CPB Memorandum

CPB Netherlands Bureau for Economic Policy Analysis

Sector	:	International Economics
Author	:	Roland de Bruijn
Number	:	140
Date	:	January 31, 2006

Scale Economies and Imperfect Competition in WorldScan[1]

Abstract

WorldScan, the CGE model for international policy analysis and long-term scenario studies, is applied regularly at the CPB. The production technology in the model is that of constant returns to scale and the market structure is characterized by perfect competition. However, it is a well known fact that many sectors such as manufacturing and service sectors feature increasing returns and firms compete imperfectly. To give the model more realism, it is therefore necessary to expand the model. Besides that, several research projects require an identification of scale economies in order to perform a sound welfare analysis. In this memorandum, I review the literature on scale economies and imperfect competition and analyze which approach is most suitable to implement in WorldScan. For the objectives at hand it appears most efficient to expand the model with an extended Dixit-Stiglitz approach. Simulations with an aggregated version of WorldScan show that the effects of incorporating scale economies are significant. Evidently, in a liberalisation scenario, sectors with increasing returns can exploit their technology more than sectors with constant returns, implying considerable increases in production and exports for these sectors. Concluding, this expansion of the model allows for an identification of formerly unidentified welfare effects.

[1] The author thanks Stefan Boeters, Arjan Lejour, Paul Veenendaal and George Gelauff for fruitful discussions and helpful comments and Gerard Verweij for his advise on programming issues.

Contents

1 Introduction

WorldScan, the CGE model for international policy analysis and long-term scenario studies, is applied regularly at the CPB. The production technology in the model is that of constant returns to scale and the market structure is characterized by perfect competition. This market structure is also applied in sectors characterized by specific fixed factors leading to decreasing returns to scale.

However, it is a well known fact that many manufacturing and service sectors feature increasing returns and firms within compete imperfectly. To give the model more realism, it is therefore necessary to expand the model. In particular, welfare effects of certain policy reforms could be underestimated, since scale economies are not identified.

For several research projects there is a demand for relaxing the assumptions of constant returns to scale and perfect competition. In the set-up fixed costs are financed from a mark-up on marginal costs. In the long run firms earn zero profit, because of free entry and exit of firms. Currently, the option for increasing returns to scale, also called economies of scale, and free entry and exit of firms is required in a project on the liberalisation of the European market for services as well as in an analysis of the European scheme for emissions trading.

In a liberalized market for services, services with fewer barriers can be supplied to foreign customers. Usually, there is an initiative for setting up local offices, leading to set-up costs implying increasing returns to scale. The welfare effects of liberalising the EU-market for services are expected to be significant. Obviously, a model which identifies and describes increasing returns can better assess these welfare effects, than a model that just assumes constant returns to scale in a sector that in fact features increasing returns.

EU member states compete in energy taxes in the sense that firms on average are taxed much lower than consumers. In addition, the largest users pay on average the least. A market structure of imperfect competition captures these aspects better than one of perfect competition and constant returns to scale.

Besides these policy issues, we would like WorldScan to be up-to-date, since comparable models can also describe economies of scale. Furthermore, adjusting WorldScan in this way, leads to a better comparison of simulation results with other models, which is useful for the analysis.

Concluding, the adjustments must enable WorldScan to identify and describe increasing returns. This memorandum discusses how the model might be adjusted.

The following two subsections briefly discuss the basics behind economies of scale and imperfect competition.

1.1 What are economies of scale?

When more units of a good or a service can be produced on a larger scale, yet with (on average) less input costs, *economies of scale* are said to be achieved. Alternatively, this means that as a company grows and production increases, a company will have a better chance to decrease its average costs. According to theory, economic growth may be achieved when economies of scale are realized.

Just as there are economies of scale, *diseconomies of scale* also exist. This occurs when growth of production is less than proportional to inputs. What this means is that there are inefficiencies within the firm or industry resulting in rising average costs.

Alfred Marshall made a distinction between *internal* and *external economies of scale*. When a company reduces average costs and increases production, internal economies of scale have been achieved. External economies of scale occur outside the firm, within the industry. The section below lists several causes for internal and external (dis)economies of scale[2].

As a firm produces more and more goods, internal economies of scale are made, so the average cost begins to fall because of

- *technical economies* made in the actual production of the good (e.g. large firms can use expensive machinery intensively),
- *managerial economies* made in the administration of a large firm by splitting up management jobs and employing specialist accountants, salesmen, etc.,
- *financial economies* made by borrowing money at lower rates of interest than smaller firms,
- *marketing economies* made by spreading the high cost of advertising on television and in national newspapers, across a large level of output,
- *commercial economies* made when buying supplies in bulk and therefore gaining a larger discount and
- *research and development economies* made when developing new and better products.

External economies of scale occur outside the firm as a result of its location and occur when

- a local skilled labor force is available,
- specialist local back-up firms can supply parts or services,
- an area has a good transport network,

[2] For a detailed discussion see also www.bized.ac.uk.

- an area has an excellent reputation for producing a particular good or
- sharing of technology or managerial expertise ('spillover').

As mentioned before, diseconomies may also occur. In internal diseconomies average costs for a firm eventually rise as production increases, because

- the disadvantages of the division of labor take effect,
- management becomes out of touch with the shop floor and some machinery becomes over-manned,
- decisions are not taken quickly and there is too much form filling,
- lack of communication in a large firm means that management tasks sometimes get done twice, or
- poor labor relations may develop in large companies.

When too many firms locate in one area, unit costs in the industry begin to rise, because

- local labor becomes scarce and firms now have to offer higher wages to attract new workers,
- land and factories become scarce and rents begin to rise or
- local roads become congested and so transport costs begin to rise.

When (dis)economies of scale are location specific, trade is used in order to gain access to efficiencies. The main reason why the presence of economies of scale can generate trade gains is because of the reallocation of resources.

In practice, the most common cause for (and way to model) increasing returns is (to allow for) a decreasing average cost curve as already said in the introduction of this chapter. The classical specification of a decreasing cost curve is given by the following reciprocal function:

$$AC = f / x + c, \qquad (1)$$

where AC denote average costs, f are the fixed costs, x is total production and c are the marginal costs of production. This type of reduced form structure can represent both internal economies of scale, see Francois (1990) and external economies of scale, see Markusen (1990). It will be the starting point of this paper.

1.2 What is imperfect competition?

There is imperfect competition as soon as an agent in the model no longer takes market prices as given.

Basically, there are three forms of imperfect competition. The simplest imperfectly competitive market structure is that of a *pure monopoly*, a market in which a firm faces no competition. The Dutch Railways is an example of a pure monopoly.

The second form of imperfect competition is *oligopoly*. The term oligopoly has Greek roots meaning *few sellers*. That is the way in which oligopoly differs both from perfect competition and monopoly: there is more than one seller, but not many more. For the small number of sellers to be stable, there presumably must be some *barriers to entry* for new competitors.

The final form of imperfect competition is *monopolistic competition*. In monopolistic competition the products sold by the firms in the industry are not homogeneous but differentiated. Thus, each firm has a “monopoly” of its own product. However, this is not a true monopoly, such as Microsoft, because the differentiated products are *close substitutes*. Monopolistic competition retains many features of perfect competition, such as the presence of many firms in the industry and the likelihood that free entry and exit of firms in response to profit would eliminate economic profit among the firms. As a result, this last form of imperfect competition offers a somewhat more realistic description of many common economic markets. Examples include automobiles, toothpaste, beer, cheese and many more.

Because competition is imperfect, there is not only competition in prices, but also in advertising and the characteristics of the good.

Perfect competition assumes homogeneity of products. This assumption is relaxed with imperfect competition and products are called heterogeneous or *differentiated*. Products are differentiated when they are not perfect substitutes, but close substitutes. Product differentiation increases variety, but it divides up the market, leading to higher prices and costs. In some industries, like the computer industry, there is a lot of competition to introduce a product that is superior to rival products. This is good for customers, because the *quality* of products improves. Unfortunately, this kind of competition could lead to overinvestment and waste of resources.

The advantage of non-price competition is that, while rivals will likely react to non-price competition, their reaction is often slower and less direct than would be the case for a price cut.

Economists have taken two main routes to address imperfect market structures according to Neary (2002). At the micro level, economists in the field of industrial organization have developed quite a few sophisticated models, which focus on strategic interaction between firms in a single market. At the more aggregate level, many research areas such as international trade, macroeconomics and economic growth, have used monopolistic competitive models to incorporate scale economies and product differentiation into a general equilibrium context. We shall elaborate a bit more on oligopoly and monopolistic competition, starting with oligopoly.

1.2.1 Oligopoly

There exist four major hypotheses about oligopoly pricing:

1. The oligopoly firms conspire and collaborate to charge the monopoly price and get monopoly profits (*cooperative*). Examples are models for *cartels* or *tacit collusion*.
2. The oligopoly firms will compete on price so that the price and profits will be the same as those of a perfect competitive industry (*non-cooperative)*. A famous example is the *Bertrand model*.
3. The oligopoly price and profits will be somewhere within the range of the previous hypotheses. Most of the literature favours this hypothesis. Examples are the *Cournot model* and the *Von Stackelberg model*.
4. Oligopoly prices and profits are "indeterminate". That is, they may be anywhere within the range and are unpredictable.

In an oligopoly, pricing is best thought of as a strategic decision. Modern study of strategy is called *game theory*, because of the analogy to strategies in a game. Game theory assumes that people are rational and self-interested, the "rules of the game" are stable, everybody knows them and each player's payoff depends on the strategies chosen by others, as well as oneself.

One possible distinction in forms of oligopoly is between cooperative and non-cooperative. The first describes the theory and practice of cartels and tacit collusion. Cartels are most likely to form when there is a relatively small number of firms (making coordination and monitoring easier), difficult entry conditions (allowing price increases to be more durable), a trade association that can coordinate output market shares, monitor prices, even allocate orders and some credible form of punishment for cheaters. Thus to be able to raise prices without inducing substantial increased competition from non-members. Moreover, expected punishment for forming a cartel must be low relative to the advantages. Finally, the cost of establishing and enforcing an agreement must be low relative to the expected gains.

When firms in an oligopoly coordinate their actions despite the lack of an explicit cartel agreement, we say that the resulting coordination is tacit collusion. Empirical evidence by for example Hay and Kelley (1974) indicates that collusion of all kinds is more likely in highly concentrated industries.

Non-cooperative oligopoly models assume few sellers, who may have similar or different costs, high entry/exit costs, products may be identical or differentiated and price or quantity may be the *strategic variable*. As said before there are three well known models of non-cooperative oligopoly, i.e. the Cournot model, Bertrand model and Von Stackelberg model. Since these models form the roots of the industrial organization theory, a brief discussion seems appropriate.

While the Cournot model features firms engaging in quantity rivalry, which seems less reasonable than direct price competition, it yields outcomes in which there is an inverse relationship between market concentration and the extent to which market outcomes yield marginal cost pricing. Cournot oligopoly models can feature identical or differentiated products, identical or heterogeneous costs, and a wide variety of number of firms. The most simple formulation is a *duopoly* with identical costs (cQ) and products. When a simple linear inverse demand function $P(Q) = a - bQ$, where $Q = q_1 + q_2$ is assumed and firms maximize profits, the best response for both firms, taking the other firm's reaction curve into consideration, is $q_1 = q_2 = (a - c)/3b$. The market output in the Cournot equilibrium falls right in between the monopoly equilibrium output level, $2(a - c)/3b$, and the perfect competitive equilibrium output level, $(a - c)/b$. When the number of firms increases, the total market output moves closer to the perfect competitive output. The other way around, when the number of firms decreases, the market output moves closer to the monopoly output. This can also be seen from the *Lerner index*, which is equivalent to the extent to which price exceeds marginal cost. In the case of the Cournot model, the Lerner index is equal to $(P - C'(Q))/P = (nE_D)^{-1}$, where n is the number of firms and E_D is the demand elasticity. Hence, when n increases the Lerner index gets smaller implying that the market structure is getting more competitive. When the assumption of identical costs is dropped, it can be proved that firms with low costs have larger market share than firms with high costs.

The Bertrand model was developed because Bertrand criticized the Cournot idea of quantity rivalry and instead argued that firms actually engage in price rivalry. If firms are identical, and if there are no binding capacity constraints, then the Bertrand non-cooperative oligopoly model has the unsatisfactory result that even with only two firms, the equilibrium is $P = c$. Price rivalry drives prices downward as firms compete for market share. Undercutting your rival's cost results in you getting the entire market for yourself. This dynamic results in the competitive outcome regardless of market structure. However, if products are differentiated, firms collude, or buyers are poorly informed of price, then the Bertrand equilibrium will not occur.

Another common non-cooperative oligopoly model was developed by Von Stackelberg. This model involves a dominant firm (the leader) and a competitive fringe (the followers). Basically, the dominant firm knows how the fringe will react to its actions and can so take action that tailors the anticipated fringe response to best suit the dominant firm's profitability. In a situation with two firms, one leader, one follower, identical costs and products, the equilibrium outcome is: $q_L = (a - c)/2b$ and $q_F = (a - c)/4b$. Hence, the market output level in equilibrium falls between the Cournot equilibrium output level and the perfect competitive equilibrium one. All oligopoly equilibrium outcomes are summarized in Table 1.1.

Table 1.1 Oligopoly equilibrium outcomes

	Bertrand	Cournot	Stackelberg	Monopoly
p	c	$(a+2c)/3$	$(a+3c)/4$	$(a+c)/2$
q_1	$(a-c)/2b$	$(a-c)/3b$	$(a-c)/2b$	$(a-c)/2b$
q_2	$(a-c)/2b$	$(a-c)/3b$	$(a-c)/4b$	n/a
π_i	0	$(a-c)^2/9b$	$(a-c)^2/8b$, $(a-c)^2/16b$	$(a-c)^2/4b$

Note: outcomes are based on the following demand and cost functions: $P(Q) = a - bQ$, where $Q = q_1 + q_2$ and $C(q_i) = cq_i$, $i = 1,2$; π_i denotes profit of firm i.

1.2.2 Monopolistic Competition

In oligopoly, the main assumption is that the market consists of just a few firms, which can compete strategically with each other. When the number of firms increases and becomes large, firms are not able anymore to compete strategically with each other and take each others actions as given. Such a market structure is called monopolistic competition. Monopolistic competition is a cross between the two extremes of perfect competition and monopoly. There is free entry and exit of firms and production technologies are those of increasing returns to scale. What I know from the literature, monopolistic competition and scale economies are often mentioned together.

Products are differentiated *horizontally* or *vertically*. With horizontal product differentiation, each consumer prefers products that have certain characteristics. These may be geographical or product-oriented. Vertical product differentiation entails people can generally agree on a preference ranking.

Products are imperfectly substitutable. Consumer demand for differentiated products is sometimes described using two distinct approaches: the *love-of-variety approach (homogeneous demand)* and the *ideal variety approach (heterogeneous demand)*. The first approach assumes each consumer has a demand for multiple varieties of a product over time, e.g. restaurant meals. If all consumers share the same love-of-variety, then the aggregate market will sustain demand for many varieties of goods simultaneously.

The second approach assumes each product consists of a collection of different characteristics and each consumer has different preferences over these characteristics. The consumer chooses a product closest to their ideal variety subject to the price of the good.

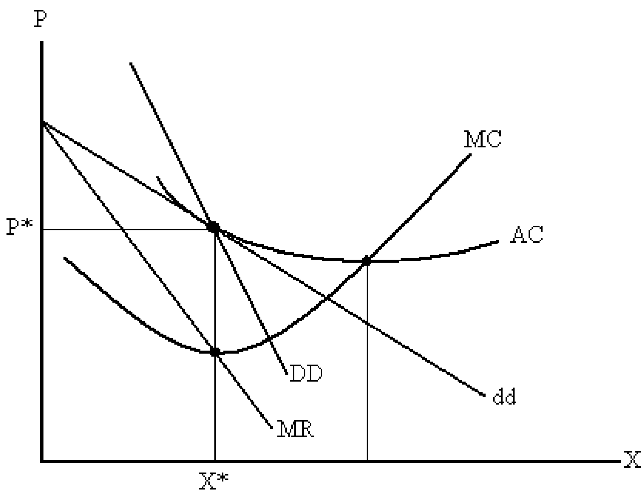
Monopolistic competition is usually conceptualized using the *representative consumer model*. The utility of the representative consumer embodies the preferences of the aggregate population of consumers and is defined over all products. In this case, preferences are symmetric. In this memorandum we only consider homogeneous demand and representative consumer models.

One particular theory finds itself at the roots of the economic literature on monopolistic competition. That is the theory by Chamberlin (1933). Edward Hastings Chamberlin is considered the “true revolutionary” (Blaug, 1997, p. 376) for proving under certain assumptions that a equilibrium exists in a market characterized by monopolistic competition and increasing returns to scale. Essentially, these assumptions include (Bishop, 1967, p. 252):

- number of sellers in group of firms is sufficiently large, such that each firm takes the behaviour of other firms in the group as given;
- group of firms is well defined and relatively small compared to the whole economy;
- products are economically differentiated and consumers love variety and;
- free entry and exit of firms.

In short, the monopolistic elements are those in the second and third assumption, which basically say that products are unique and give firms market power. Free entry and exit and the large number of firms ensure competition.

Figure 1 Chamberlinian monopolistic competition



To keep the discussion as simple as possible, I assume identical demand and cost conditions. A firm faces two individual demand curves. The first demand curve (dd) represents demand assuming all other firms in the group do not react to price changes of the firm under consideration. The second demand curve (DD) represents demand assuming other firms do react and set prices identically. DD is steeper than dd representing a lower price elasticity. Now, firms maximize profit, taking as given dd, by setting marginal revenue equal to marginal cost. It can be shown that dd is tangent to the average cost curve implying zero profit. Since all firms are identical with respect to demand and cost conditions, all firms make zero profit. Hence, there is no incentive for new firms to enter and for incumbent firms to exit the market. This is known as the Chamberlinian equilibrium, which is illustrated in figure 1.

The purpose of this paper is to compare several modelling alternatives for imperfect competition and scale economies and make a recommendation in the light of the problem issues discussed in the introduction. In section 2 I review more of the recent literature on monopolistic and oligopolistic competition in general equilibrium. Section 3 discusses the relevant properties of a selection of recently applied general equilibrium models and confronts these models with each other. Section 4 gives a recommendation for modelling imperfect competition and increasing returns to scale in WorldScan and analyzes the effects of this approach through simulations. Finally, I end with some concluding remarks.

2 Background

In the previous section I introduced oligopoly and monopolistic competition with a short description of the models from Cournot and Chamberlin. As already said earlier, economic literature shows among others two main research areas in microeconomic theory. First, there is industrial organization, which includes *small group models* focusing on strategic interaction between agents. The second mainstream in microeconomic theory is monopolistic competition. Monopolistic competition particularly works with *large group models*, hence, where the number of firms is large and strategic interaction is absent.

I will briefly discuss the background of modelling monopolistic competition and oligopoly in general equilibrium. I do not intend to give a full survey of the literature. For those interested, I refer to Brakman and Heijdra (2002) for a full survey of the literature on monopolistic competition. Recently, a review of the book was published in the "Journal of International Economics" by Redding (2005). Redding is full of praise about the book. His article starts with, quote, 'This is an important book that should be read by all research economists ...' and furthermore concludes with, quote, 'This book is a tremendous tribute to the 'second monopolistic competition revolution' initiated by the path-breaking work of Dixit and Stiglitz (1977)'. In contrast to the rich literature on monopolistic competition in general equilibrium, the literature on oligopoly in general equilibrium is very thin, as Neary (2002) makes clear. According to J. Peter Neary the development of tractable models of oligopoly in general equilibrium has been held back by a number of related problems. Nevertheless, Neary (2002) accommodates these problems and sketches a General Oligopolistic Equilibrium (GOLE) model that ensures theoretical consistency and tractability.

For the remainder of this section, I first present a simple version of the Dixit-Stiglitz (DS) (large group) model and discuss its key properties, as it is the 'workhorse model' for most currently applied models discussed in the next section. Next, I present the GOLE (small group) model of Neary (2002) and provide some discussion on the similarities and differences between DS and GOLE.

