

TWO AND A HALF THEORIES OF TRADE ^{*,†}

J. Peter Neary

University of Oxford and CEPR

December 4, 2009

Abstract

This paper discusses the place of oligopoly in international trade theory, and argues that it is unsatisfactory to ignore firms altogether, as in perfectly competitive models, or to view large firms as more productive clones of small ones, as in monopolistically competitive models. Doing either fails to account for the “granularity” in the size distribution of firms and for the dominance of large firms in exporting. The paper outlines three ways of developing more convincing models of oligopoly, which allow for free entry but do not lose sight of the grains in “granularity”: heterogeneous industries, natural oligopoly, and superstar firms.

Keywords: GOLE (General Oligopolistic Equilibrium); granularity; heterogeneous firms; international trade and market structure.

JEL Classification: F12, F10

*Address for Correspondence: Department of Economics, University of Oxford, Manor Road, Oxford OX1 3UQ, UK; tel.: +44 (0)1865-271085; e-mail: peter.neary@economics.ox.ac.uk.

†Forthcoming in *The World Economy*. This is a revised version of the 2009 Frank D. Graham Memorial Lecture at Princeton University, and the 2009 *World Economy* Annual Lecture at the University of Nottingham. For helpful comments and discussions I am grateful to participants at these lectures, at the Swiss Economics Association Annual Conference, at ETSG 2009, and at a seminar at IMT Lucca, as well as to Gene Grossman, Giammario Impullitti, and Esteban Rossi-Hansberg. The paper draws on joint work with Carsten Eckel and Kevin Roberts, and was much improved by comments from Beata Javorcik, Monika Mrázová, and Volker Nocke.

1 Introduction

The title of my paper, “Two and a Half Theories of Trade,” is a quotation from Paul Krugman, an appropriate choice just a short time after the first award of the Nobel Prize to international trade for many years. It comes not from his writings, but from the Krugman oral tradition, and in particular from one sound bite of which I got a private hearing some years ago. Paul and I were seated together at a seminar where Keith Head presented an early version of his 2002 paper on the home-market effect with Thierry Mayer and John Ries. When Keith reached the point in his presentation where he discussed the Brander (1981) and Brander and Krugman (1983) “reciprocal dumping” model of trade under oligopoly, Paul turned to me and remarked conspiratorially “I always say that there are two and a half theories of trade.” I knew immediately what he meant. International trade under oligopoly is the Cinderella of our discipline, a poor relation of the two dominant paradigms: the theory of comparative advantage based on perfect competition, and the theory of product differentiation and increasing returns based on monopolistic competition. In this paper I want to explore why this is so, to argue that it is high time this Cinderella dressed up to go to the ball, and to sketch some potential routes she might take.

Once upon a time, to continue my fairy-tale theme, there was just a single theory of international trade. From the time of Ricardo until the 1980s, models based on perfect competition dominated mainstream thinking about both positive and normative aspects of trade. During the 1980s, Paul Krugman and many others put paid to that, ended the monopoly of perfect competition you might say, and now we have two rich and insightful ways of thinking about a trading world. Indeed, especially now that it has been ennobled, and given a new lease of life by the recent explosion of theoretical and empirical work on heterogeneous firms, it sometimes seems that the theory of monopolistic competition may become the dominant approach in trade, though the resilience of perfectly competitive approaches in both teaching and research should not be underestimated.

Ironically, the perfectly and monopolistically competitive paradigms have a great deal in common with each other, and definitely much more than either has with oligopoly.

Not just in their assumptions, of free entry and exit by firms that are ex ante identical, are infinitesimal in scale, and compete non-strategically. But also in their implications. For example, gravity equations derived from the standard monopolistically competitive model are basically identical to those first derived by Anderson (1979) for a competitive exchange model with CES preferences and country-specific goods. The only difference is that the elasticity of substitution appears one extra time in the estimating equation, reflecting firm mark-ups as well as consumer tastes. At a deeper level, both paradigms imply that the production sector of the economy can be characterized as efficient, or at least constrained efficient, and so can be represented mathematically, following Dixit and Norman (1980), by a GDP function.¹ This formal equivalence between the two approaches opens up a promising research agenda, since it implies that many of the technical tools developed for perfectly competitive models can be applied to monopolistically competitive environments.²

To sum up so far, whether they are viewed as substitutes or complements, no one can dispute the status of the perfectly and monopolistically competitive approaches to trade. By contrast, the same cannot be said of the theory of trade under oligopoly. Though there have been some notable contributions that form part of the central corpus of trade theory, they are undoubtedly of lesser importance than the two dominant paradigms. This seems surprising, if only from an empirical perspective. Casual empiricism alone suggests that large firms are important in many world markets, and that in many cases their dominance has increased rather than diminished as globalisation has proceeded.³ This impression is enhanced by a growing body of empirical evidence which comes from what I call the “second wave” of micro data on firms and trade. The first wave, from the mid 1990s onwards, showed that exporting firms are exceptional, larger and more

¹This was shown by Helpman (1984) to hold for the homogeneous-firms monopolistically competitive model, building on the result in Dixit and Stiglitz (1977) that monopolistic competition with CES preferences is “constrained efficient”, where the constraint in question is that the government cannot make lump-sum transfers to firms to cover their trading losses. It has also been shown by Feenstra and Kee (2008) to hold for the heterogeneous-firms monopolistically competitive model of Melitz (2003) when the distribution of firm productivities is Pareto.

²To take one example, this is true of the methods for evaluating the gains from trade liberalization and for measuring the restrictiveness of trade policy which Jim Anderson and I surveyed in our 2005 book (Anderson and Neary (2005)).

³See, for example, “Big is Back,” *The Economist*, August 27th 2009.

productive than average. More recent data sets in the second wave have provided more disaggregated information on the activities of firms, and have highlighted the degree of heterogeneity even within exporters.

