The Economic Journal, **112** (*July*), 649–678. Published by Blackwell Publishers, 108 Cowley Road, Oxford OX4 1JF, UK and 350 Main Street, Malden, MA 02148, USA.

A GEOMETRY OF SPECIALISATION*

Joseph F. Francois and Douglas Nelson

Division of labour models have become a standard analytical tool, along with competitive general equilibrium models (Ricardian, HOS, Ricardo–Viner), in public finance, trade, growth, development and macroeconomics. Yet unlike the earlier models, these models lack a canonical graphical representation. This is because they are both new and complex, characterised by multiple equilibria, instability and emergent structural properties under parameter transformation. We develop a general framework for such models, illustrating results from current research on specialisation models, and explaining why one sub-class of these models is particularly difficult to illustrate.

One of the great traditions in the analysis of international trade is the use of canonical models: Ricardian, Ricardo–Viner, and Heckscher–Ohlin–Samuelson. Furthermore, each of these models has a simple graphical representation, useful for both intuition generation and for pedagogical purposes. Over the last fifteen years, two additional classes of model have joined the big three: strategic trade models and division of labour models.[1] The strategic trade models entered the literature with simple graphical representations developed in the industrial organisation literature, while the division of labour models have proven to be considerably more resistant to simple representation.

Recent applications work with specific functional forms, and often involve numeric simulation, obscuring for some the general properties of these models. Even so, a set of general results (low-level equilibrium traps, catastrophic adjustment, agglomeration effects) do stand out from this somewhat diverse collection of special models. Because our starting point in this paper involves examination of this class of models in the context of relatively general functional forms and technologies (linear homothetic, concave, etc.), we are able to offer a generalised treatment that links this pattern of results to the general properties of models with increasing returns due to specialisation. In the process, we demonstrate that important results in the recent literature depend critically on the stability and transformation properties that characterise the general framework highlighted here. These properties are closely related to those explored in the context of scale economy models by an earlier generation of trade and development economists.

* Thanks are due to Kym Anderson, Richard Baldwin, Jan Haaland, Mary Lovely, Anthony Venables and the participants at the CEPR sponsored European Research Workshop in International Trade (ERWIT), held in Helsinki. Special thanks are due to the reviewer and editor of this Journal for a very careful reading of the paper and the provision of several useful suggestions. While working on this paper, the authors benefited from support provided by the Tinbergen Institute (Nelson) and the Center for International and Economic Studies at the University of Adelaide (Francois) while visiting those institutions. All remaining errors are due to confusion on the part of the authors.

[1] Strategic trade models refer to essentially partial equilibrium models of strategic competition rooted in modern industrial organisation theory (Brander, 1995). Division of labour models refer to a wide class of models attempting to formally characterise and analyse the Adam Smith–Allyn Young notion of division of labour. See Buchanan and Yoon (1994) for a useful selection of papers in this literature. We focus on one member of this class of models.

The central role of division of labour models in modern economic analysis is undeniable. They are prominent in international trade theory, public economics, regional/urban economics and macroeconomics (both growth theory and business cycle theory). Following Ethier's (1979, 1982*a*) original presentation of the model as a framework for studying the interaction between national and international returns to scale, the framework diffused rapidly throughout economic analysis. The reasons why a division of labour is 'limited by the extent of the market' that were loosely discussed by Smith and examined more deeply in Young's (1928) classic analysis are here provided a simple and tractable formal structure. In international trade theory, the model has been used to study trade patterns (Ethier, 1979, 1982*a*; Markusen, 1988, 1989; van Marrewikj *et al.*, 1997), trade policy (Markusen, 1990*a*; Francois, 1992, 1994; Lovely, 1997) and factor-market adjustment to trade (Burda and Dluhosch, 1999; Francois and Nelson, 2000; Lovely and Nelson, 2000). One of the most interesting recent applications uses the multiple equilibrium property of these models to derive north–south trade structures endogenously (Markusen, 1991; Krugman and Venables, 1995; Krugman, 1995; Venables, 1996*a*; Puga and Venables, 1996; Matsuyama, 1996). Following the important work of Romer (1987, 1990), the Ethier model has also become a standard framework in endogenous growth theory (Barro and Sala-I-Martin, 1995, ch. 6) and has been used extensively in development theory (Rodriguez-Clare, 1996; Rodrik, 1996) and regional economics (Holtz-Eakin and Lovely, 1996*a,b*; Fujita *et al.*, 1999).

In this paper, we proceed as follows. To provide some structure to the exercise, we have divided the general family of specialisation models into four types, as specified in Table 1. We begin with two versions of national production externality (NPE) models. In the first, a closed-economy version of the model (unimaginatively called model Type I in Table 1), we develop our basic geometric framework in the simplest environment. Even in this simple context, we are able to illustrate basic mechanisms that have been highlighted in the literature on endogenous growth and development. From there, we develop a NPE model of trade in final goods only (called model Type II), and demonstrate that this model is operationally identical to standard models of trade with national external economies of scale. The greatest conceptual and analytical difficulties emerge with international production externalities (IPE), which surface once trade in intermediate goods is permitted (model Types III and IV). The graphical analysis makes the locus of this difficulty clear. The general treatment of IPE models is followed by an examination of trading costs (an important issue in the recent literature) in Ricardian and Heckscher–Ohlin versions of the IPE model.

1. National Production Externalities (NPEs) in Autarky: Model I

1.1. *The Basic Model*

Although there are a wide range of variants, for pedagogical purposes we start with the NPE formulation closest to the Heckscher–Ohlin–Samuelson (HOS) model beloved of trade economists. That is, we will assume that there are two factors of production, labour (L) and capital (K); and two final consumption goods, wheat (W) and manufactures (M). Wheat is taken to be produced from K and L under a

Table 1
A Classification of Specialisation Models

		Trade structure	Description
NPE Models	I	Closed economy	The properties of this type of model are those of an external scale economy model (Markusen, 1990*a*).
	II	Open economy (trade in final goods only)	*Markusen Model*: Final goods production in each region will exhibit increasing returns due to specialisation (Markusen, 1989). However, without direct trade in intermediates, trade has no effect on the production structure of the economy. Model behaves like standard model of trade under external economies of scale.
IPE Models	III	Trade in intermediates only (or intermediates and the standard good)	*Ethier model*: International economies of scale. Trade affects production conditions, so transformation functions are no longer technological facts (Ethier, 1982*a*).
	IV	Trade in intermediates and final goods	Without trading costs, this is identical to type III, and this is where the scale of one regional sector will directly effect the efficiency of other sectors. Types III and IV diverge in interpretation with trading costs.

standard neoclassical technology represented by a production function $f(K_w, L_w)$ which is twice differentiable, linear homogeneous and strictly concave. Both factors are costlessly mobile between sectors and the markets for K, L, W and M are perfectly competitive. Where demand is needed, it will be taken to be generated by a representative agent whose preferences can be represented by a twice differentiable, strictly quasi-concave, homothetic utility function defined over consumption of W and M. Division of labour models diverge from standard trade models in the technology of M production. M is produced by costless assembly of components (x). Components are produced from 'bundles' of K and L – denoted m. The market for components is monopolistically competitive and bundles production is perfectly competitive.

Ethier's key insight was that the Spence (1976)–Dixit and Stiglitz (1977) model of preference for variety, applied to international trade by Krugman (1979, 1980), when applied to production constitutes the basis of a model of division of labour. The model contains two main elements:

(1) a technology reflecting increasing returns to 'division of labour'; and
(2) something limiting the division of labour (ie 'the extent of the market').[2]

The first element is given by a CES function that costlessly aggregates components (x_i) into finished manufactures:

$$M = \left(\sum_{i \in n} x_i^{\phi} \right)^{1/\phi}. \tag{1}$$

Here, n types of components are costlessly assembled into final manufactures and ϕ is an indicator of the degree of substitutability between varieties of inputs (x_i).[3]

[2] The appendix contains a full development of the Ethier model.
[3] Note the harmless abuse of good mathematical notation: n is being used as both the label of an index set and the number of elements in that set.

In particular, note that if $x_i = x \; \forall I \in n$, (1) reduces to $M = n^{1/\phi} x$. Then, for n constant, output of manufactures is linearly related to output of components and if $0 < \phi < 1$ (as we assume it to be) there are increasing returns to the variety of inputs, i.e.

$$\frac{\partial M}{\partial n} = \frac{1}{\phi} n^{\frac{1-\phi}{\phi}} x > 1.$$ [4]

The smaller is ϕ, the stronger are the returns to the division of labour. As in the SDS formulation, fixed costs in the production of intermediates and finite resources limits the number of types of components produced and thus, since aggregate output of M is increasing in varieties of component types, economies of scale are limited by the extent of the market.

The transformation of the SDS model of preferences into a model of the division of labour, along with the trick of using 'bundles' of inputs in the production of intermediates, not only makes the model exceptionally tractable from an analytical point of view, but also lends itself to straightforward graphical representation. The first element of this representation is the *bundles transformation function*: $w = B(m)$. In each industry, 'bundles' of capital and labour are produced according to standard neoclassical production functions of K and L: $m = g(K_m, L_m)$ and $w = f(K_w, L_w)$. Since capital and labour exist in finite quantities, $\{\bar{K}, \bar{L}\}$, and if we assume, say, manufacturing is K-intensive relative to wheat at all relative factor prices, the bundles transformation function will have the usual concave shape. This is plotted in Fig. 1 in the SW quadrant. The NW and SE quadrants contain functions mapping bundles into final good outputs: $W = \psi(w)$ and $M = \theta(m)$. In the HOS case, $\psi(\cdot)$ is just a 45° line.[5]

The real core of the graphical analysis is the $\theta(\cdot)$ function. In mapping bundles into final manufactures, this function embodies the market structure assumptions in both production of intermediates (internal increasing returns due to fixed cost and monopolistic competition) and final assembly (external increasing returns due to division of labour). Following the development in the Appendix, we are able to show that

$$M = \theta(m) = A m^{\frac{1}{\phi}} \tag{2}$$

where A is a constant and, as shown in (1), ϕ is an indicator of the degree of substitutability between varieties of inputs. It will be useful in the later analysis to have expressions for the first and second derivatives of θ:

$$\begin{aligned} \theta'(m) &= \frac{1}{\phi} A m^{\frac{1}{\phi} - 1} > 0 \\ \theta''(m) &= \frac{\frac{1}{\phi} - 1}{\phi} A m^{\frac{1}{\phi} - 2} > 0 \end{aligned} \tag{3}$$

Following Mayer's (1972, Fig. 1) analysis of production and trade under increasing returns to scale, we can use the information contained in the bundles transformation function and the two mapping relations to derive the transformation relation between final manufactures and wheat, $W = T(M)$, presented in

[4] Since, as we show in the Appendix, x is constant in equilibrium, it is easy to see that the production of M is homogeneous of degree $1/\phi > 1$.

[5] That is, the 'bundles', w, used in wheat production are just the inputs from standard analysis.

