Dixit-Stiglitz (1977) approach which neglects strategic interactions between firms, we take these into consideration and derive the exact Bertrand equilibrium. Each firm i chooses the price $p(i)$ to maximize profits taking as given the price of the other firms. The first order condition for any firm i is:

$$\left\{p(i)^{-\theta} - \theta\, [p(i) - c]\, p(i)^{-\theta-1}\right\} + \frac{(\theta-1)p(i)^{-\theta}\, [p(i) - c]\, p(i)^{-\theta}}{\sum_{j=1}^{N} p(j)^{1-\theta}} = 0$$

Notice that the last term is the effect of the price strategy of a firm on the price index: higher prices reduce overall demand, therefore firms tend to internalize their impact on the price index and set higher mark ups compared to the case of monopolistic behavior. Imposing symmetry between the N firms, the equilibrium price p must satisfy:

$$\left[\theta\,(p-c)\, p^{-\theta-1} - p^{-\theta}\right] N p^{-(\theta-1)} = (\theta-1) p^{-\theta}\,(p-c)\, p^{-\theta}$$

Solving for the equilibrium we have $p = c(\theta N - \theta + 1)/(\theta - 1)(N - 1)$, which implies the following mark up:

$$\mu^P(\theta, N) = \frac{1 + \theta(N-1)}{(\theta-1)(N-1)}$$

The mark up under competition in prices is always smaller than the one obtained before under competition in quantities, as well known for models of product differentiation. As in the previous case, the mark up is decreasing in the degree of substitutability between products θ, with an elasticity $\epsilon_\theta^P = \theta N/(1-\theta+\theta N)(\theta-1)$ which is always higher than ϵ_θ^Q: higher substitutability reduces mark ups faster under competition in prices. Moreover, contrary to the case of competition in quantities, the mark up under competition in prices vanishes in case of homogenous goods $\lim_{\theta\to\infty} \mu^P(\theta, N) = 1$, a well known result in industrial organization. Finally, the mark up is again decreasing in the number of firms, with elasticity $\epsilon_N^P = N/\left[1 + \theta(N-1)\right](N-1)$, which is decreasing in the level of substitutability between goods, and approaching zero when the goods become homogenous.

In conclusion, with competition in prices the individual gross profits can be expressed as:

$$\Pi^P(\theta, N) = \frac{E}{1 + \theta(N-1)} \tag{2.39}$$

Given total expenditure, the number of firms and the degree of substitutability, it is easy to verify that the profits under competition in prices are smaller than those under competition in quantities.

If the fixed cost of entry is F, the endogenous entry condition that sets net profits equal to zero provides the following number of firms:

$$N^P = \frac{E}{\theta F} + \frac{\theta - 1}{\theta} \tag{2.40}$$

which is linearly increasing in the endowment and decreasing in the fixed cost of entry. The corresponding equilibrium markup is:

$$\mu^P(\theta, N^P) = \frac{\theta E}{(\theta - 1)(E - F)} \tag{2.41}$$

which is increasing in the fixed cost of entry and decreasing in the endowment. Notice that, given the total expenditure, the fixed costs and the degree of substitutability, competition in prices generates a smaller number of firms compared to competition in quantities. Moreover, if we take the integer constraint (on the number of firms) into account, we can verify that the equilibrium number of firms can be above the (constrained) optimal number by at most one firm.

Finally, one can easily verify that $\mu^Q(\theta, N^Q)$ is always bigger than $\mu^P(\theta, N^P)$, which means that the EMSs under competition in quantities are characterized by more firms but they preserve higher prices than competition in prices:

$$N^Q > N^P \ \text{ and } \mu^Q(\theta, N^Q) > \mu^P(\theta, N^P)$$

This shows that the index of concentration is a poor measure of the market power as an expression of the ability of firms to price above the marginal cost. When entry is endogenous, low mark ups are consistent with high concentration and *vice versa*.

Under competition in prices, the production of each firm is:

$$x^P = \frac{F(\theta - 1)(E - F)}{[E + (\theta - 1)F]\, c} \tag{2.42}$$

which is decreasing in the marginal cost and increasing in the endowment, as it was under competition in quantities. Therefore, cost and demand shocks affect the production of each firm in similar ways.

In conclusion, these EMSs provide two main differences compared to the case of monopolistic firms *à la* Dixit-Stiglitz: mark ups are reduced and individual production is increased when the size of the market increases, while they are constant in case of monopolistic firms. Moreover, under competition in prices the endogenous number of firms increases linearly with the size of the market, as in the case of monopolistic firms *à la* Dixit-Stiglitz, but under competition in quantities it increases in a less than proportional way.

2.3.3 Stackelberg competition

The EMSs can be used to study more complex forms of competition. In this section we extend the symmetric models of competition in quantities and in prices with the introduction of market leaders. In the industrial organization jargon, these are firms able to commit to their own strategies before the so-called followers. Since many markets are characterized by the presence of incumbent firms which typically have larger market shares than their rivals, taking them into account allows us to obtain a more realistic picture of the EMSs. The model of Stackelberg competition with endogenous entry has been introduced by Etro (2008,b) in a static set up as the one considered until now.[20]

For the sake of simplicity, let us consider first the case of competition in quantities and homogeneous goods, that is $\theta \to \infty$. With one leader and N followers playing simultaneously, the equilibrium mark up can be derived as:

$$\mu^S(N) = \frac{N}{N - 1/2}$$

which is lower compared to the mark up under pure Cournot competition. The profits of the leader and the representative follower are respectively larger and smaller than the profits under Cournot competition, but the impact of a change in the number of firms on the equilibrium mark up and production is qualitatively analogous to the Cournot case. In Chapter 3 we will employ also this market structure in a dynamic macroeconomic model to examine the role of market leaders over the business cycle.