2.1 Monopolistic Competition à la Dixit-Stiglitz

Where Chamberlin (1933) failed in constructing a canonical model for monopolistic competition, Dixit and Stiglitz (1977) succeeded. They formulated a model which is both easily tractable and captures the key aspects of Chamberlinian monopolistic competition. The main contributions of this seminal article by Dixit and Stiglitz are threefold. First, the definition of a sector or group of firms is simplified. In particular, all varieties are symmetric and aggregated into a composite good using a constant-elasticity-of-substitution (CES) function. In their paper DS also consider a variable elasticity case. However, the constant elasticity case has become known as *the* DS model. Second, the utility function is separable and homothetic with convex indifference curves and all commodities have unit income elasticities. Separability allows the use of a *two-stage budgeting approach*, where income is allocated to composite goods in the first stage and reallocated to product varieties in the second stage. Usually, in the first stage a Cobb-Douglas specification is applied and in the second stage a CES-utility function.[3] Third, technology features internal economies of scale and firms are identical.

The DS model provides an important tool, and is a simple formulation for the analysis of a variety of problems, from international trade to economic geography and from macroeconomics to economic growth. The book of Brakman and Heijdra (2002) contains a survey of every of these fields of research. In the international trade literature, the DS model lead to a very productive line of research, often referred to as the *new trade theory* initiated by Krugman (1979, 1980). Moreover, Ethier (1982) explained the large volume of two-way trade in producer goods by using the DS framework. The DS model was also applied for modelling multinational firms and firm heterogeneity.

Perhaps, the DS model has influenced the development of economic geography even greater. Krugman (1991) combined the DS model with *ice-berg transportation costs* (see Samuelson, 1954) and developed the so-called *core-periphery model*, also known as the Krugman model, which explains the determinants of the location of economic activity. Research by Krugman and Venables (1995) and Venables (1996) also introduced vertical linkages between firms. The DS framework was also used to shed light on the development and relationship between cities. For a full survey of the literature see e.g. Brakman, Garretsen and Van Marrewijk (2001), Fujita and Thisse (2002) and Baldwin et al. (2003).

Where the first model of monopolistic competition by Chamberlin failed to impact macroeconomic literature (amongst other causes) by the Keynesian revolution, it was just this

[3] For a detailed discussion of separability of preferences and the two stage budgeting approach, see Deaton and Muellbauer (1980), Chapter 5.

strand of literature that accepted and used the DS framework to provide microfoundations for the ideas of Keynesian macroeconomics. Many other macroeconomic models were developed with monopolistically competitive price or wage setting. A survey can be found in Chapter 17 of Brakman and Heijdra (2002).

Economic growth is closely related to international trade. Romer (1990) showed how a DS model of intermediate inputs combined with endogenous investments in research and development (R&D) to develop new varieties of inputs, could lead to long run growth by increasing specialization. Basically, the DS model was the essential building block for the new generation of growth models of Romer (1990) and others. The literature analyzes a large number of issues, including the determinants of R&D-based economic growth, the role of public policies and scale effects. In addition, there exists some literature, that views growth as a process of creative destruction related to the links between market structure, innovation and the internal organization of the firm, see Redding (2005, p. 535). For a survey of the literature, see Chapter 14 of Brakman and Heijdra (2002).

All in all, we can conclude, that the monopolistic competitive framework by Dixit and Stiglitz has had a profound impact - much more than the model of Chamberlin - on the literature in the fields of international trade, economic geography, macroeconomics and economic growth. Next follows a simple version of the DS model.

2.1.1 The basic model

Demand side

There are two industries in the economy. One, sector 0, produces a homogeneous good under constant returns to scale and perfect competition. The other, sector 1, consists of a large number of monopolistically competitive firms, which produce under scale economies at firm level. The utility function of the representative household is given as,

$$u = x_0^{\beta} y^{1-\beta}, \quad \beta \in \langle 0,1 \rangle, \tag{2}$$

where u is utility, x_0 is consumption of the homogeneous good in sector 0, y is the consumption of the composite of varieties in sector 1 and β is the Cobb-Douglas share parameter. The composite good is a bundle of imperfect substitutes. A convenient formulation for y takes the CES form,

$$y = \left\{ \sum_{i=1}^{n} x_i^{(\varepsilon-1)/\varepsilon} \right\}^{\varepsilon/(\varepsilon-1)}, \quad \varepsilon > 1, \tag{3}$$

where n is the number of varieties, x_i is the consumption of variety i, and ε is the elasticity of substitution. The interpretation of ε is, that as it increases, products become more ‘substitutable’. Another interpretation can be given if we define γ as the marginal love-of-variety which can be shown to be equal to $1/(\varepsilon - 1)$. Hence, when varieties become less substitutable, consumers prefer more varieties, γ increases, which seems logical[4].

The consumer faces the following budget constraint,

$$p_0 x_0 + \sum_{i=1}^{n} p_i x_i = I \,, \tag{4}$$

where p_0 is the price for the homogeneous good in sector 0 and p_i is the price of variety i in sector 1, and I is the consumer’s income. The consumer maximizes (2), given (3) and (4) and prices p_0 and p_i. Following the two stage budgeting approach, we obtain:

$$\begin{aligned} x_0 &= \beta \frac{I}{p_0}, \\ y &= (1-\beta)\frac{I}{p_y}, \text{ where } p_y = \left\{ \sum_{i=1}^{n} p_i^{1-\varepsilon} \right\}^{1/(1-\varepsilon)}, \\ x_i &= (1-\beta)\frac{I}{p_y}\left(\frac{p_i}{p_y}\right)^{-\varepsilon}, \ \forall i. \end{aligned} \tag{5}$$

The result is that household income is spent linear with the share parameter as usual with the Cobb-Douglas utility function. Here, p_y is the composite price of a bundle of all varieties chosen in a utility maximizing fashion. The last equation represents the individual firm’s perceived demand curve (“dd-curve”). The industry demand curve (“DD-curve”) is derived when we assume symmetry for firms. Hence, prices and consumption levels are equal for all i ($p = p_i \ \forall i$). The industry demand curve becomes (6). The dd-curve is more elastic than the DD-curve, since $E_{dd} = \varepsilon > 1 = E_{DD}$.

$$x = (1-\beta)\frac{I}{np}. \tag{6}$$

Supply side

Assume the only production factor is labour, which is perfectly mobile. Sector 0 features constant returns to scale and perfect competition. Hence technology and price are given by:

[4] Bénassy (1996) stresses that love-of-variety and ease of substitution are two different phenomena. However, in the basic DS model, these cannot be distinguished. Therefore, in the basic DS model we refer to γ as the love-of-variety.

$$x_0 = a_0 L_0 \text{ and } p_0 = \frac{w}{a_0}, \tag{7}$$

where L_0 is the amount of labour used in sector 0, a_0 is the labour productivity and w is the wage rate in the economy. From (7) we see that marginal and average cost pricing both lead to zero profit.

Sector 1 features increasing returns to scale at firm level and monopolistic competition. The economies of scale are modelled in a natural way by imposing a fixed set-up cost for the production process. The interpretation is that firms must employ a minimum amount of labour, so-called *overhead labour*, before they can produce output at all. Firms maximize profit subject to the downward sloping individual perceived demand curve (dd). We assume no barriers exist for firms to enter and exit the market. As a result, all firms earn zero profit and thus choose prices equal to average cost. Because firms have to finance their fixed cost, prices are equal to marginal costs plus a mark-up. Technology and price in sector 1 are given by:

$$\begin{aligned} &x_i = a_i(L_i - f) \text{ if } L_i \geq f, \quad 0 \text{ otherwise}, \\ &p_i = \frac{\varepsilon}{\varepsilon - 1}\frac{w}{a_i}, \end{aligned} \tag{8}$$

where f is the fixed cost and $w/a_i(\varepsilon - 1)$ is the markup over marginal cost. In the optimization, price and income effects are ignored. The income effect is also known as the Ford effect, see Ford (1922). Here, ignoring these effects is allowed, because we assume a sufficiently large number of firms in the monopolistically competitive sector. When we allow for small groups of firms, price and income effects cannot be ignored. Yang and Heijdra (1993) adjust for price effects, but neglect income effects. d'Aspremont et al. (1996) describe a model without neglecting any indirect effects. Section 2.1.2 goes deeper into the analysis of these indirect effects.

Market equilibrium

The model is fully symmetric. All active firms produce the same output, pay the same wage, use the same amount of labour and earn the same profit. Therefore, we can drop the subscript i. Before we solve the model, we have to make assumptions about the labour market. In particular, since profits are zero in equilibrium, household income is equal to income earned from (fixed) labour, say wL^S. Furthermore, we assume that the labour market clears in equilibrium.

By substituting the markup pricing rule into the profit function and using the assumption of free entry and exit of firms, implying zero profit, the output level per active firm is equal to

$$x = af(\varepsilon - 1). \qquad (9)$$

Obviously, the optimal output level depends only on the fixed cost (expressed in units of output), labour productivity and the demand elasticity. Combining (7) and the fact that household income is equal to labour income yields the output in sector 0:

$$x_0 = a_0 \beta L^S. \qquad (10)$$

From (10) and the market clearing condition on the labour market it is clear, that output for both sectors is a constant fraction of the amount of labour supply.

Combining (8), (9), (10) and the market clearing condition on the labour market yields a system of equations, from which we can solve the equilibrium number of producing firms:

$$n = \frac{(1-\beta)L^S}{f\varepsilon}. \qquad (11)$$

The equilibrium number of firms depends positively on the labour supply in the monopolistically competitive sector and negatively on the fixed cost and the demand elasticity as expected.

The output of the composite good in sector 1 is equal to[5]

$$y = \lambda L_1^{\varepsilon/(\varepsilon-1)}, \qquad (12)$$

where $\lambda = a(\varepsilon - 1)/\varepsilon\ (f\varepsilon)^{-1/(\varepsilon-1)}$ and $L_1 = (1-\beta)L^S$. Equation (12) shows increasing returns to labour, because $\varepsilon/(\varepsilon - 1)$ is larger than one.

Welfare

The previous section considered the market equilibrium. Now, one would like to know whether the model outcome is Pareto efficient and can be decentralized. In particular, does the Chamberlinian equilibrium provide too much or too little variety?

[5] $y = n^{\varepsilon/(\varepsilon-1)}x = af(\varepsilon-1)\left(\frac{(1-\beta)L^S}{f\varepsilon}\right)^{\varepsilon/(\varepsilon-1)} = \frac{\varepsilon-1}{a\varepsilon}(f\varepsilon)^{-1/(\varepsilon-1)}\left\{(1-\beta)L^S\right\}^{\varepsilon/(\varepsilon-1)} = \lambda L_1^{\varepsilon/(\varepsilon-1)}.$

In the welfare optimum the central planner allocates goods in both sectors and chooses the number of active firms in the monopolistically competitive sector, such that the household's utility is maximal. The general equilibrium model can be written as

$$\max_{\{x_0, y, n\}} u = x_0^{\beta} y^{1-\beta}, \quad \text{subject to} \quad L^S = \frac{x_0}{a_0} + \frac{n^{-\gamma} y}{a} + af. \tag{13}$$

Let the superscript '^' denote the social optimal values of the variables. Solving the model yields the following internal solution[6]:

$$\begin{aligned}
\hat{x}_0 &= \frac{a_0 \beta L^S}{1 - (1-\varepsilon)/(1-\beta)}, \\
\hat{y} &= \hat{n}^{\varepsilon/(\varepsilon-1)} \hat{x} = \hat{n}^{\varepsilon/(\varepsilon-1)} af(\varepsilon - 1), \\
\hat{n} &= \frac{(1-\beta) L^S / (\varepsilon - 1)}{\{1 - (1-\varepsilon)/(1-\beta)\} f}.
\end{aligned} \tag{14}$$

Comparing (14) to (10)-(12) shows that individual firm output in the monopolistically competitive sector is equal in both cases. It appears that aggregate output levels for both sectors are different in the market equilibrium and the social optimum. In sector 0, too much output is produced and this is not socially optimal. In sector 1, the opposite is the case. Hence, the Chamberlinian equilibrium provides too few varieties as compared to the welfare optimum.

The solution of the general equilibrium model makes clear, that for decentralization, hence marginal cost pricing, we need to impose an ad-valorem tax. Assume that each active firm receives an ad-valorem product subsidy (to cover the fixed set-up cost). The markup rule in (8) now becomes:

$$p(1+\tau) = \frac{\varepsilon}{(\varepsilon - 1)} \frac{w}{a}. \tag{15}$$

Since the social optimum requires marginal cost pricing ($p = w/a$), instead of average cost pricing, it follows from (15), that the socially optimal product subsidy, τ, is equal to the mark-up, $\varepsilon/(\varepsilon - 1)$. Note that this social optimum is a first-best decentralizable solution of our simple version of the DS model.

[6] A boundary solution can be found when imposing a lower bound on the number of firms. In that case the *business stealing effect* - output per firm decreases when the number of firms increases - is stronger than the preference-for-diversity effect, see Mankiw and Whinston (1986, p. 49).

In short, the key properties of the DS model can be summarized as follows. Firms have market power, because goods are imperfect substitutes. Firms exploit this market power by setting price above marginal cost and cover their fixed cost. This is not socially optimal, because that would require marginal cost pricing. To solve this problem, an ad-valorem tax is introduced, which is used for funding the set-up cost. Concluding, the DS model offers a tractable framework to model imperfect competition and economies of scale and captures the key insights of Chamberlin. Through the last three decades, the DS model can be seen as the foundation of more extended models in several fields of research.

2.1.2 Extensions

Market power and love-of-variety: two distinct phenomena

It has been argued by Bénassy (1996) and Broer and Heijdra (2001) that love-of-variety by households is conceptually different from market power by individual firms. This was first acknowledged by Dixit and Stiglitz in their working paper on *the* DS model in February 1975. The discussion there entails utility could also directly be a function of the range of varieties actually produced as a public good, besides the consumed amounts. Following Dixit and Stiglitz (1975), a convenient but not necessarily more restrictive function for the composite good in sector 1 is the one in (16),

$$y = n^{\theta-\varepsilon/(\varepsilon-1)}\left\{\sum_{i=1}^{n} x_i^{(\varepsilon-1)/\varepsilon}\right\}^{\varepsilon/(\varepsilon-1)}, \qquad (16)$$

where $\varepsilon/(\varepsilon\text{-}1)$ captures the market power of firms in the monopolistic competitive sector and θ captures the love-of-variety. If θ is equal to unity, then consumers are indifferent to variety. At the same time, firms do possess market power, as $\varepsilon/(\varepsilon-1) > 1$ (since $\varepsilon > 1$ by assumption). Obviously, in the basic model in section 2.1.1, θ equals $\varepsilon/(\varepsilon-1)$ and the two effects cannot be distinguished.

The two-stage budgeting approach still holds and the analysis is almost unchanged. The interior equilibrium (denoted by a superscript 'e') for the monopolistic competitive sector is then:

$$p^e = \frac{\varepsilon}{\varepsilon-1}\frac{w}{a}, \qquad x^e = af(\varepsilon-1), \qquad n^e = \frac{(1-\beta)L^S}{f\varepsilon}, \qquad y^e = \lambda' L_1^{\theta}, \qquad (17)$$

where $\lambda' = af^{1-\theta}(\theta-1)^{(\theta-1)}\theta^{-\theta}$ and $L_1 = (1-\beta)L^S$. From (17) we notice that p^e, x^e and n^e are the same as in (8), (9) and (11). However, y^e has changed. Comparing, the composite output in (17) to (12), we find that θ regulates whether there are increasing ($\theta > 1$) or constant ($\theta = 1$) returns to labour at an aggregated level. Note that, at firm level, we still have increasing returns to scale. In addition, when θ equals one there are constant returns to scale, but still we have monopolistic competition[7].

The first-best social optimum also changes in the extended model compared to the basic model. The solution is in fact the same as in (14), except that the control parameter is not the demand elasticity, ε, but the love-of-variety parameter θ[8]. Therefore, the same discussion as for the comparison of market equilibrium to the social optimum holds here. Indeed, the comparison of the social optima in the basic and extended model depends crucially on the values for θ, the love-of-variety-parameter and $\varepsilon/(\varepsilon - 1)$, the mark-up.

The first-best social (*unconstrained*) optimum can still be decentralized by imposing a product subsidy equal to $\varepsilon/(\varepsilon - 1)$, as in the basic model. However, here we additionally need a lump-sum tax to adjust firm size and the number of firms to their socially optimal values. If this lump-sum tax instrument is not available in a given context, then one could revert to a second-best social (*constrained*) optimum by finding an optimal trade-off between removing the monopoly distortion and producing an optimal amount of varieties. The optimal product subsidy is then equal to $\theta - 1$.

In the standard DS model, $\theta = \varepsilon/(\varepsilon - 1)$ holds and a product subsidy equal to $\varepsilon/(\varepsilon - 1)$ is imposed, the social welfare optimum is decentralizable. In general, however, preference-for-diversity and degree of market power are different concepts and hence take on different values: $\theta \neq \varepsilon/(\varepsilon - 1)$. Thus, the social welfare optimum is not decentralizable, unless a lump-sum tax is introduced as an additional instrument to adjust the output in the monopolistic sector to its socially optimal value. Hence, for decentralization we need two instruments. This is also known as the *Tinbergen rule*, see Tinbergen (1952).

Include indirect effects

In large group models (the number of firms, *n*, is sufficiently large), such as the basic DS model in section 2.1.1 it is convenient and allowed to neglect terms of the order $1/n$ in the computed elasticities. However, when *n* is not sufficiently large, indirect effects such as price index effects and income effects cannot be ignored. In the literature two adjustments have been explored. First, there is the approach of Yang and Heijdra (1993), who take into account the price index effect, but keeping income fixed. Hence, this approach is still an approximation since it adjusts partially for indirect effects. Second, there is the approach of d'Aspremont et al. (1996), who also take into account the income effects, and is therefore preferable to the approximation approach of Yang and Heijdra. I shall briefly elaborate on both methods.

[7] The marginal preference for variety is equal to the elasticity of the average preference-for-diversity, which can be computed by comparing the value of composite consumption, *y*, obtained if *n* varieties and x/n units per variety are chosen with the value of *y* if *x* units of a single variety are chosen ($n = 1$), see Bénassy (1996). The marginal love-of-variety is $1/(\varepsilon-1)$ in the basic DS model, where it is θ-1, in the extended model. Hence, when $\theta = \varepsilon/(\varepsilon-1)$, the extended model collapses to the basic model.

[8] Note that, this does not mean that ε can directly be replaced by θ in (14). However, ε should be replaced by $\theta/(\theta-1)$.

In the DS model, the first-order condition for profit maximization is given by equating marginal revenue and marginal cost, which yields:

$$p_i\left(1-\frac{1}{\varepsilon_i}\right)=\frac{w}{a_i}, \quad \text{with } \varepsilon_i=-\frac{dx_i/x}{dp_i/p}. \tag{18}$$

In the standard DS model in section 2.1.1, the elasticity of demand, ε_i, is calculated considering only the direct effect on demand of a change in price, neglecting the indirect effects through I and p_y. Thus using the third equation in (5), the elasticity of demand equals the *intrasector* elasticity of substitution:

$$\varepsilon_i^{DS}=\varepsilon . \tag{19}$$

Where DS only use the third equation in (5), Yang and Heijdra also use the first two equations and take into account the effect of p_i on p_y. After some algebra, the elasticity of demand is obtained as in (20).

$$\varepsilon_i^{YH}=\varepsilon+(1-\varepsilon)\left(\frac{p_i}{p_y}\right)^{1-\varepsilon}\approx\varepsilon+(1-\varepsilon)\frac{1}{n^*} \tag{20}$$

For a general overall utility function, the elasticity of demand according to Yang and Heijdra is a weighted average of the intra- and intersectoral elasticity of substitution. However, here we have chosen for a Cobb-Douglas specification, implying the intersectoral elasticity of substitution equals one. Therefore, (20) is somewhat simplified. In case of symmetry, '≈' can be replaced by '='. Obviously, when n^* is large, i.e. approaches infinity, the elasticity of demand is again equal to the intrasectoral elasticity of substitution. The disadvantage of this approach is that income I is kept fixed independently of p_i. Yang and Heijdra take into account indirect effects within a sector. However, they ignore economy wide indirect effects.

