Free entry and social inefficiency

Free entry and social inefficience, Mankiw and Whinston formally established a two-stage model to expose the conditions under which the number of entrants in a free-entry equilibrium is excessive, insufficient or optimal. In their framework, at the first stage firms make entry decisions, and at the second stage the active firms make product decisions. The important insight of their work is that in a industry with homogenous product and fixed cost of entry, imperfect competition and business stealing effect can produce excessive entry from a social stand-point of view. When integer constraints are accounted, the free-entry number of firms can be less than the socially desired number, but not by more than one firm. The tendency toward excessive entry can be reserved as a result of product differentiation. Concerning entry regulation, their analysis shows that regulation can be unnecessary, since there are cases in which fixed cost approaches zero and firms act approximately as price-taker.

By the time of Mankiw and Whinston’s work, i. e. mid-80s, there had been articles demonstrating the idea that when firms must incur fixed set up costs upon entry, the number of entrants at the equilibrium can be insufficient or excessive in the relation to the social optimum. However, the economic forces underlying these entry biases had not been fully exposed, leading to the typical presumption that free entry is desirable. To examine the conditions for establishing the presence of an entry bias, Mankiw and Whinston argue that the aspects of the postentry game played by firms should be given central roles. These aspects are imperfect competition and business-stealing effect – which they define as the effect of the increasing number of firms resulting in incumbent firms’ reduced volume of sales. In the

other extreme, business-augmenting effect means that the increase in the number of firms enhances each incumbent’s output. According to Mankiw and Whinston, when there is imperfect competition, the business-stealing effect is a critical determinant of the direction of entry biases.

Mankiw and Whinston develop a model of two stages that allows them to compare the number of entrants in a free-entry equilibrium and the socially desired number. Similar to von Weizsacker (1980) and Perry (1984), they viewed government intervention as having two types: First-best regulation is the condition in which in order to maximize social surplus, a social planner determines the number of operating firms and sets their outputs. Second-best regulation is the condition in which the planner can only determine the number of firms and not their post entry behavior. In this model, Mankiw and Whiston take as given firms’ non-competitive behavior after entry, and compare the outcomes of the second-best regulation with the outcomes under no intervention (i. e. free entry case). A planner is supposed to have the objective of maximizing total surplus in the market, while oligopolists have a tendency towards rival retaliation. The entry process have two stages: in the first stage there is an infinite number of identical firms decide whether they enter the market or not. If the potential entrant decides to enter, it must incur fixed set-up costs. At the second stage, i. e. the production period, each identical firm behaves as a quantity-setting and profit-maximizing oligopolist. Mankiw and Whinston do not model the postentry game explicitly, arguing that this approach has two advantages: (1) uncovers the reasons behind the presence of the entry biases and (2) provides a set of properties readily to be checked for other application. They

propose the assumptions which will be used throughout the paper concerning a firm’s cost function, equilibrium output and profits. In particular, each firm’s cost function specifies economies of scale, equilibrium is symmetric, and equilibrium output is not the efficient one since firms behave strategically rather than act as price-takers. The necessary and sufficient conditions for a number of entrants to be the free-entry equilibrium are that profits are not negative and if there is one more firm enters the profits per firm will be negative. The implications of this assumption are that no firm has entered would have been better-off not entering, and no firm that has not entered would have found it worthwhile to have entered. The model is developed as a partial equilibrium framework in which income effects can be ignored.

The relationship between the free-entry equilibrium number of firms and the socially optimal number of the firm is examined in two propositions, with the difference concerning the consideration of integer constraints. According to Mankiw and Whinston, when such factors are ignored, the free-entry equilibrium number of firms is not less than the number that a social planner would desire (proposition 1). When integer constraints are accounted, however, the number at equilibrium can be less than that at a social optimum, but not by more than one firm. (proposition 2). To analysis these two cases, Mankiw and Whinston propose a simple homogenous product market, with inverse market demand function for the product and the equilibrium profits per firm are determined by revenue, operating costs and set-up costs. The socially optimum number of firms is the number that solves the maximization equation of social welfare. There are three assumptions

that hold throughout the two propositions: (1) an increase in the number of the firms enhances total output, a condition that can be seen as quasi-competitiveness. They assume that this postentry equilibrium aggregate output approaches some finite bound, thus guarantees that the free-entry equilibrium number of firm is well defined; (2) a business-stealing effect is present, output per firm decreases as a result of increasing number of firms; (3) imperfect competition is viewed as a competitive mode, in which for any number of entrants the resulting equilibrium price is not below marginal cost.