Number of		U.S. 2000		France 2003	
Products	Markets	% Share of	% Share of	% Share of	% Share of
		Exporting Firms	Value of Exports	Exporting Firms	Value of Exports
1	1	40.4	0.2	29.6	0.7
5+	5+	11.9	92.2	23.3	87.3
5+	1+	25.9	98.0	34.3	90.8

Table 1: Distribution of Manufacturing Exports by Number of Products and Markets⁴

Table 1 summarizes some aspects of the distribution of manufacturing exports for the U.S. and France, adapted from Bernard, Jensen, Redding and Schott (2007) and Mayer and Ottaviano (2007). Looking at the breakdowns by number of products sold and number of foreign markets served, two features stand out. First is that the distribution of firms is bimodal, with 40.4% of U.S. firms (29.6% of French firms) exporting only one product to only one market, while 11.9% (23.3%) export five or more products to five or more markets. Second, the latter firms account for by far the bulk of the value of exports, 92.2% (87.3%), so the distribution of export sales is highly concentrated in the top exporters. If we ignore the number of destinations and simply focus on the firms that export five or more products, we find that they account for 25.9% of U.S. firms (34.3% of French firms) but an overwhelming 98.0% (90.8%) of exports. Bearing in mind that non-exporting firms are excluded, these data suggest that the largest exporting firms are different in kind from the majority. I am not the first to draw attention to these features

⁴Notes: Data are extracted from Bernard, Jensen, Redding and Schott (2007), Table 4, and Mayer and Ottaviano (2007), Table A.1. Products are defined as 10-digit Harmonised System categories.

of the data. In particular, the work of Xavier Gabaix (2005) on “granularity”, recently extended to international trade by di Giovanni and Levchenko (2009), has also highlighted the importance of large firms for aggregate behaviour. However, in their formal modelling those authors stick with the assumption of monopolistic competition between firms. They allow for large firms in one sense by assuming that the distribution of firm productivities is Pareto with a high value for the dispersion parameter, but nevertheless they continue to assume that all firms are infinitesimal in scale. I want to go further: to try and put the grains into granularity.

Of course, as already noted, there already exists an important literature on trade under oligopoly. The new trade theory revolution of the 1980s, though its major impact was on monopolistic competition and trade, also brought models of oligopolistic competition into the mainstream. Most notably, Brander (1981) and Brander and Krugman (1983) showed that oligopolistic competition provided a third reason for trade, over and above comparative advantage and product differentiation;⁵ while Brander and Spencer (1985) identified a new justification for interventionist trade policy: by committing to a trade policy, governments can change the conditions of strategic interaction between firms. However, despite the large literatures stimulated by these contributions, the theory of international trade under oligopoly has never attained the status of the two main paradigms. Why is this?

Elsewhere (see Neary (2003a), (2003b)), I have argued that one major reason why models of trade under oligopoly have had less influence than they deserve was that they were not embedded in general equilibrium. As a result, they could not deal with the interactions between goods and factor markets which are central to the “big issues” in trade. The goal of embedding oligopoly in general equilibrium seemed to generate a variety of insuperable technical problems. For example, in a widely-cited survey article, Matsuyama (1995) gave as one of the reasons why monopolistic competition was more useful than oligopoly for modelling aggregate phenomena: “it helps us to focus on the aggregate implications of increasing returns without worrying about ... the validity of

⁵This work has inspired a small number of empirical studies: see Bernhofen (1999) and Friberg and Ganslandt (2006).

profit maximization as the objective of firms.” In my earlier paper I showed how this and other difficulties of embedding oligopoly in general equilibrium could be overcome by drawing on the same kind of continuum approach pioneered by Dornbusch, Fischer and Samuelson (1977) in perfectly competitive models and Dixit and Stiglitz (1977) in monopolistically competitive ones. In particular, this meant assuming a continuum of sectors, so individual firms are “large in the small but small in the large”: significant players in their own market, interacting strategically with their local rivals, but infinitesimal in the economy as a whole, and so generating none of the esoteric technical problems which an earlier literature had suggested were unavoidable.

In previous papers I have explored some implications of this approach to what I call “General Oligopolistic Equilibrium” or just “GOLE”: literally, putting “*O*Ligopoly” into “General *E*quilibrium”. In Neary (2002b) I argued that trade under oligopoly may be interesting precisely because there is less of it rather than more, as oligopolistic entry barriers serve to reduce the degree of international specialisation. I also showed that, in general equilibrium, competition effects of trade can interact with comparative advantage differences between countries in surprising ways, so, for example, trade liberalisation can raise rather than lower the share of profits in national income. In Neary (2002a) I suggested that oligopoly models could explain why increased foreign competition can affect the behaviour of domestic firms and through them raise the relative wage of skilled labour, even when import prices and volumes change little if at all. And in Neary (2007) I showed that oligopoly models could prove useful in understanding how trade policy and other shocks can affect market structure itself, by encouraging cross-border mergers and acquisitions, the dominant mode of foreign direct investment.⁶

There is however one final issue with models of oligopoly which I believe has held back their acceptance outside the confines of industrial organisation. To quote Matsuyama (1995) again, monopolistic competition is superior to oligopoly because it allows for an “explicit analysis of entry-exit processes.” Except in the very short run, the assumption of most oligopoly models that the number of firms is given fails to account for the continual

⁶For other applications of trade models with oligopoly in general equilibrium, see Bastos and Kreickemeier (2009) and Grossman and Rossi-Hansberg (2010).

entry and exit that we observe in all empirical data sets. At the same time, this “churning” in the data does not seem, in many industries, to undermine the key position of a small number of large firms: a feature which models of monopolistic competition cannot explain. So, apart from embedding oligopoly in general equilibrium, endogenising entry and exit while retaining a role for large firms that compete strategically is I believe the key to developing oligopolistic models which can throw light on the nature of competition in today’s globalised world. In the remainder of this paper I want to discuss some ideas along these lines, drawing on work I am currently doing with Carsten Eckel and with Kevin Roberts. I first address, in the next section, the technical difficulties of modelling endogenous entry and exit without eliminating an important role for strategic interaction between firms. In later sections, I turn to sketch three possible approaches which, singly or in combination, might allow us do that. I call these: heterogeneous industries, natural oligopoly, and superstar firms.