Contrary to the case of an exogenous number of firms, the static model of Stackelberg competition with endogenous entry is characterized by a radically different market structure with only one firm active: actually, whenever the goods are homogeneous and the marginal cost of production is constant, the leader produces enough to deter entry. In our example, the equilibrium output of the leader is $x^L = (\sqrt{E} - \sqrt{F})^2/c$ and the equilibrium mark up is:

$$\mu^S = \frac{1}{\left(1 - \sqrt{\frac{F}{E}}\right)^2} \tag{2.43}$$

[20] Further extensions can be found in Maci and Žigić (2008), Žigić (2008), Tesoriere (2008,a,b), De Bondt and Vandekerckhove (2008) and Kováč, Vinogradov and Žigić (2009).

which is higher than the one emerging in the absence of the leader. Even if the EMS is radically different from the case of Cournot competition, the endogeneity of entry leads to similar comparative statics: an increase in the endowment or a reduction of the fixed cost of entry force the leader to produce more and to keep the mark up lower. The main difference compared to the Cournot case is that here the leader obtains positive profits in spite of free entry. In a recent important work, Kováč, Vinogradov and Žigić (2009) have extended the analysis to a dynamic setup: they analyze a oligopoly model in which a leader invests in process innovations facing subsequent endogenous entry by followers, and identify conditions under which it is optimal for the leader in an initially oligopoly setup with endogenous entry to undertake preemptive R&D investment (strategic predation) that eventually leads to the exit of all followers.[21]

The radical result of entry deterrence disappears when we introduce imperfect substitutability between the goods, that is when θ is low enough. Consider the general case of quantity leadership in the presence of imperfect substitutability. The Stackelberg equilibrium with endogenous entry is characterized by a larger production for the leader compared to the followers, and entry of a lower number of firms compared to the Cournot equilibrium with endogenous entry. The characterization of the equilibrium is relatively simple, with the leader selling at the monopolistic mark up, and with an endogenous number of followers adopting the same production level as under symmetric competition in quantities and the same mark up (2.36). Therefore, the equilibrium is summarized by the following mark ups for the quantity leader (index L) and for the representative follower (index F):

$$\mu^{LQ} = \frac{\theta}{\theta - 1}, \ \mu^{FQ} = \left(\frac{\theta}{\theta - 1}\right)\left(1 - \frac{2\theta F}{E + \sqrt{E^2 + 4\theta(\theta - 1)FE}}\right)^{-1} \tag{2.44}$$

[21] The technological leader adopts the accommodation strategy only when (roughly speaking) his R&D efficiency is low or/and the size of the market is relatively small. In all other cases, the leader opts for strategic predation aiming to achieve the monopoly position after certain time T. During the predation period (up to T), the leader might be willing even to incur losses in order to enjoy monopoly profit from time T onward. Thus, unlike a static game, in a fully dynamic model the costs of predation last only for a limited period and have to be contrasted to the infinite stream of monopoly profit afterwards. The time pattern of R&D investment crucially depends on the equilibrium strategy. If accommodation is the optimal strategy, then the leader chooses an R&D path which steadily increases over time towards the unique steady-state value. When, on the other hand, strategic predation is the optimal strategy, the leader first invests significantly in R&D in order to achieve the monopoly position in time T. After all rivals are eliminated, the leader may continue to increase his R&D investment as an unconstrained monopolist or to prevent the rivals from re-entering the market. Nevertheless, this investment level is still higher than in the case of accommodation. From a welfare point of view, the predation regime is optimal because it implies high R&D investments, but the target time T is usually suboptimal.

In spite of this asymmetric EMS due to the presence of a leader, the endogeneity of entry leads to similar conclusions as before: a larger endowment or a lower fixed cost attract further entry of followers, increase their individual outputs and reduce their mark ups, with a positive impact on total production.

Let us finally consider the case of price leadership with imperfect substitutability. Under competition in prices, the Stackelberg equilibrium with endogenous entry is characterized by the leader committing to a lower mark up compared to the followers. In particular, the leader adopts again the monopolistic price, and the followers adopt the same price as under symmetric competition in prices (2.41). Therefore, the respective mark ups for the price leader and the representative follower become:

$$\mu^{LP} = \frac{\theta}{\theta - 1}, \;\; \mu^{FP} = \left(\frac{\theta}{\theta - 1}\right)\left(1 - \frac{F}{E}\right)^{-1} \tag{2.45}$$

This result is in striking contrast with the usual outcome under price leadership and exogenous entry, for which leaders adopt higher prices than the followers to relax competition. When entry is endogenous, the only way for the leaders to obtain positive profits is to adopt an aggressive strategy. When the endowment increases or the cost of entry decreases, more followers are attracted in the market, and they reduce their mark ups and increase their production, while the leader maintains the lowest price.

Notice that these results on the behavior of the market leaders have substantial implications for industrial policy, since they show that large market shares by leading firms can be the result of strong entry pressure rather than of market power, and antitrust policy should be more concerned about verifying the entry conditions in a market rather than associating large market shares with dominant positions. A similar result, which we will discuss in Chapter 5, emerges in case of competition for the market, where incumbent leaders tend to invest more than their rivals only when entry is endogenous: this leads to the conclusion that also the persistence of leadership can be the consequence of strong entry pressure rather than of market power.[22]

2.3.4 Collusion, endogenous entry costs and other extensions

The framework that we adopted is tractable enough to take into account other forms of competition. We could adopt the conjectural variations approach to introduce imperfect collusion in a stylized way: in such a case, each firm adopts an exogenous conjecture on the reaction of the other firms to its

[22] For a wider discussion of the antitrust implications of this model see Etro (2008,c). On recent related advances of antitrust theory see Fumagalli, Motta and Rønde (2008), Katsoulacos (2008), Polo and Immordino (2008) and Fernández, Hashi and Jegers (2008).

strategy, and this conjecture can reproduce competitive and collusive equilibria or any intermediate case (including Cournot equilibria). We will follow this *ad hoc* model of imperfect collusion in the dynamic general equilibrium model of Chapter 3.

We could also analyze multiproduct firms which choose the production levels or the price levels of their goods to maximize the joint profits. All these and other models would lead to equilibria with mark ups $\mu(\theta, N)$ and profits $\Pi(\theta, N)$ decreasing in the number of firms, and therefore to well defined EMSs. Notice that from an empirical perspective, one could be interested in estimating these mark up and profit functions as depending on the number of firms in different markets (defined according to the degree of substitutability between products).

Vives (2008) has extended the model to endogenous costs assuming that the fixed cost of production is an investment in R&D aimed at reducing the marginal cost of production. In general, he finds that increasing the endowment increases the investment in cost reduction and the output of each firm, but with ambiguous consequences on the number of firms. For instance, consider our case of isoelastic preferences with an isoelastic cost function. In such a case a larger market size is associated with such a larger fixed investment in cost reductions that the endogenous number of firms remains the same. This result for the case of competition in quantities and homogenous goods is originally due to Dasgupta and Stiglitz (1980). They assume that demand is hyperbolic and that the marginal cost depends on the fixed R&D investment F as in $c = \varkappa F^{-\varrho}$ with $\varkappa, \varrho > 0$. The Nash equilibrium in the choice of output and R&D investment with endogenous entry implies the investment $F = E\varrho^2/(1+\varrho)^2$, the number of firms:

$$N^Q = 1 + 1/\varrho \tag{2.46}$$

the mark up (on the endogenous marginal cost):

$$\mu^Q = 1 + \varrho \tag{2.47}$$

and the production per firm:

$$x^Q = \frac{E^{1+\varrho}}{\varkappa\varrho}\left(\frac{\varrho}{1+\varrho}\right)^{2(1+\varrho)} \tag{2.48}$$

Notice that the number of firms and the mark up are now independent from the size of the market, but the individual production is still increasing in it. Similar results emerge in the case of product differentiation and also with competition in prices. Finally, an increase in the degree of product substitutability increases per-firm output and cost reduction expenditure, while reducing the number of firms as a consequence of the stronger competition.