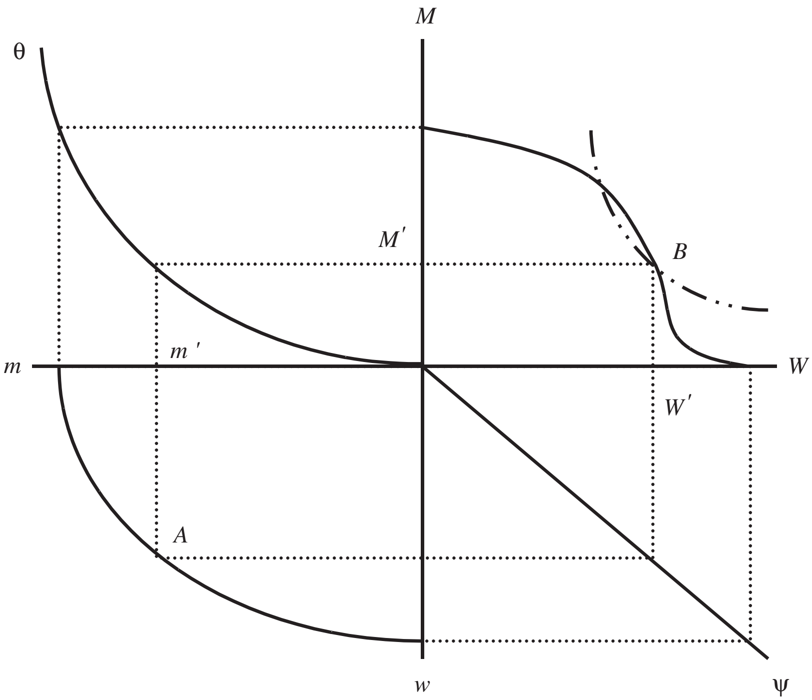

Fig. 1.

the NE quadrant in Fig. 1. That is, every point on the bundles transformation curve, $B(\cdot)$, is mapped to a point on the final goods transformation curve, $T(\cdot)$.

Herberg and Kemp (1968) and Mayer (1972) have intensively studied precisely the system we have just described for the case of variable returns to scale in both sectors (in our notation, θ and ψ are both permitted to be nonlinear). From Mayer (1972, p. 103), we have expressions for $T'(\cdot)$ and $T''(\cdot)$ in terms of θ, ψ, and $B(m)$. Note that, in these expressions, we are working with the inverses of B and T. That is, $\beta = B^{-1}$ and $\tau = T^{-1}$.[6]

$$\tau' = \frac{dM}{dW} = \left(\frac{\theta'}{\psi'}\right)\beta'$$

and

$$\tau'' = \frac{d^2M}{dW^2} = \frac{\theta'}{(w')^2}\left[\left(\frac{\theta''}{\theta'}\right)(\beta')^2 - \left(\frac{\psi''}{\psi'}\right)\beta' + \beta''\right]. \tag{4}$$

In the baseline case of HOS structure for bundles production, $\psi' = 1$ and $\psi'' = 0$, so the expressions in (4) are considerably simplified to

$$\tau' = \frac{dM}{dW} = \theta'\beta'$$

[6] Since $B(m)$ is the standard HOS production frontier, we know that it possesses a unique inverse. We adopt this both for expositional convenience and because, as graphically portrayed, the slope in W–M space is naturally seen as dM/dW.

and

$$\tau'' = \frac{\mathrm{d}^2 M}{\mathrm{d} W^2} = \theta' \left[\left(\frac{\theta''}{\theta'} \right) (\beta')^2 + \beta'' \right]. \tag{5}$$

The expressions for τ' show the interaction between $B(\cdot)$, θ, and ψ that are illustrated at any point on $T(\cdot)$ frontier in the right panel of Fig. 1. As with Herberg, Kemp and Mayer, we are particularly interested in $T''(\cdot)$ if we want to know about the curvature of $T(\cdot)$.

We can show that if $\theta''/\theta' \to \infty$ as $m \to 0$, then the transformation function must be convex in the neighbourhood of zero manufacturing output.[7] Given the derived expressions in (3), it is easy to see that

$$\frac{\theta''}{\theta'} = \frac{\frac{1}{\phi} - 1}{m} > 0 \tag{6}$$

which (since $0 < \phi < 1$) clearly approaches ∞ as m approaches zero. This equation is just a measure of local curvature (like the Arrow–Pratt measure of absolute risk aversion). Thus, because the function taking m into M is extremely (ie almost infinitely) tightly curved in the neighbourhood of zero manufacturing output, the transformation function is pulled in toward the origin. As the β'' term in the expression for τ'' suggests, the concavity of $B(m)$ works against the convexity of θ and can produce a concave portion of $T(M)$ in the neighbourhood of zero W output. In particular, it is easy to see that (6) and β' both become smaller as the output of M increases, implying that the first term in the square brackets in (5) becomes smaller. Unfortunately, while the first term should decline monotonically, unless we are willing to make some strong assumptions about the magnitude of T'', we are unable to say anything *definite* about curvature away from the neighbourhood of zero M output. This is an important point. The frontier may, in general, be characterised by multiple convexities and alternative stable and unstable regions. (Stability is discussed below.) With specific functional forms and parameter values, the approach in the literature has basically been to make implicit assumptions about where these regions occur.

1.2. *Ricardian Variations*

Given the structure that we have developed to this point, it is easy to illustrate two standard variants of the basic model: the Ricardian and Ricardo–Viner technologies for bundle production. In the Ricardian case (Chipman, 1970; Ethier, 1982*b*; Gomory, 1994), labour is the only productive factor, as a result the resource constraint takes the simple form of a straight line with a slope of negative unity in the SW quadrant. (Fig. 2*a*). Wheat is produced with a constant returns to scale production function, components are produced with a fixed and variable component (now paid entirely in labour) and manufactures are produced from components according to (1). We now have that the 'bundles' transformation

[7] This is a result of Mayer's (1972, pp. 106–9) which refines a result originally presented in Herberg and Kemp (1968).

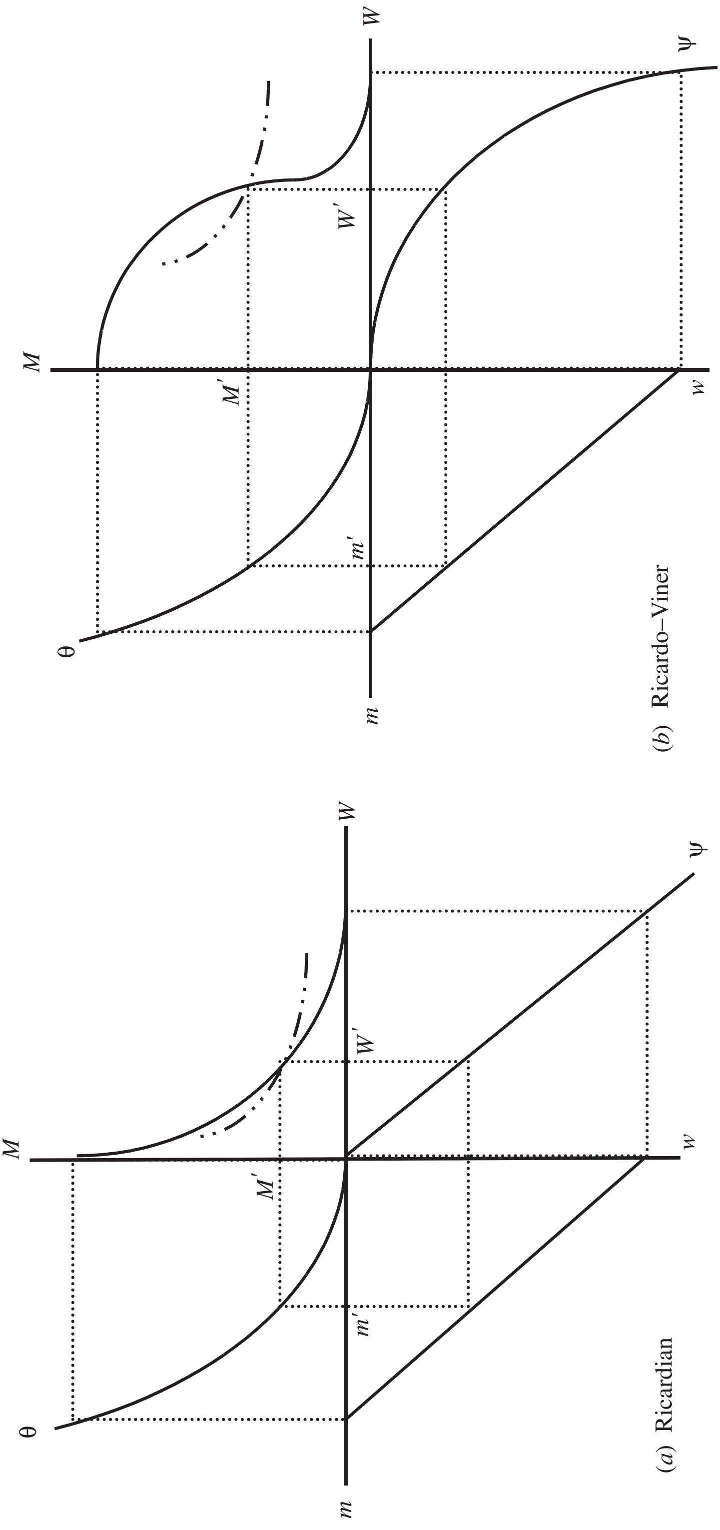

Fig. 2. *Technology for Bundles Production*

function (ie the labour constraint) is characterised by $B' = -1$ and $B'' = 0$. Since ψ is still a linear map with a slope of 1, we derive $T' = -\theta'$ and $T'' = \theta''$. That is, the shape of the transformation function between finished manufactures and wheat is defined entirely by θ, and T is concave throughout its length. The explanation of this is quite clear in Fig. 2*a* since the θ function is the only source of curvature, while both B and ψ have unit slopes.

The Ricardo–Viner structure (Fig. 2*b*) has been extensively used in an important series of papers by Markusen (1988, 1989, 1990*a*,*b*, 1991). Consider the simplest version of this model: wheat is produced with mobile labour and specific capital; components are produced with labour only (again there is a fixed and a variable part needed in production of components); and manufactures are produced by costless assembly of components. As with the Ricardian model, the resource constraint for the Ricardo–Viner model is given by the labour constraint, which will again be a straight line with a slope of negative one. The θ function, determined by monopolistic competition among component producers and the CES aggregator, has the same qualitative properties and graphical appearance as in the HOS and Ricardian cases. Unlike the two previous cases, however, the ψ function is no longer linear but, reflecting the presence of the specific capital, shows diminishing returns to mobile labour (ie $\psi' > 0$ and $\psi'' < 0$). As a result, we cannot use the expressions for τ' and τ'' in (5) but must use those in (4). On the other hand, the Ricardian resource constraint still permits us to take $B' = \beta'' = -1$ and $B' = \beta'' = 0$, so we can write

$$\tau' = \frac{\mathrm{d}M}{\mathrm{d}W} = -\left(\frac{\theta'}{\psi'}\right)$$

and

$$\tau'' = \frac{\mathrm{d}^2 M}{\mathrm{d}W^2} = \frac{\theta'}{(\psi')^2}\left[\left(\frac{\theta''}{\theta'}\right) + \left(\frac{\psi''}{\psi'}\right)\right]. \tag{7}$$

As with the HOS case, in the Ricardo–Viner case the term in square brackets contains a strictly positive term and a strictly negative term. The first term in the square brackets, still given by (6), goes to positive infinity in the neighbourhood of zero output of final manufactures. The second (negative) term is strictly finite at that point, so T will be convex at that point. We also know that the first term will decline smoothly as output of finished manufactures increases. Unfortunately, other than sign, we have very little information about the properties of the negative term, so we cannot be certain about the *structure of* T away from the neighbourhood of zero manufacturing output.[8] This is an important source of multiple equilibria in the literature.

