Endogenous market structure and the growth and welfare effects of economic integration

Pietro F. Peretto*

Department of Economics, Duke University, Room 205, P.O. Box 90097, 27708 Durham, NC, USA

Received 12 November 2000; received in revised form 1 June 2001; accepted 28 October 2001

Abstract

This paper studies the growth and welfare effects of integration in a world economy populated by global oligopolists. In economies that move from autarky to trade, growth and welfare rise because exit of domestic firms is more than compensated by entry of foreign firms so that integration generates a larger, more competitive market where firms have access to a larger body of technological spillovers that support faster growth. The effects of a gradual reduction of tariffs are different because economies start out from a situation where all firms already serve all markets. In this case, the global number of firms falls so that the variety of consumption goods and the diversity of innovation paths fall. The surviving firms, on the other hand, are larger and exploit static and dynamic economies of scale to a larger degree. These homogenization and rationalization effects work in opposite directions. Under plausible conditions, the rationalization effect dominates and growth and welfare rise.

© 2002 Elsevier Science B.V. All rights reserved.

Keywords: Market structure; Growth; R&D; Integration

JEL classification: E10; L16; O31; O40

1. Introduction

In the last 50 years the world has become more integrated as institutional and technological changes have lowered barriers to the mobility of goods, capital and

*Tel.: +1-919-660-1807; fax: +1-919-684-8974.
E-mail address: peretto@econ.duke.edu (P.F. Peretto).
people. This process of globalization affects individual economies through several mechanisms. A very powerful one is industrial restructuring, that is, the change in market structure that occurs as a result of entry/exit of firms. This paper discusses a model where industrial restructuring entails exit of domestic firms and asks whether such a change in market structure is beneficial for growth and welfare.¹

There are two ways to approach this question. One can compare autarky to free trade or one can look at the effects of an incremental liberalization of world trade. The first question is about gains from trade, the second is about how changes in global tariffs affect the exploitation of gains from trade at the margin. This paper undertakes both exercises and thus provides a comprehensive characterization of the role of trade in the determination of growth and welfare in the global economy.

In economies that move from autarky to trade—free or restricted by tariffs—integration raises growth and welfare because exit of domestic firms is more than compensated by entry of foreign firms. In other words, the fact that domestic consumers and producers gain access to foreign goods and knowledge means that integration generates a larger, more competitive market where firms have access to a more diverse body of technological spillovers that supports faster growth. The growth effect is larger the less competitive the economy is before integration, while it is negligible for economies that are very competitive to begin with. The effects of a gradual reduction of barriers to trade—a global reduction in tariffs—are different because economies start out from a situation where all firms already serve all markets. Hence, the reduction of barriers to trade makes all markets more competitive, in the sense that in each country domestic producers are less protected, and thereby triggers a reduction in the global number of firms. In other words, in each country consumers see a reduction of the variety of available goods and producers see a reduction of the diversity of spillover sources. In this case, there is a tension between internal and external increasing returns. The reduction of the global number of firms means that the variety of consumption goods and the diversity of innovation paths fall. This is the homogenization effect. On the other hand, the surviving firms are larger and exploit static and dynamic economies of

¹For a recent exposition of the view that this type of work is necessary to better understand the interactions of technology and trade, see Baldwin and Forslid (1998). One can look at industrial restructuring in two ways. In two-sector models of inter-industry trade, entry/exit is driven by a specialization effect: as a country specializes in one sector, there is entry of firms to that sector and exit from the other sector. In one-sector models of intra-industry trade, in contrast, entry/exit is driven purely by a scale effect: as a country trades with another, the market becomes larger and induces entry/exit of firms. Specialization and scale effects are not independent. In two-sector models of inter-industry trade, specialization affects market structure within each sector precisely because it changes the relative and absolute size of the two sectors and thereby triggers entry/exit within each sector. In one-sector models of intra-industry trade, the inter-industry specialization effect is absent and what is left is the pure scale effect based on market size. This paper takes this approach.
scale to a larger degree. This is the rationalization effect. A global reduction in tariffs raises growth and welfare if the rationalization effect dominates.

Two aspects of the analysis are worth emphasizing. Firstly, the notion that openness to foreign competition drives domestic firms out of business is accepted by most people. What is not often accepted, however, is that the benefits dominate the costs. An exception is the rationalization argument due to Eastman and Stykolt (1960), and formalized by Dixit and Norman (1980) and Horstmann and Markusen (1986), that protection leads to excessive entry and thus prevents firms from achieving an efficient size. Horstman and Markusen wrote partially in response to Venables (1985), who argued that protection can raise national welfare in the context of a homogeneous good model where markets are segmented, and showed that his results depended crucially on the assumption of segmented markets. Venables (1987), in turn, pointed out that the difference between segmented and integrated markets is irrelevant in a differentiated products model and showed that the benefits from protection stem from the fact that it raises the share of the market for differentiated goods that domestic firms capture. This paper provides a generalization of these arguments to a framework where the number of firms and firm size—which are jointly determined and thus interdependent—affect economic growth. Moreover, it looks at the general equilibrium effects for the global economy of multilateral actions—like the global reduction in tariffs due to the GATT-WTO—as opposed to focusing on unilateral trade policy. The framework takes into account several features of real-world industry and does not produce counterfactual predictions, like a linear scale effect, that make previous analyses of the growth effects of integration not convincing. Moreover, it sheds light on the conditions under which protection is or is not beneficial and thus clarifies important issues in the debate on the ‘new protectionism’ that trade models based on increasing returns seem to underpin intellectually.

