Agglomeration and economic development: Import substitution vs. trade liberalisation*

Diego Puga

University of Toronto, and CEPR

Anthony J. Venables

London School of Economics, and CEPR

CENTRE for ECONOMIC PERFORMANCE
Discussion Paper No. 377, November 1997

Revised, November 1998

ABSTRACT: This paper analyses a model of economic development in which international differences in industrial structure and income are caused by the agglomeration of industry in a subset of countries. Economic development may not be a gradual process of convergence by all countries, but instead involve countries moving sequentially from the group of poor countries to the group of rich countries. The role of trade policy in promoting industrialisation is studied. While both import substitution and unilateral trade liberalisation may be 'successful' in attracting industry, they attract different sectors and welfare levels are higher under trade liberalisation.

* This paper was produced as part of the Programme on International Economic Performance at the UK Economic and Social Research Council funded Centre for Economic Performance, London School of Economics.

KEY WORDS: import substitution, trade liberalisation, agglomeration, linkages.

JEL CLASSIFICATION: F12, O14, R12.

Addresses:

Diego Puga
Department of Economics
University of Toronto
150 St. George Street
Toronto, Ontario M5S 3G7
Canada

Email: dpuga@chass.utoronto.ca

http://dpuga.eco.utoronto.ca

Anthony J. Venables
Department of Economics
London School of Economics
Houghton Street
London WC2A 2AE
United Kingdom

Email: a.j.venables@lse.ac.uk

1. Introduction

The key determinants of a country's economic development are usually taken to be some combination of its factor endowment, technology, institutional structure and policy stance. While not denying the importance of these considerations, in this paper we explore a radically different view of economic development and underdevelopment, based on the idea that economic activity may agglomerate spatially. In this case it is possible that countries with similar, or even identical, underlying characteristics may nevertheless have different economic structures and income levels. Economic underdevelopment is a manifestation of the spatial pattern of agglomeration, and development occurs as this pattern changes, with industry spreading from existing concentrations to new ones.

Analysis of spatial agglomeration of industry has been formalised in recent work in economic geography (see for example Krugman and Venables (1995), Puga (1998)), and the goal of the present paper is to draw out the implications of this approach for economic development. What forces are conducive to the spatial concentration of industry, and what to its spread from country to country? If industrialisation does spread, what form does development take? What is the role of policy — in particular trade policy — in promoting industrialisation?

The basis of our analysis is a model in which there are forces which may cause industry to concentrate in a few locations. We create these forces from three main ingredients. First, there are transport costs or other trade barriers, and these create incentives for firms to locate close to customers and to suppliers. Second, firms have increasing returns to scale, which play the role of forcing firms to choose where to produce;[1] of course, with increasing returns we must handle the problem of market structure, and this we do by assuming monopolistic competition. The third ingredient is the presence of input-output linkages between firms. These linkages create an incentive for firms to locate close to other firms — their suppliers and customers.

[1] Without increasing returns, if factors are uniformly distributed, every location can become an autarkic economy producing all goods at an arbitrarily small scale (Scotchmer and Thisse (1992)) .

Krugman and Venables (1995) showed how this combination of forces creates the possibility that industry concentrates in one country, and established the dependence of the equilibrium on transport costs. Here we use the framework to study two issues which we think illuminate the process of economic development.

The first is to consider the spatial implications of growth inn world manufacturing relative to other tradeable sectors. This growth increases demand for labour in established manufacturing countries, opening up a larger and larger wage gap between these and other countries. At some point this wage gap becomes unsustainable, and industry starts to spill over to low wage economies. We analyse this process and establish that it does not lead to steady development of all low wage economies, but instead to rapid industrialisation of countries in turn. The logic of spatial agglomeration implies that development cannot proceed simultaneously in all countries. Instead there is a group of rich countries and a group of poor ones, and development takes the form of countries being drawn in turn out of the poor group, and taken through a process of rapid development into the rich group. We think that this is an insightful way of thinking about the spread of industry in a number of contexts, for example from Japan to its East Asian neighbours.

The second issue we address is the role of developing country trade policy in promoting or hindering industrialisation. While recent papers in economic geography have focussed on the location effects of reciprocal reductions in trade costs, in this paper we look at the effects of unilateral changes in trade barriers. We show that either unilateral trade liberalisation or import substitution policies may be used by the low wage economy to attract industry, but these two policies work through very different mechanisms. Although they are both superficially 'successful' in attracting industry, they have different effects on economic welfare, with trade liberalisation yielding higher welfare than import substitution policies. We analyse this in an aggregate model, and then in a multi-industry variant of the model calibrated to South Korean input-output and demand data. We use the calibration to show the different sectoral impacts of trade liberalisation and import substitution, and to confirm the different welfare outcomes generated by the two policies.

Our approach in this paper can be thought of as a formalisation of earlier ideas in development economics, in particular the role of forward and backward linkages, as emphasised by Hirschman (1958) and others. These linkages are of no particular economic significance in a perfectly competitive environment, but combined with the other ingredients sketched out above, they create pecuniary externalities between the location decisions of firms, and it is this that creates the incentives for agglomeration of industry. To see how this works, suppose that there is expansion of a downstream industry. This creates a backward linkage, expanding demand for intermediate goods, raising profits in an upstream industry, and attracting entry of upstream firms, which in turn may *decrease* the price of intermediates. How does this perverse price response occur? Entry of firms may make the industry more competitive, squeezing price cost margins and reducing price, or may lead to entry of more varieties, reducing a price index of industry output as a whole. This perverse price response constitutes a forward linkage — expanding the upstream industry reduces the costs of the downstream. Putting this together, we see a positive feedback, such that an expansion of the downstream industry makes the industry more profitable, encouraging further expansion.

The process described above is also suggestive of 'cumulative causation' — the presence of more downstream firms attracts more upstream firms which in turn attract downstream firms and so on. Again this is reminiscent of old traditions in development and regional economics (for example in the work of Perroux (1955), Myrdal (1957), Hirschman (1958), Harris (1954), and Pred (1966)), as well as some newer approaches to development economics.[2]

[2] Murphy, Shleifer and Vishny (1989) model a 'big push' in which increasing modern sector employment raises aggregate demand, thereby increasing the profitability of modern sector firms. Their model works through aggregate demand, rather than intermediate goods, so has no forward linkages. It also assumes a closed economy.

2. The model

We set out the model for the case of two countries and two sectors — manufacturing and agriculture — relegating a statement of the full multi-country and multi-industry model to the Appendix.

Agriculture

Each country is endowed with quantities L_i and K_i of labour and arable land (for countries $i = 1, 2$), both of which are internationally immobile. Agriculture is perfectly competitive and produces a homogenous output, which we choose as *numéraire*, and assume costlessly tradeable.[3] Its production function, F, is defined over labour and land, and has constant returns to scale. If manufacturing employment in country i is denoted m_i and the labour market clears, then agricultural output is $F(L_i - m_i, K_i) = K_i f\left((L_i - m_i)/K_i\right)$. The country i wage is

$$w_i = f'\left((L_i - m_i)/K_i\right) \tag{1}$$

Industry

The industrial sector is monopolistically competitive, producing differentiated manufactures under increasing returns to scale. As in Krugman and Venables (1995), we assume that the output of firms in the industry is used both as a final consumption good and as an intermediate good for use in the same industry. The set of firms in each country is endogenously determined by free entry and exit, and denoted by N_i. The cost function for an industrial firm (firm k) in country i is

$$C_i(k) = q_i^{\mu} w_i^{(1-\mu)} \left(\alpha + \beta x_i(k)\right) . \tag{2}$$

---

[3] Adding a trade cost in agriculture does not change the qualitative results of the model providing that trade in agriculture occurs in equilibrium. For details of this see Fujita, Krugman and Venables (1999).