To prove their propositions, they assume zero-profit condition in the free-entry case and the first-order condition satisfied by the socially optimum number of firms. In the case when integer constraints are ignored (i. e. proposition 1), zero-profit condition and the assumption that profits per firm declines as the number of firms grows imply that the free-entry equilibrium number is not less than the socially optimum one. When equilibrium price exceeds marginal cost, there is an excessive entry from the social stand-point. A marginal entrant produces a reduction in social surplus because he contributes directly to social welfare through his profits, but also causes other firms to contract output levels. Business-stealing effect produces the divergence between the planner’s marginal evaluation of the optimum number and the entrant’s, since the planner calculates the reduction of social surplus but the marginal entrant does not. Therefore, when entry does not lead to different (contracted) output levels, the divergence in evaluation is dismissed and the free-entry produces socially efficient number of firms. Even when entry changes the output of existing firms, however, the free-entry equilibrium number of firms can be exactly the level would be desired

by the social planner, despite the presence of the business-stealing effect. This is the case when firms act as price-takers in postentry period, therefore the output contraction no longer has any net social value. In this proposition, Mankiw and Whinston also suggest that “ business-augmenting” entry has a contrast effect, i. e. there will be an insufficient level of entry when output per firm increases as the number of firms grows.

While Mankiw and Whinston propose that in homogenous markets the presence of a business-stealing effect creates a strong tendency towards excessive entry, they also suggest that entry can be insufficient if we take account of the integer constraints (proposition 2). Although they propose that the insufficient level is never more than one firm, they notice that there are cases when welfare losses due to insufficient entry can be substantial. Consideration of integer constraints reveals the cases when no firm enters the industry even though a monopoly is the socially optimum outcome. This result relates to the common observation that a monopolist does not capture all of the social surplus generated by his product.

Mankiw and Whinston provide several examples that demonstrate their propositions. In the first example, they consider a linear market structure in which firms behave as Cournot oligopolists and show that for a given number of social optimum, there is always a higher number of entrants, and the bias towards excessive entry can be very large. Welfare losses due to free entry do not always increase as the socially optimal number increases, however, and the government can achieve welfare improvement by means of an entry tax. There are possibilities that a removal of restrictions on entry may lead to a welfare loss. Next, they propose another setting in which the market

structure is linear but firms do not behave as Cournot oligopolists. Firms can behave collusively to form a cartel, and the consequent aggregate output is invariant to the number of firms. If firms continuously enter the industry, they will do so until the collusive monopoly profits are dissolved fully into set-up costs, and there will be deadweight social losses. This is akin to the discussion by Postner (1975), who presented the argument that monopoly rents measure the social resources lost through rent-seeking activities and thus should be counted in the costs of monopoly. Both the first two examples given by Mankiw and Whinston are to demonstrate the intuition that imperfect competition and business-stealing effect produce a strong tendency toward excessive entry. Finally, at the other extreme, there can be welfare losses due to insufficient entry. A third setting was considered: in a linear market a firm acts as a monopolist but two firms act as Bertrand competitors. In this case, duopolists always earn negative profits and monopoly can be socially optimal for a certain level of set-up costs. However, when set-up costs are low, society would receive a greater surplus and there would be a net increase in social welfare with a second entrant. Intuitively, duopoly is socially optimal and welfare losses caused by monopoly can be substantially large.

Following Spence (1976), Mankiw and Whinston examine the effect of product heterogeneity on the nature of entry biases. The gross consumer benefits are specified as a function of total output level of firms, with the assumption that the function is concave, implying that consumers prefer variety and that the output of different firms are substitutes for one another. The conditions for maximizing total welfare reveal not only the business-

stealing effect but also the effect of product diversity. By increasing variety, a marginal entrant increases surplus but does not capture this gain in profits: the diversity effect is captured as the marginal entrant’s contribution to gross social plus less his revenue. Therefore, the presence of heterogeneity introduces another factor that biases entry, and the product diversity and business-stealing effect work in opposite directions. The implication is that the sign of the biases depends on the interplay between these two effects, and entry can be excessive, insufficient, or even optimal.