A more *complete* approach, which also takes into account Ford effects, is the one developed by d'Aspremont et al. The recognition that income depends on firm's prices naturally follows from splitting income into labour endowment in terms of wage and distributed profits. Since profits depend on prices, also income depends on prices. Combining, (5) and the fact that income depends on firm's prices, yields the following implicit elasticity of demand for a symmetrical equilibrium:

$$\varepsilon^*=\varepsilon+(1-\varepsilon)\frac{1}{n^*}+\frac{(1-\varepsilon^*)(1-\beta)}{\varepsilon^*-1+\beta}\frac{1}{n^*}. \tag{21}$$

The first term coincides with ε_i^{DS}. Including the second term leads to ε_i^{YH}. When income effects are also taken into account, the elasticity of demand becomes (21). (21) is an implicit solution, which can be rewritten as a quadratic equation for which a unique solution exists[9]. Since $\varepsilon > 1$ in the DS model, we have the following relation for $n^* > 1$: $\varepsilon^* < \varepsilon^{*YH} < \varepsilon^{*DS}$. It appears that taking into account the indirect effects of a price change leads to a gradual decrease in the elasticity of demand and hence an increase in monopoly power. From (11) we know that the optimal number of varieties is inversely related to the elasticity of demand. Hence, when n is small and the DS approximation is used, the equilibrium number of activities is underestimated.

Concluding, in a large group model it is convenient to use the DS approximation. However, in a small group model one should adjust for indirect effects by using the YH approximation or preferably the objective approach by d'Aspremont et al.

Variable elasticity of substitution

So far we have assumed a constant intrasectoral elasticity of substitution for the monopolistic competitive sector. In this section the implications of relaxing this assumption are discussed. We still keep the Cobb-Douglas overall utility function, but the CES function for the monopolistic competitive sector is replaced with a non-homothetic, but still separable utility function:

$$u = x_0^{\beta} \left\{ \sum_{i=1}^{n} v(x_i) \right\}^{1-\beta}, \tag{22}$$

where v is increasing and concave. This example originates from Dixit and Stiglitz (1975, section 4.5). Because the derivations and algebra are rather involved and opaque, we omit them and focus on the results. For the interested reader I refer to the Dixit and Stiglitz (1975) and Spence (1974).

The market equilibrium based on (22) is almost identical as for the basic model, except that in (6)-(12) β and ε are now functions of the equilibrium firm output level, x_e, instead of constants. In a welfare optimum price and output need not be the same as in the market equilibrium. It all depends on the sign of $\rho'(x)$, where $\rho(x) = xv'(x)/v(x)$. Dixit and Stiglitz give the following intuition behind this result:

"With our large group assumptions, the revenue of each firm is proportional to xv'(x). However, the contribution of output to group utility is v(x). The ratio of the two is ρ(x). Therefore, if ρ'(x) > 0, then at the margin each firm finds it more profitable to expand than what would be socially desirable."

[9] The quadratic equation is: $n^*\varepsilon^{*2} - \{n^*(1-\beta) + (n^*-1)\varepsilon + \beta\}\varepsilon^* + (n^*-1)(1-\beta)\varepsilon = 0$. This can be rewritten as $a\varepsilon^{*2} + b\varepsilon^* + c = 0$, which is a parabola. From mathematics we know that this equation has a solution for ε^*.

In short, the market equilibrium produces larger amounts but a smaller number of varieties, than in the social welfare optimum. Further, we can conclude that the implications are not as unambiguous as in the basic DS model.

Asymmetry

In the basic DS model we have assumed symmetry within the monopolistic competitive sector. An important modification may be to remove this assumption. This modification is also treated in the DS working paper from 1975. More generally, they ask the question: *"... will the right set of commodities be produced in the monopolistically competitive equilibrium? And if not, can we say anything about the nature of the biases?"*. They argue that the produced set of commodities and thus also the possible bias depends on a number of factors: fixed and marginal cost, own and cross-elasticities of demand and the level of the demand schedule. As in the previous paragraph on variable elasticities, here, I again omit any formulae and just present the implications of relaxing the assumption of symmetry.

The basic principle for the existence of biases between the monopolistic competitive optimal output and social optimal output is the trade-off between the ability to earn enough revenue to cover fixed costs and the desirability to take into account the consumer surplus.

DS show that while low own elasticity commodities have the potential of earning relatively large revenues over marginal costs, they may not be able to do so if there is a high cross-elasticity with a high own elasticity commodity. Even when cross-elasticities are zero, a wrong set of commodities can be produced, see figures 2 and 3.

Figure 2 Different elasticities of demand

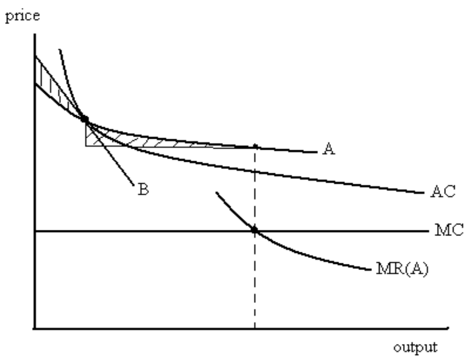


Figure 3 Different cost functions

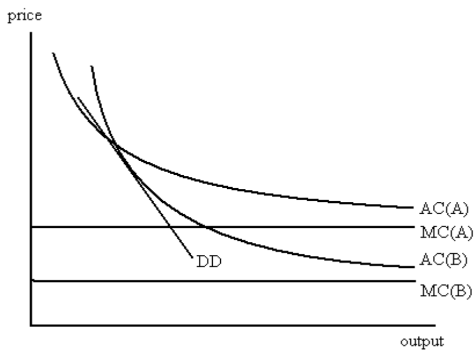


Figure 2 illustrates that high own elasticity commodity A is produced in the Chamberlinian equilibrium, where it is socially desirable to produce low own elasticity commodity B, since it is marginal. Figure 3 presents a similar analysis for the same demand curve, however, for different cost functions. Commodity A has a lower fixed cost, but a higher marginal cost than commodity B. Again, commodity A is produced where B is socially desirable, because there is a much larger consumer surplus. For both cases, we can say that welfare would increase if for example the government were to restrict production of A, so B would be viable. Nevertheless, it has to be said that the analysis of applications with asymmetry do depend on functional forms of the demand and cost schedules.

This section has provided a toolbox of possible extensions of the basic DS model. In particular, we have considered:

- market power and preference-for-diversity as two distinct concepts;
- including indirect, price index and income, effects of a price change on output level;
- making the elasticity of substitution variable instead of constant and;
- the effects of dropping the assumption of symmetry within the monopolistic competitive sector.

Next, we turn to the general oligopolistic competitive model by Neary.

2.2 Oligopolistic Competition à la Neary

In large group models, such as the DS model, price and income effects can be neglected in the computation of elasticities. However, in small group models, such as oligopolistic competitive models, these effects must be accounted for. In addition, firms are likely to interact strategically. In the literature of industrial organization, these problems have been discussed extensively for partial equilibrium. However, these issues are a lot tougher to treat in general equilibrium. The first reason is that, if firms are large in their own market, and if that market constitutes a significant segment of the economy, then the firms have a direct influence on economy-wide variables. The second reason is that, large firms influence the cost of living, and rational shareholders should take this into account in choosing the profit-maximizing level of output or price. Many authors interpret this as implying that predictions of general oligopolistic equilibrium models are sensitive to the choice of the numeraire. Last, firms in general oligopolistic equilibrium commonly have such bad-behaved reaction functions, that models cannot be solved. These problems have held back the development of tractable general oligopolistic equilibrium models.

In the literature several attempts were already made to face up with the problems at hand, but none of them were successful. Neary (2002) argues that all these problems can be solved in a simple matter: firms are *large* in their industry, but relatively *small* relative to the whole economy. This comes down to assuming a sufficiently large number of sectors. Consequently, firms do not influence economy-wide variables. Moreover, profit maximization leads to the same real allocation and without Ford effects reaction functions will likely be better behaved. In his model Neary tries to integrate the industrial organization theory with the theory of general equilibrium models.

In the remainder of this section, I discuss a simple version of the GOLE model by Neary (2002) to provide some contrast to the dominant influence of the monopolistic competitive model by Dixit and Stiglitz.

2.2.1 The model

Demand side

The key element in Neary's GOLE model is to assume a sufficiently large number of sectors, in which a few firms own the lion's share of the industry. To keep the analysis as simple as possible, we start by assuming an additively separable utility function of a continuum of goods for the representative consumer. Neary (2002) uses quadratic preferences that imply linear perceived demand curves[10]. The utility maximization problem under the budget constraint is equivalent to (23). In addition, the inverse demand functions and the marginal utility of income are shown in (24). In (24) μ_p is the mean of the prices and σ_p^2 is the uncentered variance of the prices.

$$\max_{x(i)} \int_0^1 \left[rx(i) - \tfrac{1}{2} sx(i)^2 \right] di, \quad \text{s.t.} \int_0^1 p(i)x(i)di \leq I \quad (\lambda). \tag{23}$$

$$p(i) = \tfrac{1}{\lambda}\left[r - sx(i) \right] \quad \text{and} \quad \lambda(p(i), I) = \frac{r\mu_p - sI}{\sigma_p^2}. \tag{24}$$

The key feature of the model is that the marginal utility of income in each sector is taken given by firms, and hence the perceived demand curve is linear in (7), but is endogenous in general equilibrium. Combined with the assumption of a continuum of sectors, this allows for a tractable and consistent approach for modelling oligopolistic competition in general equilibrium.

Supply side

Consider the same technology as used in the homogenous good sector for the DS model, see (7), i.e. labour is the only production factor. Combined with Cournot competition and an exogenous given number of firms, *n*, the first order condition for an individual firm in industry *i* is:

[10] Neary (2002) also considers examples of Cobb-Douglas and Dixit-Stiglitz preferences. However, Neary shows that these specifications do not provide convenient solutions and yield unattractive implications in oligopoly. In particular, Cobb-Douglas demand functions are extremely restrictive and are inconsistent with profit maximization by a monopolist. Further, DS demand functions do not allow that output are often strategic complements in Cournot competition, and reaction functions may be non-monotonic.

$$p(i)=\frac{w}{a(i)}+\frac{sx(i)}{\lambda}. \tag{25}$$

Equilibrium

Solving the system yields equilibrium output and price as in (26).

$$x(i)=\frac{r-\lambda\, w/a(i)}{s(n+1)} \qquad \text{and} \qquad p(i)=\frac{r-\lambda n\, w/a(i)}{\lambda(n+1)} \tag{26}$$

A standard property of additive preferences is that demands are homogeneous of degree zero in prices, hence *w*, and the inverse of marginal utility of income, λ^{-1}. Besides, the absolute values of these nominal variables are indeterminate and of no interest in real models. The homogeneity of degree zero ensures that scaling has no effect on the behaviour of the model. Hence, the numeraire problem is hereby solved. It is convenient to choose utility as numeraire implying a unit marginal utility of income, i.e. $\lambda = 1$.

To solve the model completely, we can combine (26) with the labour market clearing condition: $L^S = \int_0^1 nx(i)/a(i)di$, where L^S is the labour supply, and get:

$$w=\frac{1}{\sigma_a^2}\left(r\mu_a-\frac{n+1}{n}sL^S\right), \tag{27}$$

where μ_a and σ_a denote the first two moments of the Ricardian technology distribution, i.e. $\mu_a = \int_0^1 a(i)^{-1}di$ and $\sigma_a^2 = \int_0^1 a(i)^{-2}di$.

Welfare

Neary (2002) examines three issues within a welfare context. Firstly, he examines what the effect of more competition is on the functional distribution of income. It appears that when competition becomes more active in the sense that the number of firms increases, the share of wages in national income also increases. Secondly, he finds a positive effect of competition policy on overall welfare, provided $\sigma_a^2 > \mu_a^2$. Consequently, when $\sigma_a^2 = \mu_a^2$, hence all industries have the same technology, overall welfare does not increase. Last, Neary shows that aggregate welfare and the share of wages need not move together.

In addition to the welfare analysis for oligopolistic competition, Neary (2003) compares the effects with those achieved in monopolistic competition. In particular, he shows that competition policy boosts welfare, provided the variance of the technology distribution is positive.

2.3 Discussion

We have seen that the DS specification is extremely tractable and it lends itself easily to general equilibrium applications through its homotheticity properties. In addition, it allows consideration of the implications of increasing returns to scale and product differentiation in general equilibrium. However, these clean functional forms for consumer and producer behaviour can be seen as restrictive and impose several special assumptions. Here, I would like to point out some of the critiques from Neary (2000), the DS model has received over the years.

Krugman (1979) showed that product diversity alone can cause gains from trade. In particular, when two countries have identical technologies, and open their borders, the number of varieties increases, implying consumers are better off. Neary admits variety is important for consumers, but questions to what extent. As we have seen, it is possible to divide market power and preference-for-diversity as two independent parameters in the DS model. Nevertheless, *"it clearly fails to capture one of the concerns of anti-globalisation protesters: that liberalising trade may reduce rather than increase variety"*, as argued by Neary.

Another disadvantage of using the CES utility function, is that output is given for each firm if the elasticity of substitution is given. Hence, trade policy has no influence on firm output. However, all changes in industry size are taken care of by the number of firms. This problem can be solved by using a variable elasticity utility function as discussed in section 2.1.2. However, using this function has a drawback in that we lose some of the tractability. Lawrence and Spiller (1983) and Flam and Helpman (1987) relax the assumption of a homothetic production function, which leads to changes in equilibrium firm size.

From an industrial organization point of view, it is curious where entering firms come from and exiting firms go to in monopolistic competitive models at all. They imply an unlimited supply of atomistic firms, which prove empirically relatively implausible for most industries.

Finally, the most important critique from the field of industrial organization is that firms operate myopic and do not strategically interact with each other. The assumption of myopic behaviour can be treated by extending the basic DS model in including indirect effects through the discussed approaches by Yang and Heijdra (1993) and d'Aspremont et al. (1996). Nevertheless, when the number of firms in a sector is small, hence market power is concentrated with a few firms, it is likely that firms behave strategically.

Neary's concerns with the DS model are mostly characterized by restrictive assumptions and no strategic interaction. He argues what is needed is a general oligopolistic equilibrium model, i.e. the model in section 2.2. Actually, the basic idea behind Neary's GOLE model is the same as the DS model. Both models view firms as small in the whole economy, but possess monopoly power in their own market. The approach by Neary is, as he puts it himself, *"... owes a great deal to that of Dixit and Stiglitz (1977), and in particular to the clarification of its theoretical underpinnings provided by d'Aspremont, Dos Santos Ferreira, and Gérard-Varet*

(1996).” Indeed, the difference between the models is that the DS model describes a *large group* of firms combined with product differentiation and Neary’s model describes a large group of *small sectors* combined with strategic interaction.

Obviously, there is no right or wrong model. The choice for a model should depend on the application one is concerned with. For our purposes, in particular identifying and describing scale economies, we do not need to model strategic interaction between firms. Therefore, we can suffice with a tractable and easy to use large group approach like the DS model.

3 Currently applied models

In section 2, I discussed the two main strands in the imperfect competition literature when it comes to general equilibrium modelling. In particular, I presented the classical framework of monopolistic competition in general equilibrium by Dixit and Stiglitz and the innovative general oligopolistic equilibrium model by Neary. I already concluded, that for our objectives it is sufficient to follow a tractable approach like the (large group) DS model. In this section, I would like to highlight the aspects of imperfect competition and scale economies in some recent applied general equilibrium models. I focus on multi-region models, since WorldScan is a global model. In total there are six models, which I discuss in no particular order.

3.1 GTAP model

GTAP stands for Global Trade Analysis Project and is a global network of researchers and policy makers conducting quantitative analysis of international policy issues. The standard GTAP model is a multiregion, multisector, computable general equilibrium model, with perfect competition and constant returns to scale. Besides the standard model, several extensions of the model have been produced. One of them, is the inclusion of scale economies and imperfect competition in the GTAP model. The discussion that follows, is based on a GTAP working paper by Francois and Roland-Holst (FRH) (1996).

FRH focus on market power, where a class of heterogeneous goods are differentiated by country of origin. This is the Armington assumption, see Armington (1969). Further, they consider firm-level product differentiation.

In Armington models, we assume that demand from different goods is aggregated into a composite good using a CES specification:

$$y_{jr} = \left\{ \sum_{i=1}^{R} \alpha_{jir} x_{jir}^{\rho_j} \right\}^{1/\rho_j}, \tag{28}$$

where all variables have the usual interpretation as in the DS model and *j* denotes the product, *i* the index for region of origin, and *r* is the index for the region where the product is consumed. The Armington specification has almost the same properties as the DS specification[11]. Therefore we can apply the two-stage-budgeting approach to derive the demand function and hence the elasticity of demand. To save space I omit the derivation and just present the result:

$$\varepsilon_{jir} = \sigma_j + (1-\sigma_j)\left[\sum_{k=1}^{R} \left(\frac{\alpha_{jkr}}{\alpha_{jir}} \right)^{\sigma_j} \left(\frac{p_{jkr}}{p_{jir}} \right)^{1-\sigma_j} \right]^{-1}, \tag{29}$$

where $\sigma_j = 1 / (1-\rho_j)$ is the elasticity of substitution and the last term measures market share. This last term also shows that price index effects are taken into account.

Now, we can introduce imperfect competition into the system. One way is to assume that firms do not price discriminate, but operate in a single market. This means that firms charge a single mark-up. The elasticity of demand in (29) can then be aggregated over regions taking into account regional market shares:

$$\varepsilon_{ji} = \sigma_j + (1-\sigma_j)\sum_{r=1}^{R} \frac{x_{jir}}{x_{ji}} \left[\sum_{k=1}^{R} \left(\frac{\alpha_{jkr}}{\alpha_{jir}} \right)^{\sigma_j} \left(\frac{p_{jkr}}{p_{jir}} \right)^{1-\sigma_j} \right]^{-1}. \tag{30}$$

The pricing rule that FRH apply is based on the conjectural variations approach[12] and is given by (31). Conjectural variations mean that firms have expectations about other firms reactions to a change in their own behaviour. In a mathematical sense, a firm’s conjecture is the derivative of aggregate firm output with respect to a single firm’s output.

$$p_{ji} = c_{ji}\left(1 - \frac{\Omega_{ji}}{n\varepsilon_{ji}} \right)^{-1}, \tag{31}$$

where p_{ji} is price, c_{ji} is marginal cost, Ω_{ji} is the conjecture of firm *i* with respect to the change in industry output of firm *i*’s change in output and *n* is the number of firms. Actually, (31) is the oligopoly pricing rule. In case of monopoly or monopolistic competition, we have $\Omega_{ji} / n = 1$.

[11] The only differences are that with the DS model, the number of firms is endogenous and $\alpha_{jir} = 1$, while in Armington models the number of regions and thus the number of varieties is fixed and α_{jir} is in general not equal to one.

[12] For the unfamiliar reader, see Figuières et al. (2004) for a exhaustive discussion of the conjectural variations approach.

Next, we turn to firm-level product differentiation. The specification for the composite good is given by (28), where *R* in the summation should be replaced by *n*, since we now index, *i*, over *firms* instead of over regions. The specification resembles very much the DS specification. However, the FRH specification still provides a partial geographic anchor for production. The rest is analogous to the case of the Armington models: the elasticity of demand is given by (30), where *k* sums to *n* instead of *R* and *i* denotes firm *i* instead of region of origin *i*. Note, that this specification is identical to (22), the DS elasticity of demand combined with the YH approximation, when we assume full symmetry. In addition, free entry and exit of firms is assumed, in contrast to the Armington specification.

FRH also consider a case where the elasticity of demand is somewhat simplified, through *variety-scaling*, by imposing identical cost functions within a region. This means that regional firms produce the same quantity and charge the same price. The composite good is then defined by (32),

$$y_{jr} = \left\{ \sum_{i=1}^{R} n_{ji} \alpha_{jir} x_{jir}^{\rho_j} \right\}^{1/\rho_j}, \tag{32}$$

where x_{jir} is the identical consumption in region *r* of each variety produced in region *i*. The main difference with before is that the CES weights are now endogenous instead of the summation. The elasticity of demand then changes to (33). In addition, FRH show that this system of equations can easily be combined with a homothetic decreasing average cost function.

$$\varepsilon_{ji} = \sigma_j + (1-\sigma_j) \sum_{r=1}^{R} \frac{x_{jir}}{x_{ji}} \left[\sum_{k=1}^{R} n_{jk} \left(\frac{\alpha_{jkr}}{\alpha_{jir}} \right)^{\sigma_j} \left(\frac{p_{jkr}}{p_{jir}} \right)^{1-\sigma_j} \right]^{-1}. \tag{33}$$

Technology in the GTAP model follows a production ‘tree’. At the bottom of the inverted tree are the individual inputs demanded by the firm. These consist of primary production factors on the one hand and intermediate inputs produced domestically and abroad on the other. Imported inputs follow a nested CES specification. All primary and intermediate inputs are again nested in a Leontief (perfect complementarity) specification.