$x_i(k)$ is the firm's output, and the fixed and marginal input requirements, α and β, are the same for all varieties and all countries. The input is a Cobb-Douglas composite of labour, with share (1–μ), and an aggregate of the differentiated industrial goods, with price index q_i and share μ.[4] This price index takes a CES form, so is defined by

$$q_i \equiv \left[\int_{k \in N_i} \left(p_i(k) \right)^{(1-\sigma)} \mathrm{d}k + \int_{k \in N_j} \left(\tau_i \, p_j(k) \right)^{(1-\sigma)} \mathrm{d}k \right]^{1/(1-\sigma)}, \qquad j \neq i, \tag{3}$$

where $p_j(k)$ is the producer price of variety k produced in country j. Shipment of these products is subject to iceberg trade costs: (τ_i units must be shipped from country j in order that one unit arrives in i). Product differentiation is measured by the elasticity of substitution between varieties of good, σ, and captures the idea that firms benefit from access to a wider range of intermediate goods (following Ethier (1982)).

Demand

There is a single representative consumer in each country, who has quasi-homothetic preferences over agriculture (the *numéraire)* and the CES aggregate of industrial goods. Hence there is a 'love for variety' on the consumer side, as in Dixit and Stiglitz (1977). The indirect utility of the consumer in country i is

$$V_i = q_i^{-\gamma} 1^{-(1-\gamma)} (y_i - e_0) \ . \tag{4}$$

where y_i is income, and e_0 is the subsistence level of agricultural consumption. Notice that we use the same price index for varieties of industrial goods in consumption as in production. This is not necessary for our results, but does much to simplify analysis.

Each product is sold in each country, and the demand for variety k produced in country i, $x_i(k)$, can be derived from (3) and (4) as

[4] In the full multi-industry model given in the Appendix each industry uses inputs from all other industries, with value shares given by the input-output matrix of the economy.

$$x_i(k) = \left(p_i(k)\right)^{-\sigma}\left(e_i q_i^{(\sigma-1)} + e_j q_j^{(\sigma-1)} \tau_j^{(1-\sigma)}\right), \tag{5}$$

where e_i is total expenditure on manufactures in country i. Since manufactures are used both as final consumer goods and as intermediates, e_i is given by;

$$e_i = \gamma\left[w_i m_i + K_i f\left((L_i - m_i)/K_i\right) - e_0\right] + \mu \int_{k \in N_i} C_i(k)\,\mathrm{d}k \ . \tag{6}$$

The first term in (6) is the value of consumer expenditure on manufactures, and the second the value of intermediate demand. In the square brackets, the first term is wage income in manufacturing, and the second is income generated in agriculture; the consumer devotes the first e_0 of income to agriculture, and proportion γ of income above this level to expenditure on industrial products. The final term in (6) is intermediate demand, generated as firms spend fraction μ of their costs on intermediates.

Supply

Each variety is produced by at most one firm, so the firm producing variety k faces demand curves (5) and cost function (2). Since all products produced in location i are symmetric (they have the same technology and demand curves) we drop the label for individual varieties. The profits of a single representative country i firm are therefore

$$\pi_i = p_i x_i - q_i^{\mu} w_i^{(1-\mu)}\left(\alpha + \beta x_i\right). \tag{7}$$

Each firm's perceived elasticity of demand is σ, so the equality of marginal revenue to marginal cost necessary for profit maximisation takes the form

$$p_i(1 - 1/\sigma) = \beta q_i^{\mu} w_i^{(1-\mu)} \ . \tag{8}$$

We choose units of measurement for output such that $\beta\sigma = \sigma - 1$, giving an equilibrium price of $p_i = q_i^{\mu} w_i^{(1-\mu)}$. Using this pricing rule in the definition of profits, it follows that firms break even if their sales are equal to level x^* given by

$$x^* = \alpha / \sigma . \tag{9}$$

If a firm were to sell less than x^* then it would make a loss, and more than x^*, a profit. At equilibrium profits are exhausted by free entry and exit. Denoting the number (mass) of firms in region i by $n_i \equiv \#N_i$, we therefore have

$$(x_i - x^*)\, n_i = 0 \ , \quad x_i \le x^* \ , \quad n_i \ge 0 \ . \tag{10}$$

To complete characterisation of equilibrium we need only specify manufacturing labour demand. The manufacturing wage bill, $m_i w_i$, is fraction (1–μ) of costs (equal, in equilibrium, to the value of output), so

$$m_i w_i = (1 - \mu)\, n_i C_i = (1 - \mu)\, n_i p_i x^* . \tag{11}$$

We have already seen how labour market clearing determines the wage rate (equation (1)).

Equilibria

Equilibria of the model are given by equations (1)–(11). What can be said about them? This question is most easily answered if we assume that the two economies have the same factor endowments. We therefore set

$$L_1 = L_2 = A, \qquad K_1 = K_2 = A, \tag{12}$$

(where the fact that the quantities of L and K are the same is just a choice of units).

There is certainly now a symmetric equilibrium in which industry is equally divided between countries, although this equilibrium may not be stable. More interestingly, there may also be equilibria in which manufacturing is concentrated in a single country. To establish whether or not such an equilibrium exists, we *assume* that all manufacturing

is concentrated in one country (say country 1) and then see if it is profitable for any firm to start production in country 2. If not, then the hypothesised concentration of manufacturing in country 1 is an equilibrium.

Let us assume then that $n_2 = 0$. The price indices of expression (3) reduce to

$$q_1 = n_1^{1/(1-\sigma)} p_1 , \qquad q_2 = n_1^{1/(1-\sigma)} p_1 \tau_2 . \tag{13}$$

Sales of each firm in country 1 are,

$$x_1 = p_1^{-\sigma} q_1^{(\sigma-1)} (e_1 + e_2) = x^* , \tag{14}$$

where the first equation comes from using (13) in (5), and the second from the fact that country 1 industry equilibrium occurs when n_1 has adjusted to gives zero profits, so each firm sells output level $x_1 = x^*$

Suppose now that a firm starts producing in country 2. Its sales are (from (5) and (13)),

$$x_2 = p_2^{-\sigma} q_1^{(\sigma-1)} \left[e_1 \tau_1^{(1-\sigma)} + e_2 \tau_2^{(\sigma-1)} \right] \tag{15}$$

Relative goods price can be derived from (7) and (13) as

$$\left(\frac{p_2}{p_1} \right) = \tau_2^{\mu} \left(\frac{w_2}{w_1} \right)^{(1-\mu)} , \tag{16}$$

so taking the ratio of the sales equations (14) and (15) we obtain,

$$\frac{x_2}{x^*} = \left(\frac{w_1}{w_2} \right)^{\sigma(1-\mu)} \tau_2^{-\sigma\mu} \left[\frac{e_1}{e_1 + e_2} \tau_1^{(1-\sigma)} + \frac{e_2}{e_1 + e_2} \tau_2^{(\sigma-1)} \right] . \tag{17}$$

This expression provides the criterion which determines whether or not agglomeration of industry in country 1 is an equilibrium. If the expression has value greater than unity then an entrant in 2 can sell more than is required to break even ($x_2 > x^*$), so agglomeration in 1 is not an equilibrium. Conversely, if the expression is less than unity then concentration in 1 is sustainable — it is not profitable for any firm to produce in country 2.

The magnitude of this expression is determined by three forces. First, the factor market. The larger is w_1/w_2 the higher is x_2 and hence the less likely it is that agglomeration can be sustained.[5] Unsurprisingly, the larger are the wage differences associated with agglomeration, the more likely it is that production in country 2 will be profitable.

Second, forward linkages. The term $\tau_2^{-\sigma\mu}$ captures the fact that a firm setting up in country 2 would have to import all its intermediate goods, and pay τ_2 more for them than do firms in country 1. Transport costs or other barriers to country 2 imports ($\tau_2 > 1$) make this term less than unity, reducing x_2, and making it less profitable for a firm to start producing in country 2. Essentially, the term captures the forward linkages foregone by locating in country 2, away from intermediate suppliers.