2 Oligopoly plus Free Entry

It could be said that oligopoly with free entry is easy: as easy as 1, 2, 3, though unfortunately not as easy as $1, 1 + \varepsilon, 1 + 2\varepsilon, \dots$, where ε is vanishingly small. I am referring of course to the “Integer Problem”: how can we model markets with a variable but finite number of firms, when so much of our tool-kit relies on infinitesimal calculus? For something so often ignored, the integer problem gets mentioned a lot, and with good reason. On the one hand, oligopoly with free entry but ignoring the integer problem is not really a distinct market structure, since its properties are largely indistinguishable from those of monopolistic competition or even, if products are homogeneous, perfect competition. For example, with free entry and the number of firms a continuous variable, Brander and Krugman (1983) showed that trade liberalisation cannot lower welfare; Markusen and Venables (1988) showed that there is no role for strategic trade policy; and Head, Mayer and Ries (2002) showed that there can be no home-market effect. On the other hand, with relatively few exceptions, facing up to the integer problem has seemed to pose

intractable technical problems.

In ongoing work with Kevin Roberts, I have tried to make progress on this issue by restricting attention to a particular class of heterogeneous-firm oligopoly games, called “aggregative games”. In such games, each firm’s marginal cost is independent of its output, and the equilibrium profits of every firm depend only on its own cost and on a generalised mean of all firms’ costs, $\mu_n(\mathbf{c})$, as well as on the number of firms n :⁷

$$\pi^i = \pi[\underset{-}{c_i}, \underset{+}{\mu_n(\mathbf{c})}, \underset{-}{n}] \quad (1)$$

The signs under the arguments indicate their effects on profits which are very natural: a firm’s profits fall if its own costs rise or if it faces more competitors, while they rise if the generalised mean of all firms’ costs rises. In addition, we assume that the own effect of higher costs dominates the cross effect, at least when all firms have the same costs. An important special case of these games, considered by Bergstrom and Varian (1985), is where $\mu(c)$ is the arithmetic mean \bar{c} , so each firm’s profits depend only on its own costs and on the average of all firms’ costs, as well as on n . This includes the case of Cournot oligopoly with identical goods, while the more general case (1) also includes both Cournot and Bertrand oligopoly with symmetrically differentiated goods. Thus, on the one hand, (1) restricts the nature of interaction between firms a lot, relative at least to the general case where the profits of each firm depend on n variables (the costs of all n firms) rather than just three. But on the other hand it encompasses most of the oligopoly games typically considered in international trade and other branches of applied theory.

Equation (1) specifies the outcome of competition between firms that are active in the market. In addition, we need to specify the entry process. Here we follow the literature on monopolistic competition with heterogeneous firms initiated by Melitz (2003), and assume that firms’ costs are a random draw from a known distribution. Specifically, we assume that all firms’ costs c_i are drawn independently from a distribution $g(c_i)$ with

⁷This class of games was first considered systematically by Reinhard Selten. For a recent review and extension, see Acemoglu and Jensen (2009). The generalised mean $\mu_n(\mathbf{c})$ is a scalar function of the vector of all n firms’ unit costs \mathbf{c} , which is increasing and symmetric in its arguments, and equals their common value when all firms have the same cost. See Diewert (1993) for technical details.

positive support over $[\underline{c}, \infty)$. We will call a firm with the minimum cost level \underline{c} a “lean” firm. We depart from the usual setting in one, relatively minor, technical respect. We assume that there is always a “lean outsider”: there is always a potential entrant with the minimum cost \underline{c} . However, because the costs of active firms (i.e., insiders) are a random draw, some or all insiders may have costs above the minimum. Technically, the existence of a lean outsider amounts to assuming that the probability density function of costs is strictly positive at the minimum cost \underline{c} : $g(\underline{c}) > 0$.

The last step is to specify the nature of equilibrium. In a symmetric or homogeneous-firm equilibrium with the number of active firms treated as a continuous variable, equilibrium simply requires that every firm makes zero profits: $\pi(c, c, n) = 0$. In a heterogeneous-firm equilibrium where n is small, there may be no active firm with exactly zero profits. There are therefore two distinct equilibrium conditions. First, insiders cannot make losses: $\pi[c_i, \mu_n(\mathbf{c}), n] \geq 0$ for all $i = 1, \dots, n$. Second, any outsider must make a loss. Since the most profitable outsider is a lean one, and since we have assumed that there is always such a firm, this implies that, if it entered so the equilibrium became one with $n + 1$ active firms, a lean outsider would make a loss: $\pi[\underline{c}, \mu_{n+1}(\underline{c}, \mathbf{c}), n + 1] < 0$.

Given these assumptions, it is possible to prove the following result:

Proposition 1 *The number of firms in any heterogeneous-firm free-entry equilibrium of an aggregative game is the same as the integer number of firms in the corresponding lean symmetric equilibrium.*

The proof relies heavily on the assumption that there is always a lean outsider. This places bounds on the admissible range of insider equilibria. The significance of the proof comes from the fact that, with a small number of heterogeneous firms, the exact configuration of firm costs in equilibrium is not unique. Nevertheless, the proof implies that the number of firms in any such equilibrium is unique, given the values of the exogenous variables. Moreover, though solving for a heterogeneous-firms equilibrium may be difficult in general, the proof says that the unique number of firms equals that in the corresponding lean symmetric equilibrium, which is much easier to solve.

To illustrate this result, consider perhaps the simplest possible example, that of Cournot competition, with linear demands and homogeneous products. We can write the inverse demand function and the profits of a typical firm as follows:

$$p = a - s^{-1}X \qquad \pi^i = (p - c_i)x_i - f \qquad (2)$$

where p is the price, $X = \sum_i x_i$ is total demand, a is the demand intercept, and s is the size of the market. The solution for firm output and profits in the heterogeneous-firms case can be shown to equal:

$$x_i = \frac{a - (n + 1)c_i + n\bar{c}}{n + 1}s \qquad \pi_i = s^{-1}x_i^2 - f \qquad i = 1, \dots, n \qquad (3)$$

If we now allow for free entry but do not assume that n can vary continuously, it is not immediately apparent from (3) how to solve for equilibrium, or whether the equilibrium number of firms is unique given the values of the exogenous variables. However, it is clear by inspection that (3) satisfies the restrictions of an aggregative game as in (1). Hence we can invoke Proposition 1 and concentrate on the solution in the symmetric case:

$$x_i = x = \frac{a - c}{n + 1}s \qquad (4)$$

Calculating profits for this case is straightforward, and, with n treated as continuous, its unique equilibrium value as a function of the market size s is illustrated by the dashed line in Figure 1.⁸ Finally, the solid line illustrates the integer values of n .