Until now, we have limited our analysis to the case in which each firm is active for a single period only. A more realistic situation emerges when each firm is active in multiple periods, or has always a positive probability of being active in the future. In the absence of credible commitments to future strategies, we can assume that in each period the existing firms compete according to one of our static models. In such a case, the gross value of the firms would be the present discounted value of the future profits and endogenous entry would still require equalization of the initial fixed costs of entry to the gross stock market value of the same firm. This creates a dynamic behavior of the number of firms that is reflected on the equilibrium mark ups and, through them, on the aggregate behavior of the macroeconomy. Starting with Section 2.6 we will extend our analysis in this direction studying dynamic market structures. Finally, we need to notice that a multi-period framework would allow one to study dynamic models of market competition in which firms can commit to multi-period strategies or in which forms of imperfect collusion can be sustained as subgame perfect equilibria of supergames (Rotemberg and Saloner, 1986; Rotemberg and Woodford, 1992) or as Markow perfect equilibria (Maskin and Tirole, 1988).

However, before embarking in more complex analysis, we still need to extend our basic framework to account for two important aspects: intertemporal links between markets and general equilibrium considerations. Following the strategy of Chapter 1, the next section extends the static analysis of EMSs to the simplest dynamic situation, that is the one characterized by two periods only. This allows us to appreciate the potential role of EMSs in a dynamic framework.

2.4 EMSs in a Two Period Economy

In this section we follow an example by Etro (2007,a) of a two-period economy where an exogenous endowment is allocated between current and future consumption. Imperfect competition and endogenous entry in the goods market of both periods generates a novel link between exogenous shocks and real choices which works through the impact on the endogenous mark ups.

Consider a two period model of an exchange economy with logarithmic subutilities:

$$U = logC_1 + \beta logC_2 \tag{2.49}$$

where $\beta \in (0, 1)$ is the discount factor. The consumption good is homogenous and it is produced by multiple firms in each period, so that the consumption index boils down to $C_t = \sum_{j=1}^{N_t} C_t(j)$ for $t = 1, 2$. Firms compete in quantities. The interest rate r and the endowment of the agent E are assumed exogenous for simplicity. One can think of this as a small open economy facing a given international interest rate.

Given the price levels in the two periods p_1 and p_2, the corresponding budget constraints are:

$$C_1 = \frac{E - S}{p_1} \quad C_2 = \frac{S(1+r)}{p_2}$$

Utility maximization requires the demand of consumption $C_1 = E/(1+\beta)p_1$ in the first period and $C_2 = \beta(1+r)E/(1+\beta)p_2$ in the second one, which imply the inverse demand functions:

$$p_1 = \frac{E}{(1+\beta)C_1} \qquad p_2 = \frac{\beta(1+r)E}{(1+\beta)C_2}$$

In each period, N_t firms compete in quantities producing at a marginal cost c_t. For simplicity, assume $(1+r)\beta = 1$ in what follows. Defining $x_t(i)$ as the production of firm i in period t, we have the gross profit functions:

$$\Pi_t(i) = \frac{Ex_t(i)}{(1+\beta)\sum_{j=1}^{N_t} x_t(j)} - c_t x_{it}$$

In Cournot equilibrium, each firm produces $x_t(i) = E(N_t-1)/(1+\beta)N_t^2 c_t$, and the equilibrium price is $p_t = \mu_t(N_t)c_t$ where the mark up function is:

$$\mu_t(N_t) = \frac{N_t}{N_t - 1} \tag{2.50}$$

which is decreasing in the number of competitors. Therefore, we obtain the following modified Euler equation:

$$\frac{C_2}{C_1} = \frac{c_1\mu_1(N_1)}{c_2\mu_2(N_2)} \tag{2.51}$$

The traditional outcome of perfect competition emerges in case of constant returns to scale, here equivalent to the absence of fixed costs of production. In such a case, endogenous entry implies an infinite number of firms, prices are equal to the marginal cost in both periods, and relative consumption is linked to the ratio of marginal costs only: $C_2/C_1 = c_1/c_2$. Of course, under constant technology we have consumption smoothing ($C_2/C_1 = 1$). The neoclassical theory of the business cycle is largely based on this mechanism: a permanent increase in productivity does not affect the relative marginal cost and consumption, but a temporary increase in productivity (a reduction in c_1/c_2) induces an increase in relative consumption (a decline of C_2/C_1). Finally, notice that an exogenous change of the endowment does not affect prices and relative consumption.

When the markets are characterized by positive fixed costs of production, however, only few firms can be active and entry strongly affects relative prices and consumption. As a preliminary example, imagine that the fixed cost of entry in period t is F_t, and entry is endogenous. Then, in each period t we have a markup:

$$\mu_t = \frac{1}{1 - \sqrt{\frac{(1+\beta)F_t}{E}}} \tag{2.52}$$

and a number of firms:

$$N_t = \sqrt{\frac{E}{(1+\beta)F_t}}$$

This result shows that an increase in the endowment (or a reduction in the fixed cost of production) increases the number of firms and reduces the markups. Relative consumption can be calculated as:

$$\frac{C_2}{C_1} = \left(\frac{c_1}{c_2}\right)\left[\frac{\sqrt{E} - \sqrt{(1+\beta)F_2}}{\sqrt{E} - \sqrt{(1+\beta)F_1}}\right]$$

This shows two mechanisms due to the endogeneity of the market structures (and completely absent under perfect competition). The first is rather straightforward: an increase in the fixed cost of entry in one period increases the relative consumption in the other period, and *vice versa*. In particular, a reduction in the future costs of entry leads to consumption growth: for instance, the introduction of a general purpose technology that is going to reduce entry costs (say cloud computing) should exert a positive impact on growth.

The second mechanism is less intuitive: an exogenous increase in the endowment increases the relative consumption of the good produced by a lower number of firms. Suppose $F_1 > F_2$, which implies that more firms are active in the second period and $p_1/c_1 > p_2/c_2$: under these circumstances, an increase in E increases C_1 relative to C_2.