[8] Following Markusen (1989), we can obtain some additional leverage by considering specific functional forms. For example, in the Cobb–Douglas case, we have $\psi''/\psi' = (\alpha - 1)/L_w < 0$. This expression goes to negative infinity as output of wheat goes to zero. Mayer (1972), again expositing a result due to Herberg and Kemp (1968), shows that, in this case, T must be strictly concave in the relevant neighbourhood. Since both ψ''/ψ' and θ''/θ' decline smoothly with increases in L applied to W and M respectively, we can say, for this case, that T will have a single inflection (at the unique point where $-\psi''/\psi' = \theta''/\theta'$). For additional discussion of the curvature of $T(\cdot)$ in the Ricardo–Viner case, see Markusen and Melvin (1984), Herberg and Kemp (1991) and Wong (1996).

1.3. *The Closed Economy Equilibrium and Nontangencies*

We turn next to the equilibrium structure of the closed economy. This involves consumption along the *MW* frontier in Fig. 1. While under S-D-S type monopolistic competition, the closed economy produces the optimal number of varieties for a given allocation of resources to *m* production, average cost pricing and returns to specialisation mean that, even so, the relative size of the manufacturing sector will be sub-optimal.[9] As a result of average cost pricing, while autarky consumption will be at some point like *B* in Fig. 1, domestic prices will not be tangent to the $T(\cdot)$ frontier at this point (Markusen, 1990*a,b*). This leaves scope for policy interventions that target expansion of the manufacturing sector.

With the addition of Cobb–Douglas preferences, it can be shown that the production side of the economy exhibits the standard features of more classical models. In particular, the combination of Cobb–Douglas preferences (with fixed expenditure shares) and homotheticity of wheat and bundles production yields a subsystem of equations that is purely Heckscher–Ohlin. As a result, as shown in Ethier (1982*a,b*), the standard Rybczynski and Stolper–Samuelson results hold (in terms of wheat and bundles). However, the welfare calculus is complicated by aggregate scale effects in the transformation of bundles in final manufactures – which is what matters for welfare.

1.4. *Economic Growth*

In addition to the implications of returns to specialisation for the shape of the static production frontier $T(\cdot)$, such returns also carry important dynamic implications. The critical difference is captured in the θ function, which is strictly linear in the neoclassical model. With capital accumulation in the classical model, there will be an expansion of the production possibility frontier (the $T(\cdot)$ frontier), with a bias toward the capital intensive sector. With labour in fixed supply (and assuming a standard final demand system), the new equilibrium return to capital will fall. Identically, the incremental gain from an additional unit of capital will also decline. Because of these declining returns, the classical model will exhibit the dynamic property, under classical savings or Ramsey specifications, of a fixed long-run capital–labour ratio and zero growth. This process can be fundamentally altered, however, by the simple addition of returns to specialisation. Because the θ function is no longer linear, the decline in the return to capital is moderated by returns to specialisation (Grossman and Helpman, 1991, ch. 4). If returns to specialisation are sufficiently large that they effectively bound the return to capital from below, the model will produce sustained economic growth. This depends on the relative curvature of the θ function. Even if the model exhibits local long-run Solow properties (with a unique steady-state level of capital and income in the long-run), the curvature of the θ function still implies a longer period of transitional growth, and a magnification

[9] See Bhagwati *et al.* (1998) for a concise discussion of the optimal variety issue.

effect related to efficiency shocks (as may follow from policy intervention). In conjunction with average cost pricing, the externalities related to resource accumulation mean that the laissez faire equilibrium in the model exhibits not only a sub-optimal static allocation of resource, but also a sub-optimal dynamic one.

The curvature of the θ function also carries dynamic implications for the effects of learning by doing. For example, we can represent the accumulation of production knowledge in the manufacturing sector by temporal shifts in the $B(\cdot)$ frontier – simply reinterpret K as knowledge capital. Even in the neoclassical model, this may lead to sustained economic growth. This depends, critically, on whether there are diminishing returns to knowledge accumulation. Externalities following from knowledge accumulation – variations on $A(K)$-type growth – can lead to sustained growth. Specialisation economies can deliver the required externalities. It is the curvature of the θ function that proves critical to determining whether specialisation economies are sufficient to generate sustained economic growth, or whether, instead, they simply provide a magnification of static effects (and boost the Solow residual in the process).

2. NPE with Trade: with Trade: Model II

We turn next to the open economy version of the NPE model. If we are willing to permit trade in final goods only (ie in W and M, but neither in components nor in factors), $B(\cdot)$, $\theta(m)$, and $\psi(w)$ continue to be technological properties of a country's economy. (By a 'technological property', we refer to properties of an economy that are not changed by opening international trade.) Since factors are taken to be immobile (except when factor mobility is the subject of analysis), it should be clear that trade will not have any effect on the bundles transformation function. Similarly, $\psi(w)$ is defined purely in terms of a national technology. Finally, examination of (1) reveals that, as long as only nationally produced intermediates are available to producers of final manufactures, $\theta(m)$ is also determined solely in terms of national magnitudes. Thus, Figs 1–2 continue to characterise production conditions whether or not there is trade in final goods only. This is exceptionally convenient because it permits us to appropriate the substantial body of work on international trade under increasing returns to scale virtually unchanged (Helpman, 1984).

The Type II model is an extreme version of a model with local agglomeration effects. We say extreme because there are no moderating effects related to cross-border spillover of production externalities. Because the reduced form structure of the model is identical to the older external scale economy literature, we are free to stand on the shoulders of this literature when drawing policy implications about trade policy and the location of industry. One important feature of the Type II model is that, Dr Pangloss to the contrary notwithstanding, there will generally be multiple, Pareto-rankable equilibria. For small countries, in particular, there is the strong likelihood that they will specialise in wheat production and may suffer a welfare loss relative to autarky. This fact underlies the modern versions of Frank Graham's argument for protection (Panagariya, 1981; Ethier, 1982*b*).[10]

Many of the insights of the recent literature on forward linkages, development and specialisation (Rodrik, 1996; Rodriguez-Clare, 1996; Rivera-Batiz and Rivera-Batiz, 1991; Venables, 1996*b*) follow directly from this property of local agglomeration models. Basically, because specialised primary/wheat production involves a stable equilibrium, and because more developed economies, by definition, have cost advantages related to larger and more specialised upstream industries, there is a tendency for underdeveloped countries to stay that way.

3. International Production Externalities (IPEs): Models III and IV

3.1. *Introducing IPE in the Basic Model*

While, as we have seen, Ethier's model of the division of labour has provided extremely useful microfoundations for the analysis of strictly national returns to scale, in its maiden application, it was actually used to examine internationally increasing returns to scale. The notion that access to international markets permits beneficial specialisation has been an essential element of trade theoretic analysis at least since Adam Smith and David Ricardo. What is new in Ethier's formulation is the formalisation of a direct link between international trade and the technology of production: access to a wider variety of component inputs permits an increased division of labour in the production of manufactures. As we shall see, however, it is precisely the link between trade and technology that makes the analysis difficult to visualise in simple graphical form: production conditions (especially as represented by the transformation function between final goods) are no longer a 'technological fact', determined only by nationally fixed production functions and endowments, but will now be dependent on the international equilibrium.[11]

We now assume that all R countries share identical tastes, technologies for producing factor bundles ($w = f(K_w, L_w)$ and $m = g(K_m, L_m)$), technologies for producing components from factor bundles and the technology for transforming w into wheat (ie $\psi(w)$). In all countries, all markets are taken to be perfectly competitive, except the market for components which is monopolistically

[10] Where Panagariya and Ethier adopt a Ricardian model, Markusen and Melvin (1984, proposition 1) and Ide and Takayama (1993, proposition 4) present an equivalent result for the HOS case. In deriving these results, fundamental use is made of the stability properties of these models under a Marshallian adjustment process in the final goods sector. The only peculiarity, for stability analysis, of our models relative to the standard external economy models, is the monopolistic competition in the intermediate sector. However, Chao and Takayama (1990) have shown that, as long as production functions are homothetic, monopolistic competition of this sort is stable under the obvious firm entry process. Since homothetic production functions characterise all of our models in this paper, for models I/II, we can fully appropriate the stability results developed by Eaton and Panagariya (1979) and Ethier (1982*b*) for the Ricardian case, Panagariya (1986) for the Ricardo–Viner case and Ide and Takayama (1991, 1993) for the HOS case.

[11] A variation of the basic model type developed here incorporates value-added at the final assembly stage of intermediates into final goods. This leads to explicit interaction between division of labour effects and intermediate linkages. See, for example, Brown's (1994) discussion in the context of large applied general equilibrium models, and Puga and Venables' (1998) similar application in the context of smaller numerical models.

competitive. A given country, $j \in R$, assembles components into final manufactures according to the aggregator function

$$M^j = \left[\sum_{r \in R} \sum_{i \in n_r} \left(x_i^r\right)^{\phi}\right]^{\frac{1}{\phi}} \tag{8}$$

Roman subscripts and superscripts are country identifiers, Greek superscripts are numbers (ie powers). In the two-country case, the Home country will have no superscript and foreign magnitudes will be starred, ie, when $n = 2$, $n = \{\ , *\}$. With traded intermediate goods, it will no longer be the case that, at the level of a given national economy, the amount produced by a given component producer (which we now denote by y_r) will be equal to the amount of that component consumed in the country (x_r). In fact, since some strictly positive share of every component producing firm's output is exported, $x_r < y_r$. As a result, we can no longer simply substitute the expression for y – (A5) in Appendix 1 – into (8) unless we are working with global output. We can, however, exploit the fact that, under the assumption of a constant elasticity of substitution among varieties of components and zero transportation costs, if price per unit of every component is the same, every final manufacturing firm will purchase the same quantity of the intermediate from every intermediate producer in the world. Thus, we can set $x_i^r = x^r \forall i$ and r.[12] As a result, since $\sum_{i \in n_r} \left(x_i^r\right)^{\phi} = n_r x_r^{\phi}$, and letting $n^G = \sum_{r \in R} n^r$, we can write (8) as

$$M^j = \left[\sum_{r \in R} n^r \left(x^r\right)^{\phi}\right]^{\frac{1}{\phi}} = \left(n^G\right)^{\frac{1}{\phi}} x. \tag{9}$$

Furthermore, since all component producers produce the same quantity, given by (A5), and all manufacturing firms consume the same quantities of each component, it will be the case that $x^j = \delta_j \, y_r$. Since country j consumes δ_j of every variety, it is implicitly consuming δ_j of the total allocation of factors to bundle production, and denoting implicit consumption of bundles in country j by m^j , we have

$$\delta^j = \frac{m^j}{m^G}. \tag{10}$$

What we are really interested in is an expression for $\theta(m)$ incorporating the possibility of imported intermediate components. The aggregator in (9) is essentially the same as that in (1), so for national component producers the underlying competitive conditions are essentially unchanged from those underlying the analysis presented above. Thus, we can now write[13] $M^j = \theta(m^j, \mathbf{m}^{-j})$:

$$M^j = \theta^j\left(m^j, \mathbf{m}^{-j}\right) = A\left(\sum_{r \in R} m_r\right)^{\frac{1}{\phi}-1} m^j \tag{11}$$

where θ is now functions of the *global* level of component production.

[12] That is, every final assembly firm will buy the same quantity of every type of component and, since M production is produced by competitive firms under identical technologies, we can treat the economy's output as being produced by a single firm with that technology.

[13] Appendix 2 presents the analytics underlying this claim.