The second aspect worth emphasizing is that the effects of a move from autarky to free trade are different from those of an incremental liberalization of world trade. According to Baldwin and Forslid (1999, 2000), this property is shared by most trade models based on imperfect competition and stems from the fact that incremental trade liberalization has non-linear effects. This paper’s results point out that, in fact, this property is embedded in trade models of product variety

---

1Rivera-Batiz and Romer (1991a,b), for example, base their analysis on first-generation models of endogenous growth where increasing returns apply at the aggregate level. This property implies that the economy’s growth rate increases linearly with its labor endowment (population size). Recent research has shown that this linear scale effect is empirically false (see, e.g., Backus et al., 1992; Jones, 1995; Dinopoulos and Thompson, 1999) and, more importantly, is produced by implausible assumptions on the diffusion and application of knowledge (for recent reviews of this line of work, see Jones, 1999; Aghion and Howitt, 1998; Peretto and Smulders, 2001).
because these models by construction let consumers have access to all goods. This means that a reduction of trade frictions affects the quantity of each foreign good that domestic consumers buy but does not affect the wedge between the number of goods produced domestically and the number of goods that are available to domestic consumers. A comparison of autarky to trade, in contrast, focuses by construction on the wedge between goods produced domestically and goods available to consumers.

This paper studies issues that are notoriously complex. To keep the analysis tractable, it focuses on the effects of integration among identical countries. As a result, it works with a model of intra-industry trade in the tradition of Helpman and Krugman (1985) and ignores the effects that derive from comparative advantage when technologies, preferences, and endowments differ. Specifically, it models a world economy characterized by global oligopolists that produce differentiated goods, engage in worldwide Bertrand price competition and establish in-house R&D facilities to produce, over time, a constant flow of incremental, cost-reducing innovations. Welfare in the typical country depends on the level and growth of consumption and on the variety of consumption goods.

The paper is organized as follows. Section 2 presents the basic closed-economy model that in Section 3 is generalized to the case of trade restricted by ad-valorem import tariffs. Section 4 shows how a move from autarky to free trade improves growth and welfare. Section 5 studies the effects of a global reduction in tariffs. Section 6 compares the two exercises and discusses the differences in the results that they yield.

2. The basics: preferences, technology and equilibrium concept

Consider an economy with a fixed population, \( L \), of identical consumers each endowed with one unit of labor. Consumers have no preference for leisure so that labor supply is \( L \). Consumers have symmetric preferences over differentiated goods supplied by oligopolistic manufacturing firms. Each firm supplies one good. Price-setting firms maximize the present value of profits subject to their demand schedules. This guarantees that supply equals demand in the goods market. Labor and capital market clearing determine the economy’s general equilibrium.

\[ \text{A recent paper that follows a similar approach is Piermartini (1999). It considers inter-industry and intra-industry effects of incremental integration due to a reduction of iceberg transport costs.} \]

\[ \text{Like in all models of market structure and trade based on increasing returns, the pattern of intra-industry specialization (which country produces which good) is not determined (see, e.g., Helpman and Krugman, 1985). However, this pattern is irrelevant for the main results of the paper because all product lines are identical and preferences are symmetric over all goods. Hence, the aggregate performance of a country does not depend on its specialization.} \]
2.1. Preferences

The typical consumer maximizes lifetime utility

\[ U(t) = \int_{t}^{\infty} e^{-\rho(t-\tau)} \log C(\tau) \, d\tau, \quad \rho > 0, \]  

subject to the intertemporal budget constraint that the present discounted value of expenditure cannot be greater than the present discounted value of labor income plus initial wealth,

\[ \int_{t}^{\infty} e^{-\rho(t-\tau)} E(\tau) \, d\tau \leq \int_{t}^{\infty} e^{-\rho(t-\tau)} W(\tau) \, d\tau + B(t), \]

where \( \rho \) is the individual discount rate, \( E \) is per capita consumption expenditure, \( B \) is the individual’s assets holding, and \( W = 1 \) is the wage rate, which is taken as the numeraire. \( C \) is a consumption index defined as

\[ C = \left[ \sum_{i=1}^{N} C_{i}^{\epsilon^{-1}} \right]^{1/\epsilon}, \quad \epsilon > 1, \]  

where \( \epsilon \) is the elasticity of product substitution, \( C_{i} \) is the consumer’s purchase of each differentiated good, and \( N \) is the number of goods.

As is well known, in this framework consumers set

\[ \frac{\dot{E}}{E} = r - \rho \]  

and, taking as given this time-path of expenditure, maximize (2) subject to \( E = \sum_{i=1}^{N} P_{i}C_{i} \), determining

\[ X_{i} = LE \frac{P_{i}^{-\epsilon}}{\sum_{j=1}^{N} P_{j}^{1-\epsilon}}, \]  

where \( X_{i} = LC_{i} \).

2.2. Technology

Each firm produces output with the technology

\[ X_{f} = Z_{f}^{\theta}(L_{X_{f}} - \phi), \quad \phi > 0, \quad 0 < \theta < 1, \]  

where \( X_{f} \) is the firm’s output and \( L_{X_{f}} \) is the firm’s labor employment in production. There are fixed (and sunk) overhead costs \( \phi \). The firm undertakes R&D in order to accumulate cost-reducing innovations that are patented. \( Z_{f} \) is the firm’s patent
stock. Technological opportunity conditions are described by two functions. The function $Z_i^p$ states that labor productivity rises with the accumulated patent stock. The function

$$Z_i = L_i K, \quad K_i = Z_i + \sum_{j \neq i}^N \frac{\gamma}{1 + \delta(N - 1)} Z_j, \quad 0 \leq \gamma, \delta \leq 1, \quad (6)$$

states that the firm generates $Z_i$ new patents in the interval of time $dt$ by allocating $L_i$ units of labor to R&D. Patents contain knowledge that is not innovation-specific and is not appropriable. This implies that a firm’s R&D produces knowledge that can be exploited by other firms. The parameter $\gamma$ determines the fraction of knowledge that becomes public. The productivity of labor in R&D, $K_i$, is a weighted average of the firm’s own knowledge and spillovers from the other firms. $N$ is the number of firms, which is determined endogenously by the equilibrium conditions of the industry.