Assume now that the fixed cost of production is related to the marginal cost $F_t = \eta c_t$, as it typically happens when both fixed and variable costs require the same combination of inputs (for instance just labor). In such a case, we obtain a magnification effect of the technology shocks. Rewriting the optimality condition as:

$$\frac{C_2}{C_1} = \left(\frac{c_1}{c_2}\right)\left[\frac{\sqrt{E} - \sqrt{(1+\beta)\eta \cdot c_2}}{\sqrt{E} - \sqrt{(1+\beta)\eta \cdot c_1}}\right] \tag{2.53}$$

one can notice that a reduction in the marginal cost of the first period is going to increase relative consumption in the first period more than proportionally. This new propagation mechanism works through endogenous entry. A temporary shock reduces the marginal cost, which makes current consumption more attractive. Moreover, the reduction in the entry costs induces more firms to enter in the market, temporarily increasing competition. This induces a temporary reduction in the equilibrium mark up, which exhibits countercyclicality. Accordingly, the shock makes current consumption even more attractive.

In conclusion, in the presence of EMSs characterized by competition in quantities and endogenous entry, the impact of a temporary productivity shock on consumption is magnified through the impact of entry on the mark up. Notice that the result would be affected by changes in the degree of intertemporal substitution (assuming a utility function with a higher elasticity of substitution than the logarithmic one, the impact of the temporary shock on the relative consumption in the first period would be strengthened).

Moreover, if we introduce endogenous labor supply, a temporary productivity shock would generate a standard intertemporal substitution mechanism in the labor choice. This would be magnified through the competition effect: a temporary productivity shock would induce a temporary increase in the real wage, which would generate higher labor supply in the short run. In other words, the EMSs create an additional channel through which traditional intertemporal substitution mechanisms (in consumption and labor supply) work.

Finally, the model could be extended to imperfect substitutability between goods produced by different firms introducing a separate consumption index as (2.25) for each period:

$$U = log \left[\sum_{j=1}^{N_1} C_1(j)^{\frac{\theta-1}{\theta}} \right]^{\frac{\theta}{\theta-1}} + \beta log \left[\sum_{j=1}^{N_2} C_2(j)^{\frac{\theta-1}{\theta}} \right]^{\frac{\theta}{\theta-1}}$$

and examining competition in quantities and prices. While the mark up rules would change, the same logic of the results above would go through, because the modified Euler equation (2.51) still holds.

Summarizing, our outcome is dependent on two differences from the standard neoclassical set up. The first is the departure from the assumption of constant returns to scale: fixed costs of entry imply increasing returns to scale in the production function. The second difference relies on the form of competition: here we adopted standard competition in quantities, but more general models of strategic interaction as those examined in the previous sections would deliver analogous results.

Before turning to more complicated dynamic extensions, we still need to introduce our tractable static model in a general equilibrium framework to complete our overview of the EMSs approach. We will do this in the simplest possible way in the next section.

2.5 EMSs in General Equilibrium

General equilibrium analysis concerns multiple markets interacting between each other. Most of the literature on general equilibrium has been focused on the case of perfectly competitive markets with price-taking firms and without fixed costs of production or entry.[23] Only limited efforts have been dedicated to the analysis of markets with imperfect competition or with strategic interactions between few agents - see Bonanno (1990), Mas Colell, Whinston and Green (1995, Ch. 18) and Gabszewicz (2003). Even this literature has been mostly aimed at providing non-cooperative foundations of Walrasian equilibria studying the strategic behavior of the agents (rather than the firms), verifying the limit properties of the equilibria when the number of agents increases, and deriving conditions for existence, uniqueness and stability of the equilibria. Moreover, it has systematically neglected the role of fixed costs of productions or other technological non-convexities in endogenizing entry of firms in the markets.

In our view, general equilibrium theory should try to provide a deeper understanding of aggregate phenomena in the presence of strategic interactions (of different kinds) between firms and with endogenous entry in each market. The aim of this section is not to provide such a full-fledged analysis of EMSs in general equilibrium, but to introduce the simplest general equilibrium extension of our partial equilibrium static model. This example will be generalized in the next section and in future chapters.

Imagine that firms produce the goods employing labor only, which in turn is supplied in fixed quantity. One unit of labor produces A units of good. Moreover, the fixed cost of creation of a new firm corresponds to the cost of η/A units of labor, where $\eta > 0$. The nominal unit wage is W.

The representative agent provides L units of labor and maximizes the same utility as in (2.25) under the same budget constraint (2.11). However, the endowment E is now endogenous and it depends on labor income and, using the fact that the representative agent is the only shareholder of all the firms, it depends on net profits too. Summing up, the endogenous endowment becomes:

$$E = WL + \left[\sum_{j=1}^{N} \Pi(j) - N \frac{\eta W}{A} \right] \tag{2.54}$$

This allows us to derive the demand function for each good as a function of both labor income and the profits of the same firms. Therefore, the individual profits of each firm depend on the aggregate profits as well, and so on in a circular way. However, assuming that the firms take aggregate profits as given, competition takes place as before and the aggregate profits amount to zero under endogenous entry.[24] Accordingly, the endogenous endowment simplifies to $E = WL$.

[23] See Ellickson (1993) for a nice overview.

[24] This is not the case with asymmetric forms of competition. For instance, under Stackelberg competition, there are positive profits for the industry leaders even if there is endogenous entry of followers. These profits should be taken into account in the demand functions.

Using this, one can express the EMS in general equilibrium as a function of total labor supply and of the fixed cost parameter. For instance, in the case of homogenous goods and competition in quantities we have a general equilibrium number of firms given by:

$$N^Q = \sqrt{\frac{AL}{\eta}} \tag{2.55}$$

and an equilibrium markup:

$$\mu^Q = \frac{\sqrt{AL}}{\sqrt{AL} - \sqrt{\eta}} \tag{2.56}$$

Both the two markets of this economy, the one for the goods and the one for labor, are in equilibrium, and this allows us to examine the new general equilibrium effect that EMSs create in the labor market. Adopting the price of the goods as the *numeraire*, we can derive an expression for the (real) wage:

$$w^Q = \left(1 - \sqrt{\frac{\eta}{AL}}\right) A \tag{2.57}$$

This shows that the real wage is increasing in the aggregate productivity, decreasing in the size of the fixed cost, and increasing in the total labor supply. The first comparative static result is standard, and the wage is lower than the marginal productivity because of market power. The last two results derive from the impact of endogenous entry on competition in general equilibrium: larger markets or lower fixed costs attract entry, which in turn strengthens competition, reduces the mark ups and shifts resources from extra profits toward labor remuneration. These general equilibrium EMSs can be easily extended to the case of product differentiation with competition in quantities with similar implications.