Before considering the two-country case (as an approach to the *R* country case), we briefly note the analytical simplification purchased by assuming that the country in question is either the only economically large country or is economically small. In the first case, the analysis is identical to that in the closed economy case (the Type I model). In the small country case, ROW (or large country) production completely determines the magnitude of the term in parentheses on the right-hand side making $M^j = A^+ m_j$ (where A^+ is a constant that includes everything but m_j). This is, of course, a linear function, so the small country behaves like a small country under constant returns to scale.

3.2. *Allocation Curves and Policy Ranking*

Now suppose that there are two countries (Home and Foreign), both large. Ethier's allocation curves, graphed below the SW quadrant in Fig. 3, are used to identify the equilibrium quantities m and m^*.[14] At this equilibrium, m and m^* are determined and so θ is a linear function.[15] That is, θ is a linear function (shown in the NW quadrant of Fig. 3). The allocation curve diagram picks out the equilibrium point on the bundles transformation function (point *A* in the figure) which, via ψ and the linear θ, is mapped to equilibrium outputs of final goods (point *B*). If point *A* is an interior point on $B(\cdot)$, competitive conditions and technology ensure that the slope of the tangent at that point gives the equilibrium price (in units of wheat) per unit of m (which we denote p). If there is trade in intermediate goods only (ie all trade is intra-industry trade), consumption occurs at point *B* as well: $m_P^j = m_C^j$.

The same logic will also work for the case of trade in components and wheat (the Type III model), with local assembly of components into final manufactures for local consumption (the case considered in Ethier). However, if intermediate goods can be exchanged for wheat (as well as other intermediate goods), it will no longer be generally true that $m_P^j = m_C^j$. We have already seen how to find the production point on the bundles frontier (*A*) and the implicit final goods production point (*B*). The equilibrium at the intersection of the allocation curves reflects an equilibrium price of manufactures (*P*, taking wheat as the numeraire). As a result of zero-profits, full-employment of the factor-endowment and balanced trade, we know that consumption will occur on the national income line through point *B* (with a slope of $-P$). As illustrated at point *C* in Fig. 3, this will be a tangency between an indifference curve and the national income line. Using the equilibrium (linear) θ and ψ again, this time from point *C*, we can find the pair of factor bundles (m_C^j, w_C^j) needed to produce the consumption bundle of final goods. We know that the national income line tangent to the bundles frontier (at *A*) reflects the same national income as that given by the line through *B*, with an adjustment

[14] Recall that the allocation curves give, for either country, the (m, m^*) combinations that are consistent with domestic equilibrium for that country – i.e. where the domestic supply price is equal to the world demand price. The intersection of these curves identifies an (m, m^*) combination consistent with simultaneous equilibrium in both national markets and, thus, the world market.

[15] In terms of equation (B1) in Appendix 2, the equilibrium value of $k = \bar{k}$. From this we have $\theta = \bar{k}m$, a linear function in m.

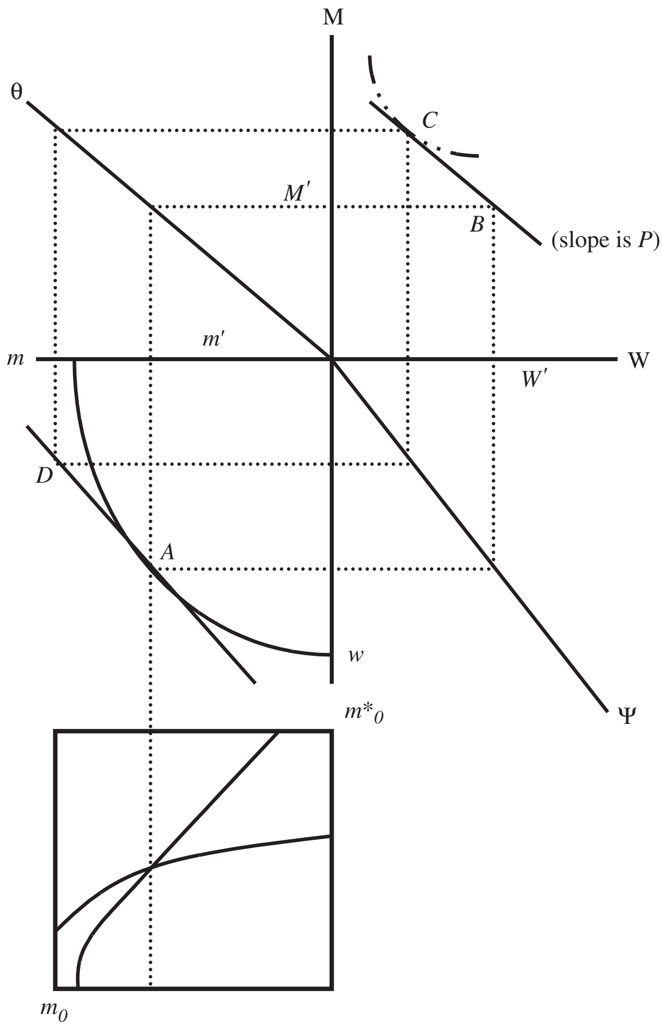

Fig. 3.

for scale.[16] Thus, *D* will lie on the national income line through *A*, the slope of which is $-p$ (i.e. the price per unit *m*). This is a full characterisation of equilibrium in the Ethier model with trade in intermediate goods and wheat (model III).

It is essential to note that we have not yet drawn a production set in the northeast quadrant. This is because of the fundamental difference between models I/II and models III/IV highlighted by the general equilibrium nature of θ in the latter case. Since θ is not a technological fact, we cannot draw a purely technological production frontier. There are only equilibrium points. In fact, at the equilibrium defined by the allocation curve intersection, we have taken the equilibrium θ as linear to draw Fig. 3. As a result, there cannot be offer curves or excess supply curves

[16] That is, we must make adjustment for the fact that $kP = p$.

of the usual sort. This, of course, is why Ethier developed the allocation curve technique.

As an aid to visualising the role of economic policy, we now construct an *experiment dependent set of production and consumption schedules.*[17] Recall that $B(\cdot)$ is a technological fact (it depends only on a fixed technology and a fixed factor endowment). Appropriate economic policy can pick out any point on the $B(\cdot)$ frontier. Consider, for example, a subsidy for home intermediate manufactures production. Such a subsidy will have a direct effect related to home output, and a second effect that captures the interaction between home production and rest of world production in the manufacturing sector. The net effect, involving changes in m and m^*, reflects the shifts in the home and foreign allocation curves that will be realised in the fifth quadrant in Fig. 3, see Ethier (1979) on this point. From (11), these, in turn, imply a shift in the efficiency of the economy in transforming m into M. As discussed more formally in the Appendix, every level of home m output is associated with a new policy dependent equilibrium characterised by a new production point in final goods space, a new relative price for manufactured goods P, and a new consumption point related back to implicit trade in bundles.

Moving to our graphic apparatus, we define Θ as the locus of all equilibrium points on the linear θ functions in mM space. This embodies the interaction between changes in the subsidy and changes in $(m + m^*)$. The Θ function can now be used to trace out the experiment dependent production frontier $T(\cdot)$ in Fig. 4, which we will refer to as the realised product transformation (RPT) frontier, defined in terms of final consumption goods.[18] This is effectively the production side of the economy. The next step involves finding the locus of all points identified by consumption of final goods at the experiment equilibria along the RPT. This follows from the imposition of final preferences and an income identity for consumption. If we impose identical homothetic preferences, we can then map the consumption locus as follows. First, along the RPT curve in Fig. 4, we have a price P associated with each point on the surface. One such price line is P_0 – associated with production point e_0. At the same time, from our imposition of homothetic preferences, this price P_0 also has associated with it an income expansion path E_0. The intersection of the price line, projected from the production point e_0, and the income expansion path E_0 projected from the origin gives us the associated consumption point C_0. We can map such consumption points for each production point on the RPT curve, yielding the consumption locus FN.

We can actually go a step further, and discuss welfare rankings along the policy-dependant consumption locus in Fig. 4. This is also discussed formally in Appendix C. The bottom line is that we are able to represent welfare ranking through shifts along social indifference curves in Fig. 4. Hence, in Fig. 4, social welfare U_1 is greater than U_0.

[17] Appendix C contains the algebra underlying the analysis of this section.
[18] In this section, we are speaking in terms of trade in intermediates (an Ethier model). Identically, the same discussion can be viewed in terms of specialised consumer goods, with M denoting the subutility index for differentiated consumer manufactures.

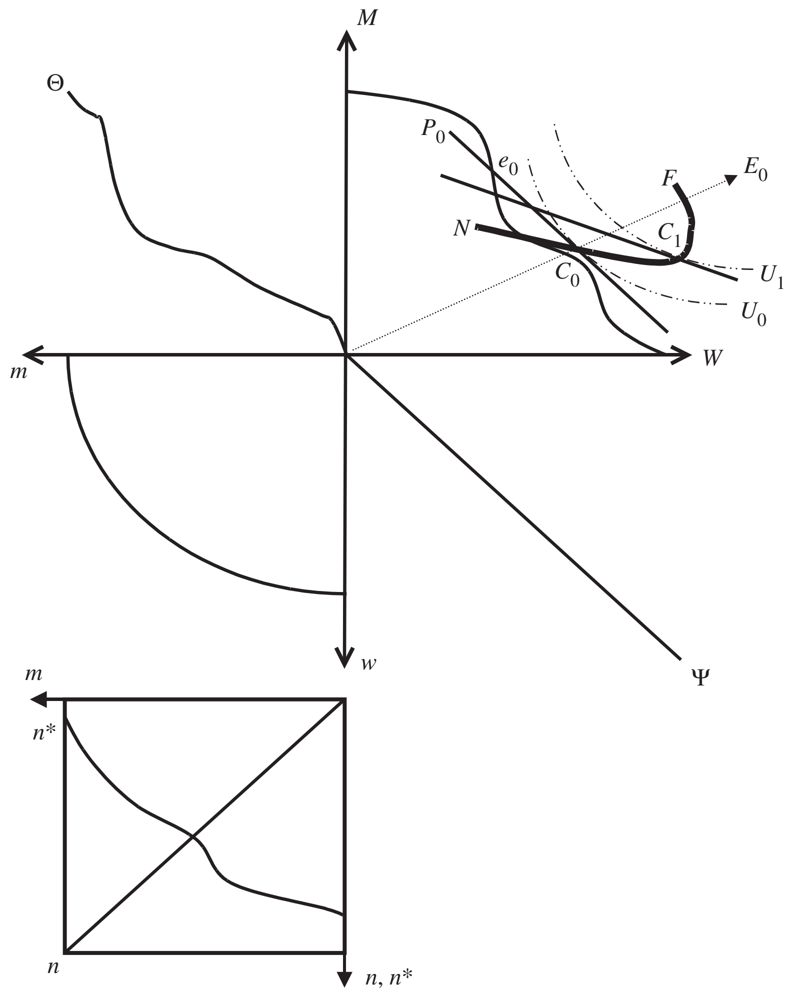

Fig. 4.

3.3. *Production and Stability*

Because the Θ function reflects a general equilibrium relationship rather than a technological fact, the same is true of the $T(\cdot)$ frontier – it is also an experiment-dependent artifact. The stability properties are also more elusive. We can demonstrate that internal equilibria sufficiently close to the vertical axis will be Marshallian stable, and that internal equilibria sufficiently close to the horizontal axis will be Marshallian unstable. The region in the middle, however is a theoretical free zone, with multiple stable and unstable equilibria allowed by the rules.[19]

[19] As Ethier (1979) makes clear, if we assume Mill–Graham preferences, there cannot be multiple equilibria. However, with more general demands, multiple equilibria emerge as a general property once again. Ethier (1979) also presents the stability analysis for this case.