FRH illustrate the workings of the model with an aggregation of version 3 of the GTAP database. As a measure for market power FRH apply a so-called cost-disadvantage-ratio (CDR), which is equivalent to fixed costs divided by total costs. CDR estimates are representative for typical reported CDR values, as presented in e.g. Pratten (1988). Furthermore, they assume that conjectural variations Ω_{ji} / n take values 0.2 and 0.5, equivalent with respectively 5 and 2 firms in Cournot competition. The mark-ups are derived from a 1992 benchmark dataset.

Concluding, FRH present a menu of relatively standard specifications of imperfect competition and scale economies. They consider two ways of differentiating products: by origin and by variety. The first is primarily based on the Armington assumption and a fixed number of firms. The second is a combination of a standard CES specification for product differentiation with an endogenous number of firms and a partial geographical anchor. In either case, price index effects are accounted for. Besides, they also considered some practical simplifications in the form of operations in a single market and variety-scaling.

3.2 MIRAGE

MIRAGE stands for Modelling International Relationships in Applied General Equilibrium and is a multi-region, multi-sector computable general equilibrium model, devoted to trade policy analysis, as is the GTAP model. It incorporates imperfect competition, product differentiation by variety and by quality, and foreign direct investment, in a sequential dynamic set-up where installed capital is assumed to be immobile. The succeeding description of the imperfect competitive properties of the model come from Bchir et al. (2002).

MIRAGE describes imperfect competition in an oligopolistic framework à la Cournot. It accounts for horizontal product differentiation linked to varieties, but also to geographical origin (nested Armington - Dixit-Stiglitz utility function). Since Harris (1984)[13], imperfect competition and horizontal product differentiation are commonly incorporated, notably based on the formalisations proposed by Smith and Venables (1988), and by Harrison, Rutherford and Tarr (1997). In fact, imperfect competition, scale economies and product differentiation are reasonably similarly modelled as in the GTAP model. The primary difference is that GTAP has been heavily engaged in data work to make these models easier accessible.

At the demand side of MIRAGE consumption is modelled in each region through a representative agent. Utility is modelled as a nested Armington - Dixit-Stiglitz function as the one used in Harrison, Rutherford and Tarr (1997)[14]. In particular, domestic demand is modelled using the DS specification, import demand is modelled using a traditional Armington specification and both demands are aggregated in a Cobb-Douglas utility function. However, this specification does not allow for vertical differentiation nor for specialisation in quality. The utility function could be extended by adding a further CES nesting level distinguishing between quality ranges. The elasticity of demand of firms in region r selling region r' is equal to the inverse of (34),

[13] Harris (1984) describes consumer demand by a blend of Armington-type and Dixit-Stiglitz preferences. Overall utility is Cobb-Douglas and is indexed over perfect and imperfect competitive sectors. Perfect competitive sectors are described by a Armington-type function and imperfect competitive sectors by a DS specification. Harris' specification introduces a certain asymmetry in consumer preferences. Perfect competitive commodities have different preference weights in the composite quantity index, while imperfect competitive commodities do not.

[14] This is where FRH en MIRAGE do differ. While MIRAGE use a nested Armington – DS function for utility, FRH use a non-nested version.

$$\frac{1}{\varepsilon_{rr'}}=\begin{cases}\theta_{rr}+\dfrac{1}{\sigma_{DD}}(n_r-1)+\dfrac{1}{\sigma_{DM}}(1-\theta_{rr}), & r=r'\\[2ex] \theta_{rr'}+\dfrac{1}{\sigma_{MM}}\left(n_r-\dfrac{\theta_{rr'}}{\theta_{r'}^{M}}\right)+\dfrac{1}{\sigma_{DM}}\left(\dfrac{\theta_{rr'}}{\theta_{r'}^{M}}-\theta_{rr'}\right), & r\neq r'\end{cases} \tag{34}$$

where $\theta_{rr'}$ is the market share of firms from region *r* in country *r'*, θ_r^M is the market share of imports in region *r*, σ_{DD}, σ_{DM} and σ_{MM} are trade elasticities of substitution between domestic and imported goods.

At the supply side of the model, technology can be of constant and increasing returns to scale and follows almost the same production tree as in the GTAP model. The only difference is that intermediate consumption here does not distinguish between domestic and imported goods. As usual, scale economies are modelled by a fixed cost implying a decreasing average cost curve. Firms behave as oligopolists and compete in a Cournot-Nash way. In principle, the pricing rule in (31) also applies here. Furthermore, as in the GTAP model, firms rule out the Ford effect. However, they do take into account their market power, that is the influence they may exert on the sectoral composite price index. Bchir et al. (2002) advocate a binary approach for the number of firms in the model. The binary approach describes the number of firms constant and profits variable in the short run, while assuming free entry and exit and zero profits in the long run.

The GTAP 5 database (see Dimanaran and MacDougall, 2002) is used as the major dataset. Although some parameters in the model can be freely chosen, several others have to be calibrated, i.e. the elasticity of substitution, economies of scale and competition intensity. The Armington elasticities are drawn from the GTAP 5 database, and are assumed to be identical across regions. The other elasticities used in the nesting for a given sector are linked to the Armington elasticity by subtracting one and dividing/multiplying by √2 when moving upward/downward and then adding one. Economies of scale are linked to the mark-up ratio, for which estimates exist for industrial sectors (see Oliveira-Martins and Scarpetta, 1999), for service sectors (see Oliveira-Martins, Pilat and Scarpetta, 1996) and for both (see Roeger, 1995). Estimates for the number of firms are based on Davies and Lyons (1996). In addition, Bchir et al. (2002) argue that firms are not direct competitors to each other within a sector. Therefore, they postulate subsectors. When assuming that each sector has one type of competition, this problem does not exist, and estimates from Gasiorek, Smith and Venables (1992) can be used. Calibration issues are discusses in more detail in Annex 2.

Summarizing, imperfect competition and scale economies are modelled in a relatively standard fashion in MIRAGE. Scale economies are described by a decreasing average cost curve. Firms behave as oligopolists and choose their output strategically, taking into account their influence on the composite price index. Furthermore, product and regional differentiation are allowed for through nested DS-Armington preferences by a representative consumer. The number of firms is controlled for in a binary fashion.

3.3 ATHENA model

For sector-specific macro-economic analyses in the Dutch economy, CPB uses the "ATHENA" model. The model consists of an explicit input-output matrix for several branches of industry and institutional sectors. The basic design of the model follows the bottom-up approach, where macro-economic aggregates are derived by adding the sectoral outcomes. The specification of each individual sector is similar to that of conventional macro-econometric models. The production process requires the use of primary production factors and intermediate inputs provided by one domestic and one foreign supplier. Factor demand equations are driven by gross output and own real prices and price information is modelled as a mark-up on marginal costs. Consumption at the macro-level is largely driven by disposable income and consumer demand at the sectoral level follows from a two-stage-budgeting allocation procedure.

Concerning calibration, the elasticities in the ATHENA model are chosen rather arbitrary. Further, the number of firms are based on a Herfindahl index and CBS data. The inverse Herfindahl index represents the number of hypothetical symmetric, equal-sized firms which would result in the same concentration level as the actually observed level. Last, the mark-ups are determined by the CDRs based on a study by Oliveira-Martins, Pilat and Scarpetta (1996), as is the case in MIRAGE.

As argued by Neary (2000), most markets feature a small number of firms competing strategically with each other and with potential entrants. The game-theoretic equilibria that result are often non-unique and highly dependent on specific details of the market structure. A very high level of disaggregation is therefore required to capture these aspects of imperfect competition. Broer et al. (2003) avoid these complications in the ATHENA model by assuming that the market operates in monopolistic competition. In particular, they adopt the basic DS model, assume free and exit of firms and distinguish between love-of-variety and market power according to section 2.1.2. Obviously, this needs no further explanation. One particular difference with the above models is that ATHENA is a national model. Hence, no Armington specification is required to model imports.

3.4 LINKAGE model

The LINKAGE model is a global, multi-region, multi-sector, recursive dynamic computable general equilibrium model, see Mensbrugghe (2003). It is a neo-classical model with both factor and goods market clearing. Trade is modelled using nested Armington and production transformation structures to determine bilateral trade flows.

All sectors are assumed to operate under cost minimization. The model allows for increasing returns to scale using fixed production costs, which are represented by some fixed combination of capital and labour. Currently, the model assumes a fixed mark-up, uniform for each region, and endogenous profits. So in fact, the LINKAGE model follows a standard DS specification with a fixed number of firms. In a future version the model will allow for free entry and exit of firms. Marginal costs are modelled by a series of nested CES production functions, that are intended to represent the different substitution and complementary relations across the various inputs in each sector. In particular, Mensbrugghe (2003) advocates the use of a sequence of CES functions, because it is easy to implement and understand. An alternative might be to use so-called flexible functional forms, which in some sense can be thought of as a functional (Taylor) approximation of the true technological relations and is calibrated to a given set of own and cross price elasticities. Technology is specified the same as in the GTAP model.

All income generated by economic activity is assumed to be distributed to consumers. A single representative consumer allocates optimally his disposable income among commodities and saving.

Imports originating in different regions are imperfect substitutes, i.e. imports are specified based on the Armington assumption. The Armington specification is implemented using two CES nests. At the top nest, domestic agents choose the optimal combination of the domestic good and an aggregate import good consistent with the agent's preference function. At the second nest, agents optimally allocate demand for the aggregate import good across the range of trading partners.

The bilateral supply of exports is specified in parallel fashion using a nesting of constant-elasticity-of-transformation (CET) functions. At the top nest, domestic suppliers optimally allocate aggregate supply across the domestic market and the aggregate export market. At the second nest, aggregate export supply is optimally allocated across each trading region as a function of relative prices.

The LINKAGE model is mainly based on the GTAP 5 version 4 database. This version has a 1997 base year. Calibration in the model occurs in a separate routine. Unfortunately, this procedure is not explicitly discussed in Mensbrugghe (2003).

In short, scale economies are modelled through a standard decreasing average cost curve and imperfect competition is again modelled in a nested Armington-DS fashion as in MIRAGE. In addition, the model allows for short-run as well as long-run Armington elasticities, i.e. it introduces the possibility of more flexible long-term responses to changes in import prices. Technology is modelled in a nested sequence of CES functions along the lines of the GTAP model. Furthermore, the number of firms is fixed and profits may vary.

3.5 Michigan model

The Michigan Model of World Production and Trade is a computable general equilibrium (CGE) modelling framework originally developed starting in the mid 1970s by Alan Deardorff and Robert Stern at the University of Michigan[15]. The model was used for a variety of purposes in addition to the Tokyo Round, including analyzing the effects of exchange rate changes, the structure of protection, and scenarios of trade liberalization leading up to the Uruguay Round.

In the late 1980's, Robert Stern collaborated with Drusilla Brown of Tufts University to construct a model of the United States and Canada for the purpose of analyzing the effects of the US-Canada Free Trade Agreement. They started with the structure of the Michigan Model, but extended its equations to include features of the New Trade Theory: imperfect competition, increasing returns to scale, and product differentiation. Shortly after, joined again by Alan Deardorff, they expanded this model to include first four and then eight countries and country groups that could be selected in different combinations from the 34-country Michigan Model database. This model, which we now call the Michigan Brown-Deardorff-Stern (BDS) Model, retains the features of the new trade theory introduced by Brown and Stern.

Turning to the market structure, some sectors are modelled as perfectly competitive but most sectors as monopolistically competitive with free entry and exit of firms. Firms charge a mark-up which is identical for all regions. In general, a standard DS specification is used to accommodate imperfect competition and scale economies. In some versions of the Michigan model, to incorporate multinational firms, nested CES functions are used to differentiate between domestic and import demand.

Consumers and producers are assumed to use a two-stage procedure to allocate expenditure across differentiated products. In the first stage, expenditure is allocated across goods without regard to the country of origin or producing firm. At this stage, the utility function is Cobb-Douglas, and the production function requires intermediate inputs in fixed proportions. In the second stage, expenditure on monopolistically competitive goods is allocated across the competing varieties supplied by each firm from all countries. In the case of sectors that are perfectly competitive, since individual firm supply is indeterminate, expenditure is allocated over each country’s industry as a whole, with imperfect substitution between products of different countries. The aggregation function in the second stage is a CES function.

The production function is separated into two stages. In the first stage, intermediate inputs and a primary composite of capital and labour are used in fixed proportion to output. In the second stage, capital and labour are combined through a CES function to form the primary composite. In the monopolistically competitive sectors, additional fixed inputs of capital and labour are required. It is assumed that fixed capital and fixed labour are used in the same proportion as variable capital and variable labour so that production functions are homothetic.

[15] For a detailed description and history of the model, see http://www.fordschool.umich.edu/rsie/model.

It is unclear how calibration is treated in the Michigan model. On their website they say that their section on the data is “seriously incomplete”. Nevertheless, they do say that elasticities are reproduced from Deardoff and Stern (1990). However, they do not describe how.

3.6 Willenbockel

Willenbockel (1994) indicated some unsatisfactory features of previous economic evaluations of the EC internal market completion programme and suggested a modified Harris-type multi-sectoral general equilibrium framework. This framework allows for the presence of imperfectly competitive markets, intra-industry product differentiation and economies of scale.

As in case of the previous discussed models, Willenbockel considers both perfect and imperfect competitive sectors. Technology are those of constant returns to scale and decreasing average cost curves and applies the same production tree as in the GTAP model. Furthermore, he considers various types of non-cooperative firm behaviour: Betrand price-setting, Cournot quantity-setting and Chamberlinian price setting.

At the demand side of the model, utility is that of a representative consumer, and follows a three-stage-budgeting approach. At the top CES nest, the representative consumer in the economy, draws utility from sectoral composite goods. The lower nest is characterized by variety-scaling as in (32). Hence, the number of firms is endogenous and indirect price index effects are taken care off.

The CGE model is calibrated to a consolidated dataset, which reflects the levels and structure of UK production, demand and trade flows for the benchmark period 1985. The number of symmetric domestic firms in imperfect competitive industries is determined by the reciprocals of *Herfindahl concentration indices* derived on the basis of the observed 1985 size distribution of firms by industry provided in Department of Trade and Industry - Business Statistics Office (1988). Just like in the GTAP model, also here, Pratten (1988) is used for estimates of the CDRs. Elasticities are chosen rather arbitrary: between 0 and 1.5 for service and primary sectors and 2.5 for manufacturing sectors.

Concluding, the general equilibrium model developed by Willenbockel is closely related to the extended GTAP model, discussed in section 3.1.

3.7 Summary

The models in section 2 are mainly theoretical, but provide the foundation for applied models. In this section, I discussed several general equilibrium models of the last five to ten years. Almost all models are multi-region models. Therefore, the specifications used in these models stem from the international trade literature.

The main characteristics are summarized in Table 3.1. We have seen that all models describe increasing returns technologies by decreasing average cost curves, assume homotheticity and follow a nested sequence of CES functions including primary and intermediate inputs. For all models we have increasing returns at the sectoral level. The ATHENA model can also distinguish between returns to diversity and ease of substitution. Because of this property, ATHENA has the ability to turn the love-of-variety effect off and have constant returns to scale at the sectoral level, while maintaining mark-up pricing due to fixed costs.

Furthermore, firms generally behave as oligopolists à la Cournot. However, monopolistic competition is considered as a border case of oligopoly, where the number of firms becomes relatively large. In the oligopoly pricing rule, this means that firms conjectures, Ω/n, equal one. Further, most models do not allow price discrimination amongst regions. In general, full price discrimination leads to numerical complexity, because of the following large number of elasticities of demand.

Consumer demand is preferably described as a blend of Armington-type and Dixit-Stiglitz preferences. In this way, goods are both differentiated with respect to origin and variety. There are a number of possible specifications:

- The Harris-type describes demand in perfect competitive industries by Armington preferences and imperfect competitive industries by DS preferences. This yields a certain asymmetry in the preference weights.
- Another specification is the one used in MIRAGE, where domestic demand is of the DS type and import demand is of the Armington type.
- Yet another specification is variety-scaling as applied in the GTAP model and Willenbockel (1994). Variety-scaling imposes symmetry on the cost functions within a region. This means that regional firms produce the same quantity and charge the same price. Furthermore, Armington and DS preferences are not nested as in MIRAGE.

Several extensions of these demand systems can be thought off. One of them is to allow for different Armington elasticities to be used for the short-run and long-run as in the LINKAGE model. This introduces the possibility of more flexible long-term responses to changes in import prices. Except for the ATHENA model, all models take into account the indirect effect of a change in price on the sectoral price index. However, they all ignore Ford effects.

Concerning calibration, the number of firms is often determined by a Herfindahl index. The degree of market power is generally determined by estimated mark-up ratios from studies of Oliveira-Martins and Scarpetta (1999), Oliveira-Martins, Pilat and Scarpetta (1996) and Roeger (1995) or cost-disadvantage-ratios from Pratten (1988). Elasticities are often chosen arbitrary or proportionate to some given input-output matrix.

In the next chapter I give a recommendation for modelling scale economies and imperfect competition in WorldScan based on the knowledge of the two previous chapters and the objectives at hand. Moreover, I run some simulations to analyze the effects of introducing a new technology and market structure.

Table 3.1 Applied Model Characteristics

	Increasing returns to scale?	Number of firms	Technology	Demand structure	Indirect effects taken into account?	Price discrimination	Imperfect competition	Relevant extensions
GTAP	sectoral level	fixed / endogenous	nested CES	Armington - DS	price index	no regional price discrimination	oligopoly / monopolistic competition	variety scaling
MIRAGE	sectoral level	short run: fixed long run: endogenous	nested CES	Armington - DS	price index	regional price discrimination	Cournot	no
ATHENA	firm level	endogenous	nested CES	DS	no	no price discrimination	monopolistic competition	market power ≠ love-of-variety
LINKAGE	sectoral level	fixed	nested CES	Armington - DS	price index	no regional price discrimination	monopolistic competition	short/long run elasticities
MICHIGAN	sectoral level	endogenous	nested CES	DS	price index	no regional price discrimination	monopolistic competition	no
WILLENBOCKEL	sectoral level	endogenous	nested CES	Armington - DS	price index	no regional price discrimination	oligopoly / monopolistic competition	variety scaling

4 Scale economies in WorldScan

4.1 Modelling approach

Previous subsections have shown that many possibilities to incorporate scale economies in WorldScan do exist. However, the objective of this memorandum is not finding the most extensive model specification for imperfect competition and scale economies, but finding one that can mainly identify and describe increasing returns. In order to quickly have an operational version of WorldScan, I follow a practical approach.

The precise model specification is discussed in section 4.1.1. Calibration issues are discussed in section 4.1.2. In short, the model adjustments are the following:

- *Dixit-Stiglitz - Armington demand specification*: we would like to model heterogeneity at the supply side of the model to represent the large trade in semi manufactured products, also to accommodate agglomeration effects. Unfortunately, it would take too much work to adjust the model in this way given time. For now, heterogeneity is modelled at the demand side of the model using a nested Dixit-Stiglitz - Armington structure.
- *Variety scaling*: uniform production technology for all firms in a certain sector for a certain region.
- *Free entry and exit of firms*: when profits are positive, this will attract new entrants to the market. When they are negative, firms will go inactive and exit the market. In the long run firms earn zero profit. Any profit is still counted as income such that the option for an exogenous number of firms can be added easily at a later time. Nevertheless, an exogenous number of firms has the odd property that some firms still produce while incurring losses.
- *Monopolistic competition*: with indirect effects of a price change on the price index and income ignored, since the number of firms is assumed to be sufficiently large. Strategic interaction between firms is not needed to show the effects of increasing returns, and thus modelling an oligopolistic market structure is not needed. Monopolistic competition is easier to model.
- *Distinguish between love-of-variety and ease of substitution*: it has been argued that these are two distinct phenomena. Therefore we would like to model them separately. This is easily accomplished. Furthermore, in this way, the welfare effects of the number of varieties can be ‘turned off’.
- *Mill-pricing*: firms demand a fixed producer price independent of the region of destination.

Summarizing, we could say that the approach above is a mix of the imperfect competitive model aspects from ATHENA and MIRAGE. Although the model could be extended in many ways, the current approach is sufficient to meet our demands. The following two subsections discuss the model and calibration of the parameters.