Eq. (6) exhibits constant returns to scale to knowledge that sustain constant endogenous growth in steady state. Spillovers imply increasing returns to the number of firms. The degree to which these apply depends on the interaction of knowledge specialization and R&D replication. Firms produce differentiated products and develop specialized technologies. They thus learn from their similarities as well as from their differences in the sense that the knowledge produced by one firm benefits the R&D effort of another firm according to how close they are in technology space and how little they replicate each others’ R&D. Eq. (6) summarizes these factors with the assumption that the weights assigned to spillovers of knowledge from the other firms are symmetric and decreasing in the number of firms. The first is a simplifying assumption that preserves the overall symmetry of the model; the second captures the notion that ‘idea congestion’ increases with the number of firms in the industry.\(^5\) The parameter $\delta$ determines how quickly congestion sets in and increasing returns to the number of firms vanish. To see this, notice that in symmetric equilibrium

$$\frac{Z_i'}{Z_i} = L_i \frac{1 + (\gamma + \delta)(N - 1)}{1 + \delta(N - 1)},$$

where firm-level variables without subscripts denote industry averages. This expression exhibits positive, diminishing returns to the number of firms $N$ that are bounded from above.

\(^5\)Peretto and Smulders (2001) provide a microfoundation for these assumptions. They construct a model where the spillovers function that I use in the text is obtained as a result of the interaction among firms.
2.3. Market structure in general equilibrium

The market equilibrium of the economy sketched above has three building blocks. Firstly, value maximization determines the price and investment strategies of active firms. Next, entry/exit decisions determine the number of active firms. Finally, these two results combined with labor and market clearing conditions determine the general equilibrium of the economy.

The typical firm maximizes the present discounted value of net cash flow,

$$V_t = \int_0^\infty e^{-r(\tau)} d\Pi(\tau) d\tau.$$

Instantaneous profits are

$$\Pi_t = P_t X_t - L_t - L_{Z_t},$$

where $L_{X_t}$ is total production costs and $L_{Z_t}$ is R&D expenditure. With perfect foresight, $V_t$ is the stock market value of the firm. Peretto (1996) provides a formal definition of the Nash equilibrium with free entry/exit for this set-up. In words, the Nash equilibrium requires that each firm chooses time-paths of price and R&D investment that maximize $V_t$ and that such maximized value be driven to zero by entry/exit. The Nash equilibrium can thus be described as the outcome of the interaction between two relationships. One represents the optimal R&D strategy of firms given the number of firms; the other represents the number of firms that the market supports in zero-profit equilibrium given that firms play the optimal R&D strategy. By construction, these relationships incorporate the conditions that the capital and labor markets clear.

This set-up reduces the determination of the economy’s general equilibrium to the determination of the equilibrium of its manufacturing sector and focuses the analysis on the role of market structure. The next section adapts this structure to a model of international trade where global oligopolists sell differentiated goods in all countries.

3. The trade model with import tariffs

There are $c$ identical countries. Let $N_k$ be the number of domestic firms in country $k$. Let $M = \sum_{k=1}^c N_k$ be the global number of firms. All countries charge an

---

6The driving assumption here is that the cost of entry/exit is zero. Peretto (1996) shows that in this case the no-arbitrage condition on the value of the firm reduces to the zero-profit condition. This is an extreme assumption made only in order to simplify the analysis and the presentation of the model. Its main implication is that the economy jumps to the steady state instantaneously. See Peretto (1998) for an example of a model where costly entry generates transitional dynamics.
ad-valorem tariff $\tau$ on foreign goods. Tariff revenues are rebated to households in a lump-sum fashion. The part of the model that describes producers and consumers does not change with the exception that now the number of goods in (2) and the number of firms in (6) is the global number of firms (domestic and foreign). In other words, the typical consumer in the typical country consumes all available goods independently of tariffs while the typical producer draws spillovers from all firms. This turns out to be important when distinguishing between the exercise that compares autarky to trade and the one that considers the incremental effects of tariffs.

3.1. Preliminaries

Most of the paper’s results depend on the structure of demand and its implications for the pricing behavior of firms. It is useful to emphasize this aspect of the model. Formally, (2) and (6) are modified as follows:

\[
C_k = \left[ \sum_{i=1}^{N_k} C_{ik}^{e-1} + \sum_{s \neq k, i=1}^{N_s} C_{is}^{e-1} \right]^{\varepsilon/\varepsilon - 1}, \tag{7}
\]

\[
K_{ik} = Z_{ik} + \sum_{j \neq k}^{N_k} \frac{\gamma}{1 + \delta(N-1)} Z_{jk} + \sum_{s \neq k}^{N_s} \sum_{j=1}^{N_j} \frac{\gamma}{1 + \delta(N-1)} Z_{js}. \tag{8}
\]

The first equation posits that the typical consumer likes all goods symmetrically. The second posits that the firm draws spillovers from the global population of firms and that spillovers from domestic and foreign firms get the same weight.

In country $k$, consumers maximize (7) subject to

\[
E_k = \sum_{i=1}^{N_k} P_{ik} C_{ik} + \sum_{s \neq k}^{N_s} \sum_{i=1}^{N_s} (1 + \tau)P_{is} C_{is},
\]

where $C_{ik}$ is consumption of good $i$ produced in country $k$, $C_{is}$ is consumption of good $i$ imported from country $s$, $P_{ik}$ is the price of good $i$ produced in country $k$ and $P_{is}$ is the price of good $i$ imported from country $s$. The price index of consumption goods in country $k$ is

\[
P_k = \left[ \sum_{i=1}^{N_k} P_{ik}^{1-\varepsilon} + \sum_{s \neq k}^{N_s} \sum_{j=1}^{N_j} (1 + \tau)^{1-\varepsilon} P_{js}^{1-\varepsilon} \right]^{1/1-\varepsilon}.
\]

Solving the optimization problem, one finds that consumers demand goods according to the following schedule:

\[
C_{ik}^D = \frac{E_k P_{ik}^{\varepsilon}}{P_k^{\varepsilon}},
\]
The first equation is the demand for good $i$ when good $i$ is produced domestically, the second is the demand for good $i$ when good $i$ is imported. Using these results, the total demand faced by firm $i$ in country $k$ can be written

$$X_{ik} = X_{ik}^d + X_{ik}^e = LE_k \frac{p_{ik}^{-\epsilon}}{p_k^{1-\epsilon}} + \sum_{s \neq k} LE_s \frac{(1 + \tau)^{-\epsilon} p_{ik}^{-\epsilon}}{p_s^{1-\epsilon}},$$

(9)

where $X_{ik}^d$ and $X_{ik}^e$ denote, respectively, the firm’s domestic and foreign sales and $p_s$ is the price index of consumption goods in country $s$.