Their existence will depend on the relative curvature of home and foreign $B(\cdot)$ frontiers and the relative importance of returns to specialisation.

Recall that we have drawn an experiment-dependent $T(\cdot)$ frontier in Fig. 4, where we add additional information about the home and foreign industries for equilibria along the RPT. In drawing the figure, we have forced the economy to move along its $B(\cdot)$ frontier, while allowing the rest of the global economy to adjust and clear all markets. Each point on the $T(\cdot)$ frontier, therefore, represents an equilibrium level of production (though a tax-cum-subsidy scheme may be required to sustain the equilibrium.) We have also represented, in the box at the lower left, the relative size of home and foreign industry, as indexed by the number of intermediate firms. Because we are mapping the implications of movements along the $B(\cdot)$ frontier, the number of home firms will be a linear function of m. (Recall the properties of the model, where m expansion involves entry of identical firms.) We cannot make such a statement about foreign firms, since their entry and exit (or identically the size of the m^* sector) is driven by the nature of the general equilibrium system, which will include the relative and absolute curvatures of the home and foreign bundles frontiers, the relative importance of specialisation economies, and the underlying preference structure. When we have a foreign region made up of many countries, the nonlinearity of n^* will be even more evident. It is, in fact, the nonlinearity of the mapping of n to n^* (or identically from m to m^*) that leads directly to the varied curvature of the $\Theta(m)$ function. This is immediately evident from inspection of (19).[20]

4. The IPE Model and Trading Costs

4.1. *The Ricardian IPE Model and Trading Costs*

An important application of Type IV models involves the implications of trading costs (Venables, 1996*a,b*; Krugman and Venables, 1995; Venables and Krugman, 1996; Fujita *et al.*, 1999). Trading costs are used, alternatively, to represent actual trading costs (transport, paperwork etc.) and government imposed costs, like tariffs and non-tariff barriers.

What do trading costs look like in the generic system? For expositional purposes, we start with a simple Ricardian model with identical home and foreign technologies, as illustrated in Fig. 5*a*. The autarky transformation frontier $T(\cdot)$ is represented in the upper right quadrant by the curve *142*. Consider next the integrated

[20] It is useful to note that, for the example developed in Fig. 4, we have shown an economy gaining from squeezing its own manufacturing sector out. From the box in the lower left, we see that this is accompanied by some relocation of industry (indexed by n and n^*) from home to foreign. Along with this, there is an increase in the price of M, implying that home had some natural comparative advantage in m. As drawn, this price increase is sufficient to generate terms of trade gains. The point highlighted by the diagram is as follows: the fact that a country can relocate industry to within its own borders, including the capture of associated agglomeration benefits, is insufficient to justify such a move on welfare grounds. In actuality, if the country is a net exporter of manufactured goods, terms of trade effects may justify intervention that squeezes the manufacturing sector out. In other words, in general equilibrium, the benefits of agglomeration effects must be weighed against potential terms-of-trade effects. Net exporters of manufactured goods are likely to gain from forcing prices up instead of down. Net importers of manufactures are more likely to benefit from a forced increase in global supply (with consequent agglomeration effects). See Francois (1994).

equilibrium (Dixit and Norman, 1980). The equilibrium level of M production is represented by the horizontal dashed line. With trade, the $T(\cdot)$ frontier (now represented by *1432*) will be linear up to the point where the home level of M exactly matches the M that we would observe in the integrated equilibrium.[21] In this range, there is a one-for-one displacement of home and foreign firms, with the allocation of firms between countries being indeterminate. This linear relationship between n and n^* is represented in the box in the lower left corner. The linear region of the $T(\cdot)$ frontier also corresponds to a linear section of the $\Theta(m)$ function.[22] Beyond point *4*, the home $T(\cdot)$ frontier will correspond to the autarky frontier (and the trade-based Θ function will rejoin the autarky one). It is also beyond this point that the home industry completely displaces the foreign industry. Interestingly, if production in the convex region of the home $T(\cdot)$ frontier improves home welfare, it will also improve foreign welfare. This is because of the basic non-tangency condition first discussed in Section 2, which will also characterise the integrated equilibrium. In the absence of terms-of-trade effects (which we have sterilised with our assumption of identical Ricardian countries), the world is actually better off if a single country (or set of countries) can capture the complete industry and then introduce a nationally optimal subsidy strategy. In this case, the nationally optimal subsidy will correspond to the globally optimal value.[23]

Next, consider the effect of transport costs as represented in Fig. 5*b*. We will have an inward shift of the $T(\cdot)$ frontier in regions where both countries produce m.[24] Whether or not either country specialises will depend on relative trading costs for intermediate and final goods. If the foreign country does specialise in W production at some point on the $T(\cdot)$ frontier, then, beyond that point, the frontier will again correspond to the autarky frontier. We will also have a shift in the n to n^* mapping, as represented in the lower left box. The complete specialisation point on the n^* curve will correspond to the point on the $T(\cdot)$ frontier where it rejoins the autarky frontier. There will also be an associated inward shift in the Θ function (see the upper left

[21] In terms of the analytics in Appendix B, the expression for k' in the region of an internal equilibrium, with identical Ricardian technologies, collapses to zero for the Type III/IV model.

[22] The linear section of Θ is determined where k remains constant as long as we are reproducing the integrated equilibrium.

[23] For groups of countries, the nationally optimal subsidy will be proportional to their share of the global industry, *if* we sterilise terms of trade effects (Francois, 1992).

[24] While beyond the scope of this paper, an alternative way to represent the model graphically in product space is with variety-scaled output. In particular, if we represent the introduction of trading costs as a break in the symmetry of weights on the regional CES aggregation functions, then we can represent the effects of regional variety changes in the productivity of m when used in production through a scaling term based on the size of local industry. Viewed this way, the production side of the economy, in terms of variety-scaled intermediates (Francois and Roland-Holst, 1997) and W, collapses to the Type II class of models. We then have Armington demand for intermediates indexed over regions (which reflects differential CES weights in different regions). These are produced regionally under increasing returns. The full effect of variety can again be represented by the Θ function, where this now transforms m into variety-scaled intermediates. Increases in foreign variety boost productivity of domestic varieties through marginal product increases in the CES aggregation function. The only non-concave feature of the model, represented this way, follows from the technological Θ relationship. For the Ricardian example developed here, the frontier for variety-scaled output (call it Z) and W will be convex over its entire length. The basic non-tangency result will still obtain along this frontier. Local agglomeration effects will be reflected in the CES weights, which will be higher for local goods. Hence, specialisation economies (and related convexities) will carry downstream to local M production.

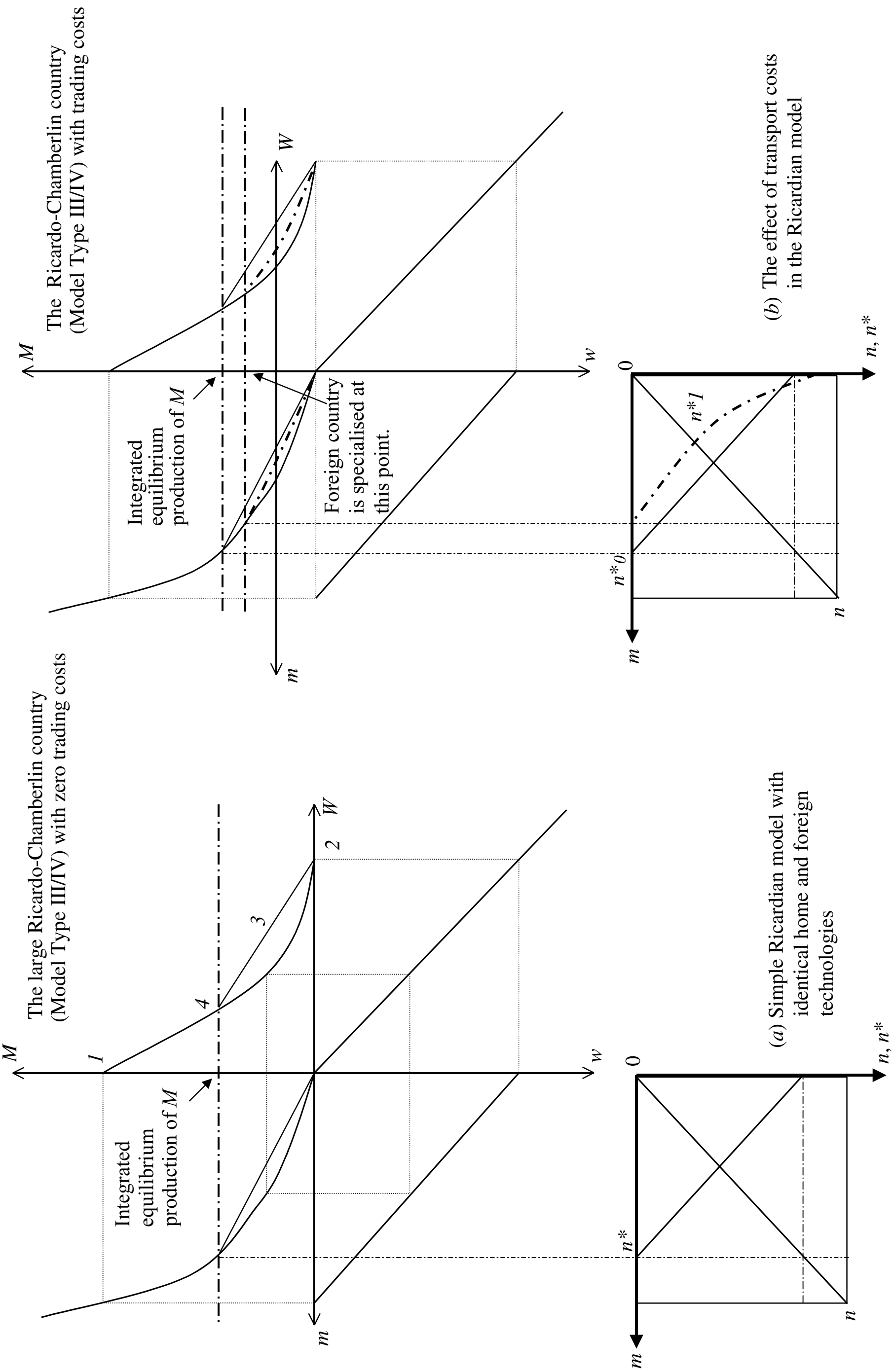

Fig. 5. Ricardian Models with IEPS

quadrant), with the Θ mapping also rejoining the autarky mapping at the point where the foreign country specialises.

The appearance of the $T(\cdot)$ frontier, with trading costs, closely resembles the frontier for the Type II class of models. In particular, in the present example of the Ricardian model, we have the type of frontier examined by Kemp (1964) and Kemp and Negishi (1970).[25] The critical difference is that, while Kemp was able to take world prices as fixed, we are unable to do so. From the point of view of the home country, border prices will shift as we move along the $T(\cdot)$ frontier, with the expectation that prices will be flatter as we move closer to the horizontal axis. The similarity to Kemp-type external scale economies leads us directly to a generalisation, in our generic framework, of a basic result of the location literature. In trade equilibria, there are again good and bad equilibria, and there may be instances of catastrophic collapse due to instability of internal equilibria along regions of the $T(\cdot)$ frontier. The possibility of foreign collapse also means that the $T(\cdot)$ frontier mapping from the horizontal axis to the point of foreign specialisation will not necessarily be continuous (hence also for the Θ mapping).