4.1.1 The Model

Let us define the following variables and indices:

p = price	v = firm level
x = quantity	0 = raw / semi-manufacturing level
n = number of firms	i = firm i
ε = elasticity of demand	j = sector of production
f = fixed costs	r = region of production
τ = tariff and transportation costs	s = region of consumption

Consider a composite input, x_{jr}, that is a CES-aggregate of n_{jr} different varieties x^v_{jir} originating from a specific industry

$$x_{jr} = n_{jr}^{\theta_j - \varepsilon_j/(\varepsilon_j - 1)} \left\{ \sum_{i=1}^{n_{jr}} x^v_{jir}{}^{(\varepsilon_j - 1)/\varepsilon_j} \right\}^{\varepsilon_j/(\varepsilon_j - 1)}, \quad \text{with } \varepsilon_j > 1. \tag{35}$$

Note that, we distinguish between love-of-variety and ease of substitution by including the additional parameter θ_j, which controls for love-of-variety; $\theta_j = 1$ and $\theta_j = \varepsilon_j/(\varepsilon_j - 1)$ correspond to the extreme cases 'no love-of-variety' and 'maximum love-of-variety'.

Customers buy the brand aggregate (35) at minimal costs. Hence, they minimize unit costs of (35), p_{jr}, with respect to individual brands, x^v_{jir}, with prices, p^v_{jir}:

$$\min_{\{x^v_{jir}\}} \quad p_{jr} = \sum_{i=1}^{n_{jr}} p^v_{jir} x^v_{jir} \quad \text{s.t.} \quad n_{jr}^{\theta_j - \varepsilon_j/(\varepsilon_j - 1)} \left\{ \sum_{i=1}^{n_{jr}} x^v_{jir}{}^{(\varepsilon_j - 1)/\varepsilon_j} \right\}^{\varepsilon_j/(\varepsilon_j - 1)} = 1. \tag{36}$$

It is straightforward to show that the resulting demand for variety i with corresponding price index, p_{jr}, is equal to

$$x^v_{jir} = \left(\frac{n_{jr}{}^{1-\theta_j} p^v_{jir}}{p_{jr}} \right)^{-\varepsilon_j} n_{jr}{}^{-\theta_j} x_{jr} \quad \text{and} \quad p_{jr} = n_{jr}^{\varepsilon_j/(\varepsilon_j - 1) - \theta_j} \left(\sum_{i=1}^{n_{jr}} p^v_{jir}{}^{1-\varepsilon_j} \right)^{1/(1-\varepsilon_j)}, \tag{37}$$

where the brand aggregate, x_{jr}, is determined elsewhere in the model. Let us assume firms are identical. Hence, firms produce the same quantity, x^v_{jr}, and charge the same price, p^v_{jr}. Therefore we drop index i. The result is that (37) greatly simplifies to (38).

$$x_{jr} = n_{jr}{}^{\theta_j} x^v_{jr} \quad \text{and} \quad p_{jr} = n_{jr}{}^{1-\theta_j} p^v_{jr} \tag{38}$$

Let us assume that firms within the industry produce a so called 'raw output', which has the interpretation of a semi-manufactured output, $x^{v,0}_{jr}$, under constant-returns-to-scale, using composite inputs and factor services at a unit cost, p^0_{jr}. Note that, sectoral raw output is equal to firm raw output multiplied by the number of firms: $x^0_{jr} = x^{v,0}_{jr} n_{jr}$. Next, firms use $f^v_{jr} + x^v_{jr}$ units of this raw output to produce x^v_{jr} units of a single variety. In other words, raw output denotes what all firms produce in terms of costs. This output is not sold. The raw output is an input for a very simple production process, where a part is used as a set-up cost and the rest is sold on the market. The part that is sold is obviously less than raw output in terms of quantity. However, in terms of value, these are exactly equal because the quantity on the market is sold at a producer price, p^0_{jr}, plus a mark-up, $p^0_{jr}/(\varepsilon_j - 1)$ that covers the raw output that is used as set-up cost. Summarizing, firms have identical technology

$$x^v_{jr} = \begin{cases} x^{v,0}_{jr} - f^v_{jr}, & \text{if } x^{v,0}_{jr} \geq f^v_{jr}, \\ 0 \quad , & \text{if } x^{v,0}_{jr} < f^v_{jr}. \end{cases} \tag{39}$$

It can be shown, that if n is large, the demand elasticities are approximately fixed and equal to -ε_j for all varieties in sector j. Assuming a large number of firms, a monopolistic firm can ignore strategic behaviour by competitors. Each firm determines the price of its brand by maximizing profits

$$\pi^v_{jr} = p^v_{jr} x^v_{jr} - p^0_{jr} x^{v,0}_{jr} = p^v_{jr} x^v_{jr} - p^0_{jr}\left(x^v_{jr} + f^v_{jr}\right) \tag{40}$$

yielding

$$p^v_{jr}\left(1 + \frac{x^v_{jr}}{p^v_{jr}} \frac{\partial p^v_{jr}}{\partial x^v_{jr}}\right) = p^v_{jr}\left(1 - \frac{1}{\varepsilon_j}\right) = p^0_{jr} \quad \Leftrightarrow \quad p^v_{jr} = \frac{\varepsilon_j}{\varepsilon_j - 1} p^0_{jr}\,. \tag{41}$$

Thus, firms set a price equal to marginal costs plus a mark-up.

Let us assume firms can enter and exit an industry freely. Then, as long as profits are positive, firms offering new varieties will enter the market. On the other hand, if profits are negative, firms incur a loss, and exit the market. In the long run, these dynamics drive profits to zero. Hence, in equilibrium price equals average costs

$$\pi_{jr}^{v} = p_{jr}^{v} x_{jr}^{v} - p_{jr}^{0} x_{jr}^{v,0} = p_{jr}^{v} x_{jr}^{v} - p_{jr}^{0}\left(x_{jr}^{v} + f_{jr}^{v}\right) = 0 \quad \Leftrightarrow$$
$$\Leftrightarrow \quad p_{jr}^{v} = p_{jr}^{0}\left(1 + f_{jr}^{v} / x_{jr}^{v}\right). \qquad (42)$$

Substituting (41) in (42) yields the equilibrium firm output level

$$x_{jr}^{v} = f_{jr}^{v}\left(\varepsilon_{j} - 1\right). \qquad (43)$$

Because firms have identical technology, the number of firms in the industry can be solved for as $n_{jr}x^{v}_{jr} = x^{0}_{jr} - n_{jr}f^{v}_{jr}$:

$$n_{jr} = \frac{x_{jr}^{0}}{x_{jr}^{v} + f_{jr}^{v}} = \frac{x_{jr}^{0}}{f_{jr}^{v}\left(\varepsilon_{j} - 1\right) + f_{jr}} = \frac{x_{jr}^{0}}{f_{jr}^{v}\varepsilon_{j}}. \qquad (44)$$

This solves the imperfect competitive model. Summarizing, the first order conditions of the model are

firm level: → *number of firms:* → *industry level*:

firm level:	*number of firms:*	*industry level*:
$p_{jr}^{v} = \frac{\varepsilon_{j}}{\varepsilon_{j} - 1} p_{jr}^{0}$ $x_{jr}^{v} = f_{jr}^{v}\left(\varepsilon_{j} - 1\right)$	$n_{jr} = \frac{x_{jr}^{0}}{f_{jr}^{v}\varepsilon_{j}}$	$p_{jr} = n_{jr}^{1-\theta_{j}} p_{jr}^{v}$ $x_{jr} = n_{jr}^{\theta_{j}} x_{jr}^{v}$

This is combined with the Armington specification as follows. Prices in region of consumption *s* of products produced in region *r* are equal to

$$p_{jrs} = p_{jr}\left(1 + \tau_{jrs}\right), \qquad (45)$$

where *p* times τ represent the tariff and transportation costs. Consequently, the Armington specification follows via

$$p_{js} = \left\{\sum_{r} \alpha_{jrs} p_{jrs}^{1-\sigma_{j}}\right\}^{1/\left(1-\sigma_{j}\right)}. \qquad (46)$$

Two important remarks should be made here. Firstly, note that individual varieties are transported and not the aggregate of all brands together. The increasing-returns-to-scale effects occur in the country where varieties are consumed and the transport costs should not be affected by the love-of-variety of consumers in the destination country. Secondly, the Armington elasticity in (46) should be smaller than the DS elasticity: $\sigma_j < \varepsilon_j$, for the model to have an interior solution. Ignoring this restriction may lead to corner solutions or numerical problems when solving the model.

4.1.2 Calibration

Issues concerning calibration are heavily discussed in Annex 1. Calibration for this imperfect competitive model is particularly easy. Only two parameters have to be calibrated. These are the sectoral elasticities of demand and the fixed costs[16].

Elasticities are given by rewriting estimated mark-ups in Oliveira-Martins, Pilat, Scarpetta (1996a, 1996b). If these elasticities are lower than the Armington elasticities they are set slightly higher than the relevant Armington elasticity.

Fixed costs are determined by (44), where raw output is given by GTAP data and the number of firms is set equal to one for the EU. For the other regions it is determined in such a way, that fixed costs are equal over regions. Basically, we do not know anything about differences between regions, so therefore we equalize them. Using Herfindahl indices to calibrate the number of firms is also possible. However, Eurostat data show that the sectoral numbers of firms can be very large. This causes numerical problems in the calibration procedure not to mention the peculiar side effects, see Erp, et al. (2001). Therefore we scale the number of firms. This alters the interpretation of the variable *n*, though. Now, *n* is interpreted as an index rather as the absolute number of firms. The same holds for the fixed costs and brand quantity, since these variables change accordingly. However, a time series of these three variables still shows the relative change.

4.2 Simulations

Before we implement the imperfect competitive modelling approach chosen in the previous section in WorldScan, it may be useful to analyze the effects in an aggregated version. This will be the objective of this section.

This section is structured as follows. Firstly, I analyze the effects, when increasing returns to scale at firm level are introduced compared to a base run with constant returns to scale[17]. Here, I

[16] The love-of-variety parameter, θ, is not calibrated in the model. We just set it equal to the mark-up for pure love-of-variety or equal to 0 to turn it off. It might be interesting to also consider other values to change the degree of the economies of scale. However, as far as I know, there are no empirical studies, yet, which estimate the love-of-variety by consumers. Therefore, I stick with the two cases of pure and no love-of-variety.

allow for the love-of-variety effect, hence we set $\theta_j = \varepsilon_j/(\varepsilon_j - 1)$ in (35) and follow the basic DS model. Therefore, we also have increasing returns at the sectoral level.

Secondly, I turn the love-of-variety effect off, by setting $\theta_j = 1$ in (35). In this scenario, we have constant returns to scale at the sectoral level, while maintaining mark-up pricing due to fixed costs.

And finally, I simulate a liberalisation variant by gradually reducing import tariffs over time until they are non-existent. Here, it is interesting to see where specialisation patterns occur and how these differ between constant and increasing returns to scale technologies.

This aggregated version of WorldScan is exactly the same as the 'core-version'[18] of WorldScan. However, here we consider only five sectors, i.e. agriculture, energy, manufacturing, services and transport and three regions, i.e. EU, rest of OECD and rest of the world. Calibration is done for 2001 and I simulate yearly until 2040. The analysis focuses on the long run effects, say the effects in 2040.

4.2.1 Standard Dixit-Stiglitz: pure love-of-variety

To profoundly understand the workings of increasing returns, I go step by step with introducing scale economies. Firstly I simulate four scenarios in which one sector features increasing returns in the EU. The agricultural sector features constant returns, because this is more realistic. Changing the production technology for one particular sector in one particular region is the smallest possible variation on the case of constant returns. Hence, in this way we are able to analyze the direct effects of introducing increasing returns. Furthermore, we can analyze the differences between sectors.

Secondly, I simulate increasing returns in a single sector for different regions. Manufacturing is chosen arbitrarily here. Besides the EU I also consider increasing returns in the rest of the OECD, all OECD regions together and OECD and non-OECD regions together to look at regional effects. Note that I do not differentiate mark-ups between regions. Still, regional differentiation is possible.

Thirdly, I look at a variant where all sectors but agriculture and all regions feature increasing returns. This variant corresponds most with reality and is thus the one that will be most likely implemented in WorldScan.

Fourthly, I look at the effects of halving the elasticity of demand in the manufacturing sector in the EU in the one-sector-increasing-returns case. Fixed costs are calibrated once more. This test serves as a sensitivity analysis to the parameters.

---

[17] We simulate constant returns to scale with the DS model, by using a demand elasticity equal to ten billion. Differences with true constant returns to scale are approximately zero. Therefore, DS technology with demand elasticities equal to ten billion is a good proxy for constant returns to scale.

[18] The core version of WorldScan excludes research and development and does not distinguish between short and long term Armington elasticities.

Finally, it is interesting to see what the effects of giving firms more market power will be, by again halving the elasticity of demand but now holding fixed costs fixed.

The simulations will be discussed in turn. I begin with the observations about the simulations of the one-sector-increasing-returns case. All results are summarized in Annex 2.

Single region, single sector case

In the constant returns case we see that the number of firms is not equal to one as you initially would expect. A time path of the simulation with constant returns shows that the number of firms is equal to one in the base year, 2001. From 2001 onwards the number of firms increases gradually to a value of 1.79. Because firm output is fixed, raw output increases. This rise is caused by lower input prices, which again are the result of technological progress and increased labour supply.

Because the mark-up approaches zero in the constant-returns-case, firm price is equal to sectoral price and in turn is equal to marginal costs.

Increasing returns are introduced by lowering the elasticity to a value of 7.5 yielding higher fixed costs. On the one hand, firms lower their producer price (– 3%), because they can produce more efficiently due to increasing returns. On the other hand, they tend to produce more because they have to cover their fixed costs (+ 12%). Higher production leads to a higher demand for production factors and therefore to higher factor prices (+ 22%).

The effect on the number of firms is determined by respectively firm output (0%), fixed costs (813) and raw output[19] (+ 18%). Notice that total production (read output plus fixed cost) of the firm increases by (813 + 5285) / 5285 – 1 ≈ + 16%. The difference in growth of raw output and growth of firm production is growth in the number of firms, + 2%.

The manufacturing sector not only demands more production factors but also more intermediate inputs from the other sectors. These sectors react by increasing their producer prices and de facto a decreasing raw producer price. De facto, the aggregate raw producer price falls (– 8%). Because of the mark-up (7.5 / 6.5 ≈ 1.15) the firm price increases by about 6%. Essentially, the message is that increasing returns affect producer prices: it lowers the input price of the “increasing-returns-sector” and increases all other producer prices, especially factor prices.

The third column considers a further decrease in the elasticity of demand. Hence mark-ups over marginal costs increase and all effects on producer prices are even stronger than before. Here, we see that the number of firms does not increase much, although mark-ups have increased dramatically. The reason is that fixed costs have more than doubled while raw output increased by 40%.

[19] From now on, when we talk about raw output, we talk about raw output at the sectoral level, hence x^0_{jr} in (44).

Table 4.1 Simulation results for manufacturing sector in EU

Year 2040	constant returns	increasing returns in manufacturing EU		
			half ε	half ε, keep μ fixed
elasticity of demand	1E+10	7.5	3.75	3.75
fixed cost per firm	5.3E-7	813	1922	813
number of firms	1.79	+ 2%	+ 3%	+ 143%
firm output	5285	0%	0%	– 58%
composite output	9486	+ 12%	+ 29%	+ 76%
raw output	9486	+ 18%	+ 40%	+ 40%
firm output price	0.71	+ 6%	+ 15%	+ 15%
composite price	0.71	– 3%	– 8%	– 32%
raw producer price	0.71	– 8%	– 15%	– 15%
factor price	1.00	+ 22%	+ 59%	+ 59%
producer price agricul.	0.67	+ 6%	+ 15%	+ 15%
producer price energy	1.53	+ 5%	+ 13%	+ 13%
producer price manuf.	0.71	– 3%	– 6%	– 6%
producer price services	1.22	+ 6%	+ 14%	+ 14%
producer price transport	0.81	+ 8%	+ 19%	+ 19%

Note: composite variables are CES aggregates (including the love-of-variety effect). Raw output denotes semi-manufactured sectoral output of a constant-returns-to-scale production process. Raw output is equal to firm output plus the set-up cost and the outcome of that multiplied with the number of firms in the sector.

The last column shows the results of a test which examines what happens when firms set higher mark-ups over marginal costs holding fixed costs in the sector at the same level (compared to column 2) Intuitively, firms get more market power. The number of firms increases dramatically (+ 143%), because new firms enter the market to profit from the relatively high mark-ups. Consequently, firms individually produce less (– 58%) since demand has not increased that much (+ 40% – + 18% = + 22%). Because of the high mark-up raw producer prices are reduced further (– 15% – – 8% = – 7%) and production is increased leading to higher input prices.

Alternatively, column 4 can be seen as a test of lowering the fixed costs holding the elasticity of demand at the same level (compared to the third column). Now, only two variables change in value. Lowering fixed costs in the sector attracts entering firms leading to a dramatic increase in firms (+ 143%). The elasticity of demand is the same as before, thus, demand has not changed. Therefore, output per firm falls to compensate for the increase in firms. Furthermore, composite output and price also change because of the change in the number of firms. The love-of-variety effect becomes stronger.

Concluding, an increase in the demand elasticity changes the demand structure and therefore the cost structure. Effects of increasing returns get stronger, because firms get more market power. A decrease in fixed costs leads to an increase in the number of firms and reduces firm output implying economies of scale.

Because of trade there are spillovers to other regions. However, in general these effects are small.

Single region, multi sector case

Since demand elasticities are much lower in the services and transport sector, mark-ups are much higher. Hence, firms have more market power. Fixed costs for the energy and transport sector are much smaller, respectively 62 and 257, while for the services sector much larger, 2780, compared to the manufacturing sector, see Table A1 in Annex 2. These differences in fixed costs determine for the most part the differences for the other variables. The energy and transport sector attract more firms while the services sector attracts less firms relative to the manufacturing sector because firms in the energy and transport sector have far lower set-up costs than in the services sector. Because firm output does not change, the same effects hold for sectoral raw output. The energy and transports sector also have the highest increase in composite output. Hence, although firms in the services sector set the same mark-ups as the transport sector and far higher mark-ups compared to the energy sector, the difference in fixed cost plays a far more important role in the love-of-variety effect (scaling firm output with $n^{\varepsilon/(\varepsilon-1)}$), since the number of firms increases dramatically. This, we already noticed in Table 4.1.

Furthermore, notice that the effects in the energy sector and the transport sector do not differ much from each other. One part of the explanation is the difference in fixed cost, i.e. for the transport sector 257 and for the energy sector only 62. The other part is a difference in cost shares in the raw producer price for different sectors. Indeed, the cost shares are highest for the energy sector. Sometimes this cost share is twice as large as for other sectors. Hence, the energy price is not only important for the raw output price in the energy sector itself, but also very important for other sectors.

Multi region, single sector case

There is no differentiation between regions in elasticities and fixed costs. Introducing increasing returns in all regions shows that the effects are largest for the rest of the world. For example factor prices are several times larger than in other regions, see column “manufac. all” in Table A7 in Annex 2. The main reason is that economic growth in the rest of the world is on average three times as large as for the EU and rest of OECD. In fact, when we target economic growth at a 1% growth rate for all regions, variation between regions gets significantly smaller.

Although raw output prices fall and firms receive more market power, this does not lead to entering firms to the market. In fact, when increasing returns is also introduced in the rest of the world the number of firms decreases for all three increasing returns regions for the manufacturing sector, see column “manufac. all” in Table A10 in Annex 1. Why?

The number of firms depends on three things, raw output, firm output and fixed costs. Since firm output is constant for all variants, only raw output and fixed costs matter. Therefore, it has to hold that raw output does not increase enough (or even decreases) to compensate the increase in firm production caused by the fixed costs. What is the underlying process?

Because of fixed costs, firms have to produce more to ensure zero profit. Because of this, firms demand more inputs, i.e. production factors and intermediate inputs. Higher demand leads to higher prices. Therefore, inputs become more expensive. As a result, the raw output price may even increase due to increasing returns, because other inputs besides your own become more expensive. Because firms have to make zero profit, and face high input prices, firms cannot produce enough to compensate for the fixed costs and exit the market until firms in the market make zero profit.

Thus, when increasing returns is introduced in the rest of the world, demand for inputs increases dramatically, because it is such a large region. Moreover, the rest of the world has an economic growth rate which is three times larger than that of the EU and rest of OECD. Hence, other sector's producer prices and especially factor prices shoot up causing losses for firms. This leads to exiting firms until the market is in equilibrium.