The following lemma derives one of the crucial properties of the model. To ease the exposition, this and other proofs are relegated to a Technical Appendix available on request. The text focuses on their implications.

**Lemma 1.** The demand schedule (9) yields the following price elasticity of demand for a global price-setting oligopolist:

$$\xi = \epsilon - (\epsilon - 1) \frac{1 + (1 + \tau)^{-\epsilon} (c - 1)}{1 + (1 + \tau)^{-\epsilon} (c - 1)} \frac{c}{1 + (c - 1)(1 + \tau)^{-\epsilon}} \frac{1}{cN}.$$  

(10)

**Proof.** See the Technical Appendix. □

The expression for the price elasticity of demand summarizes the effect of trade barriers on market competition: it is monotonically decreasing in the tariff, meaning that trade barriers raise the market power of firms in their own market more than they reduce it in foreign markets.

### 3.2. Market structure, growth and welfare in the global economy

It is useful to characterize the general equilibrium of the global economy in terms of *country-level variables*. It is also useful to proceed in three steps. Firstly, solve for the equilibrium number of firms. Next, evaluate growth at the equilibrium number of firms. Finally, evaluate welfare. The following proposition accomplishes the first task. Its proof contains the bulk of the solution of the model.

**Proposition 2.** The following equation determines the number of domestic firms in each country in the general equilibrium of the global economy where the global industry is in symmetric Nash equilibrium with free entry/exit, all domestic labor markets clear, and the global capital market clears:

$$\phi = \frac{L}{N} - \frac{\theta}{\xi(cN)} - 1 + \frac{\rho}{\alpha(cN)}.$$  

(NN\text{\textsubscript{trade}})
where the price elasticity of demand $\xi(cN)$ is given by (10) and

$$\alpha = \frac{1 + (\gamma + \delta)(cN - 1)}{1 + \delta(cN - 1)} \tag{11}$$

is the productivity of R&D. The number of domestic firms is monotonically increasing in the tariff.

**Proof.** See the Technical Appendix. □

Because of the role that it plays in the analysis, it is worth emphasizing that the number of domestic firms is *monotonically increasing* in the tariff. Moreover, the relation $M = cN$ implies that the number of goods available to households in each country is also monotonically increasing in the tariff.

Since Eq. (NN\text{trade}) determines the number of domestic firms in each country, one can determine the other variables recursively. Step two concerns growth. In symmetric equilibrium all firms advance at the same pace. Moreover, in each country per capita consumption expenditure, $E$, and the number of domestic firms, $N$, are constant. Under these conditions, the rate of cost reduction,

$$g = \theta \frac{\dot{Z}}{Z} = \theta \alpha L_Z,$$

determines the rate of growth of output, which is equal to the rate of growth of consumption. Each firm follows the R&D strategy

$$L_Z = \frac{L(\xi - 1)}{\xi N} - \frac{\rho}{\alpha}. \tag{12}$$

The zero-profit condition is

$$\frac{L}{\xi N} = \phi + L_Z. \tag{13}$$

Substituting (13) into (12) and using the definition of $g$ yields

$$g = \theta \frac{\phi(\xi - 1)\alpha - \rho}{1 - \theta(\xi - 1)}. \tag{GG\text{trade}}$$

To interpret this expression, consider the two steps taken to construct it.7 Firstly, each firm takes the number of firms as given and solves an intertemporal optimization problem that yields the optimal price and R&D strategy. The symmetry of the model and the general equilibrium conditions imply that this strategy can be written as a function that maps the number of firms into the firm’s

---

7The detailed derivation of (12) and (13) is in the proof of Proposition 2 in the Technical Appendix.
optimal level of R&D spending which, in turn, yields the rate of growth. This is Eq. (12). Given the number of firms, this equation describes the Nash equilibrium of the oligopoly game that is the first building block of the model. Of course, the number of firms is endogenous and must be determined by the zero-profit condition. In its usual interpretation, this equation determines the number of firms given that they undertake the optimal level of R&D spending. The second step in the construction of \((\text{GG})\) is thus to substitute (13) into the R&D strategy obtained in the first step. The result is a modified version of the R&D strategy that takes into account the fact that firms have perfect foresight about the effects of a parameter change on the zero-profit condition and correctly anticipate entry/exit.

The determination of welfare follows a similar logic. The consumption index (7), the demand curve (9), the firm’s price policy derived in the appendix and integration of (1) yield

\[
U = \frac{1}{\rho} \left[ \log(cN)^{\epsilon/\epsilon-1} + \log \left( \frac{\xi - 1}{\xi cN} \right) + \frac{\xi}{\rho} \right].
\]

The first two terms in the bracket capture the fact that utility is increasing in the number of consumption goods and the quantity of each consumption good; the term \(g/\rho\) captures the fact that utility is increasing in the rate of growth of consumption. Using \((\text{GG})\) yields

\[
U = \frac{1}{\rho} \left[ \frac{1}{\epsilon - 1} \log (cN) + \log \frac{\xi - 1}{\xi} + \frac{\theta}{\rho} \frac{\phi \theta (\xi - 1) \alpha - \rho}{1 - \theta (\xi - 1)} \right].
\]

This is the equivalent of \((\text{GG})\) in \((N, U)\) space.

3.3. Properties of the equilibrium

At this point, one should notice the importance of this paper’s characterization of the international economy as a global, differentiated oligopoly. In the literature on market structure and trade it is customary to posit monopolistic competition—as opposed to deriving it as the limit of the oligopoly model when the number of firms becomes very large—by assuming that there is a continuum of goods. With such an assumption, the price elasticity of demand faced by firms is always \(\epsilon\) regardless of the tariff. Inspection of \((\text{NN})\), \((\text{GG})\) and \((\text{UU})\) then reveals that the tariff has no effects. One way to recover effects of the tariff is to allow for non-tradeable goods. A reduction of the tariff then leads to an expansion of the tradeable sector with respect to the non-tradeable one. Such effects, however, do not stem from changes in market structure within industries. This
points out that the model set up in this paper successfully isolates the effects of the
tariff on competition in international oligopolies.