The third force is backward linkages, and these are captured by the term in square brackets. To interpret this, suppose that $\tau_1 = \tau_2 > 1$ and that $e_1 + e_2$ is constant. Since $\tau_1^{(1-\sigma)} < \tau_2^{(\sigma-1)}$, an increase in e_1 and reduction in e_2 reduces the size of this term, meaning simply that a transfer of expenditure from market 2 to market 1 reduces the sales of a firm in 2. With manufacturing concentrated in country 1 we have $e_1 > e_2$, making for a relatively small value of this term. The effect therefore captures the backward linkages foregone by not being close to industrial consumers.

Equation (17) has endogenous variables on the right hand side, but these can be found as follows. By construction, all of country 2's labour force is in agriculture, so $m_2 = 0$. Country 1's agricultural labour force must therefore adjust to equate world supply and demand for agriculture, that is to satisfy,

---

[5] The relative wage term enters with exponent $\sigma(1-\mu)$ because labour accounts for $(1-\mu)$ of costs, and a price increase reduces sales according to elasticity σ. Relative wages are endogenous, and determined in equations (18) and (19) below.

$$2e_0 + (1-\gamma)\left[m_1 f'(1-m_1/A) + Af(1-m_1/A) + Af(1) - 2e_0\right] = Af(1-m_1/A) + Af(1) \qquad (18)$$

The right hand side of this expression is food production and the left hand side is demand, coming from the subsistence requirement of the representative consumer in each country, e_0, plus proportion $(1-\gamma)$ of world income in excess of this subsistence requirement. This equation gives m_1 as a function of parameters of the model. We can then find the wage rates from equation (1)

$$w_1 = f'(1-m_1/A), \qquad w_2 = f'(1) \ , \qquad (19)$$

and manufacturing expenditure levels,

$$e_1 = \gamma\left[w_1 m_1 + Af(1-m_1/A) - e_0\right] + \mu(e_1+e_2), \qquad e_2 = \gamma\left[Af(1) - e_0\right] . \qquad (20)$$

(Derived from (6), noting that manufacturing costs equal total expenditure on manufactures, and all are incurred in country 1).

The dependence of x_2/x^* on some of the parameters of the model is illustrated by the lines in figure 1, in which the vertical axis gives A/e_0 and the horizontal the country 2 import barrier, τ_2.[6] Lines correspond to different values of the input-output linkage, μ, and each line is the locus along which $x_2/x^* = 1$. We call this relationship the sustain curve, because it delimits the region of parameter space where agglomeration is sustainable. Below the sustain curve $x_2/x^* < 1$, so that concentration of manufacturing in country 1 is an equilibrium, while above it we know that concentration is not an equilibrium (it is profitable for a firm to start production in country 2). As the economy passes through this line a bifurcation occurs — the qualitative structure of equilibria changes — and in remaining sections of the paper we explore this transition.

Figure 1 provides a framework for the analysis of the rest of the paper, and in subsequent sections we shall discuss the shape of the $x_2/x^* = 1$ curve in greater detail.

---

[6] The figure is computed using solutions of (18)-(20) in (17). Key parameters are μ, and σ which we set at 5, corresponding to a price - marginal cost mark up of 25%. Full details of parameters used in this and subsequent figures are given in the Appendix.

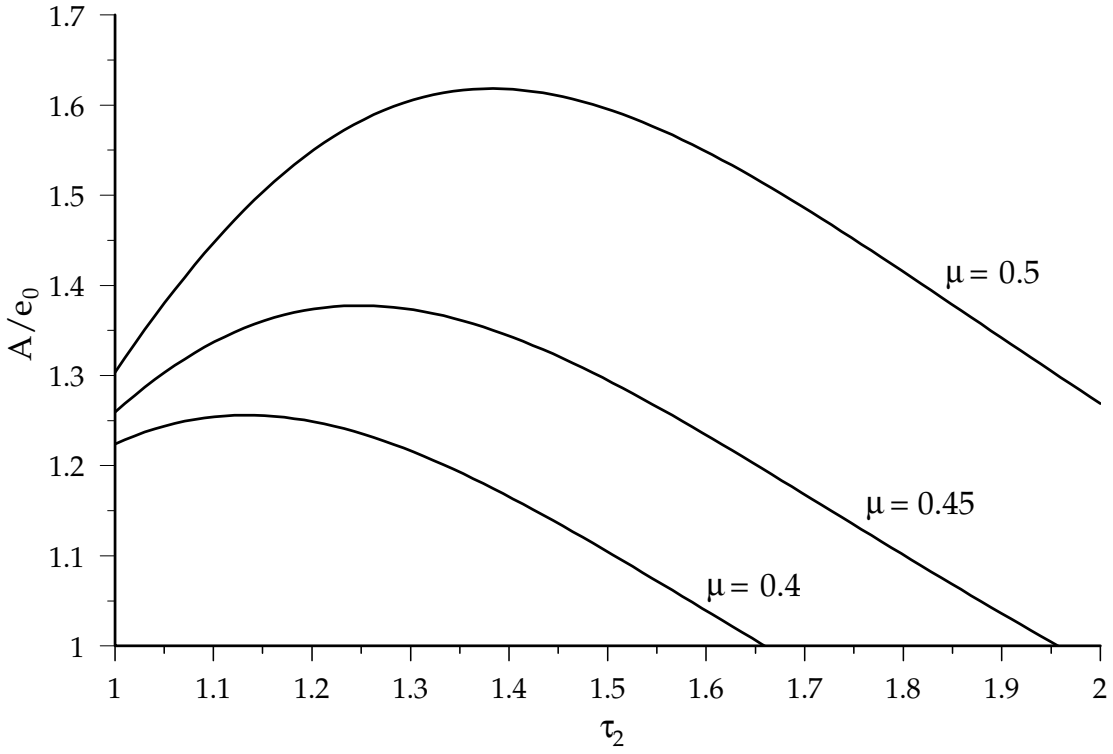

FIGURE 1
Sustainable agglomeration

In the next section we consider the effects of increasing A/e_0, and show what happens as we move upwards through the bifurcation set. In section 4 we look at developing country trade policy — that is, at horizontal movements across the figure. In both these sections we shall at various points move to a more general model, with more countries and more industries. However, much of the intuition for our results comes from equation (17) and its illustration on figure 1.

3. Growth and the spread of industry

We now turn to investigating the implications of growth in the world economy as a whole for the spatial location of economic activity. Since we do not seek to explain growth, we simply assume that exogenous technical progress augments the productivity

of all primary factors in all countries equally. What are the implications of such technical progress for the location of manufacturing production?

A completely homogenous process of economic growth—raising supply and demand for each sector in each country in the same proportion—will not have any spatial effects. But if demand for manufactures rises faster than demand for agriculture, then relative price changes will occur and, as we shall see, this can trigger industrial relocation.[7] We capture a relatively rapid growth of demand for manufactures by using the linear expenditure system with a positive level of subsistence expenditure on agriculture, e_0, so that growth of household income is associated with proportionately faster growth of demand for manufactures.[8]

In terms of the model, we assume then that A increase through time, and reinterpret w_i and m_i as wages and employment of efficiency units of labour. Starting from a situation in which manufacturing is concentrated in country 1, we see from the expression for country 1 manufacturing employment, (expression (18)) that such growth causes an equiproportionate increase in m_1 only if $e_0 = 0$, while if $e_0 > 0$, then m_1 increases more than proportionately with A. Turning to the wage equations (19), if m_1/A increases, then so too does the country 1 wage per efficiency unit, both absolutely and relatively to the country 2 wage.

The implications of this for the location of industry can be seen from equation (17). The increase in w_1/w_2 raises x_2, working against sustainability of the agglomeration in country 1. But as the wage in country 1 increases, so does country 1's share of world expenditure on manufactures, $e_1/(e_1 + e_2)$, tending to decrease x_2 and make 1 a relatively more profitable location. The net effect depends on parameters of the model. If the share of manufactures in consumption, γ, is very small then the effect on wages will be small and the agglomeration will always be sustainable. However, if γ is large enough then the wage effect will come to dominate, and at a high enough value of A the

[7] Between 1960 and 1990 world value added in manufacturing increased fourfold while world GDP increased threefold.