Although Mankiw and Whinston assert that the presence of product differentiation can reverse the tendency towards excessive entry, they argue that this effect is always dominated by the business-stealing effect. This suggestion is opposite to that of Spence, who argues for the presence of insufficient entry resulted from free entry case. Spence takes into account a parameter that determines the ratio of maximized profits to maximized contribution to total welfare, asserting that this parameter is crucial in determining biases in product selection. In Spence’s model, when firms choose their quantities they can act as price-takers, so that each firm’s equilibrium profits are exactly equal to its net contribution to consumer’s benefits minus costs that firms must incur. Under this set of assumptions, Spence’s assertion is that there are more products at the optimum than at the equilibrium. For any number of firms, therefore, the product diversity effect always dominates the business-stealing effect. Contrary to this result, Mankiw and Whinston give examples in which one can replace Spence’s postentry price-taking assumption (with the functional form, i. e. constant elasticity of substitution, unchanged), or find other functional forms (with the

price-taking assumption remained) to expose the cases in which entry is excessive.

The importance of entry regulation is examined by Mankiw and Whinston with a consideration of set-up costs. Their given examples show that small set-up costs do not always imply that entry regulation becomes unimportant. Particularly, in the linear market structure where firms do not act as Cournot oligopolists but have a tendency to form a cartel, the loss due to free entry does not fall as the set-up costs decline. Therefore, with the purpose of establish a limiting result, Mankiw and Whinston assume that firms behave as price-takers as the number of entrants increases. They remain the three assumptions that hold for their propositions concerning the relation between the free-entry equilibrium number of firms and the socially optimal number of firms (i. e. proposition 1&2), with some generalizations to a heterogeneous product setting. Specifically, the new set of assumptions includes: (1) each firm’s equilibrium price declines as the number of firms grows, and the aggregate output is bounded by a finite value; (2) business-stealing effect exists; (3) equilibrium price is larger than marginal cost for any finite number of firms; and (4) when the number of firms grows infinitely large, price approaches marginal cost.

For a given level of set-up costs, the welfare associated with the free-entry equilibrium number of firms is not larger than that associated with the socially optimum number of firms, which is in turn equal or less than the socially optimal level of welfare. When set-up costs decline to zero, the number of firms approaches infinity in both the free-entry equilibrium and the optimum, therefore price approaches marginal cost (assumption 4).

Excessive entry produces the only difference between the social optimum and free-entry welfare. Because of assumption 1 and 4, operating profits and total set-up costs at free-entry equilibrium also approach zero, implies that there is no welfare difference between the equilibrium and the social optimum. Mankiw and Whinston therefore introduce their third proposition: When set-up costs approach zero, the free-entry welfare approaches the social optimum welfare. Thus, in this case government regulation on entry would be unnecessary since the welfare cost of excessive entry is diminished. However, they notice that this is the case only if assumption 3 holds strictly (i. e. price exceeds marginal cost for all finite number of firms), since if price falls to marginal cost and set-up costs do not grow large, the loss due to excessive entry can persist. Their results are similar to those shown by Hart (1979) and Novshek (1980), but there are two key differences concerning the nature of postentry interaction between firms and the sequential character of the posited models. Specifically, Mankiw and Whinston do not assume firms' Cournot behaviour and propose a sequential entry process as opposed to the simultaneous one in the other authors' model.

Conclusion

The paper " Free entry and social inefficiency" by Mankiw and Whinston examines the fundamental and intuitive economic forces underlying the entry biases in homogenous product markets given the presence of fixed set-up costs, imperfect competition and business-stealing effect. From the social stand-point, the authors argue for government restrictions on entry, but notice that such regulations can be unnecessary if set-up costs are small.

In heterogeneous setting, the direction of entry biases is determined by the interplay between the product diversity and business-stealing effect.

Two decades after their article, there have been new insights in the industrial organization literature concerning firms' entry and imperfect competition. For example, Amir and Lambson (2003) construct a stochastic model of entry in which the tendency towards excessive entry need not hold. There are also critiques about Mankiw and Whinston's use of partial equilibrium analysis as the resulted insight would be somewhat paradoxical (Konishi 1989). Despite those possible contrast ideas, however, today Mankiw and Whinston's model is still inspiring economists in examining market structures and the associated elements.