Growth and welfare are determined in the upper and lower panels of Fig. 1 by
evaluating \( GG_{\text{trade}} \) and \( UU_{\text{autarky}} \) at the equilibrium number of firms. There exists a critical value of the number of firms, \( N_0 \), such that growth is positive only if there are more than \( N_0 \) domestic firms in the market, which means that there are overall \( cN_0 \) firms in the global economy. Moreover, growth is increasing in the number of domestic firms and is bounded from above. Similar features character-

---

Alternatively, one could follow Venables (1987) and study the effects of unilateral trade policy. This, however, would not help us understand the effects of globalization as driven by the worldwide reduction in trade frictions due to the GATT-WTO or to reductions in transport costs (see, e.g., Piermartini, 1999).
ize the relation between welfare and market structure. When the global industry equilibrium supports zero growth, utility is increasing in the number of firms because this yields a larger variety of consumption goods. As the number of firms increases, the price elasticity of demand, $\xi$, increases, prices fall, and consumption increases. As the number of firms reaches the critical level $N_0$ and firms start spending on R&D, entry leads to faster growth thus realizing further utility gains.

To illustrate how the equilibrium is determined, consider the effects of country size. As population, $L$, increases in each country, the number of domestic firms, $N$, increases and each economy moves up along the $(GG_{\text{trade}})$ and $(UU_{\text{trade}})$ curves. Growth is thus an increasing function of population size that is bounded from above. In other words, the scale effect on growth is positive but eventually vanishes (see Peretto and Smulders (2001) for a general discussion of this property in this class of models). It is large for small economies with few firms and much market power and is negligible for large economies with many firms that approach monopolistic competition where the price elasticity of demand approaches the upper bound $\epsilon$ and spillovers approach the upper bound $\gamma + \delta/\gamma$.

The economic mechanism driving this property is the following. The larger population of the typical country generates a larger global market that raises firms’ gross profits. This attracts entry, raises competition and induces firms to spend more on R&D. However, as the number of firms becomes large, the positive effect of the larger market on the incentives to undertake R&D is fully offset by the fact that entry reduces the gross profits that firms earn on their innovations (this is a market share effect). Thus, when population is very large, a further increase in population size has no effect on the growth rate while it has a large effect on the number of firms. Similar results apply to welfare, with the difference that welfare is everywhere increasing in the number of firms (goods) because it includes the scale effect in levels due to the love-of-variety property exhibited by the consumption index (7).^9

4. Autarky versus free trade

This section compares autarky to free trade, the next studies the role of incremental trade liberalizations. Two definitions of autarky are possible in the context of this paper. One is that autarky entails no trade and no flows of knowledge. The other is that autarky entails no trade while flows of knowledge are allowed. In the former case, one obtains autarky by setting $c = 1$ in $(NN_{\text{trade}})$, $(GG_{\text{trade}})$ and $(UU_{\text{trade}})$. In the latter, one sets $c = 1$ in $(NN_{\text{trade}})$, $(GG_{\text{trade}})$ and $(UU_{\text{trade}})$ but leaves (11) unchanged. The two approaches produce similar results.

^9A discussion of the effects of the other parameters is available on request.
The formal analysis that follows focuses on the first one. When appropriate, it will be specified how results differ if one adopts the second definition of autarky.

The equations that characterize the general equilibrium of the autarchic economy where the industry is in symmetric Nash equilibrium with free entry/exit and all markets clear are as follows:

\[ \phi = \frac{L}{\xi N} \left[ 1 - \theta(\xi - 1) \right] + \frac{\rho}{\alpha}; \quad (\text{NN}_{\text{autarky}}) \]

\[ g = \theta \frac{\rho}{1 - \theta(\xi - 1)} \]

\[ U = \frac{1}{\rho} \left[ \frac{1}{e - 1} \log N + \log \frac{\xi - 1}{\xi} + \theta \frac{\rho}{1 - \theta(\xi - 1)} \right], \quad (\text{UU}_{\text{autarky}}) \]

where

\[ \xi = \frac{e(N - 1) + 1}{N} \quad (14) \]

is the price elasticity of demand and

\[ \alpha = \frac{1 + \gamma \delta(N - 1)}{1 + \delta(N - 1)} \quad (15) \]

is the productivity of R&D.

Comparing \((\text{GG}_{\text{trade}})\) and \((\text{UU}_{\text{trade}})\) to \((\text{GG}_{\text{autarky}})\) and \((\text{UU}_{\text{autarky}})\), one notices the following differences. Firstly, the price elasticity of demand is given by \((10)\), which is larger than \((14)\) for all \(N\). Secondly, spillovers are given by \((11)\), which is larger than \((15)\) for all \(N\). Thirdly, in \((\text{UU}_{\text{trade}})\) there is the extra term \(1/e - 1 \log c\), which captures the fact that trade allows access to a larger variety of goods. As a result, \((\text{GG}_{\text{trade}})\) and \((\text{UU}_{\text{trade}})\) are everywhere above \((\text{UU}_{\text{autarky}})\) and \((\text{GG}_{\text{autarky}})\).

Fig. 1 illustrates the market equilibrium of the global economy in comparison to the market equilibrium of the autarchic economy. Let \(N^*\) be the equilibrium number of domestic firms in autarky. A move from autarky to trade—free or restricted by tariffs—corresponds to an exogenous increase in the price elasticity of demand, \(\xi\), and R&D productivity, \(\alpha\), for all levels of the number of domestic firms, \(N\). That is, the terms \(\xi(N)\) and \(\alpha(N)\) are replaced by \(\xi(cN)\) and \(\alpha(cN)\). Alternatively, it takes the form of a change in the number of firms from \(N\) to \(M = cN\) that preserves the shape of the functions \(\xi(\cdot)\) and \(\alpha(\cdot)\). As it is apparent from the figure, the effects of such a move are potentially ambiguous because the number of domestic firms falls. To resolve the ambiguity and obtain sharp predictions, it helps to look first at the case of free trade (zero tariffs).