4.2. *The Heckscher-Ohlin IPE Model and Trading Costs*

We turn next to the characterisation of the $T(\cdot)$ frontier for the Type IV version of the economy developed in section 2. Recall that the economy is characterised by a Heckscher–Ohlin structure underlying the $B(\cdot)$ frontier. In autarky, the structure of the economy can be represented as in Fig. 6*a*. The Bundles frontier $B(\cdot)$ is strictly concave to the origin, while the existence of specialisation economies, imply increasing returns in production of M, and hence we have a realised product transformation frontier $T(\cdot)$ in the upper right quadrant that is characterised by concave and convex regions. We have represented the number of intermediate firms n (which is a strictly linear function of m) in the box at the lower left.

In developing the RPT frontier for the Type IV version of the Heckscher–Ohlin model, Fig. 6 proves to be an important reference case. Another useful reference case is the integrated equilibrium (Dixit and Norman, 1980). Recall from Helpman and Krugman (1985, ch. 7) that, for a global set of factor endowments $\boldsymbol{F}$ within the factor price equalisation set (which is bounded by the relative factor intensities for m and w in the integrated equilibrium), a trading economy will replicate the integrated equilibrium.[26] Hence, within $\boldsymbol{F}$, we know that the production of m and w will be that necessary to allow reproduction of the integrated equilibrium level of output. This defines one equilibrium point on

[25] The effective collapse of Type III/IV models to a complex version of the Type II class of models in the presence of trading costs is not limited to the Ricardian case, but will instead hold as long as bundle cost functions are homothetic. The proof, however, is beyond the scope of this paper. See the discussion in note 24.

[26] If we introduce additional sectors, the pattern of production and trade becomes indeterminate within $\boldsymbol{F}$. This implies flat regions on the RPT hypersurface, where production patterns are indeterminate nationally but the global level of production of $(m + m^*)$ will be fixed at the integrated equilibrium value.

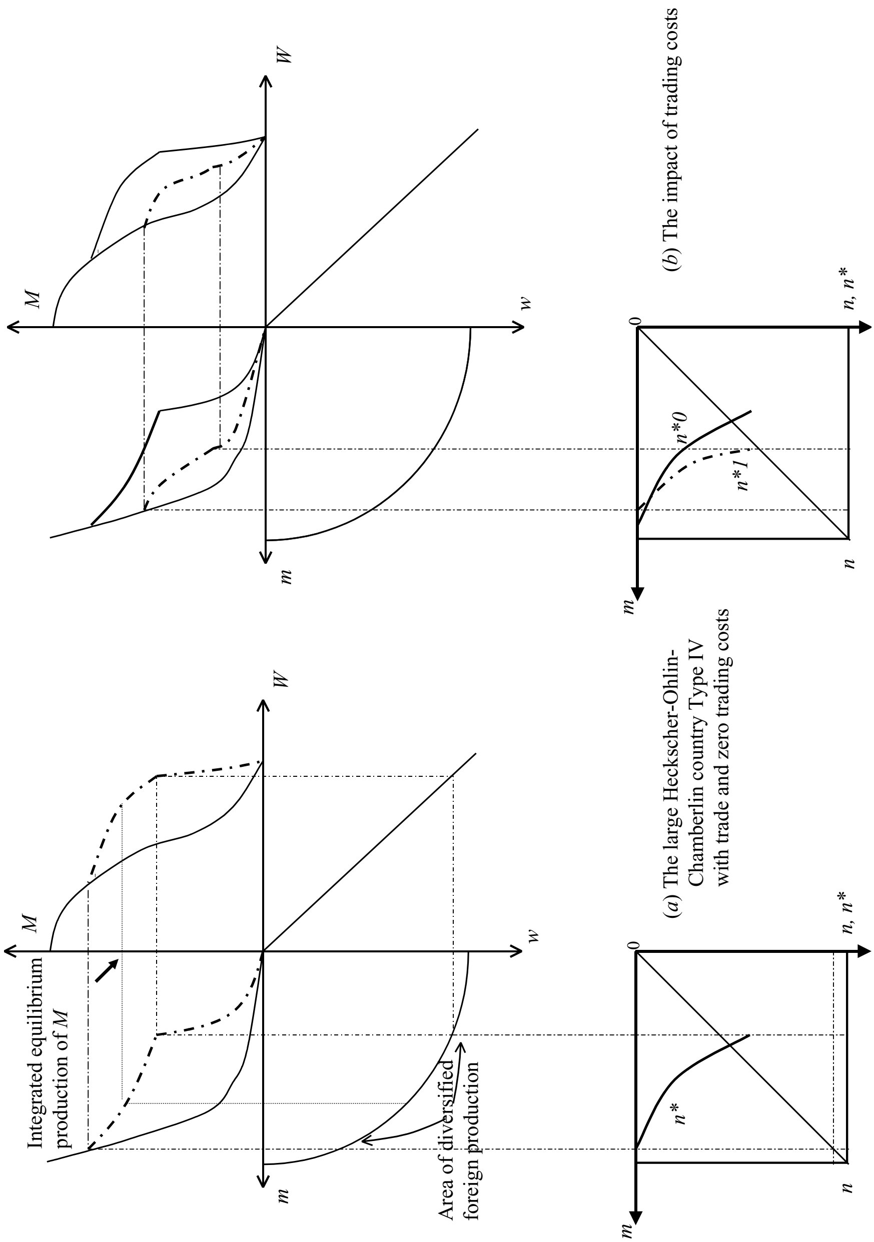

Fig. 6. *Type IV Heckscher–Ohlin–Chamberlin Country*

the $B(\cdot)$ frontier, and also on the RPT curve, as represented in Fig. 6*a*. When we move left from this point on the $B(\cdot)$ frontier, we will induce exit of foreign firms (though the total $n+n^*$ will be increasing). Movements left/right in the region of the integrated equilibrium will map the RPT curve in the region of the integrated equilibrium level of production on the $T(\cdot)$ frontier, through Θ. If the home country is sufficiently large, sufficient movement along the $B(\cdot)$ frontier will induce specialisation of the foreign country, either in m^* or in W^*. In the case of full foreign specialisation in W, the properties of the Θ function are determined strictly by home production of m, so that the relevant Θ function collapses to the corresponding region of the Type I θ function. As a consequence, the RPT curve rejoins the autarky RPT curve past the point of foreign specialisation in W. In terms of the mapping of n to n^* in the box at lower left, this is also the point where n^* reaches zero. Alternatively, at the point of foreign specialisation in m^*, the contribution of m^* to the Θ function becomes linear, and we are in a situation of strictly national returns to scale (where such returns are greater than in the autarky case). This is represented in the lower left quadrant of the figure, defining the regions of the $B(\cdot)$ frontier where we will observe diversified foreign production.

Consider next the implications of trading costs for intermediate manufactures. Formally, we can represent this by a shift in relative weights in the CES aggregator function for the national producers of M. Graphically, this means that Θ will be strictly lower (unless the foreign region is specialised in W) than without trading costs, and hence that Θ in Fig. 6*b* will shift in. Again, for points beyond where the foreign region specialises in W, the RPT will correspond to the autarky $T(\cdot)$ frontier. For those regions where the RPT remains above the autarky $T(\cdot)$ frontier, we have the general result that with increasing trading costs, the RPT curve will converge on the autarky $T(\cdot)$ frontier.

It should be evident by now that an important implication of trading costs is that the RPT curve is endogenous with respect to trading costs, including tariffs. Hence, a country can affect the efficiency of its national economy through commercial policy. In general, trade protection targeting trade in manufactures will reduce the efficient set for the economy, because of its impact on international agglomeration effects. This may be welfare improving if the country benefits from the consequent increase in manufactures prices (though it does not come close to being a first-best option for industrial policy, which would involve production subsidies and taxes). Transport costs will also have an important effect on the RPT curve, with declining transport costs boosting the national frontier.

The graphical analysis of tariff policy and transport costs in Fig. 6*b*, when compared to a classic Heckscher–Ohlin setting, or even to the setting of the Type II Heckscher–Ohlin model, is complicated by the endogeneity of the $T(\cdot)$ frontier. It is also complicated by the endogeneity of price along the frontier. Depending on where we locate along the old and new RPT frontiers, we may also observe catastrophic agglomeration (or collapse) of the foreign region (such that n^* reaches zero or its upper bound in the lower box), meaning that net m exporters may become net importers or vice versa.

5. Summary

Division of labour models have become a standard analytical tool, along with competitive general equilibrium models (like the Ricardian, Heckscher–Ohlin–Samuelson, and Ricardo–Viner models), in public finance, trade, growth, development and macroeconomics. Yet unlike these earlier general equilibrium models, specialisation models so far lack a canonical representation. This is because they are both new, and also highly complex. Typically, they are characterised by multiple equilibria, instability and emergent structural properties under parameter transformation.

Given the prominence of specialisation models in modern economics, the value of a generic representation seems considerable. In this paper, we have developed such a framework. In the process, we demonstrate that important results in the recent literature depend critically on the stability and transformation properties that characterise the generic model. We have also highlighted why one sub-class of these models is particularly difficult to illustrate easily.

Tinbergen Institute and CEPR
Tulane University and Leverhulme Centre for Research on Globalisation and Economic Policy

Date of receipt of first submission: April 2000
Date of receipt of final typescript: August 2001

Appendix A: Structure of Ethier's division of Labour Model

Recall from the text that the first element of the model is given by a CES function that costlessly aggregates components (x_i) into finished manufactures:[27]

$$M = \left(\sum_{i \in n} x_i^{\phi}\right)^{1/\phi} \tag{A1}$$

Here n types of components are costlessly assembled into final manufactures and ϕ is an indicator of the degree of substitutability between varieties of inputs (x_i).[28] In particular, note that if $x_i = x\ \forall I \in n$, (A1) reduces to $M = n^{1/\phi} x$. Then, for n constant,

[27] Ethier's original formulation is actually slightly different:

$$M = n^{\alpha}\left(\sum_{i \in n} \frac{x_i^{\phi}}{n}\right)^{1/\phi} \tag{A1a}$$

Note that, in both formulations, the elasticity of substitution between varieties is given by $1/(1-\phi)$ – ie the higher is ϕ, the more easily can the x_i be substituted for one another in production. In (A1a) there are effects operating both through market power (reflected in imperfect substitution among components – $0 < \phi < 1$) and returns to variety ($\alpha > 1$). It should be clear, though that the formulations in (A1) and (A1a) are identical if $x_i = x\ \forall I \in n$ and $\alpha = 1/\phi$. Since the distinct effects of market power and returns to variety are not essential to our analysis, we will use the simpler form in (A1), where returns to division of labour emerge directly from the imperfect substitutability between varieties of inputs. For an analysis of policy that exploits the distinct effects in (A1a) see Holtz-Eakin and Lovely (1996a,b).

[28] Note the harmless abuse of good mathematical notation: n is being used as both the label of an index set and the number of elements in that set.

output of manufactures is linearly related to output of components and if $0 < \phi < 1$ (as we assume it to be) there are increasing returns to the variety of inputs, i.e.[29]

$$\frac{\partial M}{\partial n} = \frac{1}{\phi} n^{\frac{1-\phi}{\phi}} x > 1$$

The smaller is ϕ, the stronger are the returns to the division of labour.

Component production is the final essential element of the standard division of labour model – and the locus of the actual division of labour. Where production of final manufactures is characterised by external economies of scale, components are produced under internal decreasing costs. Specifically, we assume (again following Ethier) that production of x_i units of components requires the purchase of *bundles* of capital and labour (denoted m) according to the relation

$$m_i = ax_i + b \tag{A2}$$

Note that both the fixed (b) and marginal (a) costs are paid in bundles and are constant across firms in the component producing sector. If we let $m = \sum m_i$, and $x_i = x_i \; \forall i \in n$, then $m = n(ax + b)$. As a result of decreasing costs, no two firms will produce the same type of component. Fig A1 shows the overall production relations in manufacturing and wheat.