In case of competition in prices the general equilibrium number of firms becomes:

$$N^P = \frac{AL}{\theta\eta} + \frac{\theta - 1}{\theta} \tag{2.58}$$

and the equilibrium markup is:

$$\mu^P = \frac{\theta AL}{(\theta - 1)(AL - \eta)} \tag{2.59}$$

Again, we cannot derive equilibrium prices without a normalization. However, in case of product differentiation, it is convenient to express the real wage as the ratio between the nominal wage and the price index, which in the symmetric equilibrium is $P = pN^{1/(1-\theta)}$. Therefore, the real wage becomes:

$$w^P = \frac{(\theta-1)(AL-\eta)}{\theta AL}\left(\frac{AL}{\theta\eta}+\frac{\theta-1}{\theta}\right)^{\frac{1}{\theta-1}} A \tag{2.60}$$

which is again a fraction of the aggregate productivity, and is decreasing in the size of the entry cost, but increasing in the size of the labor force. Larger and more accessible markets attract entry, which reduces mark ups on one side and increases the number of different varieties on the other side. Now, both effects increase the real wage: the former because it leads to a reduction in the average price, the latter because it increases the purchasing power for a given income (because of the love for variety effect).

The competition effect of market size and entry costs on mark ups and real wages appears extremely simple and possibly trivial in such a static model with a single market, but it will be at the source of a number of crucial results in the presence of multiple markets or multiple periods, in particular when intertemporal substitution mechanisms are available (as in Chapter 3) or when intra-industry trade between countries occurs (as in Chapter 4). Loosely speaking, the fundamental reason is that, contrary to what happens in the neoclassical approach, here prices depend on both the marginal cost of production and the mark up, and the latter is affected by shocks through the entry mechanism. In turn, the impact on the mark ups is transmitted to the real wages through the general equilibrium mechanims shown in this section, and any change in mark ups and wages has an impact on consumption and labor supply choices. In turn, this feeds back on profits and affects the entry decisions and with them the mark ups and the real wages. In such a way, the EMSs create a new mechanism of propagation of shocks in general equilibrium.

As of now we have largely discussed microfounded EMSs in a static framework in partial and general equilibrium and in a simple dynamic framework with two periods only. It is time to approach a more ambitious task and to build a fully dynamic general equilibrium model, which should give to the reader the ultimate flavor of the EMSs approach to macroeconomics.

2.6 EMSs in an Infinite Periods General Equilibrium Economy

In this section we provide an application of the EMSs approach to a dynamic production economy with an infinite horizon both for the representative agent and the firms. These have to pay an initial fixed cost to enter in the market, and subsequently they compete *à la* Cournot in the production of a homogenous good. Production occurs with a single input, labor, which is inelastically provided by the agent, and business creation is driven by savings, that are inelastically provided as well.[25]

[25] In Chapter 3 we extend this same model to endogenous savings, endogenous labor supply, imperfectly substitutable goods and competition in prices.

We adopt the simplifying assumption used by Solow (1956) for which savings are a constant fraction of income. However, in our model income includes both the remuneration of inputs and the profits. This allows us to obtain a dynamic model in which it is not investment in physical capital to generate the accumulation of the reproducible input over time, as in the neoclassical Solow model, but it is entry of new firms to generate the 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 neoclassical model. Here, however, it is entry that strengthens competition and reduces the marginal profitability of subsequent entry. Therefore, both models generate a gradual convergence toward a steady state: in the Solow case through a decreasing growth rate of the capital stock, in our case through a decreasing rate of business creation.

2.6.1 A model of business creation with Cournot competition

Consider a representative market for a homogenous good with N_t firms active in each period t.[26] Each firm i produces $x_t(i)$ according to a linear production function:

$$x_t(i) = A_t l_i \tag{2.61}$$

where A_t is the exogenous productivity of labor (or the total factor productivity in this case without other inputs), which is common to all firms, and l_i is the labor input used by firm i. Given the nominal wage W_t, the constant marginal cost of production is $c_t = W_t/A_t$. Total expenditure in the sector is:

$$E_t = p_t C_t = p_t \sum_{j=1}^{N_t} x_t(j)$$

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:

$$\begin{aligned} \Pi_t(i) &= \left[p_t(i) - \frac{W_t}{A_t}\right] x_t(i) = \\ &= \frac{x_t(i) E_t}{\sum\limits_{j=1}^{N_t} x_t(j)} - \frac{W_t x_t(i)}{A_t} \end{aligned} \tag{2.62}$$

[26] In the next chapter we will explicitly introduce multiple sectors of this kind, which adds realism to the description without changing the main insights of the representative sector model.

We assume that the 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 $x_t(i)$ to maximize its profits taking as given the production of the other firms, the equilibrium generates individual output $x_t = (N_t - 1)E_t A_t / W_t N_t^2$. Substituting, one obtains the equilibrium price at time t:

$$p_t = \frac{N_t}{N_t - 1}\left(\frac{W_t}{A_t}\right) \tag{2.63}$$

which is associated with the usual equilibrium mark up $\mu(N_t) = N_t/(N_t - 1)$. This equilibrium generates individual profits $\Pi_t(N_t) = E_t/N_t^2$ in nominal terms. Since the equilibrium price of the consumption good is p_t, 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 $\pi_t(N_t) \equiv \Pi_t(N_t)/p_t$ become:

$$\pi_t(N_t) = \frac{C_t}{N_t^2} \tag{2.64}$$

and the real wage $w_t = W_t/p_t$ can be derived from the equilibrium pricing relation as:

$$w_t = \frac{N_t - 1}{N_t} A_t$$

This implies that each firm produces $x_t = (N_t - 1)C_t A_t / w_t N_t^2 = C_t/N_t$.

When the number of firms increases, the equilibrium price goes down and the wage goes up, with the former approaching the marginal cost and the latter approaching the productivity of labor for $N_t \to \infty$. However, here we do not want to approach the neoclassical paradigma with infinite firms, but we want to endogenize the number of firms. One way to do it is to assume as usual that there is a fixed cost of production in each period and that free entry occurs at all times. Such an assumption, however, would exclude any interesting dynamics because profits would be zero at any time. Another way to endogenize entry, which is more realistic and interesting for dynamic models, is to assume that entry is constrained by the expectations on future profitability and by a one-shot fixed cost of entry. This is the approach that we will adopt from now on.

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

$$N_{t+1} = (1 - \delta_N)(N_t + N_t^e) \tag{2.65}$$

which is analogous to the equation of motion of capital in the Solow model.[27]

[27] This equation of motion for the number of firms is borrowed from Ghironi and Melitz (2005). Analogous results would emerge with a more traditional version as $N_{t+1} = (1 - \delta_N)N_t + N_t^e$, in which new firms are always active for at least one period.