Multi region, multi sector case

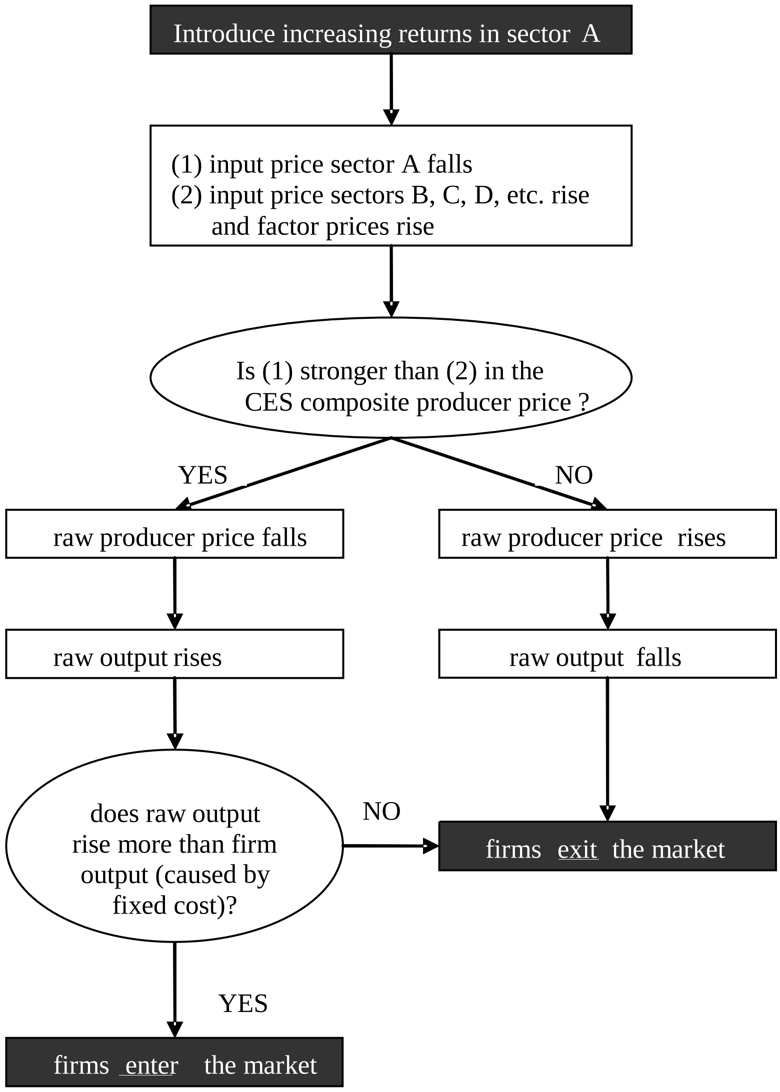
In this variant where increasing returns is introduced everywhere where possible, all previously discussed effects hold. In particular, the underlying processes are the same.

The fall in the number of firms as discussed in the previous paragraph also holds here. In fact, the numbers are more extreme. In all sectors for all regions the number of firms fall as a result of increasing returns (see column "all sectors" in Table A10 in Annex 2). Furthermore, all factor prices shoot up dramatically with tens sometimes hundreds of percentages (see same column in Table A7 in Annex 2). The effects on raw producer prices vary across sectors. They increase for agriculture, energy and manufacturing. It appears the mark-up over marginal cost is not enough to compensate for the high input prices next to the fixed costs. The other sectors show a decrease in the raw producer prices. The mark-ups for those sectors are much higher and can compensate the high input prices.

Summary

The main effects of the introduction of increasing returns appear in the cost structure: "own input" price decreases and "other intermediate input" prices and factor prices increase. Depending on which effect is stronger, this leads to entering/exiting firms to/from the market. Figure 4 illustrates what happens when increasing returns are introduced. In general, sectors with large market power, i.e. demanding a high mark-up, have firms entering the market. Sectors with small market power, i.e. demanding a small mark-up, have firms exiting the market. Increasing returns in many sectors and/or regions can cause a drain on production factors and intermediate inputs and therefore an outflow of firms in all sectors and regions. The degree of economies of scale is influenced by the mark-up and the fixed cost. In particular, halving the fixed cost induces relatively more increasing returns than halving the demand elasticity. Concluding, the effects of introducing increasing returns to WorldScan are significant!

Figure 4 Introducing increasing returns



4.2.2 Different degrees of love-of-variety

This section addresses the love-of-variety effect. In particular, how does tuning love-of-variety affect the results obtained from simulations with the basic DS model? Unfortunately, there exist no empirical studies that estimate the love-of-variety parameter. Therefore, as a sensitivity exercise, here we consider the extreme case of no love-of-variety.

From theory we then know the following (see sections 2.1.2 and 4.1.1). The composite brands price is equal to the individual brand price and the brands output aggregate is equal to the number of firms in the sector multiplied by the firm output. So, in fact we have (38), with $\theta_j = 1$. In the basic model, $\theta_j = \varepsilon_j/(\varepsilon_j - 1)$, the love-of-variety effect implies a more than proportional increase in the composite output when firm output increases, see (38). Inserting (41), (43) and (44) in (38) yields, dropping indices *j* and *r*,

$$\hat{x} = \frac{\varepsilon - 1}{\varepsilon}\left(f^{v}\varepsilon\right)^{1-\theta}\left(\hat{x}^{0}\right)^{\theta} \quad \text{and} \quad \hat{p} = \frac{\varepsilon}{\varepsilon - 1}\left(f^{v}\varepsilon\right)^{\theta-1}\left(\hat{x}^{0}\right)^{1-\theta}\hat{p}^{0}. \tag{47}$$

A '^' denotes the value of a certain variable in equilibrium. When we turn the love-of-variety on and off, we can see what happens in (47). Table 4.2 presents the results.

Table 4.2 Equilibria with and without love-of-variaty

	love-of-variety on: θ = ε/(ε–1)	love-of-variety off: θ = 1	love-of-variety effect
x^{e}	$\frac{\varepsilon-1}{\varepsilon}\left(f^{v}\varepsilon\right)^{-1/(\varepsilon-1)}\left(\hat{x}^{0}\right)^{\varepsilon/(\varepsilon-1)}$	$\frac{\varepsilon-1}{\varepsilon}\hat{x}^{0}$	$\hat{n}^{1/(\varepsilon-1)} = \left(\hat{x}^{0} / f^{v}\varepsilon\right)^{1/(\varepsilon-1)}$
p^{e}	$\frac{\varepsilon}{\varepsilon-1}\left(f^{v}\varepsilon\right)^{1/(\varepsilon-1)}\left(\hat{x}^{0}\right)^{-1/(\varepsilon-1)}\hat{p}^{0}$	$\frac{\varepsilon}{\varepsilon-1}\hat{p}^{0}$	$\hat{n}^{-1/(\varepsilon-1)} = \left(\hat{x}^{0} / f^{v}\varepsilon\right)^{-1/(\varepsilon-1)}$

Table 4 shows that, when love-of-variety is turned off, the equilibria simplify. The difference between raw output price/quantity and final output price/quantity is just the mark-up. What happens is that raw output falls with exactly the fixed set-up cost (= $x^{0,e}/\varepsilon$). Because firms earn zero profit in the long run, prices increase with a mark-up (= $p^{0,e}/(\varepsilon - 1)$). Where we first had increasing returns at the sectoral level affecting intermediate input and factor prices, we now have constant returns at the sectoral level, yet retaining monopolistic competition through mark-up pricing.

A simulation with $\theta = 1$ confirms these expectations. Compared to the base run, almost nothing changes. We notice that output quantities rise by a factor equal to the mark-up and raw output prices fall by a factor equal to the mark-up. All other variables, such as the number of firms, input prices, firm and sectoral output prices and quantities remain unchanged[20].

Concluding, the model without love-of-variety does not seem appropriate to describe the welfare effects of the introduction of the new services directive. Because in this way, the welfare effects of introducing mark-ups are non-existent, since we have constant returns at the sectoral level. In other words, this analysis shows that any welfare effects that are existent in the DS model come from love-of-variety and not from mark-up pricing! Actually, this confirms what we already know from section 2.1.2, in particular (17), where the effects on welfare are nullified when θ is set equal to 1.

[20] Although the raw output price falls with the mark-up, input prices stay the same. This is because the location parameters in the calibration have changed.

4.2.3 Liberalisation

As a final experiment, it is interesting to see how the imperfect competitive model reacts in liberalisation scenario. One would expect, when import tariffs are abolished, that sectors with increasing returns technologies profit relatively more than sectors featuring constant returns technologies.

In this section, I compare two simulations. One considers a reduction of import tariffs under constant returns and the other under increasing returns[21]. This reduction is done gradually to circumvent numerical problems with solving the model. The first year of reduction is 2005 and the last year is 2015. The import tariffs are fully abolished from 2015 onwards. Because we only consider the long term effects of liberalisation, it does not matter much for the results if tariff abolishment occurs instantly or gradually. Here, I follow the basic DS approach. Hence, I assume consumers prefer diversity in products.

The import tariffs are shown in table 4.3. Import tariffs for the services and transport sector are zero. Notice that the import tariffs are relatively largest for non-OECD countries. Furthermore, the agricultural sector shows the largest import tariffs. Therefore, one would expect for instance, larger exports to the agricultural sector in the rest of the world. Note that, we simulate with a very aggregated sector structure. For instance, manufacturing of food or other agricultural products is aggregated under manufacturing, which accounts for only a small part of total manufacturing. In that respect, this analysis does fully reflect the gains from liberalisation for third world countries.

Table 4.3 Import tariffs in percentages of import

	Agriculture			Energy			Manufacturing		
from - to	EU	rOECD	ROW	EU	rOECD	ROW	EU	rOECD	ROW
EU	0	8	11	0	1	7	0	3	11
rOECD	1	0	14	0	0	7	2	0	9
ROW	3	5	0	1	1	0	4	5	0

Source: GTAP database

Table 4.4 presents the simulation results for the two scenarios. The first thing we recognize is that raw output has increased most in both scenarios in the agricultural sector for the OECD countries and decreased for the non-OECD countries. Because OECD countries can now export more to non-OECD countries, they produce more. Furthermore, we see that the transport sector profits in all regions from liberalisation, although there were no tariff barriers in the first place.

[21] Here, all sectors except agriculture feature increasing returns. In particular, we assume demand elasticities as they are given in Table A1 in Annex 2.

It is reasonable that regions trade more after liberalisation, which leads to increased transports. Indeed, the services sector shows almost no changes.

Table 4.4 Long term macroeconomic effects of trade liberalisation

		CRTS		IRTS	
sector	region	net trade position	value added	net trade position	value added
agriculture	ROW	– 64	– 0.1%	– 316	– 4.2%
	EU	32	4.9%	109	12.0%
	rOECD	35	3.3%	219	13.7%
energy	ROW	– 115	– 1.5%	– 135	– 1.3%
	EU	39	5.1%	52	6.9%
	rOECD	70	4.6%	77	4.6%
manufact.	ROW	443	2.6%	551	5.5%
	EU	– 127	2.2%	– 132	– 0.4%
	rOECD	– 342	0.3%	-461	– 4.2%
services	ROW	38	0.2%	– 11	0.1%
	EU	– 40	0.1%	– 33	0.1%
	rOECD	3	0.1%	46	0.1%
transport	ROW	25	3.2%	2	3.1%
	EU	– 22	2.3%	– 14	2.7%
	rOECD	– 3	1.7%	12	2.4%

Source: WorldScan simulations. CRTS stands for Constant Returns To Scale and IRTS for Increasing Returns To Scale. Numbers are changes compared to the base run in the year 2040: liberalisation under crts compared to crts and liberalisation under irts compared to irts. EU, rOECD and ROW denote respectively European Union, rest of OECD and Rest Of World. All values are changes with respect to the baserun and measured in quantities.

If we compare value added over the two scenarios, we see that under increasing returns to scale for the agricultural sector in OECD countries produces relatively more and in non OECD countries relatively less. Hence, if we assume increasing returns, it is more efficient that the OECD countries produce even more than under constant returns and export that to the other countries. Notice that the agricultural sector has constant returns technology in both scenarios. Therefore, these changes have to originate from higher input demands from other sectors which do have increasing returns. To confirm these specialisation patterns, one could consider changes in the net trade position. Table 4.4 shows that OECD countries indeed have increased their export more than their import and non-OECD countries vice versa. Consequently, the EU and the rest of the OECD countries specialize in agriculture. In addition, the EU specializes in energy and withdraws means from the other sectors. The other OECD countries clearly withdraw from the manufacturing sector and spent their inputs in the other sectors. Finally, the rest of the world apparently specializes in manufacturing, which includes amongst others manufacturing of agricultural products.

Table 4.5 Consumption annual growth, 2040		
	CRTS	IRTS
Rest of the World	0.2%	0.7%
European Union	0.5%	0.5%
Rest of the OECD	0.2%	0.1%

Source: WorldScan simulations. CRTS stands for Constant Returns To Scale and IRTS for Increasing Returns To Scale. Numbers represent differences in annual growth of consumption between the baserun and the liberalisation variant.

As for the general welfare effects, we can look at annual growth of consumption. Table 4.5 shows that as a result of liberalisation consumption grows for all regions under constant as well as increasing returns. In general, the welfare effects are almost the same under constant and increasing returns to scale. However, the rest of the world does benefit significantly more assuming scale economies.

Concluding, specialisation patterns are the same in both scenarios. Nevertheless, specialisation is much stronger under increasing returns than under constant returns, because firms can exploit their technology better and reallocate resources efficiently.

5 Conclusions

The purpose of this memorandum is to find a suitable modelling approach to identify and describe scale economies which can be incorporated in WorldScan.

From section 1.1 we know that scale economies are commonly modelled through a fixed set-up cost in the production process. In this way, producers set a mark-up on marginal costs to at least cover their fixed costs. From section 1.2 we know that as soon as an agent does not take prices as given, we have imperfect competition. Therefore, scale economies introduce imperfect competition.

The second chapter reviews the literature on scale economies and especially imperfect competition in general equilibrium. Economic literature shows among others two main research areas in microeconomic theory. First, there is industrial organization, which includes *small group models* focusing on strategic interaction between agents. The second mainstream in microeconomic theory is monopolistic competition. Monopolistic competition particularly works with *large group models*, hence, where the number of firms is large and strategic interaction is absent. We conclude, that a large group model is more suitable in our case, since we do not need strategic interaction between firms to describe scale economies. In particular, the extended Dixit-Stiglitz model should accommodate our goals very well. The third chapter shows that in many currently applied models, the Dixit-Stiglitz specification for domestic demand is combined with the Armington specification for import demand. Hence, we follow this in our approach.

The expansion of WorldScan with the approach as discussed in section 4.1 appears to describe scale economies well as the simulations show. In particular, we understand that, first of all, increasing returns do not naturally lead to higher output prices. Rather, increasing returns affect the cost structure in the model due to additional demands for inputs. The brand price is the outcome of the producer price multiplied by the mark-up.

Secondly, producer prices can decrease as well as increase. This depends on which price effect is stronger, the fall of the 'own input' price (caused by efficiency in production) or the rise in all other input prices and factor prices (caused by additional input demand).

Thirdly, introducing increasing returns through a fixed cost does not necessarily lead to entering firms to the market. The number of firms depends on raw output and the set-up cost. If raw output does not increase enough to compensate for the set-up cost, firms even exit the market.

Fourthly, sectors with strong increasing returns technologies can demand high mark-ups on producer prices and consequently attract possible entrants to the sector. On the other hand, sectors with low mark-ups cause firms to exit the market, since they cannot cover their fixed cost.

Fifthly, the love-of-variety effect determines whether or not we have constant returns at a sectoral level. Assuming customers prefer diversity in products leads to increasing returns to scale at an aggregated level. Assuming consumers are indifferent leads to constant returns to scale at an aggregated level. In conclusion, any welfare effects in this model originate from love-of-variety of consumers.

Finally, a liberalisation exercise shows that specialisation patterns are stronger under increasing returns, because production is more efficient. In particular, producing more causes averages costs to drop.

Overall, we can conclude that the effects of this expansion of WorldScan can be called significant!

Annex 1: Calibration

Estimation of mark-ups

Over the past decades, a substantial body of literature has been devoted to the empirical identification of market power (e.g. Schmalensee, 1989; Bresnahan, 1989). This literature focused on the identification of monopoly pricing, i.e. whether there is evidence of pricing above marginal costs. In theory, it is possible to define the degree of monopoly power of a given producer as the mark-up over marginal cost normalized by the producer price, which is called the Lerner index. The empirical measurement of the Lerner index and related measures is quite difficult, particularly at an aggregate level. As a result, there have been few empirical studies identifying market power at the aggregate level (cf. Geroski, et al., 1995). The main problem arises from the fact that while prices can be measured, marginal costs are not directly observable. Therefore, indirect measures have to be developed. Hall (1986, 1988) proposed a methodology to estimate mark-ups by using the short-run fluctuations of output and production inputs by sector. This method has become popular and is widely applied in the empirical literature.

Hall's approach

The approach by Hall (1986, 1988) to the estimation of mark-ups is based on ideas on productivity measurement contained in Solow (1957). The marginal cost of a firm that uses capital and labour as inputs for technical progress, can be approximated by

$$MC = \frac{w\Delta L + r\Delta K}{\Delta Q - \theta Q}, \qquad \text{(A.1)}$$

where Q is real value added, L is labour, K is capital, w is the wage rate, r is the rental price of capital and θ is the rate of technical progress. In the denominator, the change in output is adjusted for the amount by which output would rise if there were no increase in the production inputs. Rewriting (A.1) and assuming that mark-ups are constant over time and the rate of technological progress is described by a random deviation from a constant rate yields the following:

$$\Delta(q_t - k_t) = \mu\alpha_t\Delta(l_t - k_t) + \theta + u_t, \qquad \text{(A.2)}$$

where lowercase variables indicate log-levels, μ is the mark-up, α is equal to wL/PQ and u_t is a random deviation[22,23]. Unfortunately, (A.2) cannot be estimated directly, since imperfect

[22] It is assumed that $wL + rK = PQ$, i.e. constant-returns-to-scale. As a results $rK/PQ = 1 - \alpha$, which simplifies (A.2). However, this implies a bias in case of economies or diseconomies of scale.

competition would imply a correlation between the labour/capital ratio and the productivity term. Usually, this can be solved by using the instrumental variable estimation approach. However, the instruments used for the labour/capital ratio have been criticised to be rather implausible. In addition, in small samples the advantage of using the instrumental variable approach is with respect to Ordinary Least Squares (OLS) is not clear. As argued by Caballero and Lyons (1989), a very small correlation between the instruments and productivity growth may prove more problematic than biases emerging from the OLS procedure. Concluding, Hall's approach is stuck with an identification problem.

Roeger's approach

Roeger (1995) proposes an alternative way of computing mark-ups. Here, it is convenient to make the relation between the Solow residual and the Lerner index (*B*) more explicit. By definition:

$$B = \frac{P - MC}{P} = 1 - \frac{1}{\mu} \quad \text{or} \quad \mu = \frac{1}{1 - B}. \tag{A.3}$$

Including (A.3) in the Solow residual yields:

$$SR = \Delta q - \alpha \Delta l - (1 - \alpha)\Delta k = B(\Delta q - \Delta k) + (1 - B)\theta \ . \tag{A.4}$$

An equivalent expression can be derived for the price-based Solow residual (SRP):

$$SRP = \alpha \Delta w + (1 - \alpha)\Delta r - \Delta p = -B(\Delta p - \Delta r) + (1 - B)\theta \ . \tag{A.5}$$

Subtracting (A.5) from (A.4) and adding an error term gives Roeger's equation from which *B* can be estimated using OLS:

$$\begin{aligned} &\Delta y_t = B\Delta x_t + \varepsilon_t, \quad \text{where} \\ &\Delta y_t = (\Delta q_t + \Delta p_t) - \alpha_t(\Delta l_t + \Delta w_t) - (1 - \alpha_t)(\Delta k_t + \Delta r_t), \\ &\Delta x_t = (\Delta q_t + \Delta p_t) - (\Delta k_t + \Delta r_t). \end{aligned} \tag{A.6}$$

Estimation of (A.6) compared to (A.2) has two advantages: 1. B can be estimated directly using OLS and 2. only nominal variables appear in (A.6), which helps overcome some data availability problems. Furthermore, (A.6) can be extended straightforwardly by incorporating

---

[23] If we rewrite (A.2) a bit, the Solow residual (SR) is obtained: $SR = \Delta q - \alpha\Delta l - (1-\alpha)\Delta k = (\mu-1)\alpha(\Delta l - \Delta k) + \theta$. Under Solow's assumptions the following should hold: *The productivity residual is uncorrelated with any variable that is uncorrelated with the rate of growth of true productivity* (Hall, 1990). However, data often reject this theorem.

intermediate inputs and defining mark-up ratios over gross output instead of value added (Oliveira-Martins, Pilat, Scarpetta, 1996). Accordingly, Δy_t and Δx_t become:

$$\begin{aligned} \Delta y_t &= \left(\Delta q_t^0 + \Delta p_t^0\right) - \alpha_t \left(\Delta l_t + \Delta w_t\right) - \beta_t \left(\Delta m_t + \Delta p_t^m\right) - \left(1 - \alpha_t - \beta_t\right)\left(\Delta k_t + \Delta r_t\right), \\ \Delta x_t &= \left(\Delta q_t^0 + \Delta p_t^0\right) - \left(\Delta k_t + \Delta r_t\right), \end{aligned} \tag{A.7}$$

where q^0 and p^0 are respectively gross output and its price, and m and p^m are respectively intermediate inputs and their prices.