The result discussed in this section suggests that we can handle the integer problem after all, at least if we are prepared to limit our results to some special functional forms. However, a model of oligopoly with potential entry still needs some mechanism which ensures that in equilibrium the number of oligopolistic firms remains small, since otherwise the model reverts to one that, for all intents and purposes, is indistinguishable from perfect or monopolistic competition. In the remainder of the paper I consider three possible mechanisms which allow for free entry without in effect eliminating any role for

⁸ $a - c$ is normalised to equal one in the simulations.

strategic behaviour.

3 Heterogeneous Industries

The first model I want to sketch extends to multiple industries the Melitz (2003) model of firm heterogeneity in monopolistic competition. In this model there is a continuum of ex ante identical potential firms, each of which must pay a sunk cost f_E to find out their unit cost c . Incurring the sunk cost f_E entitles an entering firm to draw c from a known distribution of unit costs $g(c)$ which, as in the previous section, has positive support over $[\underline{c}, \infty)$. On learning its cost c , each firm then calculates its expected profits and chooses to produce or exit depending on whether or not they are positive: i.e., a firm will exit if $c > c^e$, where c^e is the cost level which yields zero expected profits, or, equivalently, which equates revenue $r(c)$ to the sunk cost f_E , conditional on a rational expectation of all other firms' behaviour: $E[\pi(c^e)] = 0$ or $E[r(c^e)] = f_E$. Finally, if exporting or some other activity requires an additional fixed cost, only lower-cost entrants will engage in them.

This model and its many extensions have made possible a rich research agenda in international trade theory in recent years, which has developed in parallel with the availability of large firm-level data sets. However, although firms differ in size in this model, they do not differ in kind. In particular, no matter how productive any firm is, it remains of measure zero in its sector and does not engage in strategic behaviour. To see how the model can be extended to allow for such behaviour, assume that the sector which firms contemplate entering is made up of a continuum of sub-sectors or industries, indexed by z , which we can arbitrarily restrict to the unit interval: $z \in [0, 1]$. Assume in addition that, when firms pay the sunk cost of entry, they learn two pieces of information rather than one: both their unit cost of production and their specific capabilities, which make them best suited to enter a particular industry. Thus in effect each firm is assigned a $\{c, z\}$ pair. The significance of the industry assignment is that industries differ in their fixed costs, which we can describe by a function $f(z)$; without loss of generality, we can

rank sectors by increasing fixed cost, so that: $f' > 0$.⁹

Figure 2 illustrates the kinds of outcomes which are now possible. In equilibrium, industries have different expected firm numbers. Moreover, provided the restrictions of an aggregative game are satisfied, we can apply Proposition 1: the dashed line shows the equilibrium number of firms as a continuous function of the sector z and hence of the fixed costs $f(z)$. This can be calculated for lean symmetric equilibria, which from the theorem gives the number of firms in any free-entry equilibrium. Finally, the solid line gives the integer number of firms in each industry in equilibrium. In every industry, the equilibrium number of firms is finite. However, in industries with relatively low fixed costs (i.e., low z) there will be a large number of firms, so large that the industry can reasonably be characterised as a monopolistically competitive one. By contrast, in industries with relatively high fixed costs (i.e., high z) the equilibrium number of firms may be very small. Figure 2 illustrates a case where industries with very high values of z are monopolised, whereas those with intermediate values are characterised by oligopolistic interaction between a small number of firms.

4 Natural Oligopoly

The previous section viewed fixed costs as exogenous, but a different approach views them as endogenous. This is not just more realistic, but it opens up a new possibility: even though there is free entry of firms into a single industry, the number of firms may not grow without bound as the market expands.¹⁰

To see how this is possible, write the free-entry condition in terms of equilibrium profits as a reduced-form function of firm numbers and market size:

$$\pi(n, s) = 0 \tag{5}$$

(For simplicity we concentrate on the case of symmetric firms, but we can invoke Propo-

⁹There is a slight loss in generality in assuming that the function $f(z)$ is continuous.

¹⁰See, for example, Dasgupta and Stiglitz (1980), Gabszewicz and Thisse (1980), Shaked and Sutton (1983), and Sutton (1991, 1998).

sition 1 where necessary to allow for firm heterogeneity.) Now, define the Market-Size Elasticity of Market Structure \tilde{E} as the proportional rise in the equilibrium number of firms in response to a proportional increase in market size:

$$\tilde{E} \equiv \frac{s}{n} \frac{dn}{ds} = -\frac{s\pi_s}{n\pi_n} \quad (6)$$

It is clear that stability of equilibrium requires that the denominator must be negative: $\pi_n < 0$, meaning that entry by an additional firm for a *given* market size s must lower profits. Hence the sign of \tilde{E} is the same as that of the numerator π_s , the effect on profits of a rise in market size for a given number of firms. It seems natural to expect this to be positive, in which case \tilde{E} too is presumptively positive. Moreover, it is clearly so in simple entry games, such as those of Cournot or Bertrand competition. For example, in the linear Cournot game of Section 2, it is easy to check that \tilde{E} equals $\frac{n+1}{2n}$, which is always positive, falling from one to $\frac{1}{2}$ as successive increases in market size s cause the equilibrium number of firms to rise from one towards infinity.¹¹

However, the market-size elasticity of market structure is not necessarily positive in multi-stage games, opening up the possibility of what Shaked and Sutton (1983) call “natural oligopoly”: the equilibrium number of firms does not increase as market size rises. To see this, reinterpret the profit function in (5) as the equilibrium outcome of a two-stage game in which each firm first chooses its level of investment in some variable that raises its own profits but may raise or lower that of its competitors. We can call this R&D, though it could equally well be capital stock or product quality. Denote firm i ’s level of investment in R&D by k_i and the sum of all other firms’ investments as K_{-i} . In the first stage of the game, firm i ’s profits depend on both k_i and K_{-i} , as well as on the exogenous variables n and s , and they can be written as $\hat{\pi}^i(k_i, K_{-i}; n, s)$. The reduced-form profit function already introduced equals this profit function subject to the constraint that each firm has chosen its R&D optimally to maximise its profits. Formally:

¹¹To see this, differentiate the expression for profits from (2) where output is at its symmetric equilibrium level given by (4), to obtain: $\pi_s = s^{-2}x^2$ and $\pi_n = -\frac{2}{n+1}s^{-1}x^2$.

$$\pi(n, s) \equiv \hat{\pi}^i(k_i, K_{-i}; n, s) \quad \text{when:} \quad k_j = \arg \max_{k_j} \hat{\pi}^j \quad \forall j \quad (7)$$

Rewriting profits in this way is helpful because it allows us to show that the derivative of profits with respect to market size need not be positive. To see this, consider the different channels by which an increase in market size affects profits. First, there is a direct effect at given levels of R&D, denoted by $\hat{\pi}_s^i$. Second is an indirect effect as the greater market size encourages all firms to adjust their investment in R&D, which in turn impacts on profits. The firm's own R&D has been chosen optimally, so changes in this do not affect profits, but changes in the R&D levels of each of the $n - 1$ rival firms do affect it. Hence the full effect of the increase in market size on profits is given by:

$$\pi_s = \hat{\pi}_s^i + (n - 1) \hat{\pi}_K^i \frac{dk}{ds} \quad (8)$$

There is a clear presumption that an increase in market size encourages all firms to raise their R&D, so dk/ds is positive.¹² Hence a necessary condition for natural oligopoly (a negative value of π_s and so of \tilde{E}) is that R&D investments are “unfriendly” in the sense that a firm's profits are reduced when rival firms raise their levels of R&D investment: $\hat{\pi}_K^i < 0$.

An example which illustrates these possibilities is the linear Cournot model from Section 2, extended to allow for investment in R&D. The demand function is as before, while profits and unit costs now become:

$$\pi^i = (p - c_i) x_i - \frac{1}{2} \gamma k_i^2 - f, \quad c_i = c_{i0} - \theta k_i \quad (9)$$

Firm i 's investment in R&D incurs quadratic fixed costs and reduces its production costs by θ .¹³ Assume firms first choose their R&D levels, and then compete in quantities. Taking account of optimal choice of quantities, profits at the beginning of the first stage

¹²The full expression for dk/ds is: $-\left[\hat{\pi}_{kk}^i + (n - 1) \hat{\pi}_{kK}^i\right]^{-1} \hat{\pi}_{ks}^i$. The expression in square brackets must be negative from the stability condition for oligopoly games due to Seade (1980), while it seems reasonable to expect an increase in market size to raise the marginal profitability of R&D at given levels of R&D by all firms, implying that the final term $\hat{\pi}_{ks}^i$ positive. Hence, we expect dk/ds to be positive.

¹³Details of the solution to this model can be found in Neary (2002a) and Leahy and Neary (2009).

are:

$$\hat{\pi}^i(k_i, K_{-i}; n, s) = s^{-1}x_i^2 - \frac{1}{2}\gamma k_i^2 - f, \quad x_i = \frac{a - (n+1)c_{i0} + \Sigma c_{j0} + n\theta k_i - \theta K_{-i}}{n+1} s \quad (10)$$

Clearly, investments are unfriendly: firm i 's output and profits fall when its rivals increase their R&D. Solving for optimal investment in R&D, this turns out to be proportional to output: $k_i = \mu \frac{\theta}{\gamma} x_i$ or $\theta k_i = \mu \eta x_i$. Here $\mu \equiv \frac{2n}{n+1}$ is greater than one, reflecting the strategic over-investment by each firm relative to the benchmark case of investment that is productively efficient (i.e., where $\mu = 1$); while $\eta \equiv \frac{\theta^2}{\gamma}$ is a kind of benefit-cost ratio for investment, which can be interpreted as the relative efficiency of R&D for a unit market size. Profits and output can now be calculated explicitly:

$$\pi^i = \left(1 - \frac{1}{2}\eta s\right) s^{-1}x_i^2 - f, \quad x_i = \frac{(1 - \mu\eta s)a - (n+1 - \mu\eta s)c_{i0} + n\bar{c}_0}{(1 - \mu\eta s)(n+1 - \mu\eta s)} s \quad (11)$$

where $\bar{c}_0 \equiv \Sigma_i c_{i0}$. Once again, it is convenient to invoke Proposition 1 and focus on the solution in the symmetric case (where $c_{i0} = c_0$ for all firms):

$$x_i = \frac{a - c_0}{n+1 - \mu\eta s} s \quad (12)$$

The dashed lines in Figure 3 illustrate the equilibrium number of firms as a continuous variable for a range of values of the relative efficiency of investment η . (The case of $\eta = 0.0$ is the model without R&D as in Figure 1.) Finally, the solid lines illustrate the corresponding integer numbers of firms. It is clear that \tilde{E} eventually turns negative for all strictly positive values of η .¹⁴ There are also wide ranges of s where incumbents' R&D response to increases in market size ensures that no additional firms enter and the number of firms remains constant.

¹⁴The tendency towards eventual monopoly in this example reflects in part the fact that firms produce identical goods, so competition in the final stage is particularly intense. Allowing for product differentiation would reduce this effect and allow for more persistence of equilibria with more than one firm.

5 Superstar Firms

The title of my final approach to combining oligopoly with free entry has echoes of Rosen (1981), but the mechanism I have in mind is quite different. Return to the standard Melitz framework, with no sub-sectors as in Section 3 and no mechanism enforcing natural oligopoly as in Section 4. The new feature is that firms face two choices rather than one. First they choose whether or not to enter, paying a fixed cost f to discover their unit cost c in the usual way. In a second stage they choose to either remain as a “small” firm or to pay a further cost f_L to invest and become a “large” firm: a “superstar”. The investment can be in any one of a number of firm attributes: capacity, R&D, the adoption of a superior technology, or an extended product range. The key feature is that it involves their becoming of finite mass. Finally, in a third stage, there is competition in either quantities or prices, with strategic competition between a small number of large firms on the one hand, and a competitive fringe (if products are differentiated, a monopolistically competitive fringe) of infinitesimal firms on the other.