Let \(N^{**}\) be the equilibrium number of domestic firms in free trade. Fig. 2 and the following two propositions characterize the differences between autarky and free trade.
Proposition 3. The global number of firms in free trade is larger than the number of domestic firms in autarky, which is larger than the number of domestic firms in free trade, \( M^{**} = cN^{**} > N^* > N^{**} \).

Proof. By definition, \( N^* \) and \( N^{**} \) are such that:

\[
\phi = \frac{L}{N^*} \left( 1 - \frac{\xi(N^*) - 1}{\xi(N^*)} \right) + \frac{\rho}{\alpha(N^*)}, \quad (16)
\]

\[
\phi = \frac{L}{N^{**}} \left( 1 - \frac{\xi(cN^{**}) - 1}{\xi(cN^{**})} \right) + \frac{\rho}{\alpha(cN^{**})}. \quad (17)
\]

The right-hand side of these equations is decreasing in \( N, \xi(\cdot) \) and \( \alpha(\cdot) \). Comparing the two equations, one sees that for a given \( N \) integration raises \( \xi(\cdot) \) and \( \alpha(\cdot) \). Hence, \( N^{**} < N^* \). Now replace \( cN^{**} \) with \( M^{**} \) in the second equation to obtain

\[
\phi = \frac{cL}{M^{**}} \left( 1 - \frac{\xi(M^{**}) - 1}{\xi(M^{**})} \right) + \frac{\rho}{\alpha(M^{**})}. \quad (18)
\]

The right-hand side of this expression is decreasing in \( M \). Comparing (16) and (18), one sees that the effect of free trade is to multiply \( L \) by \( c \) while the shape of \( \xi(\cdot) \) and \( \alpha(\cdot) \) is preserved. It follows that \( M^{**} > N^* \). □

The proof of this proposition makes clear that domestic firms are squeezed out by a combination of falling cash flow and escalating R&D spending due to tougher competition. This illustrates the crucial property of the model: a move from autarky to free trade leads to exit of domestic producers while domestic
consumers gain access to foreign goods and the surviving domestic producers gain access to foreign knowledge. This change in market structure is beneficial.

**Proposition 4.** Growth and welfare in autarky are lower than in free trade, $g^{**} > g^*$ and $U^{**} > U^*$.

**Proof.** Consider (GG) and (UU), set $\tau = 0$, and replace $cN$ with $M$. This yields:

$$g = \theta \frac{\phi(\xi - 1)\alpha - \rho}{1 - \theta(\xi - 1)},$$

$$U = \frac{1}{\rho} \left[ \frac{1}{\epsilon - 1} \log M + \log \frac{\xi - 1}{\xi} + \frac{\theta \phi(\xi - 1)\alpha - \rho}{1 - \theta(\xi - 1)} \right],$$

with

$$\xi = \epsilon - (\epsilon - 1) \frac{1}{M}$$

and

$$\alpha = \frac{1 + (\gamma + \delta)(M - 1)}{1 + \delta(M - 1)}.$$

These equations have the same shape as (NN) and (UU). Hence, to compare growth and welfare in autarky and free trade one simply needs to evaluate them at, respectively, $N^*$ and $M^{**}$. Since both equations are increasing in their arguments, $g^{**} > g^*$ and $U^{**} > U^*$.

The proof of this proposition—perhaps not surprisingly—emphasizes that the growth and welfare gains due to a move from autarky to free trade stem from the increase in the number of firms that operate in each market. It is the increase in competition, together with the increase in the variety of goods and sources of knowledge spillovers, that drives the faster growth and higher lifetime utility attained in free trade.

An important aspect of this result is that the scale effect behind it—the relation between the growth rate and the size of the market—is non-linear. For economies that do not support growth in autarky, integration produces the conditions for

---

A move from autarky to free trade entails global exit in the sense that the global number of firms falls. In a world where all countries are autarchic there are $cN^*$ firms, while in a world with free trade there are $cN^{**}$ firms. Since $N^* > N^{**}$, the global number of firms in a world of free trade is lower than in an autarchic world.
positive growth if the number of integrating countries is sufficiently large. Consider, for example, an economy with fundamentals such that \( N^* < N^\alpha \), which yields \( g(N^*) = 0 \). It is possible to find a number of trading partners sufficiently large so that \( g(cN^{**}) > 0 \). This is because the scale effect is stronger the smaller the economy. Notice, moreover, that this scale effect does not run from the country’s population to its growth rate because the country does not have access to foreign labor and thus its resources endowment is unchanged. This means that the positive growth effect of integration is due to the firm’s access to a larger, more competitive market where larger means that the number of customers is larger (a demand side effect), not the number of workers (a supply side effect). Moreover, the transmission channel is tougher competition and the stronger knowledge spillovers due to the larger number of firms that the larger market supports. Finally, the growth effect of integration is bounded from above. It is large for small economies and is negligible for large economies.

To understand better what drives these results it is useful to investigate further the role of spillovers and competition in speeding up growth and the role of growth and product variety in raising welfare. Suppose that there are no knowledge spillovers. The increase in product variety then does not speed up innovation since the productivity of R&D is unchanged. Does integration raise growth in this case? Yes, because tougher price competition increases the incentives to undertake R&D. Regardless of spillovers, how much of the increase in welfare stems from larger product variety as opposed to faster growth? It depends on the pre-integration degree of competition. For economies that are very competitive and highly specialized to begin with, where \( \xi \) and \( \alpha \) are close to their upper bounds, the growth effect is small and most of the welfare gains stem from the larger product variety. For economies that start out with very little competition and very little specialization, in contrast, the growth effect can be substantial.