Under constant-returns-to-scale, Roeger’s equation provides an unbiased estimate of the Lerner index B. However, decreasing-/increasing-returns-to-scale induce an upward/downward bias in the estimate of the mark-up. Correcting for this bias would imply lower/higher mark-ups than those following from (A.6) or (A.7). Thus, mark-up estimates from Roeger’s approach represent lower bounds under economies of scale.

Estimates for 14 OECD countries

Oliveira-Martins, Pilat, Scarpetta (1996a, 1996b) estimate mark-up ratios using Roeger’s approach with intermediate inputs, i.e. they estimate B using (A.16). They cover 36 manufacturing and 7 service sectors over 14 OECD countries in the period 1970-1992.

Two modifications are made. First, since no sector-specific information was available, the rental price of capital is simplified. In particular, the rental price is defined as the sum of the expected real cost of funds for the firm and the discard rate of gross capital stock multiplied by the economy-wide deflator for fixed business investment. Second, (A.7) is estimated using nominal output data. Often these include net indirect taxes. Including these in the estimation would cause an upward bias in the mark-up ratios. Therefore, one usually uses output data at factor costs. When prices include taxes, this can be adjusted by dividing the estimated mark-up by the net indirect tax rate.

The estimated mark-up ratios for services are more tentative than those for manufacturing for two reasons. First, the service industries represent much broader aggregates than the industry detail used for the manufacturing sector, hence, firms operating in (some of) these service sectors are likely to be quite heterogeneous. Second, the quality of statistical information for the service sectors is poorer than that available for manufacturing industries. The main results are the following:

- In general, in all countries, and almost all manufacturing industries considered, the estimated mark-ups are positive and statistically significant at the 5%-level.

- Compared to previous studies such as Hall (1990) and Roeger (1995)[24], the estimated mark-ups are substantially lower and more in line with observed profit rates. Overall, the level is between 5 and 25%.
- Mark-ups in the service sectors are generally higher than those in manufacturing, suggesting that departures from perfect competition are even more frequent in these sectors than in manufacturing. In several services, entry restricting regulations are likely to contribute to high mark-ups.
- Where manufacturing sectors are concerned, the level of the mark-ups appears related to the market structure of a certain sector. In particular, they are substantially lower in industries with small establishments' size, where the number of firms typically grows in line with the size of the market, than in industries with large establishments, where concentration remains relatively stable.
- There is considerable variation of mark-ups across countries and industries. Some of this variation may be due to the impact of specific policies. Across time the mark-ups remain relatively stable. Furthermore, it appears that high mark-ups (over 40%) occur less often since 1970.

Calculation of number of firms

Herfindahl-Hirschman-index-method

The most common method for calculating the number of firms in an applied general equilibrium models is the Herfindahl-Hirschman-index (HHI) method. This section explains this method, analyses the pros and cons and provides some alternatives. The discussion is greatly inspired by Erp, et al (2001).

HHI is a commonly accepted measure of market concentration. It is calculated by squaring the market share of each firm competing in the market and then summing the resulting numbers. For example, for a market consisting of four firms with shares of thirty, thirty, twenty and twenty percent, the HHI is 2600 ($30^2 + 30^2 + 20^2 + 20^2 = 2600$). Possible values for the HHI range from 0 to 10,000.

The HHI takes into account the relative size and distribution of the firms in a market and approaches zero when a market consists of a large number of firms of relatively equal size. The HHI increases both as the number of firms in the market decreases and as the disparity in size between those firms increases.

In the AHTENA model the HHI was formerly calculated based on CBS-data, which classifies the number of firms in intervals. Hence, the HHI is calculated as follows:

[24] The difference between the results of Roeger (1995) and Oliveira-Martins, Pilat, Scarpetta (1996) is due to the adjustment for intermediate inputs in the latter study. This tends to lower mark-ups substantially.

$$HHI = \sum_{j=1}^{m} k_j \zeta_j^{\,2} = \sum_{j=1}^{m} \frac{1}{k_j} \varphi_j^{\,2}, \quad \text{where } \varphi_j = \frac{k_j \mu_j}{\sum_{i=1}^{m} k_i \mu_i}, \tag{A.17}$$

and k_j is the number of firms in class j, μ_j is the midmark and ζ_j is the market share. The reciproce of HHI is then used as measure for the number of firms. When firms are assumed identical, all firms are in the same class (φ_j = 1) and k = 1 / HHI.

Two problems

Calculations following the approach above can yield some peculiar results. Erp, et al. (2001) find for example that in half of all industries the development following the HHI is the oppposite of growth according to CBS-data. These results seem to follow from two problems.

The first problem is that in (A.17) the weights are determined by the number of employees instead of the number of workers. In particular, one-man businesses are not accounted for.

The HHI is actually a measure for concentration and not much of a weighting factor. As an illustration, suppose there exist 10 firms with each one employee. The weight for the only class is ζ_1 = 1/10 and thus HHI = $10(1/10)^2$ = 1/10. Hence, the number of firms is 10. Now, suppose there is one firm with 10 employees, one with 9, etc. The weights equal, $\zeta_i = i/55$ and HHI = $\Sigma_i\, i(i/55)^2$ = 0.127. The number of employees is 7.86. Thus, when the actual number of firms does not change, only the size, the HHI shows that the number firms has decreased. This displacement-effect is a direct effect of the definition of the HHI as a measure for the degree of concentration.

Erp, et al. (2001) conclude by saying that few alternatives exist for the HHI-method. Nevertheless, they pose a few alternatives without a theoretical foundation.

Alternatives

One possible alternative is to calculate the number of firms based on a weighted average. First, the weighted average number of firms per class is calculated. Second, the total number of firms is computed by multiplying the weighted average by the number of classes. The disadvantage of this method is that a change in the size of firms still influences the calculated number of firms. Hence, the displacement-problem yet remains.

Another possibility is to actually use the number of firms according to the CBS-data, ignoring the differences in size. Each company that entries the market, no matter the size, increases the total number of firms in the market. The displacement-effect is solved. Perhaps, that a change in sales changes the average production per firm. Furthermore, there is no problem concerning the mid-mark of a class. Basically, the only disadvantage is that this method implicitly assumes symmetry between firms. This might be in contrast with the empirics. Nevertheless, this second alternative is the approach which is now modelled in a new version of ATHENA, because most data problems are solved.

Approach for WorldScan

In the approach for WorldScan, we set the number of firms in the base year equal to one, because of numerical problems caused by the high numbers of firms in the data. I argued that this alters the interpretation, because the number of firms is interpreted as an index rather than the absolute number of firms. The same holds for the fixed costs and brand quantity, since these variables change accordingly. However, a time series of these three variables still shows the relative change.

Annex 2: Long term effects of introduction of increasing returns to scale

Table A1 Model parameters

Sector	Elasticity of demand	Fixed cost
Agriculture	10^{10}	3.6×10^{-8}
Energy	7.7	62
Manufacturing	7.5	813
Services	4.0	2780
Transport	4.0	257

Source: Oliveira-Martins, Pilat, Scarpetta (1996a, 1996b) and WorldScan calculations. Fixed costs are in terms of quantities. In terms of firm output the fixed cost is equal to the reciproce of the demand elasticity

Table A2 Input prices, Agriculture, 2040

		CRTS	IRTS								
sector	region		all sectors	manufac. all	manufac. EU + rOECD	energy EU	manufac. EU	services EU	transport EU	man II EU	man III EU
agricult.	ROW	0.71	73%	31%	− 6%	0%	0%	1%	0%	0%	0%
	EU	0.67	40%	7%	1%	2%	6%	6%	1%	15%	15%
	rOECD	0.72	43%	11%	0%	0%	0%	1%	0%	0%	0%
energy	ROW	0.71	73%	31%	− 6%	0%	0%	1%	0%	0%	0%
	EU	0.67	40%	7%	1%	2%	6%	6%	1%	15%	15%
	rOECD	0.72	43%	11%	0%	0%	0%	1%	0%	0%	0%
manufac.	ROW	0.71	73%	31%	− 6%	0%	0%	1%	0%	0%	0%
	EU	0.67	40%	7%	1%	2%	6%	6%	1%	15%	15%
	rOECD	0.72	43%	11%	0%	0%	0%	1%	0%	0%	0%
services	ROW	0.71	73%	31%	− 6%	0%	0%	1%	0%	0%	0%
	EU	0.67	40%	7%	1%	2%	6%	6%	1%	15%	15%
	rOECD	0.72	43%	11%	0%	0%	0%	1%	0%	0%	0%
transport	ROW	0.71	73%	31%	− 6%	0%	0%	1%	0%	0%	0%
	EU	0.67	40%	7%	1%	2%	6%	6%	1%	15%	15%
	rOECD	0.72	43%	11%	0%	0%	0%	1%	0%	0%	0%

Source: WorldScan simulations. CRTS denotes Constant-Return-To-Scale and IRTS denotes Increasing-Returns-To-Scale. EU, rOECD and ROW denote respectively European Union, rest of OECD and Rest Of World. “man II” and “man III” denote respectively results for a test of halving the demand elasticity with fixed costs endogenous and keeping fixed costs fixed. The percentages are relative changes compared to the CRTS numbers.

Table A3 Input prices, Energy, 2040

sector	region	CRTS	IRTS all sectors	manufac. all	manufac. EU + rOECD	energy EU	manufac. EU	services EU	transport EU	man II EU	man III EU
agricult.	ROW	1.36	3%	12%	− 3%	− 6%	1%	1%	0%	2%	2%
	EU	1.53	4%	6%	0%	− 16%	5%	2%	2%	13%	13%
	rOECD	1.05	4%	6%	1%	− 2%	0%	1%	0%	1%	1%
energy	ROW	1.36	3%	12%	− 3%	− 6%	1%	1%	0%	2%	2%
	EU	1.53	4%	6%	0%	− 16%	5%	2%	2%	13%	13%
	rOECD	1.05	4%	6%	1%	− 2%	0%	1%	0%	1%	1%
manufac.	ROW	1.36	3%	12%	− 3%	− 6%	1%	1%	0%	2%	2%
	EU	1.53	4%	6%	0%	− 16%	5%	2%	2%	13%	13%
	rOECD	1.05	4%	6%	1%	− 2%	0%	1%	0%	1%	1%
services	ROW	1.36	3%	12%	− 3%	− 6%	1%	1%	0%	2%	2%
	EU	1.53	4%	6%	0%	− 16%	5%	2%	2%	13%	13%
	rOECD	1.05	4%	6%	1%	− 2%	0%	1%	0%	1%	1%
transport	ROW	1.36	3%	12%	− 3%	− 6%	1%	1%	0%	2%	2%
	EU	1.53	4%	6%	0%	− 16%	5%	2%	2%	13%	13%
	rOECD	1.05	4%	6%	1%	− 2%	0%	1%	0%	1%	1%

Source: WorldScan simulations. The percentages are relative changes compared to the CRTS numbers.

Table A4 Input prices, Manufacturing, 2040

sector	region	CRTS	IRTS all sectors	manufac. all	manufac. EU + rOECD	energy EU	manufac. EU	services EU	transport EU	man II EU	man III EU
agricult.	ROW	0.64	19%	− 15%	− 6%	0%	0%	1%	0%	− 1%	− 1%
	EU	0.71	18%	− 6%	− 8%	1%	− 3%	5%	1%	− 6%	− 6%
	rOECD	0.70	22%	− 8%	− 8%	0%	0%	1%	0%	0%	0%
energy	ROW	0.64	19%	− 15%	− 6%	0%	0%	1%	0%	− 1%	− 1%
	EU	0.71	18%	− 6%	− 8%	1%	− 3%	5%	1%	− 6%	− 6%
	rOECD	0.70	22%	− 8%	− 8%	0%	0%	1%	0%	0%	0%
manufac.	ROW	0.64	19%	− 15%	− 6%	0%	0%	1%	0%	− 1%	− 1%
	EU	0.71	18%	− 6%	− 8%	1%	− 3%	5%	1%	− 6%	− 6%
	rOECD	0.70	22%	− 8%	− 8%	0%	0%	1%	0%	0%	0%
services	ROW	0.64	19%	− 15%	− 6%	0%	0%	1%	0%	− 1%	− 1%
	EU	0.71	18%	− 6%	− 8%	1%	− 3%	5%	1%	− 6%	− 6%
	rOECD	0.70	22%	− 8%	− 8%	0%	0%	1%	0%	0%	0%
transport	ROW	0.64	19%	− 15%	− 6%	0%	0%	1%	0%	− 1%	− 1%
	EU	0.71	18%	− 6%	− 8%	1%	− 3%	5%	1%	− 6%	− 6%
	rOECD	0.70	22%	− 8%	− 8%	0%	0%	1%	0%	0%	0%

Source: WorldScan simulations. The percentages are relative changes compared to the CRTS numbers.

Table A5 Input prices, Services, 2040

sector	region	CRTS	IRTS all sectors	manufac. all	manufac. EU + rOECD	energy EU	manufac. EU	services EU	transport EU	man II EU	man III EU
agricult.	ROW	1.20	− 8%	24%	− 6%	0%	0%	0%	0%	0%	0%
	EU	1.22	4%	1%	0%	1%	6%	− 12%	3%	14%	14%
	rOECD	1.13	− 5%	1%	2%	0%	0%	0%	0%	0%	0%
energy	ROW	1.20	− 8%	24%	− 6%	0%	0%	0%	0%	0%	0%
	EU	1.22	4%	1%	0%	1%	6%	− 12%	3%	14%	14%
	rOECD	1.13	− 5%	1%	2%	0%	0%	0%	0%	0%	0%
manufac.	ROW	1.20	− 8%	24%	− 6%	0%	0%	0%	0%	0%	0%
	EU	1.22	4%	1%	0%	1%	6%	− 12%	3%	14%	14%
	rOECD	1.13	− 5%	1%	2%	0%	0%	0%	0%	0%	0%
services	ROW	1.20	− 8%	24%	− 6%	0%	0%	0%	0%	0%	0%
	EU	1.22	4%	1%	0%	1%	6%	− 12%	3%	14%	14%
	rOECD	1.13	− 5%	1%	2%	0%	0%	0%	0%	0%	0%
transport	ROW	1.20	− 8%	24%	− 6%	0%	0%	0%	0%	0%	0%
	EU	1.22	4%	1%	0%	1%	6%	− 12%	3%	14%	14%
	rOECD	1.13	− 5%	1%	2%	0%	0%	0%	0%	0%	0%

Source: WorldScan simulations. The percentages are relative changes compared to the CRTS numbers.

Table A6 Input prices, Transport, 2040

sector	region	CRTS	IRTS all sectors	manufac. all	manufac. EU + rOECD	energy EU	manufac. EU	services EU	transport EU	man II EU	man III EU
agricult.	ROW	0.73	12%	38%	− 5%	− 1%	0%	1%	− 1%	0%	0%
	EU	0.81	− 1%	5%	3%	− 2%	8%	5%	− 27%	19%	19%
	rOECD	0.85	2%	6%	6%	0%	0%	0%	− 2%	1%	1%
energy	ROW	0.73	12%	38%	− 5%	− 1%	0%	1%	− 1%	0%	0%
	EU	0.81	− 1%	5%	3%	− 2%	8%	5%	− 27%	19%	19%
	rOECD	0.85	2%	6%	6%	0%	0%	0%	− 2%	1%	1%
manufac.	ROW	0.73	12%	38%	− 5%	− 1%	0%	1%	− 1%	0%	0%
	EU	0.81	− 1%	5%	3%	− 2%	8%	5%	− 27%	19%	19%
	rOECD	0.85	2%	6%	6%	0%	0%	0%	− 2%	1%	1%
services	ROW	0.73	12%	38%	− 5%	− 1%	0%	1%	− 1%	0%	0%
	EU	0.81	− 1%	5%	3%	− 2%	8%	5%	− 27%	19%	19%
	rOECD	0.85	2%	6%	6%	0%	0%	0%	− 2%	1%	1%
transport	ROW	0.73	12%	38%	− 5%	− 1%	0%	1%	− 1%	0%	0%
	EU	0.81	− 1%	5%	3%	− 2%	8%	5%	− 27%	19%	19%
	rOECD	0.85	2%	6%	6%	0%	0%	0%	− 2%	1%	1%

Source: WorldScan simulations. The percentages are relative changes compared to the CRTS numbers.

Table A7 Factor prices, 2040

sector	region	CRTS	IRTS all sectors	manufac. all	manufac. EU + rOECD	energy EU	manufac. EU	services EU	transport EU	man II EU	man III EU
agricult.	ROW	0.05	588%	201%	− 8%	2%	− 1%	2%	0%	− 1%	− 1%
	EU	0.54	97%	11%	21%	9%	31%	45%	6%	85%	85%
	rOECD	0.24	113%	17%	25%	0%	0%	1%	0%	0%	0%
energy	ROW	0.09	111%	59%	− 6%	1%	0%	0%	0%	− 1%	− 1%
	EU	0.22	33%	4%	5%	4%	11%	10%	4%	28%	28%
	rOECD	0.16	33%	6%	7%	0%	0%	0%	0%	0%	0%
manufac.	ROW	0.07	313%	129%	− 7%	2%	− 1%	1%	0%	− 1%	− 1%
	EU	1.00	65%	7%	14%	6%	22%	29%	5%	59%	59%
	rOECD	0.80	71%	12%	17%	0%	0%	1%	0%	0%	0%
services	ROW	1.02	50%	32%	− 6%	1%	0%	0%	0%	− 1%	− 1%
	EU	2.61	17%	1%	1%	2%	7%	1%	4%	18%	18%
	rOECD	4.67	12%	2%	3%	0%	0%	0%	0%	0%	0%
transport	ROW	0.27	198%	92%	− 7%	1%	− 1%	1%	0%	− 1%	− 1%
	EU	1.61	48%	5%	9%	5%	17%	19%	5%	44%	44%
	rOECD	2.91	48%	8%	12%	0%	0%	0%	0%	0%	0%

Source: WorldScan simulations. The percentages are relative changes compared to the CRTS numbers.

Table A8 Raw output prices, 2040

sector	region	CRTS	IRTS all sectors	manufac. all	manufac. EU + rOECD	energy EU	manufac. EU	services EU	transport EU	man II EU	man III EU
agricult.	ROW	0.69	76%	32%	− 6%	0%	0%	1%	0%	0%	0%
	EU	0.63	38%	5%	2%	2%	8%	8%	1%	20%	20%
	rOECD	0.70	40%	9%	2%	0%	0%	1%	0%	0%	0%
energy	ROW	1.23	4%	14%	− 3%	− 5%	1%	1%	0%	2%	2%
	EU	0.87	2%	6%	0%	− 12%	6%	2%	2%	14%	14%
	rOECD	0.91	0%	6%	1%	− 2%	0%	1%	0%	1%	1%
manufac.	ROW	0.58	32%	1%	− 6%	0%	0%	1%	0%	0%	0%
	EU	0.71	8%	− 13%	− 14%	2%	− 8%	6%	1%	− 15%	− 15%
	rOECD	0.69	12%	− 12%	− 12%	0%	0%	1%	0%	0%	0%
services	ROW	1.16	2%	25%	− 6%	0%	0%	0%	0%	− 1%	− 1%
	EU	1.20	− 15%	1%	0%	1%	6%	− 26%	3%	15%	15%
	rOECD	1.11	− 18%	1%	2%	0%	0%	0%	0%	0%	0%
transport	ROW	0.71	25%	41%	− 6%	− 1%	0%	1%	0%	0%	0%
	EU	0.83	− 11%	4%	3%	− 2%	10%	5%	− 26%	24%	24%
	rOECD	0.83	− 9%	5%	6%	0%	0%	0%	0%	0%	0%

Source: WorldScan simulations. The percentages are relative changes compared to the CRTS numbers.