Figure 4 illustrates the equilibrium outcome of this game. The curve $G(c)$ represents the cumulative distribution of unit costs, starting from the minimum cost corresponding to a lean firm \underline{c} .¹⁵ Firms’ decisions on whether or not to enter lead to a threshold for entry indicated by $\gamma(f)$, so all firms with unit costs less than this enter. Of these, an infinitesimal subset chooses to acquire the superstar technology, becoming of finite size in the market. The threshold for acquisition of the superstar technology, denoted by $\gamma(f_L)$, is less than $\gamma(f)$, but can be higher than the lean threshold \underline{c} . For example, if we follow the approach of Section 2 and assume that the firms which adopt the superstar technology are a random draw from the subset of all entrants, then some or all of them may have costs above \underline{c} .

So far, the advantage of superstar firms has not been specified exactly. One interesting and important case is where the superstar technology involves the ability to produce a large number of products. In that case, the small number of superstar firms are multi-

¹⁵Reflecting the crucial assumption made in Section 2 that there is always a lean outsider, the gradient of the cumulative distribution is strictly positive at \underline{c} : $G'(\underline{c}) > 0$.

product firms, while the remaining insiders which constitute the competitive fringe are single-product firms. This configuration is consistent with the empirical evidence quoted in the Introduction, especially when applied at the level of a single industry, rather than of manufacturing as a whole. In addition, modelling superstar firms in this way solves a technical problem: if multi-product firms produce a continuum of products, then large firms are of finite measure, while small ones remain of zero measure. A number of models of multi-product firms producing a continuum of products have recently been introduced in the literature on international trade, and there seems to be huge potential for applying them in the present context.¹⁶

6 Conclusion

In this paper I have discussed the role of oligopoly models in international trade, and have sketched a number of approaches which, singly or in combination, could be used to integrate strategic behaviour with entry and exit. For the most part, the discussion has been non-technical, outlining a road-map rather than a set of fully-specified models. It has also focused on modelling issues *per se* rather than on the detailed predictions of such models for substantive questions in trade. Much work remains to be done to implement these approaches. Nevertheless, I hope that the discussion has implications for a variety of issues, ranging from theory to policy and empirics.

At a theoretical level, the goal of the models sketched here is to contribute to reconciling the “Two Faces of IO”. Within the field of industrial organisation, models of small-group strategic competition are at least as important as models of the size distribution of firms. Both clearly capture important aspects of the real world, and cross-fertilisation between them seems desirable in itself. In the theory of international trade, a lot more attention has been devoted to the implications of monopolistic competition, greatly strengthened in recent years by its ability to encompass firm heterogeneity. Casual

¹⁶Theoretical models of firms which produce a continuum of products are presented in Allanson and Montagna (2005), Bernard, Redding and Schott (2009), Eckel and Neary (2010) and Nocke and Yeaple (2006). Empirical applications include Bernard, Redding and Schott (2009), Goldberg, Khandelwal, Pavcnik and Topalova (2010) and Eckel, Iacovone, Javorcik and Neary (2009).

empiricism as well as the evidence cited in the introduction suggests that the distinctive features of large firms also merit serious consideration. Given their dominance in exporting, it seems very likely that they matter for more than just reciprocal dumping.

Turning to policy, there are a great many issues which an oligopolistic perspective is likely to illuminate. At the upper end of the size distribution of firms, whether or not a country hosts any superstar firms is likely to matter for many questions. Further work on this topic may throw light, for example, on the contrasting experiences of Finland, with a large domestic multinational, and Ireland, which has no home-owned firms of comparable size but has served as a successful export platform for foreign-owned multinationals.¹⁷ Different policy issues arise at the lower end of the size distribution of firms. Governments everywhere devote great effort to fostering entrepreneurship and promoting entry by new firms, but such a focus may be inappropriate if oligopolistic market structures are dominant. It seems clear too that other policy questions, such as competition policy in general equilibrium and the effects of trade liberalisation on industries with relatively few large firms, can only be satisfactorily addressed in models which take oligopolistic market structures seriously.

Finally, turning to empirics, one of the most exciting developments in the study of international trade in recent years has been the increased availability of firm-level data sets, which lend themselves naturally to the application of models of monopolistic competition with heterogeneous firms. This interplay between theory and empirics has proved enormously fruitful, and is the hallmark of a genuinely scientific methodology. However, the approaches outlined here suggests some notes of caution. One issue is that a good explanation is not necessarily one which explains the firm data themselves: to the extent that large firms are dominant in exporting and other activities, errors in predicting the upper part of the distribution will matter far more for explaining aggregate behaviour than those at the lower part. A second issue is that most firm-level data sets cover a large swathe of manufacturing. From the point of view of any single firm, such data sets therefore contain many firms with which it does not compete directly, and exclude at

¹⁷In 2004 Nokia's share of Finnish GDP was 3.5% and in 2003 it accounted for almost a quarter of Finland's exports. See <http://en.wikipedia.org/wiki/Nokia>.

least some foreign firms with which it competes head to head. The perspective of the heterogeneous-industries model of Section 3 suggests that it may be better to think of industries at the world level, with head-to-head competition between a relatively small number of firms, rather than at the national level, with symmetric competition between a continuum of firms as in models of monopolistic competition.

To conclude, in this paper I have argued for “Trade Theory 3.0”. The Big Two of trade theory can defend themselves, and for many questions they will remain the appropriate way to model global phenomena. But I have argued that for some questions it is not enough to ignore firms altogether, as in the theory of perfect competition, or to model large firms as merely more productive clones of small ones, as in the theory of monopolistic competition. Doing either fails to account for the “granularity” in the size distribution of firms and for the dominance of large firms in exporting. Developing more convincing models of oligopoly, in particular models which allow for free entry but do not lose sight of the grains in “granularity”, seems sure to help our understanding of many issues, and I have outlined some of the ways in which we might progress in this direction.