[8] Econometrically estimated values of e_0 are positive, see for example Lluch and Powell (1975).

agglomeration will become unsustainable.[9] This corresponds to the case illustrated on figure 1, where as A increases so the economy crosses the $x_2 / x^* = 1$ locus.

When $x_2 / x^* = 1$, entry of a manufacturing firm in country 2 is profitable. What then happens as A increases further? The presence of some manufacturing in country 2 creates forward and backward linkages (reduces q_2 and increases e_2), but also has the effect of narrowing the wage gap between the countries (raising w_2/w_1). If the linkage effects are very powerful compared to the wage effects then there may be discontinuous change — production in country 2 becomes profitable enough that the two economies jump to the symmetric equilibrium. Discontinuities are avoided if the wage effects are relatively strong, in which case as A increases further so the two economies converge smoothly to the symmetric equilibrium.[10]

To draw out the economics of the process we illustrate it not for our two country example, but by numerical simulation of the model for a world of four identical countries. Figure 2 plots real wages per efficiency unit of labour in each economy relative to the average for all economies as A increases. The vertical axis is this relative real wage per efficiency unit, ω_i, and the horizontal axis the exogenous technological progress parameter A.

At low A, all industry is in country 1, but growth in A causes an increase in demand for manufactures and hence a divergence of wages — the country 1 wage reaching, for our parameter values, nearly one and a half times the level of wages elsewhere in the world. Despite this wage gap it is not profitable for any firm to move out — the forward and backward linkages received by being in country 1 compensate for the higher wage. But as pressure builds up in country 1, so A reaches the bifurcation point at which production in the other countries becomes profitable. Industrialisation commences in all of them, but as their volume of manufacturing increases so do the associated linkages and pecuniary externalities. There comes a point at which simultaneous industrialisation

---

[9] In the former case the curve is unbounded for some intermediate values of τ_2. Puga (1998) investigates further.

[10] Details of the nature of the bifurcation (whether or not it is discontinuous) are examined by Puga (1998) and Krugman, Fujita and Venables (1999) in similar models, and we do not pursue them here.

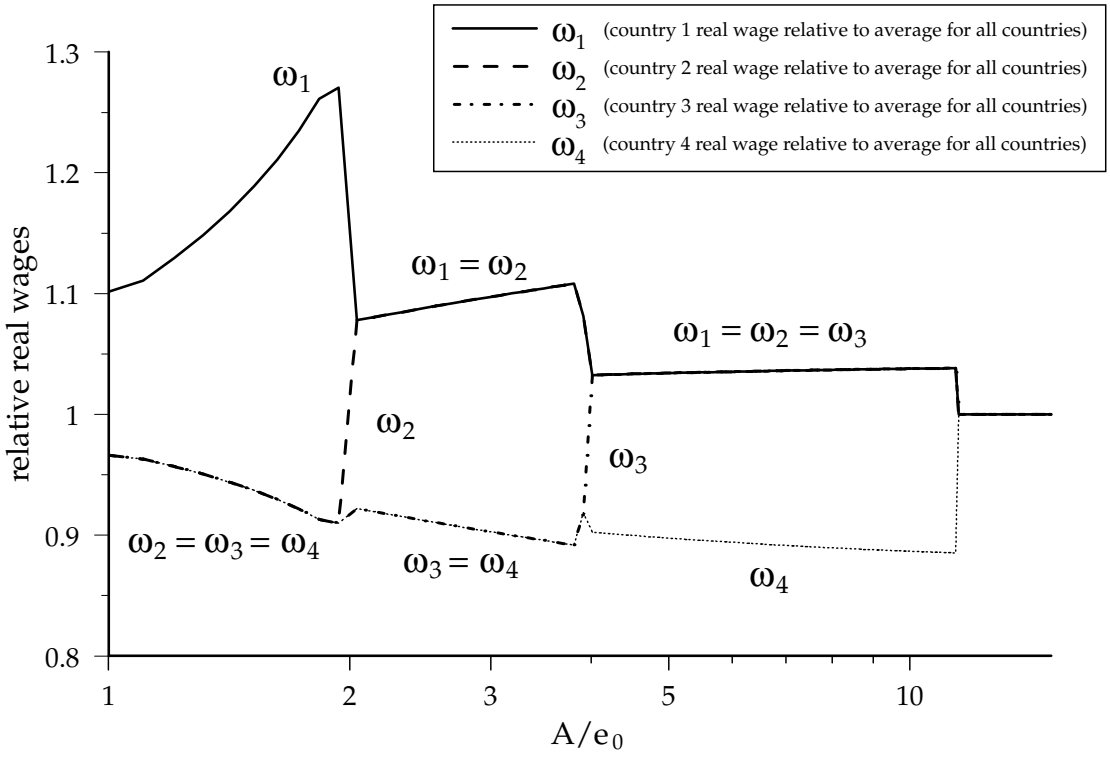

FIGURE 2
Waves of industrialisation

in *all* of them ceases to be a stable equilibrium — if one got slightly ahead of the others then its lead would cumulate.[11] This means that just one of the countries (call it country 2) gains manufacturing; this country's wage path is illustrated by the first dashed line, and we see very rapid convergence with country 1. Country 1 suffers both a relative and absolute real wage decline, as it loses a share of its manufacturing to the newly industrialised country.

The other countries remain specialised in agriculture following this first transition, but continuing growth now starts to raise real wages in 1 and 2 relative to these countries. This continues until another critical value is reached at which point industry

[11] Entry and exit of firms occurs in response to instantaneous profits, so the dynamic system is $dn_i/dt = k\pi_i$, and stability is defined with respect to this system. Analysis of the stability of models of this type is undertaken in Fujita, Krugman and Venables (1999), and analytical results on the instability of simultaneous development by two economies are available on request from the authors.

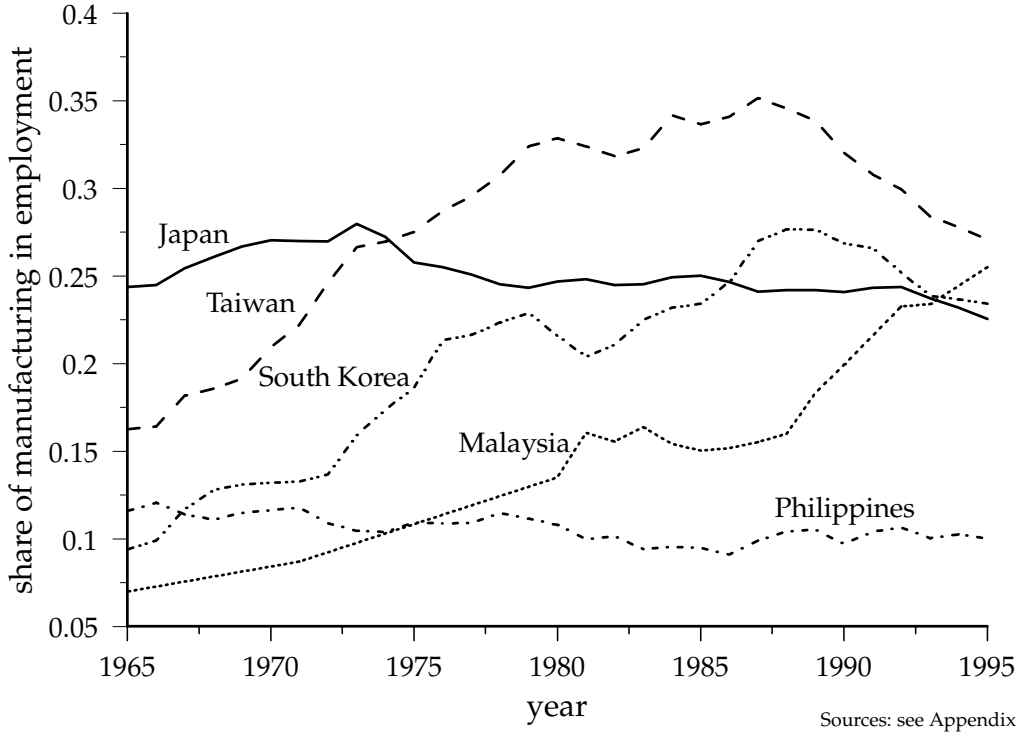

FIGURE 3
Waves of industrialisation in East Asia

spreads to a third country, and so on. What we see then is industrialisation spreading, in a series of waves, from country to country. The model predicts that economic development is not a smooth process of many countries catching up with the rich. It is instead the coexistence of a rich and a poor group of nations, but with growth of world demand for manufactures causing successive poor countries to join the rich club.