5. The effects of a global reduction in tariffs

Fig. 3 illustrates the effects of a global reduction in tariffs. The \((GG)_{trade}\) curve shifts up because the higher price elasticity of demand raises the incentives to engage in cost-reducing R&D. The number of domestic firms, on the other hand, is smaller and this entails a movement down along the \((GG)_{trade}\) curve. As a result, the growth effect is ambiguous. The intuition is as follows. When \( \tau \) decreases, firms experience tougher price competition; this induces them to rise R&D spending. The direct effect of lower tariffs is thus to boost growth. However, the fall in the number of domestic firms, which corresponds to a fall in the global number of firms, reduces global competition and weakens spillovers in the sense that each firm faces fewer rivals in each market in which it sells and faces fewer sources of knowledge. This induces firms to reduce R&D effort while the
productivity of R&D is lower. Hence, the indirect effect of lower tariffs tends to reduce growth. A similar story applies to welfare. The (UU\textsubscript{trade}) curve shifts up because lower tariffs raise price competition. Holding constant the number of domestic firms, this lowers the mark-ups that firms charge and this allows consumers to buy larger quantities of goods. Moreover, the tougher price competition induces firms to speed up growth. However, because some firms go out of business, the variety of consumption goods and of sources of spillovers falls and this has a negative effect on welfare and a negative effect on growth, which has a further negative effect on welfare. These forces show up as a movement down along the new, higher (UU\textsubscript{trade}) curve.
Overall, there is a tension between the benefits of product and knowledge variety that tariffs protect and the costs of smaller firm size that leads to less exploitation of economies of scale. In other words, there is a tension between external and internal increasing returns. Tariffs raise the number of firms and thus foster external economies—love of variety for consumers and technological spillovers for producers. On the other hand, the larger number of firms prevents the individual firm from fully realizing internal economies of scale due to fixed management costs and dynamic increasing returns in R&D. The effects of a reduction in tariffs can be called, respectively, the homogenization and rationalization effects. It is important to notice that the homogenization effect described here is opposite to the diversity effect (Helpman and Krugman, 1985, p. 264) that drives the gains that countries realize when moving from autarky to free trade. To summarize:

**Proposition 5.** The growth and welfare effects of a global reduction in tariffs are negative (positive) if increasing returns external to the firm due to love of variety and spillovers dominate (are dominated by) increasing returns internal to the firm due to fixed management and R&D costs.

The discussion so far has focused on the incremental effects of the tariff. One, however, would like to know the global shape of the relation between growth and welfare and the tariff. Autarky and free trade—which were compared in the previous section—provide two anchors that help pin down the behavior of the model’s equilibrium as the tariff ranges from zero to infinity. The following limiting results also help.

**Proposition 6.** For \( \tau \to \infty \) \((NN_{\text{trade}})\) yields \( N = N_\infty \), \( g = g_\infty \), and \( U = U_\infty \) with \( N^{**} < N_c < N^* < cN^{**} < cN_\infty \), \( g^* < g_\infty < g^{**} \) and \( U^* < U_\infty \).

**Proof.** See the Technical Appendix. \( \square \)

Recall that autarky can be defined as either no trade with knowledge spillovers or no trade and no knowledge spillovers. The two cases yield qualitative similar results. If one defines autarky as neither trade nor knowledge spillovers, then one finds that even as the tariff goes to infinity the growth rate of the trading economy remains above the growth rate of the autarchic economy, \( g_\infty > g^* \), simply because of the presence of knowledge spillovers which are not affected by the tariff. If one defines autarky as no trade but allows for knowledge spillovers, then one quite naturally finds \( g_\infty = g^* \).

More remarkable is the result that welfare does not converge to the autarky level.
when the tariff goes to infinity, \( U^* < U_\infty \), even if one defines autarky as no trade with knowledge spillovers and the growth rate does converge to the autarky one. This is because there is a wedge between the \((UU_{\text{autarky}})\) and \((UU_{\text{trade}})\) curves due to the fact that consumers have access to the entire menu of goods produced in the world for all possible levels of the tariff. In other words, the relation \( M = cN \) prevents the \( UU \) locus converging to the one that applies in autarky. There is thus a crucial difference between the exercise that compares autarky to free trade and the exercise that looks at an incremental reduction in tariffs.\(^{11}\) The next section discusses the implications of this property in detail. The remainder of this section uses Proposition 6 to impose restrictions on the overall relation between growth and welfare and the tariff.

Since the number of domestic firms is monotonically increasing in the tariff, one can conclude that for all values of the tariff the number of domestic firms and the global number of firms are, respectively, smaller and larger than the number of domestic firms in autarky. In other words, autarky yields the largest number of domestic producers and the smallest number of available goods, that is, the smallest number of firms operating in the domestic market. This provides the intuition behind the next result: regardless of whether autarky entails spillovers or not, trade dominates autarky for all values of the tariff in the sense that growth and welfare are always higher. The reason is that, regardless of the value of the tariff, trade allows a country to consume a larger variety of goods than it produces, while the larger number of firms generates tougher competition and stronger spillovers that support faster growth. The limiting result that \( g^* \leq g_\infty \) (with strict inequality if autarky rules out knowledge spillovers) allows one to conclude that the relation between growth and the tariff is at most hump-shaped—growth could be increasing for small values of the tariff but eventually must be decreasing in the tariff. In fact, inspection of \((GG_{\text{trade}})\) suggests that growth is monotonically decreasing in the tariff if the direct effect of the tariff on the price elasticity of demand dominates the indirect effect,

\[
\frac{dg}{d\tau} = \frac{\partial g}{\partial \tau} + \frac{\partial g}{\partial N} \frac{\partial N}{\partial \tau} < 0.
\]

This sufficient condition is likely to hold in reality.