The real gross value of a new firm V_t is the present discounted value of its future expected profits, which, using the expectations operator $E[\cdot]$, and taking into account the exit probability in each period, becomes:

$$V_t = (1-\delta_N)E\left[\frac{\pi_{t+1}(N_{t+1})}{1+r_{t+1}}\right] + (1-\delta_N)^2 E\left[\frac{\pi_{t+2}(N_{t+2})}{(1+r_{t+1})(1+r_{t+2})}\right] + ... =$$
$$= (1-\delta_N)E\left[\frac{\pi_{t+1}(N_{t+1}) + V_{t+1}}{1+r_{t+1}}\right]$$

where r_{t+k} is the real interest rate at time $t+k$, whose expectation is taken as given by the firms, and the second line rearranges the first one in a recursive form. In each period entry occurs until the real value of the representative firm 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 $w_t = (N_t - 1)A_t/N_t$, the endogenous entry condition $V_t = F_t$ amounts to:

$$V_t = F_t = \eta \frac{(N_t - 1)}{N_t} \tag{2.66}$$

Notice that this endogenous entry condition determining the investment in business creation can be re-interpreted in terms of the Tobin (1969) approach, for which additional investment takes place if the (stock) market value of a unit of capital is higher than its replacement cost, or in other words if their ratio, known as the *Tobin's* q, is larger than one. In our framework, additional investment in business creation takes place when the stock market value of a firm is larger than the entry cost, that is when the *Tobin's* q defined as:

$$q_t \equiv \frac{V_t}{F_t} \tag{2.67}$$

is larger than one - augmenting the model with adjustment costs would generate a gradual (dis)investment for $q_t > (<)1$.

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:

$$I_t = N_t^e F_t = \frac{\eta(N_t - 1)N_t^e}{N_t} \tag{2.68}$$

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 in real terms:

$$Y_t = N_t \pi_t(N_t) + w_t L = \frac{C_t}{N_t} + w_t L_t$$

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 standard approach of Solow (1956) we assume that savings are an exogenous fraction $s \in (0,1)$ of income, $S_t = sY_t$. From the aggregate resource constraint derived above, this implies:

$$\begin{aligned} Y_t &= \frac{(1-s)Y_t}{N_t} + w_t L_t = \\ &= \frac{(1-s)Y_t}{N_t} + \frac{(N_t-1)}{N_t} A_t L_t \end{aligned}$$

where we used the equilibrium expression for the wage. Solving for income we obtain:

$$Y_t = \frac{(N_t-1)A_t L_t}{N_t-(1-s)} \tag{2.69}$$

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 as for the traditional Keynesian multiplier). Given the number of active firms, a stronger propensity to consume increases aggregate demand and total profits,[28] which in turn increases total output. 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:

$$S_t = sY_t = \frac{s(N_t-1)A_t L_t}{N_t-(1-s)}$$

[28] Total profits can be derived as:

$$N_t \pi_t(N_t) = \frac{(1-s)(N_t-1)A_t L_t}{N_t\left[N_t-(1-s)\right]}$$

which is increasing in productivity but decreasing in the number of firms and in the savings rate. One can notice similarities with the "big push" story by Murphy, Shleifer and Vishny (1989).

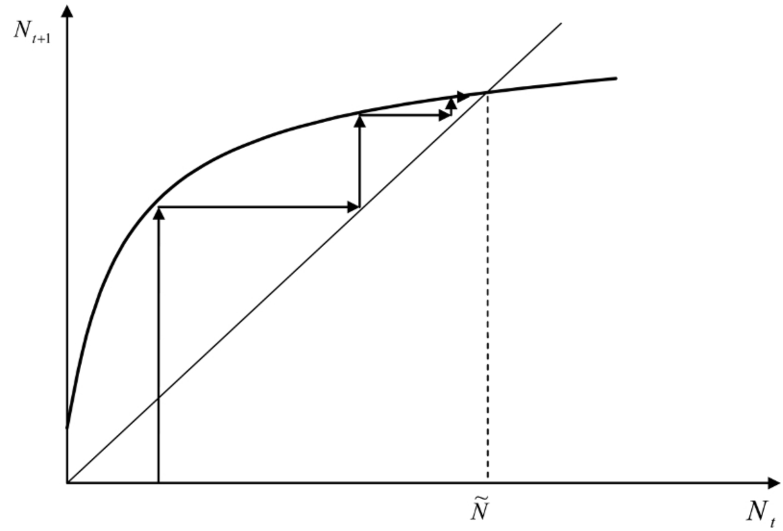

Fig. 2.1. Dynamic Creation of Firms in (N_t, N_{t+1}).

with investments I_t as defined in (2.68), we can solve for the equilibrium number of new firms:

$$N_t^e = \frac{sN_tA_tL_t}{\eta\,(N_t - 1 + s)}$$

Plugging the above expression in the equation of motion for N_t we have our final result for the evolution of the number of firms:

$$N_{t+1} = N_t\,(1 - \delta_N) + \frac{s(1-\delta_N)A_tL_t}{\eta - \frac{\eta(1-s)}{N_t}} \tag{2.70}$$

Assume that the labor force is constant at the level L at each point in time, and that the aggregate productivity is fixed at $A_t = A$. Then, the dynamic adjustment of the number of firms toward its steady state value is shown in Figure 2.1, which clearly resembles the dynamic adjustment of capital in the Solow model toward zero growth.

From the dynamics of the number of firms one can reconstruct the path of all the other variables. Two remarks are in order. First, the value of the stock market can be expressed as the value of all the firms $N_tV_t = \eta(N_t - 1)$, which follows the same dynamic path of N_t. For this reason, the aggregate behavior of the economy (consumption and output) is strictly related to the behavior of the stock market. Second, the model provides a dynamic path for income distribution, because the labor share $1 - \alpha_t$ is procyclical. This can be derived from:

$$1-\alpha_t = \frac{w_t L}{Y_t} = 1 - \frac{1-s}{N_t}$$

which is increasing in the savings rate and in the number of firms (i.e.: the fraction of income distributed as labor income is procyclical). Contrary to the neoclassical approach, in which the labor share was constant (at least under a standard Cobb-Douglas technology), the EMSs approach is able to generate more complex dynamics for the distribution of income between capital remuneration (in the form of dividends) and labor income.