The supply side of the model is closed by the resource constraint. With a fixed endowment of factors of production $\{\bar{K}, \bar{L}\}$ this is summarised by the transformation function between bundles of factors used in the production of components and wheat. For this purpose, we assume that bundles used in the production of components are produced according to a standard neoclassical production function, $m = g(K_m, L_m)$, which is linear homogeneous, twice differentiable, and strictly concave. Along with the equivalent assumptions on the production of wheat, and no factor-intensity reversal, we know that the function $W = B(m)$ is the strictly concave transformation function of the HOS model. In particular, we know that $B'(m) < 0$ and $B''(m) \leq 0$. Since wheat just is bundles of K and L, one might also think of the transformation process as involving bundles used in components and bundles used in wheat, where the former are transformed linearly into wheat. Where it is not confusing, we adopt the purely rhetorical simplification of referring to as 'wheat' both wheat and bundles of K and L used in producing wheat. Where necessary, we will denote bundles of K and L in wheat production as $w = f(K_w, L_w)$ and relate bundles to wheat via $W = \psi(w)$, where

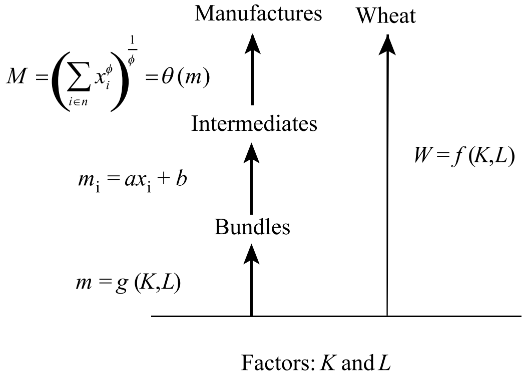

Fig. A1. *Schematic Representation of Production Structure*

[29] Since, as we shall see, x is constant in equilibrium, it is easy to see that the production of M is homogeneous of degree $1/\phi > 1$.

ψ is a linear relationship, ie $\psi(w) = a_{wW}\ w$. We will usually take the technical coefficient a_{wW} to be unity. The exception will be in the Ricardo–Viner case, where $\psi(w)$ is a nonlinear function.

Furthermore, given perfect competition in wheat and bundles (ie in m), and taking wheat as the numeraire, the relative cost of factor bundles for component production is

$$P_m = -B'(m). \tag{A3}$$

Now we want to link bundles to final production of manufactures: $M = \theta(m)$.[30] θ will serve as a basis for our graphical analysis, so we will need to pursue its properties, which will depend on both technology and the monopolistic competition among component producers, in some detail. Since component producers purchase bundles under competitive conditions, total cost for a representative component producing firm is $-B'(m)(ax_j + b)$ and total revenue is just p_jx_j, so the condition that marginal revenue equals marginal cost can be rearranged to obtain an expression for p_j:

$$p_j = -B'(m)\frac{a}{\phi} \tag{A4}$$

In deriving (A4), we substitute

$$\frac{\mathrm{d}x_j}{\mathrm{d}p_j} = \frac{1}{(1-\phi)}\frac{x_j}{p_j}$$

which can be derived from the demand curve for component j (Ethier, 1982, eq. (4)). The profit of each component producing firm is $\pi_j = p_jx_j + B'(m)(ax_j + b)$ which will be driven to zero by free entry and exit (abstracting from integer problems). Thus, setting $\pi_j = 0$ and substituting for p_j from (A4) gives

$$x_j = \frac{b\phi}{a(1-\phi)}. \tag{A5}$$

Since this is made up entirely of parameters that are constant across component producing firms, (A5) underwrites our treatment of x as identical across firms. Thus, from the fact that $m = n(ax + b)$, we can solve for the number of firms as a function of aggregate output of bundles:

$$n = \frac{(1-\phi)}{b}m. \tag{A6}$$

Note the implication that m and n are linearly related. Ethier works with a function $k(m)$, given by $M = km$. We can use (A5) and (A6), along with the fact that $M = n^{1/\phi}x$, to obtain an expression for $k = M/m$:[31]

$$k = \left\{\left[\frac{(1-\phi)}{b}\right]^{\frac{1}{\phi}-1}\frac{\phi}{a}\right\}m^{\frac{1}{\phi}-1}. \tag{A7}$$

Furthermore, since the expression in parentheses is made entirely of parameters, we can write this as $k = Am^{\frac{1}{\phi}-1}$. Since $0 < \phi < 1$, k is clearly an increasing function of m, ie $k' > 0$, and since $k'' > 0$, θ is a strictly convex function, reflecting strictly increasing

[30] This type of analysis begins with Herberg and Kemp (1968), where the equivalent function is denoted h^{-1}. Mayer (1972), whose graphical analysis we build on, also refers to this function as θ. In both of these papers, the θ function is derived from a multiplicative relationship involving a scale multiplier and a kernel constant returns production function: $M = \gamma(M)g(K_m,L_m)$. Ethier (1982) works with a slightly different formulation which, as we shall see, yields an even simpler form for the θ function. See Helpman (1984) for a clear review of models with variable returns to scale.

[31] This expression is equation (8) in Ethier (1982). Note that while $k(m)$ is explicitly a function of bundles in Ethier, the equality to M/m makes it clear that it serves the same purpose as $\gamma(M)$ in the Herberg–Kemp/Mayer analysis.

returns in the production of M. From this, it is easy to define the function $\theta(m)$ that maps bundles into final outputs as

$$M = \theta(m) := km = \left(Am^{\frac{1}{\phi}-1}\right)m = Am^{\frac{1}{\phi}} \tag{A8}$$

which is (2) in the text.

Summarising the resource constraint and the bundles part of the model by the bundles transformation function, we can derive the transformation function between W and M via a pair of mapping relations from bundles to outputs. This system is

$$\begin{aligned} w &= B(m) \\ W &= \psi(w) \\ M &= \theta(M) := Am^{\frac{1}{\phi}}. \end{aligned} \tag{A9}$$

Given the production structure assumed above, and summarised in (10), the supply side of the model has a simple graphical representation (Mayer, 1972, Fig. 1). We plot the bundles transformation curve with the usual smoothly concave curvature. The SE quadrant contains the ψ function, a ray from the origin with a slope of 1. The NW quadrant contains the strictly convex θ function. This information can be used to plot the transformation function between W and M ($W = T(M)$): every point on $B(\cdot)$ is mapped to a point on $T(\cdot)$ by ψ and θ.

Appendix B

As we note in the text, the aggregator in (9) is essentially the same as that in (1), so, for national component producers, the equilibrium marginal conditions are still given by (A4), the volume of output of each component producing firm is still a constant determined by parameters as in (A5), and the number of component producers (and, thus, varieties of components) in country r (n_r) is still a linear function of the volume of bundle production (A6). Since all producers of final manufactures have access to all varieties, we can denote the global number of varieties by

$$n^G = \frac{1-\phi}{b}\left(\sum_{r\in R} m_r\right).$$

Using these facts, we can derive an expression for k^j in exactly the same way as we derived (A7):

$$\begin{aligned} k^j &= \left[\left(\frac{1-\phi}{b}\right)^{\frac{1}{\phi}-1}\frac{\phi}{a}\right]\left(m^G\right)^{\frac{1}{\phi}-1} \\ &= A\left(m^G\right)^{\frac{1}{\phi}-1}. \end{aligned} \tag{A10}$$

Note that this, in fact, is precisely the same expression that we derived in (A7). Thus, we can still write $M^j = \theta(m^j, \mathbf{m}^{-j}) = km^j$:

$$M^j = \theta^j\left(m^j, \mathbf{m}^{-j}\right) := km^j = A\left(\sum_{r\in R} m_r\right)^{\frac{1}{\phi}-1} m^j \tag{A11}$$

where k and θ are now functions of the *global* level of component production, which is (11) in the text. It is also useful to note that, since k is the same for all countries in equilibrium, we can rewrite (10) as

$$\delta^j = \frac{m^j}{m^G} = \frac{km^j}{km^G} = \frac{M^j}{M^G}. \tag{10'}$$

Appendix C: Policy Ranking and the *FN* Schedule

In this section, we provide a more formal treatment of the ranking of economic policy, as discussed in section 3.2. Consider, for example, a subsidy for home intermediate manufactures production. From (11), we then have

$$\begin{aligned} M &= km = \Theta(m) \\ k &= A(m+m^*)^{\frac{1}{\phi}-1}. \end{aligned} \tag{A12}$$

With the introduction of (or change in) a production subsidy, the change in k will be

$$\begin{aligned} \frac{\mathrm{d}k}{\mathrm{d}s} &= \left(\frac{1}{\phi}-1\right)(m+m^*)^{\frac{1}{\phi}-2}\left[\frac{\partial m}{\partial s}+\left(\frac{\partial m^*}{\partial m}\right)\left(\frac{\partial m}{\partial s}\right)\right] \\ &= \left[\frac{1-\phi}{\phi k(m+m^*)}\right]\left[1+\left(\frac{\partial m^*}{\partial m}\right)\right]\left(\frac{\partial m}{\partial s}\right). \end{aligned} \tag{A13}$$

In (A13), the first term in square brackets captures the direct effect of subsidies on home output given ROW output, while the second captures the interaction between home production and ROW production in the manufacturing sector.

We move from this functional relationship to our graphic apparatus as follows. We have effectively defined Θ in (A12) as the locus of all equilibrium points on the linear q functions in mM space. The Θ function can now be used to trace out the experiment dependent production frontier $T(.)$ in Fig. 4. From there, the next step involves finding the locus of all points identified by consumption of final goods at the experiment equilibrium along the RPT. This follows from the imposition of final preferences and an income identity for consumption. Formally, we first impose identical homothetic preferences, such that we can specify a social welfare function of the form

$$\begin{aligned} U^i &= \Pi^i f(P) \\ \Pi^i &= M^i P + W^i \end{aligned} \tag{A14}$$

where Π denotes national income measured in units of wheat, and the term P is the price of final manufactures. With identical homothetic preferences, we can specify demand as

$$P = \eta\left(\frac{W^h + W^{\text{row}}}{M^h + M^{\text{row}}}\right) \tag{A15}$$

where $\eta > 0$. In addition, relative world supply of M and W can be specified in reduced form as

$$\left(\frac{W^h + W^{\text{row}}}{M^h + M^{\text{row}}}\right) = \varsigma(P, s) \qquad \varsigma_1 = \partial\varsigma/\partial P < 0, \varsigma_2 = \partial\varsigma/\partial s < 0. \tag{A16}$$

Taken together, (A15) and (A16) imply a mapping from the subsidy rate s to the price level P. In particular, we will have

$$P = \eta\varsigma(P, s) \Rightarrow \frac{\mathrm{d}P}{\mathrm{d}s} = \frac{\eta\varsigma_2}{(1-\eta\varsigma_1)}. \tag{A17}$$

Consider next the welfare rankings of policy – in this case a subsidy. Given (A14)–(A17), we can show that the welfare effect of shifting our subsidy – and hence moving consumption along *FN* – will be

$$\frac{\mathrm{d}U^h}{\mathrm{d}s} = f(P)[(A+BD)+CD] \tag{A18}$$

where

$$A = \left(\frac{\partial M^h}{\partial s}\right)P + \left(\frac{\partial W}{\partial s}\right)$$
$$B = \left(\frac{\partial M^h}{\partial P}\right)P + \left(\frac{\partial W}{\partial P}\right)$$
$$C = M + \left(\frac{\Pi}{f(P)}\right)\left(\frac{\partial f}{\partial P}\right)$$
$$D = [\eta\varsigma_2/(1 - \eta\varsigma_1)].$$

(See the precise derivation of this expression in Francois (1992).) In (A18), the term D is identical to the derivative found in (A17), and represents the general equilibrium shift in P as we induce movements along the bundles frontier through a subsidy. The first term in square brackets $(A+BD)$ is zero at the level that maximises national income in terms of the numeraire good W. The second term then measure the interaction of shifts in world prices with welfare as consumption moves along the *FN* frontier.