This, we think, provides a useful way of thinking about the spread of industry from Japan to several of its East Asian neighbours over the last three decades. Figure 3 plots the share of manufacturing in total employment in Japan, Taiwan, South Korea, Malaysia and Philippines between 1965 and 1995. Throughout this period about 25% of Japanese workers have been employed in manufacturing; Taiwan, South Korea and Malaysia have gained industrial employment in that sequence, while the fraction of workers in manufacturing has remained below 12% in Philippines.

4. Import substitution vs. trade liberalisation

Can the less developed country use trade policy to attract industry? Figure 1 suggests that the answer is affirmative — changing τ_2 can move the economy out of the region in which country 2 has no industry — and we now investigate this in more detail.[12]

Consider first the effects of an import substitution policy, raising trade barriers. τ_2 effects x_2/x^* in two ways. An increase in τ_2 makes imported inputs more expensive, this reducing the term $\tau_2^{-\sigma\mu}$ and making it less attractive for a firm to establish production in country 2. But pulling in the opposite direction, an increase in τ_2 switches country 2 expenditure on manufactures towards production in country 2 (the term $e_2\tau_2^{(\sigma-1)}$)[13]. Which of these effects is more powerful? Letting $\tau_2 \to \infty$ we see that $x_2 \to \infty$ providing that there is some manufacturing expenditure in country 2 ($e_2 > 0$) and that $(\sigma-1)/\sigma > \mu$. We shall assume that this restriction on parameters is satisfied — without it we have the curious result that even under autarky it is not profitable to set up production to meet local demand.[14] In this case then, raising trade barriers can always be successful in attracting industry (we return in a moment to seeing how much industry and of what type).

What about trade liberalisation? Can a reduction in τ_2 cause industrialisation to commence? We see from figure 1 that the answer depends on the values of A/e_0 and other parameters. We can get some more information on this by looking at the derivative $dx_2/d\tau_2$. Differentiating (17) this is,

---

[12] We look only at unilateral changes in trade costs. A multilateral reduction in trade costs can cause relocation of industry to non-industrialised economies, as shown in Krugman and Venables (1995). In a three country set up, the relocation might not be to all countries simultaneously (see Puga and Venables (1997), who also study the effects of regional integration).

[13] None of the endogenous variables in (17) depend directly on τ_2, as may be seen by inspection of equations (18) - (20).

[14] This condition is standard in models of this type, and rules out unbounded agglomeration. Fujita, Krugman and Venables (1999) have labelled it the 'no-black-hole' condition.

$$\frac{dx_2}{d\tau_2}.\frac{\tau_2}{x_2} = \left[\frac{(\sigma - 1 - \sigma\mu)e_2\tau_2^{(\sigma-1)} - \sigma\mu e_1\tau_1^{(1-\sigma)}}{e_1\tau_1^{(1-\sigma)} + e_2\tau_2^{(\sigma-1)}}\right]. \qquad (21)$$

The derivative is positive for large enough τ_2 (providing $(\sigma-1-\sigma\mu)e_2 > 0$), reflecting our earlier discussion about import substitution. The numerator switches sign as τ_2 becomes small enough, and it is this that generates the hump of the sustain curve illustrated on figure 1, and suggests that reducing τ_2 will attract industry. However, it is not necessarily the case that this change in the sign of $dx_2/d\tau_2$ occurs at $\tau_2 > 1$. From inspection of (21) we see that this is more likely the stronger are linkages (larger μ, increasing the forward linkage benefits from trade liberalisation), and the larger is e_1/e_2 (reducing the value of protecting the country 2 domestic market).

Although both import substitution and unilateral liberalisation can be effective in attracting industry, they have different welfare implications. This is explored in Figure 4, which plots country 2 welfare as a function of its import barrier, τ_2. The figure is computed from a two country numerical example, details of which are given in the Appendix, but two aspects of which need to be explained here. First we assume that country 2 is quite small — one third of the size of country 1 in endowments (and less in income); the reason is that we want to think in terms of a developing country importing from the rest of the world. Second, we split the trade barriers into natural and tariff barriers, and set the natural trade barriers at 15%, so define $\tau^* = 1.15$. Country 1 has no further barriers, so $\tau_1 = \tau^*$, but country 2 has tariff barriers over and above τ^*. The horizontal axis on figure 4 is this extra country 2 barrier, $\tau_2 - \tau^*$, and the vertical is country 2 real income.[15]

In the interval $\tau_2^- - \tau^*$ to $\tau_2^+ - \tau^*$ there is no manufacturing in country 2, and these values of τ_2 are the two solutions of $x_2/x^* = 1$ from the sustain curve. Import substitution draws industry in when $\tau_2 > \tau_2^+$, as does trade liberalisation when $\tau_2 < \tau_2^-$.

The two curves illustrated on the figure plot country 2 welfare under different assumptions about the social value of tariff revenue. The lower solid line gives the case when its social value is zero, and in the upper line we suppose that the tariff revenue

[15] Welfare is the long run equilibrium utility V_i as defined in equation (4).

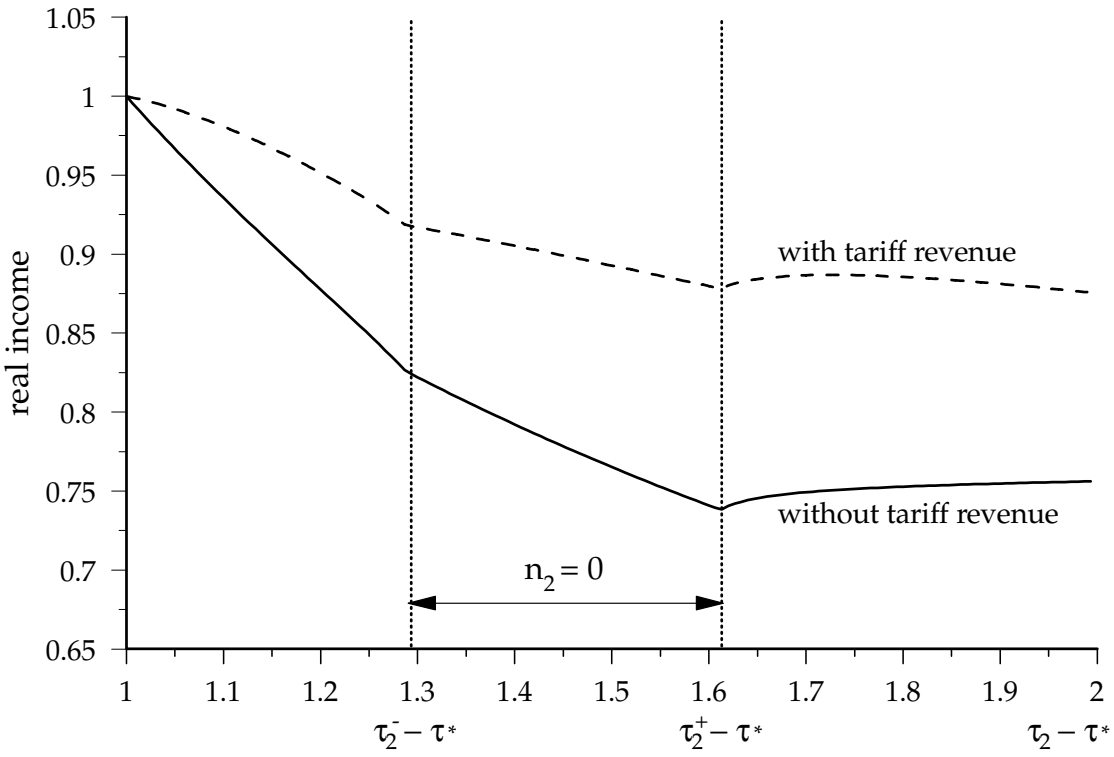

FIGURE 4
Import substitution vs. trade liberalisation

accrues to country 2 citizens and is added into the welfare measure.[16]