Table A9 Raw output quantities, 2040

sector	region	CRTS	IRTS all sectors	manufac. all	manufac. EU + rOECD	energy EU	manufac. EU	services EU	transport EU	man II EU	man III EU
agricult.	ROW	9901	− 56%	− 35%	1%	0%	0%	0%	0%	0%	0%
	EU	1012	1%	16%	− 17%	− 1%	− 20%	− 28%	− 3%	− 44%	− 44%
	rOECD	1613	− 5%	14%	− 21%	0%	0%	2%	1%	0%	0%
energy	ROW	2849	− 11%	− 14%	2%	− 4%	1%	0%	0%	2%	2%
	EU	1186	− 7%	9%	− 8%	41%	− 11%	− 11%	1%	− 24%	− 24%
	rOECD	2248	− 11%	0%	− 8%	− 6%	1%	1%	0%	2%	2%
manufac.	ROW	44058	− 36%	− 3%	− 6%	1%	− 2%	1%	1%	− 5%	− 5%
	EU	9486	− 4%	− 23%	12%	− 7%	18%	− 26%	− 3%	40%	40%
	rOECD	18875	− 22%	− 13%	11%	1%	− 2%	2%	1%	− 6%	− 6%
services	ROW	21185	− 8%	− 19%	1%	0%	0%	− 1%	0%	0%	0%
	EU	12460	16%	0%	− 4%	− 1%	− 5%	25%	0%	− 10%	− 10%
	rOECD	31115	16%	− 2%	− 4%	0%	0%	− 1%	0%	0%	0%
transport	ROW	5294	− 36%	− 34%	1%	0%	1%	0%	− 5%	1%	1%
	EU	1627	12%	8%	− 10%	2%	− 14%	− 17%	70%	− 30%	− 30%
	rOECD	3088	1%	− 1%	− 12%	0%	1%	1%	− 8%	3%	3%

Source: WorldScan simulations. The percentages are relative changes compared to the CRTS numbers.

Table A10 Number of firms, 2040

sector	region	CRTS	IRTS all sectors	manufac. all	manufac. EU + rOECD	energy EU	manufac. EU	services EU	transport EU	man II EU	man III EU
agricult.	ROW	27.2	− 56%	− 35%	1%	0%	0%	0%	0%	0%	0%
	EU	2.78	1%	16%	− 17%	− 1%	− 20%	− 28%	− 3%	− 44%	− 44%
	rOECD	4.43	− 5%	14%	− 21%	0%	0%	2%	1%	0%	0%
energy	ROW	6.85	− 23%	− 14%	2%	− 4%	1%	0%	0%	2%	2%
	EU	2.85	− 19%	9%	− 8%	23%	− 11%	− 11%	1%	− 24%	− 24%
	rOECD	5.41	− 22%	0%	− 8%	− 6%	1%	1%	0%	2%	2%
manufac.	ROW	8.34	− 45%	− 16%	− 6%	1%	− 2%	1%	1%	− 5%	− 5%
	EU	1.79	− 17%	− 33%	− 3%	− 7%	2%	− 26%	− 3%	3%	143%
	rOECD	3.57	− 33%	− 25%	− 4%	1%	− 2%	2%	1%	− 6%	− 6%
services	ROW	2.54	− 31%	− 19%	1%	0%	0%	− 1%	0%	0%	0%
	EU	1.49	− 13%	0%	− 4%	− 1%	− 5%	− 6%	0%	− 10%	− 10%
	rOECD	3.73	− 13%	− 2%	− 4%	0%	0%	− 1%	0%	0%	0%
transport	ROW	6.85	− 52%	− 34%	1%	0%	1%	0%	− 5%	1%	1%
	EU	2.11	− 16%	8%	− 10%	2%	− 14%	− 17%	27%	− 30%	− 30%
	rOECD	4.00	− 25%	− 1%	− 12%	0%	1%	1%	− 8%	3%	3%

Source: WorldScan simulations. The percentages are relative changes compared to the CRTS numbers.

Table A11 Composite price, 2040

sector	region	CRTS	IRTS all sectors	manufac. all	manufac. EU + rOECD	energy EU	manufac. EU	services EU	transport EU	man II EU	man III EU
agricult.	ROW	0.69	76%	32%	− 6%	0%	0%	1%	0%	0%	0%
	EU	0.63	38%	5%	2%	2%	8%	8%	1%	20%	20%
	rOECD	0.70	40%	9%	2%	0%	0%	1%	0%	0%	0%
energy	ROW	1.23	− 7%	14%	− 3%	− 5%	1%	1%	0%	2%	2%
	EU	0.87	4%	6%	0%	− 16%	6%	2%	2%	14%	14%
	rOECD	0.91	− 7%	6%	1%	− 2%	0%	1%	0%	1%	1%
manufac.	ROW	0.58	20%	− 15%	− 6%	0%	0%	1%	0%	0%	0%
	EU	0.71	17%	− 3%	− 8%	2%	− 3%	6%	1%	− 8%	− 32%
	rOECD	0.69	13%	− 6%	− 9%	0%	0%	1%	0%	0%	0%
services	ROW	1.16	12%	25%	− 6%	0%	0%	0%	0%	− 1%	− 1%
	EU	1.20	4%	1%	0%	1%	6%	− 12%	3%	15%	15%
	rOECD	1.11	− 26%	1%	2%	0%	0%	0%	0%	0%	0%
transport	ROW	0.71	12%	41%	− 6%	− 1%	0%	1%	0%	0%	0%
	EU	0.83	− 2%	4%	3%	− 2%	10%	5%	− 29%	24%	24%
	rOECD	0.83	− 16%	5%	6%	0%	0%	0%	0%	0%	0%

Source: WorldScan simulations. The percentages are relative changes compared to the CRTS numbers.

Table A12 Composite quantity, 2040

sector	region	CRTS	IRTS all sectors	manufac. all	manufac. EU + rOECD	energy EU	manufac. EU	services EU	transport EU	man II EU	man III EU
agricult.	ROW	9901	− 56%	− 35%	1%	0%	0%	0%	0%	0%	0%
	EU	1012	1%	16%	− 17%	− 1%	− 20%	− 28%	− 3%	− 44%	− 44%
	rOECD	1613	− 5%	14%	− 21%	0%	0%	2%	1%	0%	0%
energy	ROW	2849	− 1%	− 14%	2%	− 4%	1%	0%	0%	2%	2%
	EU	1186	− 9%	9%	− 8%	48%	− 11%	− 11%	1%	− 24%	− 24%
	rOECD	2248	− 4%	0%	− 8%	− 6%	1%	1%	0%	2%	2%
manufac.	ROW	44058	− 30%	15%	− 6%	1%	− 2%	1%	1%	− 5%	− 5%
	EU	9486	− 12%	− 31%	6%	− 7%	12%	− 26%	− 3%	29%	76%
	rOECD	18875	− 23%	− 19%	7%	1%	− 2%	2%	1%	− 6%	− 6%
services	ROW	21185	− 16%	− 19%	1%	0%	0%	− 1%	0%	0%	0%
	EU	12460	− 6%	0%	− 4%	− 1%	− 5%	5%	0%	− 10%	− 10%
	rOECD	31115	29%	− 2%	− 4%	0%	0%	− 1%	0%	0%	0%
transport	ROW	5294	− 29%	− 34%	1%	0%	1%	0%	− 5%	1%	1%
	EU	1627	2%	8%	− 10%	2%	− 14%	− 17%	77%	− 30%	− 30%
	rOECD	3088	9%	− 1%	− 12%	0%	1%	1%	− 8%	3%	3%

Source: WorldScan simulations. The percentages are relative changes compared to the CRTS numbers.

Table A13 Brand price, 2040

sector	region	CRTS	IRTS all sectors	manufac. all	manufac. EU + rOECD	energy EU	manufac. EU	services EU	transport EU	man II EU	man III EU
agricult.	ROW	0.69	76%	32%	− 6%	0%	0%	1%	0%	0%	0%
	EU	0.63	38%	5%	2%	2%	8%	8%	1%	20%	20%
	rOECD	0.70	40%	9%	2%	0%	0%	1%	0%	0%	0%
energy	ROW	1.23	19%	14%	− 3%	− 5%	1%	1%	0%	2%	2%
	EU	0.87	18%	6%	0%	1%	6%	2%	2%	14%	14%
	rOECD	0.91	15%	6%	1%	− 2%	0%	1%	0%	1%	1%
manufac.	ROW	0.58	52%	16%	− 6%	0%	0%	1%	0%	0%	0%
	EU	0.71	25%	0%	0%	2%	6%	6%	1%	15%	15%
	rOECD	0.69	29%	1%	2%	0%	0%	1%	0%	0%	0%
services	ROW	1.16	36%	25%	− 6%	0%	0%	0%	0%	− 1%	− 1%
	EU	1.20	13%	1%	0%	1%	6%	− 2%	3%	15%	15%
	rOECD	1.11	9%	1%	2%	0%	0%	0%	0%	0%	0%
transport	ROW	0.71	67%	41%	− 6%	− 1%	0%	1%	0%	0%	0%
	EU	0.83	19%	4%	3%	− 2%	10%	5%	− 2%	24%	24%
	rOECD	0.83	21%	5%	6%	0%	0%	0%	0%	0%	0%

Source: WorldScan simulations. The percentages are relative changes compared to the CRTS numbers.

Table A14 Brand quantity, 2040

sector	region	CRTS	IRTS all sectors	manufac. all	manufac. EU + rOECD	energy EU	manufac. EU	services EU	transport EU	man II EU	man III EU
agricult.	ROW	364	0%	0%	0%	0%	0%	0%	0%	0%	0%
	EU	364	0%	0%	0%	0%	0%	0%	0%	0%	0%
	rOECD	364	0%	0%	0%	0%	0%	0%	0%	0%	0%
energy	ROW	416	0%	0%	0%	0%	0%	0%	0%	0%	0%
	EU	416	0%	0%	0%	0%	0%	0%	0%	0%	0%
	rOECD	416	0%	0%	0%	0%	0%	0%	0%	0%	0%
manufac.	ROW	5285	0%	0%	0%	0%	0%	0%	0%	0%	0%
	EU	5285	0%	0%	0%	0%	0%	0%	0%	0%	− 58%
	rOECD	5285	0%	0%	0%	0%	0%	0%	0%	0%	0%
services	ROW	8340	0%	0%	0%	0%	0%	0%	0%	0%	0%
	EU	8340	0%	0%	0%	0%	0%	0%	0%	0%	0%
	rOECD	8340	0%	0%	0%	0%	0%	0%	0%	0%	0%
transport	ROW	773	0%	0%	0%	0%	0%	0%	0%	0%	0%
	EU	773	0%	0%	0%	0%	0%	0%	0%	0%	0%
	rOECD	773	0%	0%	0%	0%	0%	0%	0%	0%	0%

Source: WorldScan simulations. The percentages are relative changes compared to the CRTS numbers.

References

Armington, P. (1969). A Theory of Demand for Products Distinguished by Place of Production. *IMF Staff Papers*, 16: 159-178

d'Aspremont, C., Dos Santos Ferreira, R., Gerard-Varet, L. (1996). On the Dixit-Stiglitz Model of Monopolistic Competition. *The American Economic Review*, 86: 623-629

Baldwin, R., Forslid, R., Martin, P., Ottaviano, G. and Robert-Nicoud, F. (2003). *Economic Geography and Public Policy*. Princeton University Press, Princeton and Oxford

Bchir, M.H., Decreux, Y., Guérin, J. and Jean, S. (2002). MIRAGE, a Computable General Equilibrium Model for Trade Policy Analysis. CEPII, working paper, 2002-17

Bénassy, J.P. (1996). Taste for Variety and Optimum Production Patterns in Monopolistic Competition. *Economic Letters*, 52: 41-47

Bishop, R.L. (1967). Monopolistic Competition and Welfare Economics, in Kuenne, R.E. (ed.). *Monopolistic Competition Theory: Studies of Impact, Essays in Honor of Edward H. Chamberlin*. John wiley & sons, New York

Blaug, M. (1997). *Economic Theory in Retrospect*, 5th edition. Cambridge University Press, Cambridge

Brakman, S., Garretsen, H. and Van Marrewijk, C. (2001). *An Introduction to Geographical Economics*. Cambridge University Press, Cambridge

Brakman, S. and Heijdra, B.J. (Eds.) (2004). *The Monopolistic Competition Revolution in Retrospect*. Cambridge University Press, Cambridge

Bresnahan, T. (1989). Empirical Studies of Industries with Market Power, in: R. Schmalensee and R. Willig (eds.), *Handbook of Industrial Organization*, North Holland, Amsterdam

Broer, D.P., Ederveen, J.P., Van Erp, F.A.M., Van Hien, A.E., Lunsing, J.J. and Vromans, M.W.A.M. (2003). Production and Market Structure in the Athena Model. CPB, Interne notitie, April 8, 2003

Broer, D.P. and Heijdra, B.J. (2001). The Invesment Tax Credit under Monopolistic Competition. *Oxford Economic Papers*, 53: 318-351

Caballero, R.J. and Lyons, R.K. (1989). The role of external economies in U.S. manufacturing. *NBER Working Paper Series*, 3033, Cambridge, MA

Chamberlin , E.H. (1933). *The Theory of Monopolistic Competition: A Re-orientation of the Theory of Value*, 8th ed. (1962). Cambridge, MA, Harvard University Press

Davies, S. and Lyons, B. (1996). *Industrial Organisation in the European Union*. Oxford University Press, Oxford

Deardorff, A.V. and Stern, R.M. (1990). *Computational Analysis of Global Trading Arrangements*. Ann Arbor: The University of Michigan Press

Dixit, A.K. and Stiglitz, K.E. (1975). Monopolistic competition and optimum product diversity. Research Papers in Economics, 64, University of Warwick

Dixit, A.K. and Stiglitz, K.E. (1977). Monopolistic competition and optimum product diversity. *The American Economic Review*, 67: 297-308

Dixit, A.K. and Stiglitz, K.E. (1979). Monopolistic Competition and Optimum Product Diversity: Reply. *The American Economic Review*, 69: 961-963

Dixit, A.K. and Stiglitz, K.E. (1993). Monopolistic competition and optimum product diversity: reply. *The American Economic Review*, 83: 302-304

Ethier, W. (1982). National and international returns to scale in the theory of international trade. *American Economic Review*, 72: 389-405

Erp, van F., Hien, van A., Smid, B. (2001). Aantal Bedrijven. CPB Memo, 23 juli 2001, CPB
Figuières, C., Jean-Marie, A., Quérou, N. and Tidball, M. (2004). *Theory of conjectural variations*. World Scientific Computing

Flam, H. and Helpman, E. (1987). Industry Policy under Monopolistic Competition. *Journal of International Economics*, 22: 79-102

Ford, H. (1922). *My Life and Work*, in collaboration with S. Crowther. London: William Heinemann

Francois, J.F. (1990). Trade in Producer Services and Returns due to Specialization. *Canadian Journal of Economics*, 21

Francois, J.F. and Roland-Holst, D.W. (1996). Trade Policy, Scale Economics, and Imperfect Competition in Applied Models. Working paper ERAD-1996-001, The World Trade Organization

Fujita, M. and Thisse, J.F. (2002). *Economics of Agglomeration: Cities, Industrial Location and Regional Growth*. Cambridge University Press, Cambridge

Gasiorek, M., Smith, A. and Venables, A. (1992). 1992: Trade and Welfare, a General Equilibrium Model. CEPR, discussion paper, 672

Geroski, P., Gregg, P. and Van Reenen, J. (1995). Market Imperfections and Employment. *The OECD Jobs Study: Working Papers*, 5, OECD, Paris

Hall, R.E. (1986). Market Structure and Macroeconomic Fluctuations. *Brookings Papers on Economic Activity*, 2: 285-338

Hall, R.E. (1988). The relation between price and marginal cost in U.S. industry. *Journal of Political Economy*, 96, 5: 921-947

Hall, R.E. (1990). The Invariance Properties of Solow's Productivity Residual, in P. Diamond (ed.), *Growth, Productivity, Unemployment*, MIT Press, Cambridge MA

Harris, R. (1984). Applied General Equilibrium Analysis of Small Open Economies with Scale Economies and Imperfect Competition. *The American Economic Review*, 74: 1016-1032

Harrison, G.W., Rutherford, T.F. and Tarr, D.G. (1997). Quantifying the Uruguay Round. *The Economic Journal*, 107: 1405-1430

Hay, G. A. and Kelley, D. (1974). An empirical survey of price-fixing conspiracies. *Journal of Law and Economics*, 17: 13-38

Krugman, P. (1979). Increasing Returns, Monopolistic Competition, and International Trade. *Journal of International Economics,* 9: 469-480

Krugman, P. (1980). Scale Economies, Product Differentiation, and the Pattern of Trade. *The American Economic Review,* 70: 950-959

Krugman, P. (1991). Increasing Returns and Economic Geography. *Journal of Political Economy*, 99: 483-499

Krugman, P. and Venables, A. (1995). Globalization and the Inequality of Nations. *Quarterly Journal of Economics*, 110: 857-880

Lawrence, C. and Spiller, P.T. (1983). Product Diversity, Economies of Scale, and International Trade. *Quarterly Journal of Economics*, 98: 63-83

Mankiw, N.G. and Whinston, M.D. (1986). Free entry and social inefficiency. *Rand Journal of Economics*, 17: 48-58

Markusen, J.R. (1990). Micro-foundations of External Scale Economies. *Canadian Journal of Economics*, 23: 495-508

Mensbrugghe, D. van der (2003). LINKAGE Technical Reference Document. Available at: http://www.worldbank.org

Neary, J.P. (2000). Monopolistic Competition and International Trade Theory. ", in S. Brakman and B.J. Heijdra (eds.): *The Monopolistic Competition Revolution in Retrospect*, Cambridge: Cambridge University Press, 2003, 159-184

Neary, J.P. (2002). The Road Less Travelled: Oligopoly and Competition Policy in General Equilibrium, in Arnot, R., Greenwald, B., Kanbur, R. and Nalebuff, B. (eds.), *Economics for an Imperfect World: Essays in Honor of Joseph Stiglitz,* Cambridge, Massachusetts: MIT Press

Neary, J.P. (2003). International Trade in General Oligopolistic Equilibrium. unpublished, avalaible at: http://www.ucd.ie/economic/staff/pneary/neary.htm

Oliveira-Martins, J. and Scarpetta, S. (1996). Mark-up Ratios in Manufacturing Industries: Estimates for 14 OECD countries. OECD Economics Department working papers, 162

Oliveira-Martins, J., Pilat, D. and Scarpetta, S. (1996). Mark-up Pricing, Market Structure and the Business Cycle. *OECD Economic Studies*, 27, 71-105

Pratten, C. (1988). A Survey of the Economies of Scale. *Research on the Cost of Non-Europe*, 2. Brussels: Commission of the European Communities

Redding, S. (2005). Book review, The Monopolistic Competition Revolution in Retrospect. *Journal of International Economics*, 63: 531-537

Roeger, W. (1995). Can Imperfect Competition Explain the Difference Between Primal and Dual Productivity Measures? Estimates for US Manufacturing. *Journal of Political Economy*, 103, 316-330

Romer, P. (1990). Endogenous technological change. *Journal of Political Economy*, 98: 71-102.
Samuelson, P. (1954). The Transport Problem and Transport Costs: Analysis of Effects of Trade Impediments. *Review of Economic Studies*, 46, 2

Schmalensee, R. (1989). Inter-industry Studies of Structure and Performance, in: R. Schmalensee and R. Willig (eds.), *Handbook of Industrial Organization*, North Holland, Amsterdam

Smith, A. and Venables, A. (1988). Completing the Internal Market in the European Community: Some Industry Simulations. *European Economic Review*, 32, 1501-1525

Solow, R. (1957). Technical changes and the aggregate production function. *Review of Economics and Statistics*, 39: 312-320

Spence, A.M. (1974). Product Selection, Fixed Costs and Monopolistic Competition. manuscript; subsequently published in *Review of Economic Studies*, 43: 217-235, 1976

Tinbergen, J. (1952). *On the Theory of Economic Policy*. Amsterdam, North-Holland.

Venables, A. (1996). Equilibrium Locations of Vertically Linked Industries. *International Economic Review*, 37: 341-359

Willenbockel, D. (1994). *Applied General Equilibrium Modelling: Imperfect Competition and Economic Integration*. John Wiley & Sons Ltd, Series in Financial Economics and Quantitative Analysis

Yang, X. and Heijdra, B.J. (1993). Monopolistic Competition and Optimum Product Diversity: Comment. *American Economic Review*, 83: 295-301