References

Barro, R. and Sala-I-Martin, X. (1995). *Economic Growth.* New York: McGraw Hill.

Bhagwati, J., Panagariya, A. and Srinivasan, T. N. (1998). *Lectures in International Trade.* Cambridge MA: MIT Press.

Brander, J. (1995). 'Strategic trade policy', in (G. Grossman and K. Rogoff, eds.), *Handbook of International Economics*, Vol. III. Amsterdam: Elsevier, pp. 1395–455.

Brown, D. (1994). 'Properties of applied general equilibrium trade models with monopolistic competition and foreign direct investment', in (J. Francois and C. Shiells, eds.), *Modeling Trade Policy: Applied General Equilibrium Assessments of North American Free Trade*, Cambridge: Cambridge University Press, pp. 124–50.

Buchanan, J. and Yoon, Y. eds (1994). *The Return to Increasing Returns.* Ann Arbor: University of Michigan Press.

Burda, M. and Dluhosch, B. (1999). 'Globalization and European labour markets', in (H. Siebert, ed.), *Globalization and Labour*, Tübingen: Mohr Siebeck, pp. 181–207.

Chao, C. C. and Takayama, A. (1990). 'Monopolistic competition, non-homotheticity, and the stability of the Chamberlinian tangency solution', *International Economic Review*, vol. 31, pp. 73–86.

Chipman, J. (1970). 'External economies of Acale and competitive equilibrium', *Quarterly Journal of Economics*, vol. 84, pp. 347–85.

Dixit, A. and Norman, V. (1980). *Theory of International Trade.* Cambridge: Cambridge University Press.

Dixit, A. and Stiglitz, J. (1977). 'Monopolistic competition and optimum product diversity', *American Economic Review*, vol. 67, pp. 297–308.

Eaton, J. and Panagariya, A. (1979). 'Gains from trade under variable returns to scale, commodity taxation, tariffs and factor market distortions', *Journal of International Economics*, vol. 9, pp. 481–501.

Ethier, W. (1979). 'Internationally decreasing costs and world trade', *Journal of International Economics*, vol. 9, pp. 1–24.

Ethier, W. (1982*a*). 'National and international returns to scale in the modern theory of international trade', *American Economic Review*, vol. 72, pp. 388–405.

Ethier, W. (1982*b*). 'Decreasing costs in international trade and Frank Graham's argument for protection', *Econometrica*, vol. 50, pp. 1243–68.

Francois, J. (1992). 'Optimal commercial policy with international returns to scale', *Canadian Journal of Economics*, vol. 25, pp. 184–95.

Francois, J. (1994). 'Global production and trade: factor migration and commercial policy with international scale economies', *International Economic Review*, vol. 35, pp. 565–81.

Francois, J. and Nelson, D. (2000). 'Victims of progress: globalization, specialization, and wages for unskilled labour', CEPR discussion paper 2527.

Francois, J. and Roland-Holst, D.W. (1997), 'Scale economies and imperfect competition', in (J. F. Francois and K. A. Reinert, eds.), *Applied Methods for Trade Policy Analysis: A Handbook*, Cambridge: Cambridge University Press, pp. 331–63.

Fujita, M., Krugman, P. and Venables, A. (1999). *The Spatial Economy: Cities, Regions, and International Trade*. Cambridge: MIT Press.

Gomory, R. (1994). 'A Ricardo model with economies of scale', *Journal of Economic Theory*, vol. 62, pp. 394–419.

Grossman, G. and Helpman, E. (1991). *Innovation and Growth in the Global Economy*. Cambridge, MA: MIT Press.

Helpman, E. (1984). 'Increasing returns, imperfect markets and trade theory', in (R. Jones and P. Kenen, eds.), *Handbook of International Economics*, vol. 1, Amsterdam: North-Holland, pp. 325–65.

Helpman, E. and Krugman, P. (1985). *Market Structure and Foreign Trade*. Cambridge, MA: MIT Press.

Herberg, H. and Kemp, M. (1968). 'Some implications of variable returns to scale', *Canadian Journal of Economics*, vol. 2, pp. 403–15.

Herberg, H. and Kemp, M. (1991). 'Some implications of variable returns to scale: the case of industry specific factors', *Canadian Journal of Economics*, vol. 24, pp. 703–4.

Holtz-Eakin, D. and Lovely, M. (1996*a*). 'Scale economies, returns to variety, and the productivity of public infrastructure', *Regional Science and Urban Economics*, vol. 26, pp. 105–23.

Holtz-Eakin, D. and Lovely, M. (1996*b*). 'Technological linkages, market structure, and production policies', *Journal of Public Economics*, vol. 61, pp. 73–86.

Ide, T. and Takayama, A. (1991). 'Variable returns to scale, paradoxes and global correspondences in the theory of international trade', in (A. Takayama, M. Ohyama, and H. Ohta, eds.), *Trade, Policy, and International Adjustments*, San Diego: Academic Press, pp. 108–54.

Ide, T. and Takayama, A. (1993). 'Variable returns to scale, comparative statics paradoxes, and the theory of comparative advantage', in (H. Herberg and N. V. Long, eds.), *Trade, Welfare and Economic Policies*, Ann Arbor: University of Michigan Press.

Jones, R. (1968). 'Variable returns to scale in general equilibrium theory', *International Economic Review*, vol. 9, pp. 261–72.

Kemp, M. C. (1964). *The Pure Theory of International Trade*. Englewood Cliffs, NJ: Prentice-Hall.

Kemp, M. C. and Negishi, T. (1970). 'Variable returns to scale, commodity taxes, factor market distortions and their implications for trade gains', *Swedish Journal of Economics*, vol. 72, pp. 1–11.

Krugman, P. (1979). 'Increasing returns, monopolistic competition, and international trade', *Journal of International Economics*, vol. 9, pp. 469–79.

Krugman, P. (1980). 'Scale economies, product differentiation, and the pattern of trade', *American Economic Review*, vol. 70, pp. 950–9.

Krugman, P. (1995). 'Complexity and emergent structure in the international economy', in (J. Levinsohn, A. Deardorff, and R. Stern, eds.), *New Directions in Trade Theory*,' Ann Arbor: University of Michigan Press, pp. 23–46.

Krugman, P. and Venables, A. (1995). 'Globalization and the inequality of nations', *Quarterly Journal of Economics*, vol. 110, pp. 857–80.

Lovely, M. (1997). 'Playing by the new subsidy rules: capital subsidies as substitutes for sectoral subsidies', *Journal of International Economics*, vol. 43, pp. 463–82.

Lovely, M. and Nelson, D. (2000). 'Marginal intraindustry trade and labour adjustment', *Review of International Economics*, vol. 8, pp. 436–47.

Markusen, J. (1988). 'Production, trade, and migration with differentiated, skilled workers', *Canadian Journal of Economics*, vol. 21, pp. 492–506.

Markusen, J. (1989). 'Trade in producer services and in other specialized inputs', *American Economic Review*, vol. 79, pp. 85–95.

Markusen, J. (1990*a*). 'Micro-foundations of external economies', *Canadian Journal of Economics*, vol. 23, pp. 495–508.

Markusen, J. (1990*b*). 'Derationalizing tariffs with specialized intermediate inputs and differentiated final goods', *Journal of International Economics*, vol. 28, pp. 375–83.

Markusen, J. (1991). 'First mover advantages, blockaded entry, and the economics of uneven development', in (E. Helpman and A. Razin, eds.), *International Trade and Trade Policy*, Cambridge MA: MIT Press, pp. 245–69.

Markusen, J. and Melvin, J. (1984). 'The gains from trade theorem with increasing returns to scale', in (H. Kierzkowski ed.), *Monopolistic Competition and International Trade*, New York: Oxford University Press, pp. 10–33.

Matsuyama, K. (1996). 'Why are there rich and poor countries? Symmetry-breaking in the world economy', *Journal of the Japanese and International Economies*, vol. 10, pp. 419–39.

Mayer, W. (1972). 'Homothetic production functions and the shape of the production possibility locus', *Journal of Economic Theory*, Vol. 8, pp. 101–10.

Panagariya, A. (1980). 'Variable returns to scale in general equilibrium once again', *Journal of International Economics*, vol. 10, pp. 221–30.

Panagariya, A. (1981). 'Variable returns to scale in production and patterns of specialization', *American Economic Review*, vol. 71, pp. 221–30.

Panagariya, A. (1986). 'Increasing returns, dynamic stability, and international trade', *Journal of International Economics*, vol. 20, pp. 43–63.

Puga, D. and Venables, A. (1996). 'The spread of industry: spatial agglomeration in economic development', *Journal of the Japanese and International Economies*, vol. 10, pp. 440–64.

Puga, D. and Venables, A. (1998). 'Trading arrangements and industrial development', *World Bank Economic Review*, vol. 12, pp. 221–49.

Rivera-Batiz, F. and L. (1991). 'The effects of direct foreign investment in the presence of increasing returns due to specialization', *Journal of Development Economics*, vol. 34, pp. 287–307.

Rodríguez-Clare, A. (1996). 'The division of labour and economic development', *Journal of Development Economics*, vol. 49, pp. 3–32.

Rodrik, D. (1996). 'Coordination failures and government policy: a model with applications to East Asia and Eastern Europe', *Journal of International Economics*, vol. 40, pp. 1–22.

Romer, P. (1987). 'Growth based on increasing returns due to specialization', *American Economic Review*, vol. 77, pp. 56–62.

Romer, P. (1990). 'Endogenous technical change', *Journal of Political Economy*, vol. 98, pp. s71–102.

Spence, A. M. (1976). 'Product selection, fixed cost and monopolistic competition', *Review of Economic Studies*, vol. 43, pp. 217–35.

van Marrewikj, C., Stibora, J., de Vaal, A. and Viaene, J. M. (1997). 'Producer services, comparative advantage, and international trade patterns', *Journal of International Economics*, vol. 42, pp. 195–220.

Venables, A. (1995). 'Economic integration and the location of firms', *American Economic Review*, vol. 85, pp. 296–300.

Venables, A. (1996*a*). 'Trade policy, cumulative causation, and industrial development', *Journal of Development Economics*, vol. 49, pp. 179–97.

Venables, A. (1996*b*). 'Equilibrium location of vertically linked industries', *International Economic Review*, vol. 37, pp. 341–58.

Venables, A. and Krugman, P. (1996). 'Integration, specialization, and adjustment', *European Economic Review*, vol. 40, pp. 959–67.

Wong, K. Y. (1996). 'A comment on ''Some implications of variable returns to scale: The case of industry-specific factors''', *Canadian Journal of Economics*, vol. 29, pp. 240–4.

Young, A. (1928). 'Increasing returns and economic progress', ECONOMIC JOURNAL; vol. 38, pp. 527–40.