References

- [1] Acemoglu, Daron and Martin Kaae Jensen (2009): “Aggregate Comparative Statistics,” CEPR Discussion Paper No. 7254, London.
- [2] Allanson, Paul and Catia Montagna (2005): “Multiproduct Firms and Market Structure: An Explorative Application to the Product Life Cycle,” *International Journal of Industrial Organization*, 23, 587-597.
- [3] Anderson, James E. (1979): “A Theoretical Foundation for the Gravity Equation,” *American Economic Review*, 69, 106-116.
- [4] Anderson, James E. and J. Peter Neary (2005): *Measuring the Restrictiveness of International Trade Policy*, Cambridge, Mass.: MIT Press.
- [5] Bastos, Paulo and Udo Kreickemeier (2009): “Unions, Competition and International Trade in General Equilibrium,” *Journal of International Economics*, 79, 238-247.
- [6] Bergstrom, Theodore C. and Hal R. Varian (1985): “When Are Nash Equilibria Independent of the Distribution of Agents’ Characteristics?,” *Review of Economic Studies*, 52, 715-718.
- [7] Bernard, Andrew B., J. Bradford Jensen, Stephen J. Redding, and Peter K. Schott (2007): “Firms in International Trade,” *Journal of Economic Perspectives*, 21, 105–130.
- [8] Bernard, Andrew B., Stephen J. Redding and Peter K. Schott (2009): “Multi-Product Firms and Product Switching,” forthcoming in *American Economic Review*.
- [9] Bernhofen, Daniel M. (1999): “Intra-Industry Trade and Strategic Interaction: Theory and Evidence,” *Journal of International Economics*, 47, 225-244.
- [10] Brander, James A. (1981): “Intra-Industry Trade in Identical Commodities,” *Journal of International Economics*, 11, 1-14.

- [11] Brander, James A. and Paul R. Krugman (1983): “A ‘Reciprocal Dumping’ Model of International Trade,” *Journal of International Economics*, 15, 313-321.
- [12] Brander, James A. and Barbara J. Spencer (1985): “Export Subsidies and International Market Share Rivalry,” *Journal of International Economics*, 18, 83-100.
- [13] Dasgupta, Partha and Joseph Stiglitz (1980): “Industrial Structure and the Nature of Innovative Activity,” *Economic Journal*, 90, 266-293.
- [14] Diewert, W. Erwin (1993): “Symmetric Means and Choice Under Uncertainty,” in W.E. Diewert and A.O. Nakamura (eds.): *Essays in Index Number Theory*, Volume I, Amsterdam: Elsevier, 355-433.
- [15] di Giovanni, Julian and Andrei A. Levchenko (2009): “International Trade and Aggregate Fluctuations in Granular Economies,” Discussion Paper No. 585, Gerald R. Ford School of Public Policy, University of Michigan.
- [16] Dixit, Avinash K. and Victor D. Norman (1980): *Theory of International Trade: A Dual, General Equilibrium Approach*, London: Cambridge University Press.
- [17] Dixit, Avinash K. and Joseph E. Stiglitz (1977): “Monopolistic Competition and Optimum Product Diversity,” *American Economic Review*, 67, 297-308.
- [18] Dornbusch, Rudiger, Stanley Fischer and Paul A. Samuelson (1977): “Comparative Advantage, Trade and Payments in a Ricardian Model with a Continuum of Goods,” *American Economic Review*, 67, 823-839.
- [19] Eckel, Carsten, Leonardo Iacovone, Beata Javorcik and J. Peter Neary (2009): “Multi-Product Firms at Home and Away,” mimeo., Department of Economics, University of Oxford.
- [20] Eckel, Carsten and J. Peter Neary (2010): “Multi-Product Firms and Flexible Manufacturing in the Global Economy,” forthcoming in *Review of Economic Studies*.

- [21] Feenstra, Robert C. and Hiau Looi Kee (2008): “Export Variety and Country Productivity: Estimating the Monopolistic Competition Model with Endogenous Productivity,” *Journal of International Economics*, 74, 500-518.
- [22] Friberg, Richard and Matthias Ganslandt (2006): “An empirical assessment of the welfare effects of reciprocal dumping,” *Journal of International Economics*, 70, 1-24.
- [23] Gabaix, Xavier (2005): “The Granular Origins of Aggregate Fluctuations,” mimeo, MIT.
- [24] Gabszewicz, Jean Jaskold and Jean-Francois Thisse (1980): “Entry (and Exit) in a Differentiated Industry,” *Journal of Economic Theory*, 22, 327-338.
- [25] Goldberg, Pinelopi K., Amit Khandelwal, Nina Pavcnik and Petia Topalova (2010): “Multi-Product Firms and Product Turnover in the Developing World: Evidence from India,” forthcoming in *Review of Economics and Statistics*.
- [26] Grossman, Gene and Esteban Rossi-Hansberg (2010): “External Economies and International Trade Redux,” forthcoming in *Quarterly Journal of Economics*.
- [27] Head, Keith, Thierry Mayer and John Ries (2002): “On the Pervasiveness of Home Market Effects,” *Economica*, 69, 371-390.
- [28] Helpman, Elhanan (1984): “Increasing Returns, Imperfect Markets, and Trade Theory,” in R.W. Jones and P.B. Kenen (eds.): *Handbook of International Economics*, Volume I, Amsterdam: North-Holland, .
- [29] Leahy, Dermot and J. Peter Neary (2009): “Multilateral Subsidy Games,” *Economic Theory*, 41, 41-66.
- [30] Markusen, James R. and Anthony J. Venables (1988): “Trade Policy with Increasing Returns and Imperfect Competition: Contradictory Results from Competing Models,” *Journal of International Economics*, 24, 299-316.
- [31] Matsuyama, Kiminori (1995): “Complementarities and Cumulative Processes in Models of Monopolistic Competition,” *Journal of Economic Literature*, 33, 701-729.