2.6.2 EMSs in the long run

Let us consider the stationary situation to characterize the long run EMSs. Since the right hand side of (2.70) 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:

$$\tilde{N} = 1 + s\left[\frac{(1-\delta_N)\,AL - \eta\delta_N}{\eta\delta_N}\right] \tag{2.71}$$

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_N 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:

$$\tilde{\mu} = \frac{s\,(1-\delta_N)\,AL + (1-s)\eta\delta_N}{s\,(1-\delta_N)\,AL - s\eta\delta_N} \tag{2.72}$$

which is characterized by the opposite comparative statics of the number of firms.

Notice that dynamic inefficiency holds, since a better allocation of resources could be achieved through a reduction of the number of firms and an increase in the production of each firm (so as to reduce the waste in fixed costs of production). As we will see in the next chapters, this inefficiency result is a particular case of a much more general result that holds under different market conditions and also when firms produce differentiated goods.

Of 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:

$$\tilde{Y} = AL - \frac{\eta\delta_N}{(1-\delta_N)} \tag{2.73}$$

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.

In its simplicity, this model can be used for multiple purposes, and in the next sections we will provide a short overview of those that will be at the core of the following chapters.

2.7 Business Cycle

The EMSs approach can be used to study business cycles in an environment where, contrary to the neoclassical approach (Lucas and Rapping, 1969; Kydland and Prescott, 1982) competition in the market is not perfect, and, contrary to the New-Keynesian approach (Blanchard and Kiyotaki, 1987), the market structure is not exogenous. Our characterization of the market structure and of the incentives to create new firms gives raise to a new mechanism of propagation of the shocks that has nothing to do with the process of capital accumulation, with phenomena of intertemporal substitution (of consumption or labor supply) or with price rigidities, all elements that are absent here. 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. We are mainly interested in temporary shocks, because permanent ones would simply lead to monotonic convergence to a new steady state. Therefore, 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 mark up, 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.

More formally, we can derive the impulse response function of the number of firms by log-linearizing around the steady state the equation of motion (2.70). Taking the logs of both sides, differentiating with respect to the time-varying variables, and evaluating them at their steady state levels, we obtain:

$$\hat{N}_{t+1} = \left(1 - \frac{\tilde{N}\delta_N}{\tilde{N} - 1 + s}\right)\hat{N}_t + \delta_N \hat{A}_t \tag{2.74}$$

$$= \frac{s\left(1-\delta_N\right)^2 AL - (1-s)\delta_N^2\eta}{s\left(1-\delta_N\right)AL}\hat{N}_t + \delta_N \hat{A}_t \tag{2.75}$$

where $\hat{X}_t \equiv dX_t/\tilde{X}$ is the percentage distance from the steady state value of a variable X_t. Log-linearizing (2.69) around the steady state we obtain also:

$$\hat{Y}_t = \left\{1 - \frac{\eta\delta_N}{s\left[(1-\delta_N)AL - \eta\delta_N\right]}\right\} \frac{\eta\delta_N}{(1-\delta_N)AL}\hat{N}_t + \hat{A}_t$$

The response functions show that a one-shot increase in productivity increases the number of firms on impact. Afterward, even if productivity goes back to its initial value, the number of firms and output remain above their steady state values. They gradually decreases over time because of the increased competition and lower mark ups. 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 $\eta \to 0$) any additional propagation mechanism would disappear.

In case of a temporary but persistent technology shock, the effects are much 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. 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. Nevertheless, the impact on output and consumption follows closely the behavior of the technology parameter - for this reason the performance of the model can be improved introducing endogenous consumption and labor choices.

Analogous effects would derive from temporary shocks to the size of the entry cost (which could be interpreted as product market reforms for liberalization or deregulation or as introduction of cost reducing general purpose technologies) or even to the savings rate (which could be interpreted as demand shocks). A positive shock to the exit rate could be interpreted as a crisis leading to a chain of bankruptcies, and would have the consequence of reducing the number of firms and the output level, which would return only gradually to their steady state levels.

In Chapter 3 we will augment this same model with endogenous savings decision and also with endogenous labor supply: the former will introduce a new propagation mechanism based on the positive effect of competition on demand (as we have already seen in our two period partial equilibrium model), the latter will strengthen the propagation of the shock through a mechanism of intertemporal substitution of labor supply due to the impact of shocks on real wages through a general equilibrium effect (already seen in the previous section). In this context, we will also analyze fiscal policy and, under nominal price rigidities, monetary policy: we will suggest that the optimal policies should be aimed at stabilizing the business creation process around the efficient allocation of resources, with countercyclical tax rates and interest rate rules aimed at equity price stabilization.

2.8 Trade

A second way to use the EMSs approach is to augment closed economy models as the one used until now with trade with other countries. When countries open up, the international EMSs are affected in an interesting way that leads to a new source of gains from trade. This depends on the reduction of the mark ups and of the prices, contrary to what happens in the neoclassical approach and in the new trade theory with monopolistic behavior (Krugman, 1980), in which prices are not affected by trade between identical countries.

We can use our simple model of business creation to verify the impact of opening up to trade with another country and evaluate the effect of increasing the size of the market. Traditional models of intra-industry trade based on monopolistic behavior of the firms usually emphasize the impact of openness on the number of varieties produced and traded across countries and determine the gains from trade on the basis of this variety effect. When strategic interactions play a role, however, openness has the additional effect of strengthening competition and reducing the mark ups. This phenomenon leads to a reduction of the international prices which creates a second form of gains from trade.

Imagine that the closed economy considered above opens up to trade with another identical economy in the absence of trade frictions. Since the size of the market doubles, the new steady state number of firms from both countries in the joint market becomes:

$$\tilde{N} + \tilde{N}^{*} = 1 + s\left[\frac{2\left(1-\delta_N\right)AL - \eta\delta_N}{\eta\delta_N}\right] \tag{2.76}$$

The substantial increase in the number of firms strengthens competition and reduces the global mark up to the following steady state level:

$$\tilde{\mu} = \frac{2s\left(1-\delta_N\right)AL + (1-s)\eta\delta_N}{2s\left(1-\delta_N\right)AL - s\eta\delta_N} \tag{2.77}$$

which corresponds to an increase in steady state output in both countries.

In conclusion, in this model the gains from trade do not derive from the variety effect, as in the model of Krugman (1980): this effect is absent here because goods are homogenous. The gains from trade derive uniquely from the reduction in the price level. Of course, if we introduce product differentiation the gains from trade would derive from both sources: lower prices and more varieties (a similar point is made by Devereux and Lee, 2001). Notice that these dynamic models can be used to examine the reaction to shocks in a dynamic open economy framework and to explain a number of stylized facts without explanation in the neoclassical approach. Ghironi and Melitz (2005) have provided a fundamental contribution in this direction which will be discussed in Chapter 4.