Looking first at the case in which tariff revenue is included in welfare (the upper curve), we see four main points. First, there is a range of tariffs within which import substituting industrialisation raises welfare, but this range is of limited width (between τ_2^+ and approximately $\tau_2 = 1.73$), and higher tariffs lead to welfare reduction. This arises from the trade-off between the beneficial linkages created by industry, and the loss of gains from trade. Second, in the region in which there is no manufacturing, $\tau_2 \in (\tau_2^-, \tau_2^+)$, reducing the tariff raises welfare. This is simply because it reduces the distortion on imported manufactures. Third, reducing the tariff further, into the region in which country 2 gains industry ($\tau_2 < \tau_2^-$), increases the rate at which welfare increases, because

[16] Tariff revenue is of no value if all rents are dissipated, or accrue to foreigners. When we add tariff revenue back we assume that it is all spent on agriculture or on leisure, and do not allow it to change manufacturing expenditure levels e_1 or e_2.

of the linkage benefits that are being achieved. The overall message is therefore clear. While import substitution may be locally welfare raising, it yields lower welfare levels than industrialisation via unilateral trade liberalisation.

If the tariff does not generate income for domestic consumers, then the lower line applies. The qualitative conclusions of the previous paragraph apply, with two quantitative qualifications. The gains from attracting industry by unilateral liberalisation are now larger — simply because of the real trade costs now being saved. However, at trade barriers above τ_2^+ welfare is increasing in τ_2. And essentially trade barriers are so costly that — within this interval — the best a country can do is go on raising the trade barriers to create further import substitution and drive imports to zero.

Figures 4 is of course, just from a numerical example. How are things changed as parameters of the model change? The effects of stronger industrial linkages (higher μ) can be seen from figure 1. Higher μ has the effect of increasing the interval of trade barriers within which agglomeration occurs (for a given value of A/e_0) raising the point at which import substitution commences, τ_2^+, and reducing the point at which unilateral liberalisation causes industrialisation, τ_2^-. However, it is noteworthy that the upwards shift of τ_2^+ is much larger than the downwards shift of τ_2^-, indicating that strong linkages make inward-looking industrialisation more difficult. Reducing the size of country 2 has a similar effect, again with τ_2^+ rising more than τ_2^- falls; in other words an import substitution policy is more difficult to implement the smaller is the economy. Raising developed country import barriers also shifts τ_2^+ up and τ_2^- down, but the relative magnitude of the shift is now reversed; the fall in τ_2^- is large, reflecting the difficulty that developed country trade barriers create for an outward looking development strategy.[17]

[17] This point suggests that if developed countries reacted to import substituting policies by raising their own tariffs, then the effectiveness of these policies would be reduced.

5. Trade policy and industrial structure.

So far we have assumed a single manufacturing sector. We now disaggregate this in order to see how import substitution and trade liberalisation policies affect the industrial structure of the developing economy.

We base our investigations on a five sector model, with input-output structure aggregated from a South-Korean input-output matrix. Sector 1 is the aggregate of all primary sectors and is assumed to be perfectly competitive and tradeable. Sectors 2-4 are manufacturing sectors, all of them monopolistically competitive. Sector 2 gathers all labour intensive manufacturing activities (those with an above average labour share). Labour un-intensive manufacturing sectors are split by consumer (sector 3) versus industry (sector 4) orientation (consumer oriented being those with an above-average ratio of final to total demand). Finally, Sector 5 is made up of services and is assumed to be monopolistically competitive and non-tradeable. (The mapping from the 19-sector transaction table in the 1980 input-output tables for South Korea to our five sectors is detailed in the Appendix).

To implement the linear expenditure system we combine the 'subsistence' levels of consumer expenditure calculated for South Korea by Lluch and Powel (1975) and the consumer expenditure shares in the South Korean input-output tables, as detailed in the Appendix. In all respects other than technical coefficients and demand parameters, we leave the manufacturing sectors identical. Thus, we do not attempt to estimate how product differentiation varies across industries, instead leaving $\sigma = 6$ in all imperfectly competitive sectors.

We assume that there are three primary factors, arable land, internationally mobile capital and labour. There are three countries, one (on which we shall focus) having one quarter of the world endowment, and the other two having the rest divided equally between them. We continue to abstract from traditional sources of comparative advantage, so relative endowments of each factor are the same in all countries.

Trade barriers and tariffs are the same for all manufacturing sectors, with developed country import barriers for manufactures set at 20%. Our experiment is to change the less developed country's trade barriers, but we keep these barriers equal across sectors

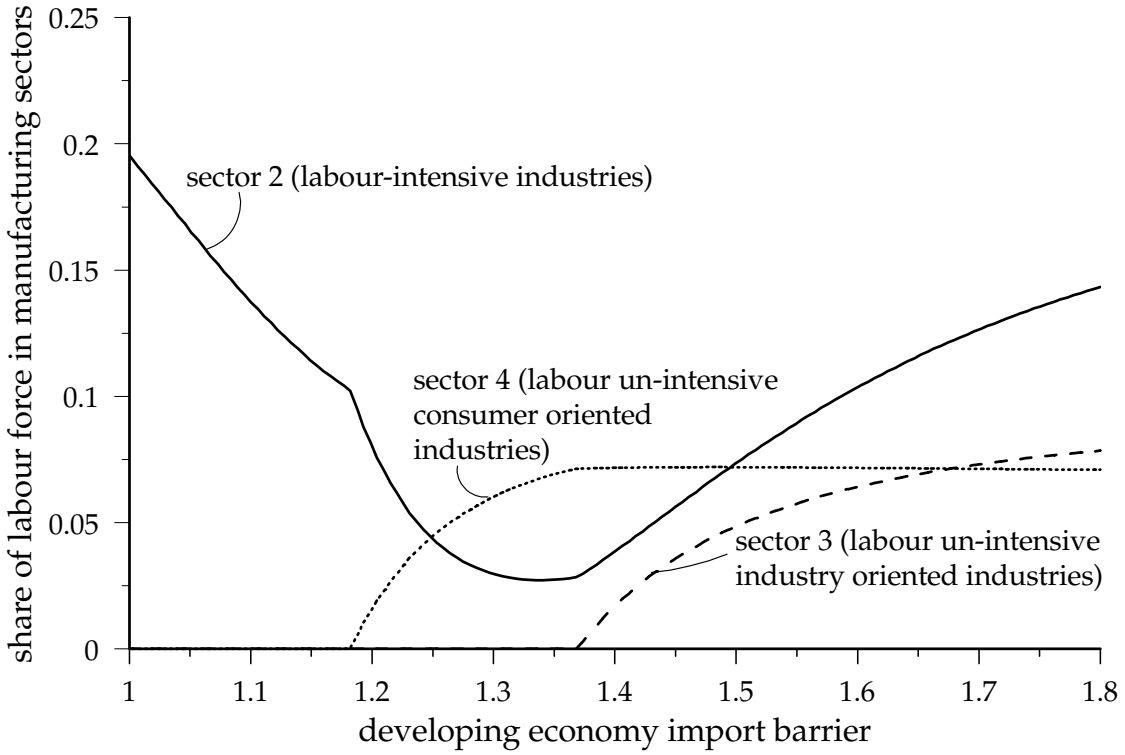

FIGURE 5
Shares of developing economy labour force in manufacturing

in order to show how the same trade policy affects different sectors differently, and to reflect the evidence that both Korea and Taiwan shows 'relatively low variances in protection across sectors' (Pack, 1992).