Unfortunately, there is not a similar result for welfare because of the extra term \(1/\epsilon - 1 \log (cN)\) in \((UU_{\text{trade}})\), which does not allow one to prove that \( U_\infty < U^{**} \). In other words, the relation between welfare and the tariff could be monotonically increasing. This, however, would require implausibly strong external increasing

\(^{11}\)This property does not depend on the definition of autarky as a situation where knowledge spillovers across countries are zero. Rather, it follows from the property of the CES preferences that the marginal utility of each good is infinity for a quantity that approaches zero. This means that domestic consumers buy all foreign goods even if the tariff is very large so that the effect of the tariff is absorbed by the quantity of each good, not by the number of goods.
returns. Overall, the model suggests that the relations between growth and welfare and the tariff are either monotonically decreasing or hump-shaped. Fig. 4 illustrates this conclusion.\textsuperscript{12}

\textsuperscript{12}These results establish an important difference between this paper and the contributions of Rivera-Batiz and Romer (1991b) and Baldwin and Forslid (1999, 2000) who obtain U-shaped relationships between the ad-valorem tariff and growth. Firstly, in their models the tariff affects only a firm's level of sales in foreign markets and not the markup of price over marginal cost. Secondly, the number of firms is fixed so that there is no additional feedback due to the endogenous structure of the market. The second channel is most important. If the number of firms is fixed, growth increases linearly with population size and decreases monotonically with the tariff. The latter result stems purely from the direct effect of the tariff on the price elasticity of demand since this version of the model shuts down the possibly offsetting indirect effect.
6. Summary and conclusions

This paper has discussed the role of economic integration in a model of endogenous growth where the size of the firm and the interactions between growth and the endogenous structure of the market play a crucial role. In contrast to previous work based on first-generation models of endogenous innovation (e.g. Grossman and Helpman, 1991; Rivera-Batiz and Romer, 1991a,b; Baldwin and Forslid, 1999, 2000), the paper focused on a model of Schumpeterian innovation undertaken in-house and characterized by partial knowledge spillovers. This allowed analysis of the effects of integration on competition, R&D activity, and ultimately on growth and welfare, in a world economy populated by global oligopolists.

The paper has provided a characterization of the conditions under which integration improves growth and welfare. The results differ quite drastically according to whether one compares autarky to free trade or looks at the effects of an incremental trade liberalization.

Consider first a move from autarky to trade. Each firm faces a market where the number of customers and the number of competing firms have increased in the same proportion, so that sales per firm do not change, but where tougher competition reduces firms’ market power and yields lower mark-ups and profit margins. As a result, in each country the domestic number of firms falls. From the viewpoint of the consumer, however, exit of domestic producers is more then compensated by entry of foreign firms that are now free to export to the consumer’s country. More generally, integration produces a major change in industrial market structure: the integrated market is larger and concentration is lower. Firms in a larger, more competitive market spend more on R&D—which is also more productive because spillovers are stronger—so that growth is faster. Welfare increases because the global number of firms in the integrated market is larger than the number of domestic firms in autarky, while tougher competition forces firms to charge lower mark-ups and prices. This mechanism is at work also in the case of a move from autarky to trade restricted by tariffs. Hence, one can conclude:

Summary 7. (Gains from trade) A move from autarky to trade—free or restricted by ad-valorem tariffs—raises growth and welfare because it generates a larger, more competitive market where producers have access to a larger number of customers and a larger body of knowledge spillovers, while consumers have access to a larger variety of cheaper goods. Hence, the quantity, variety and growth of consumption rise.

The effects of an incremental reduction in tariffs are drastically different. The competitive effect of lower tariffs reduces firms’ market power and yields lower mark-ups and profit margins. The crucial result is that the global number of firms
falls. The consumer now sees exit of domestic and foreign firms. Integration, therefore, leads to a market that is of the same size—because the number of customers has not changed—while concentration is higher. There is thus a tension between the effect of the lost protection in the domestic market and the effect of the larger share of the global market that each firm now commands. On the other hand, spillovers are now weaker because some sources of knowledge are lost due to exit. In principle, the growth effect is ambiguous. Similarly, the welfare effect is ambiguous because the variety of consumption goods is smaller while, partially compensating this, firms have less market power and charge lower mark-ups. One can thus conclude:

Summary 8. (Incremental trade liberalization) A global reduction in tariffs triggers a trade-off between internal and external economies of scale. The variety of sources of knowledge spillovers and of consumption goods falls. This homogenization effect tends to reduce growth and welfare. On the other hand, firms become larger while tougher competition induces them to reduce mark-ups, which allows consumers to buy larger quantities of the available goods, and raise R&D spending. This rationalization effect tends to raise growth and welfare. The balance of these forces is such that growth and welfare generally increase with the reduction in tariffs except at sufficiently low values of the tariff where it is possible that they decrease.

These exercises suggest that to determine the benefits of integration one needs to be very clear about the question that one is asking. Where does the difference between them come from? In the case of a global reduction in tariffs, there cannot be entry of foreign firms in the domestic market because all firms operate in all markets to begin with. Hence, the only effect that the model can capture is the loss of market power of domestic firms due to lower tariffs with the consequent fall in net profits that squeezes out some firms. As a consequence, looking at an incremental trade liberalization gives a different answer from comparing autarky to trade.

The first exercise produces results that match conventional intuition on the benefits of international trade: larger markets, more competition, more choice. This is because one is looking at a shock that entails an increase in the number of firms that serve domestic consumers. The second exercise shuts down this channel. In so doing, it focuses attention on the opposite case where the shock entails a reduction of the number of firms that serve domestic consumers. In other words, the results for the case of incremental liberalization depend crucially on the property that the wedge between the number of domestic producers and the number of goods available to consumers does not depend on the tariff. This means that the number of goods produced and the number of goods consumed are constrained to move in the same direction and thus from the viewpoint of the domestic consumer exit of domestic firms is bound to yield exit of domestic and foreign firms. As a result,
the growth and welfare effects of integration depend on the relative importance of increasing returns internal to the firm and increasing returns external to the firm. A global reduction in tariffs is beneficial when the love of variety effect is weak and technological spillovers do not change much with the number of firms. To some extent, these results also match conventional intuition expressed in complaints like ‘globalization is making the world more homogeneous’ that it is common to hear today.

An interesting direction for future research is to develop a model that reconciles these different perspectives. Such a model should have the property that the wedge between the number of goods produced domestically and the number of goods consumed be a function of trade policy. To my knowledge there is no model with this property in the literature.

Acknowledgements

Many thanks to Art Benavie, Michelle Connolly, Kent Kimbrough, Mike Salemi, Steve Meardon, Peter Thompson, workshop participants at UNC Chapel Hill and Duke, and two anonymous referees for their comments. Special thanks to Jim Herrigan, Co-Editor of this journal, for provocative questions that led to a better paper and for his editorial advice. I am solely responsible for all remaining errors.

References