Another consequence of the reduction in the mark ups is that gross profits increase less than proportionally with the increase in the size of the integrated market, therefore the endogenous number of firms active in each country tends to decrease. This phenomenon is quite evident in a static environment as the one considered in the new international trade literature associated with Krugman (1980). In such a case, globalization generates business destruction due to the reduction in the global prices. This phenomenon depicts a well known fear associated with our times (which has induced widespread support to novel forms of protectionism against globalization).

In Chapter 4 we will examine open economy issues and we will also use the EMSs approach to analyze the role of trade policy for globalized markets. Such an analysis is crucial to understand a world where competition takes place at the global level and most firms are active in domestic and foreign markets. Most importantly, our analysis will emphasize a result in sharp contrast with the traditional results. In particular, contrary to a standard outcome of neoclassical trade policy for which export taxes are always optimal to improve the terms of trade, the EMSs approach shows that export subsidies are always the optimal unilateral policy because they are the only way to provide a strategic advantage to the domestic firms active in international markets where entry is endogenous. This happens not only in certain markets characterized by Cournot competition as noticed in the literature on strategic trade policy starting with Brander and Spencer (1985), but under any form of competition including Bertrand competition. The optimal unilateral policy always requires policies that turn domestic exporters into aggressive leaders conquering larger market shares abroad. The result has also implications for exchange rate policy and R&D policy.

2.9 Growth

Finally, we can switch our attention to the process of business creation as a source of growth. As we have seen, our simple model confirms that the growth rate should be declining toward its steady state level (because of the decreasing marginal incentive to enter), and that only an exogenous growth of total factor productivity could generate long run growth, exactly as in the neoclassical approach of Solow (1956). However, endogenous growth can emerge when the creation of new firms is associated with an increase in total factor productivity.

Long run growth can be seen as the result of externalities in the accumulation of knowledge, as in Romer (1986). For instance, imagine that the productivity parameter 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 (possibly thanks to the investment in sunk costs of production, which could be seen as an investment in R&D). In particular, assume that $A_t = BN_t$ with $B > 0$. Then, the equation of accumulation of the number of firms (2.70) becomes:

$$N_{t+1} = (1 - \delta_N) N_t + \frac{s(1-\delta_N)BLN_t^2}{\eta N_t - \eta(1-s)} \tag{2.78}$$

which implies a process of perpetual business creation in which the growth rate of N_t converges to a constant and stable steady state level. This process is associated with a dynamic of the growth rate of income that converges to:

$$\tilde{g} = \frac{s(1-\delta_N)BL}{\eta} - \delta_N \tag{2.79}$$

which is positive as long as the savings rate is high enough or the rate of exit is low enough. Notice that, as in the Romer model with externalities in capital accumulation with growth rate (1.39), also here the long run growth rate is increasing in the size of the labor force: scale effects take place here (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 (Romer, 1986), the endogeneity of the market structure generates a gradual convergence of the growth rate to its long run level, which is empirically 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).

There is a deeper way in which the creation of new business augments total factor productivity. As suggested by the recent revival of the Schumpeterian tradition (Aghion and Howitt, 1992), 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. Such a mechanism requires a system of intellectual property rights which preserves the incentives to undertake R&D investments with an uncertain return, and its evolution relies on the market structure of the innovative activity (rather than the market structure of the productive activity, on which we focused until now). In Chapter 5 we will study a model of endogenous growth of this kind and we will analyze the EMS of the innovative sector with particular reference to the role of technological leaders and policy issues.

2.10 Conclusions

In this chapter we have introduced the concept of EMSs in a partial equilibrium context built on the basis of the industrial organization literature. We have adopted a demand structure that derives from the Dixit-Stiglitz utility function, a framework that we will keep using in the following chapters for its tractability and large generality, but many of our conclusions would apply to different frameworks. We have also extended the basic model to a simple two-periods framework to show a basic mechanism of propagation of the shocks in the presence of EMSs, and to a general equilibrium framework to show the mechanism of transmission of a shocks to the labor market. Finally, we have developed a fully dynamic model inspired by the Solow model but augmented with imperfect competition and gradual business creation. The reader should keep in mind that the models we presented in this chapter were largely explorative and they may serve mainly as prototypes. In the next chapters we will introduce more complex and realistic models concerning the main fields of the macroeconomy.

The aim of this chapter was to provide a simple introduction to these topics, and to support the idea that it is important to introduce in new fields the study of market structures characterized by strategic interactions between a limited number of firms and endogenous entry determining this limited number of firms. In this book we will focus on the growing literature introducing EMSs in the study of business cycle, trade and growth, but the EMSs approach can be useful also in other microeconomic and macroeconomic fields that are not discussed here. Etro (2007,a) has reviewed a large number of applications to industrial organization and industrial policy. A lot of work has taken place independently in the last years also in the theory of auctions with endogenous entry of strategic bidders,[29] in the theory of competition under asymmetric information with endogenous entry,[30] in the theory of tax incidence and optimal taxation for markets with endogenous

[29] Palfrey and Pevnitskaya (2008) have presented a model of endogenous entry in first-price auctions with heterogeneous risk averse bidders and have tested its predictions with an experimental study.

[30] The important work by Creane and Jeitschko (2009) shows that endogenous entry overturn the collapse of markets with adverse selection. They consider a market in which each firm can pay an observable fixed costs of entry that generates a product of a quality that becomes known only to the firm. Entry has the tendency to lower prices, which may lead to exit of high quality products. However, the implied price collapse endogenously limits the amount of entry, so that high mark ups are supported in the market equilibrium.

entry,[31] in the theory of regulation with endogenous entry,[32] in the theory of political competition[33] and in other fields as well. Hopefully, these further applications confirm the utility of the joint analysis of strategic interactions and endogenous entry, and the necessity, that cannot be postponed further, of introducing EMSs in a systematic way within mainstream economic theory.

[31] See Wu and Zhang (2000) and Tamai (2006) on taxation in dynamic models and McCracken and Stähler (2007) on international tax competition. Katsoulacos and Xepapadeas (1995) is an early application to environmental policy.

[32] Gautier, Dam and Mitra (2007) have introduced the first analysis of endogenous entry in a model of regulated competition in differentiated retail goods and services between an incumbent leader, who owns a network good (an essential input) and potential entrants, whose cost of production is private information. The regulator sets the retail prices and the access charge that the entrant pays to the incumbent, but entry is endogenous.

[33] See Mulligan and Tsui (2009).

http://www.springer.com/978-3-540-87426-3

Endogenous Market Structures and the Macroeconomy
Etro, F.
2009, XX, 346 p., Hardcover
ISBN: 978-3-540-87426-3