Figures 5 and 6 give, for the less developed economy on which we focus, the simulated shares of the labour force in each manufacturing industry, and the associated level of welfare.

We see from figure 5 that manufacturing employment as a whole is lowest at intermediate levels of the country's trade barriers, although at no point is manufacturing employment zero. At these intermediate trade barriers the country has a small presence in labour intensive manufactures (sector 2), and also in labour un-intensive but consumer oriented manufactures (sector 4). Sector 2 is active because of low wages, and sector 4 because of consumer demand. However, the developing country is a net importer of all manufacturing products.

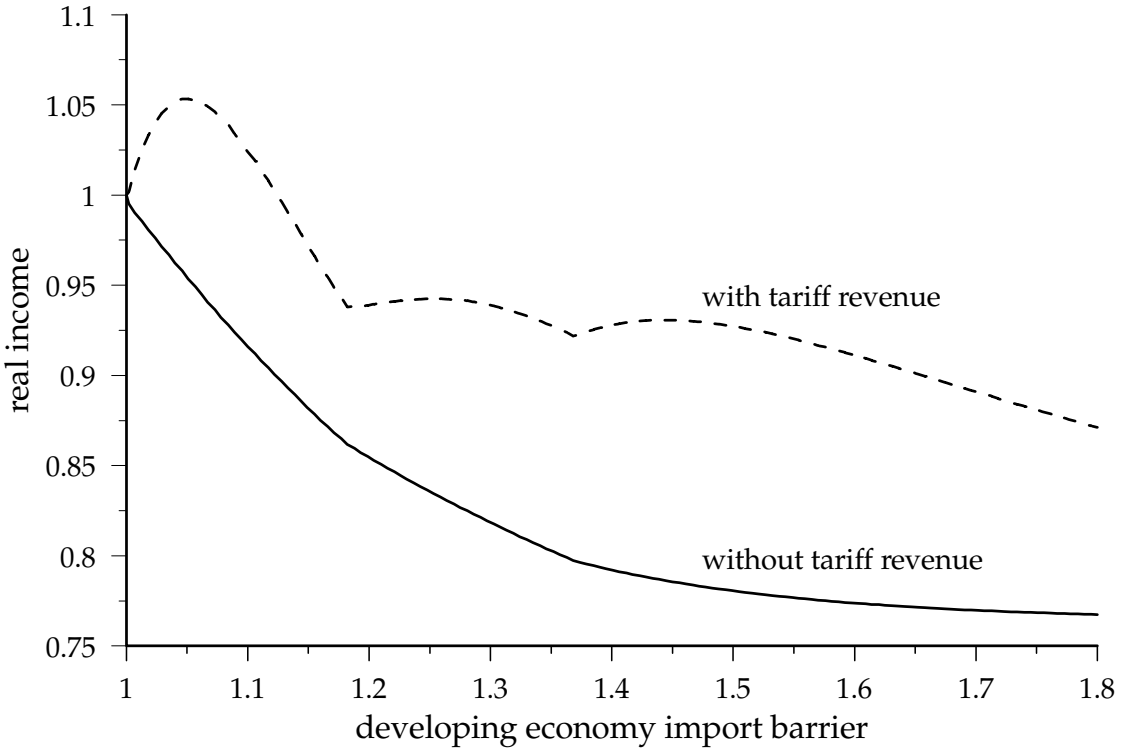

FIGURE 6
Import substitution vs. trade liberalisation with several manufacturing sectors

Raising import barriers leads to import substituting industrialisation. There is expansion of consumer oriented manufactures (sector 4) in order to meet final demand, and beyond some level the labour industry oriented sector 3 becomes active, driven by intermediate demand from the other sectors. As we raise import barriers at no point does the country become a net exporter in any of these sectors.

Trade liberalisation leads to a quite different industrial structure, as would be expected. The low wages of the developing lead to rapid expansion of labour intensive industries, which become significant net exporters. This employment expansion starts to raise wages, forcing out the other two sectors.

In interpreting these results it is important to bear in mind that we have assumed that there are no differences in relative factor endowments, and that the trade liberalisation is a unilateral import liberalisation by the developing country. How then does this import liberalisation lead to such a dramatic expansion of the labour intensive sector?

The reason is that import barriers make intermediate goods expensive in the developing country, and this is one of the factors inhibiting industrial development. Reducing import barriers removes this obstacle. All sectors benefit from this effect but, since wages are low in the developing country, it is the labour intensive sector that is best placed to expand output in response to the change. Interestingly then, even without assuming comparative advantage differences, open developing countries will export labour intensive products.

All this conforms with the empirical evidence from the newly industrialising East Asian economies which have liberalised trade over the last 20 years. For example, Korea reduced its average tariffs from about 32% in 1982 to 22% in 1985 and to about 10% in 1992 (World Bank, 1994). These countries have seen labour intensive products grow most as exports have exploded. Little (1994) looks at the sectoral detail of this process, comparing the patterns of industrial development of newly industrialising East Asian economies with the standard norms for less developed economies calculated by Syrquin and Chenery (1989). He finds that textiles, clothing, and metal products and machinery (all of them included in sector 2 in our numerical example) have grown much faster than normal in these economies, while chemicals and primary metal manufactures (aggregated into sector 3 in our example) remained well below normal in their shares of GDP.

The real income effects in this case confirm those that we saw in the aggregate case. Results are on figure 6 and as before the solid line is with real trade barriers, and the dashed line gives the case in which barriers create tariff revenue. Concentrating on the dashed line we see that real income is higher with trade liberalisation than with import substitution. The welfare maximum at low tariffs derives from the optimal tariff argument. This arises primarily on exports of labour intensive products, in which the developing country has a significant share of world production. (It is absent on figure 4 because with a single manufacturing sector the less developed economy never achieves a large enough world market share for terms of trade effects to be significant).

6. Conclusions

In the analysis of this paper we have abstracted from many of the differences between countries which are the focus of traditional development economics — for example, we have assumed that all countries have the same amounts of capital and land per unit labour. Clearly, we do not think that international differences in these factors are unimportant, but our objective is to see how much can be explained without them.

We think the answer is a great deal. Industrial linkages create agglomeration of manufacturing sectors, and this results in substantial real income differences between countries.[18] Yet firms do not move to the low wage economy, because if they were to do so they would forego the benefits of proximity to suppliers of intermediate goods and to their industrial customers.

Industrial centres may however become too large to be constricted in their initial set of locations. Rising relative demand for manufactures will widen the wage gap between those countries that have industry and those that do not, and at some point industry will spread to other locations. But the logic of agglomeration dictates that this is not spread evenly over developing countries. Instead industry spreads from one country to the next in a series of steps. Development takes place not as a process of smooth convergence of countries, but instead by countries in turn making the transition from a low level of development to the rich country club.

The approach also provides insights into the effects of trade policy. Unilateral trade policy can be used to attract industry by import substitution. More surprisingly, unilateral trade liberalisation can also be successful in attracting industry, as lower cost intermediate goods remove a barrier to industrial development. Comparison of import substitution and trade liberalisation indicates that while the former leads to a presence in a wider range of sectors, the latter yields higher levels of welfare. And a simple calibrated example of the model indicates a sectoral development pattern that fits well with that observed in many newly industrialising countries. Even though we abstract

[18] Of course, the real wage differentials suggested by our simulations are nothing like as large as real world income differentials, presumably reflecting the fact that we have assumed the same labour quality and levels of social capital in all countries.

from comparative advantage, we see that an open newly industrialising country will tend to export labour intensive manufactures.