- [32] Mayer, Thierry and Gianmarco I.P. Ottaviano (2007): *The Happy Few: The Internationalisation of European Firms. New Facts Based on Firm-Level Evidence*, Brussels: Bruegel.
- [33] Melitz, Marc J. (2003): “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity,” *Econometrica*, 71, 1695-1725.
- [34] Neary, J. Peter (2002a): “Foreign Competition and Wage Inequality,” *Review of International Economics*, 10, 680-693.
- [35] Neary, J. Peter (2002b): “International Trade in General Oligopolistic Equilibrium,” mimeo., University College Dublin; revised 2009.
- [36] Neary, J. Peter (2003a): “Globalization and Market Structure,” *Journal of the European Economic Association*, 1, 245-271.
- [37] Neary, J. Peter (2003b): “Competitive versus Comparative Advantage,” *The World Economy*, 26, 457-470.
- [38] Neary, J. Peter (2007): “Cross-Border Mergers as Instruments of Comparative Advantage,” *Review of Economic Studies*, 74, 1229-1257.
- [39] Nocke, Volker and Stephen Yeaple (2006): “Globalization and Endogenous Firm Scope,” NBER Working Paper No. 12322, Cambridge, Mass.
- [40] Rosen, Sherwin (1981): “The Economics of Superstars,” *American Economic Review*, 71, 845-858.
- [41] Seade, Jesus (1980): “The Stability of Cournot Revisited,” *Journal of Economic Theory*, 23, 15-27.
- [42] Shaked, Avner and John Sutton (1983): “Natural Oligopolies,” *Econometrica*, 51, 1469-1483.
- [43] Sutton, John (1991): *Sunk Costs and Market Structure: Price Competition, Advertising, and the Evolution of Concentration*, Cambridge, Mass.: MIT Press.

- [44] Sutton, John (1998): *Technology and Market Structure: Theory and History*, Cambridge, Mass.: MIT Press.

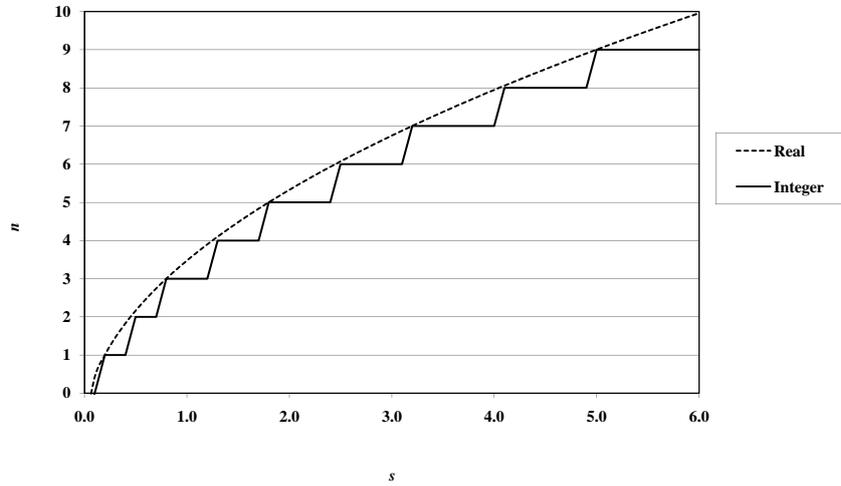


Figure 1: Cournot Competition: Equilibrium Firm Numbers as a Function of Market Size

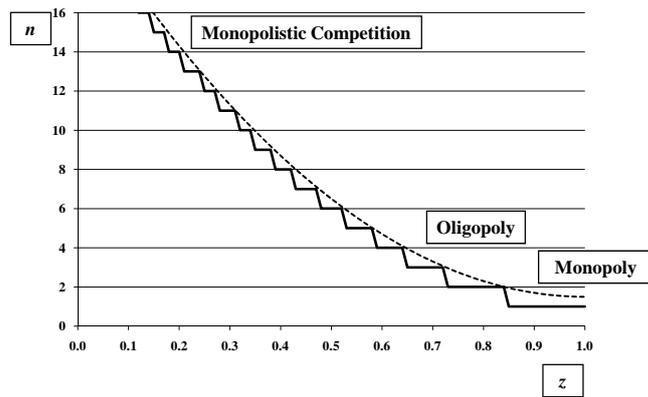
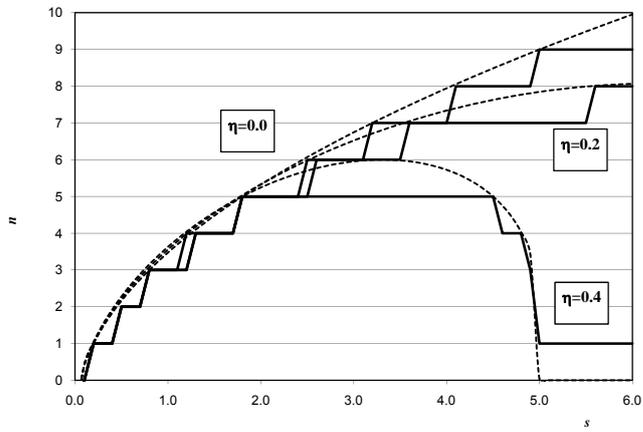


Figure 2: Equilibrium Market Structures



**Figure 3: R&D and Cournot Competition:
Equilibrium Firm Numbers as a Function of Market Size**

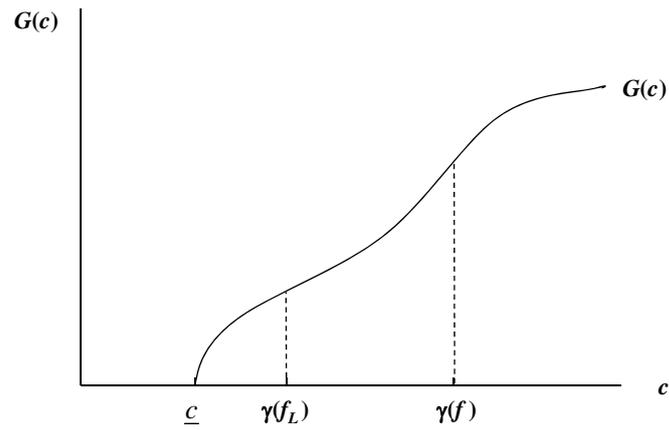


Figure 4: Equilibrium with Superstar Firms