Appendix

The multi-industry model: To the two internationally immobile primary factors of the single sector model, labour and arable land, we add capital which is perfectly mobile across the M countries. There is a perfectly competitive primary sector and a number of monopolistically competitive tradeable manufacturing and non-tradeable service sectors. The price index for sector s in the set of industrial sectors I takes the form:

$$q_i^s = \left[\sum_{j=1}^{M} n_j^s (p_j^s \tau_i)^{(1-\sigma)}\right]^{1/(1-\sigma)}, \quad s \in I . \tag{A.1}$$

where the superscript s denotes the sector. The price index for sector s in the set of service sectors S is :

$$q_i^s = \left[n_i^s (p_i^s)^{(1-\sigma)}\right]^{1/(1-\sigma)}, \quad s \in S . \tag{A.2}$$

The cost function of a single firm in sector s at location i is:

$$C_i^s = (\alpha + \beta x_i^s)\, r^{\rho^s} w_i^{(1-\eta^s-\rho^s-\sum_{k\in I\cup S}\mu^{k,s})} 1^{\eta^s} \prod_{k\in I\cup S} (q_i^k)^{\mu^{k,s}}, \quad s \in I\cup S , \tag{A.3}$$

where the share of primary sector output in the s industry is η^s, the share of industry k in the s industry is $\mu^{k,s}$, r is the rate of return on capital, ρ^s is the share of capital in costs and η^s is the share of primary sector output, and other parameters and variables are as

in the single-industry version of the model. The cost function in the perfectly competitive primary sector P is:

$$C_i^P = z_i r^{\rho^P} t^{\theta} w_i^{(1-\eta^P-\rho^P-\sum_{k\in I\cup S}\mu^{k,P})} 1^{\eta^P} \prod_{k\in I\cup S} (q_i^k)^{\mu^{k,P}} , \qquad (A.4)$$

where z_i is primary sector production in country i, t is the rental price of arable land, and θ is land share in agriculture. Preferences are given by the following indirect utility function:

$$V_i = 1^{-(1-\sum_{s\in I\cup S}\gamma^s)} \prod_{s\in I\cup S} (q_i^s)^{-\gamma^s} (y_i - \sum_{s\in\{P\}\cup I\cup S} e_0^s) . \qquad (A.5)$$

Demand for primary factors comes from the cost functions in the usual way, and their prices are determined by market clearing. Final and derived demands for output come from cost functions and indirect utility functions. Given these demand functions and cost functions (A3) firms maximise profits, and the number of firms in each sector is determined by free entry and exit.

Simulation parameters: ***Section 2***: $\sigma = 5$, $\tau_1 = 1.1$, $\gamma = 0.5$, and $\theta = 0.3$, where θ is the land share in the Cobb-Douglas agricultural production function.

Section 3: $\sigma = 5$, $\tau_1 = \tau_2 = 1.3$, $\gamma = 0.5$, $\mu = 0.4$, and $\theta = 0.1$.

Section 4: $\sigma = 5$, $\tau_1 = \tau^* = 1.15$, $\gamma = 0.5$, $\mu = 0.5$, and $\theta = 0.3$.

Section 5: Sectors are the aggregate of the following sectors from the 19-sector 1980 Korean input-output tables (pp. 56-57 in Bank of Korea (1983)):

Sector 1(primary sectors): agriculture, forestry and fishing; and mining.

Sector 2(labour intensive manufactures): textiles and leather; lumber and wood products; paper, printing and publishing; non-metal mineral products; metal products and machinery; and miscellaneous manufactures.

Sector 3 (labour un-intensive and industry oriented manufactures): chemicals and chemical products; and primary metal manufacturing.

Sector 4 (labour un-intensive and consumer oriented manufactures): food and beverages.

Sector 5 (services): all service sectors.

The input output matrix takes the following form:

η^s		0.085	0.058	0.268	0.498	0.010
$\mu^{2,s}$		0.018	0.340	0.041	0.030	0.111
$\mu^{3,s}$		0.093	0.201	0.397	0.029	0.116
$\mu^{4,s}$	=	0.056	0.005	0.004	0.145	0.016
$\mu^{5,s}$		0.050	0.135	0.102	0.070	0.203
θ		0.373	-	-	-	-
ρ^s		0.192	0.125	0.129	0.177	0.281
$1 - \eta^s - \rho^s - \Sigma\, \mu^{r,s}\ (-\theta)$		0.132	0.135	0.059	0.050	0.262

'Subsistence' consumer expenditure shares are adapted from those calculated for South Korea by Lluch and Powel (1975):

$e_0^s / \sum e_0^s$	0.308	0.198	0.000	0.308	0.186

Marginal consumer expenditure shares are calculated from the subsistence levels of expenditure and the actual consumer expenditure shares in the South Korean input-output tables:

γ^s	0.060	0.190	0.057	0.215	0.477

$\sigma = 6$, $\tau_1 = 1.2$.

***Data sources for figure 3*:**

International Labour Office. Annual issues. *Year Book of Labour Statistics*. Geneva: International Labour Office.

Republic of China, Directorate General of Budget, Accounting and Statistics. Annual issues. *Year Book of Statistics*.

References

Dixit, Avinash K. and Joseph E. Stiglitz. 1977. 'Monopolistic competition and optimum product diversity.' *American Economic Review*, 67: 297–308.

Bank of Korea. 1983. *1980 Input-Output Tables of Korea*. Seoul: The Bank of Korea

Ethier, Wilfred J. 1982. 'National and international returns to scale in the modern theory of international trade.' *American Economic Review*, 72: 389–405.

Fujita, Masahisa, Paul R. Krugman and Anthony J. Venables. 1999. *The Spatial Economy: Cities, Regions and International Trade*, MIT Press.

Harris, C. 1954. 'The market as a factor on the localization of industry in the United States.' *Annals of the Association of American Geographers*, 64: 640-656.

Helpman, Elhanan and Paul R. Krugman. 1985. *Market Structure and Foreign Trade*. Cambridge: MIT Press.

Hirschman, Albert O. 1958. *The Strategy of Economic Development*. New Haven: Yale University Press.

Krugman, Paul R. and Anthony J. Venables. 1995. 'Globalization and the inequality of nations.' *Quarterly Journal of Economics*, CX: 857-880.

Little, Ian. 1994. 'Trade and industrialisation revisited.' *Pakistan Development Review*, 33: 359–389.

Lluch, Constantino and Alan Powel. 1975. 'International comparisons of expenditure patterns.' *European Economic Review*, 5: 275-303.

Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny. 1989. 'Industrialization and the big push.' *Journal of Political Economy*, 97: 1003-1026.

Myrdal, Gunnar. 1957. *Economic Theory and Under-developed Regions*. London: Duckworth.

Pack, Howard. 1992. 'New perspectives on industrial growth in Taiwan.' In Gustav Ranis (ed.). *Taiwan, from Developing to Mature Economy*. Boulder, Colorado: Westview Press

Perroux, François. 1955. 'Note sur la notion de pôle de croissance.' *Economique appliquée*, 1-2: 307-320.

Pred, Alan R. 1966. *The Spatial Dynamics of US Urban-Industrial Growth, 1800-1914: Interpretive and Theoretical Essays*. Cambridge: MIT press.

Puga, Diego. 1998. 'The rise and fall of regional inequalities.' *European Economic Review*, 43: 303-334.

Puga, Diego and Anthony J. Venables. 1996. 'The spread of industry: agglomeration in economic development.' *Journal of the Japanese and International Economies*, 10: 440-464.

Puga, Diego and Anthony J. Venables. 1997. 'Preferential trading arrangements and industrial location.' *Journal of International Economics*, 43: 347-368.

Scotchmer, Susan and Jacques-François Thisse. 1992. 'Space and competition: A Puzzle.' *Annals of Regional Science*, 26: 269-286.

Syrquin, Moshe and Hollis B. Chenery. 1989. 'Patterns of development: 1950-83.' Discussion Paper 41, The World Bank.

World Bank. 1994. *East Asia's Trade and Investment: Regional and Global Gains from Liberalization*. Washington, DC: World Bank.