A Geometry of Specialization

Joseph F. Francois
Tinbergen Institute and CEPR

Douglas Nelson
Tulane University and CREDIT

January 1998
(revised)

Abstract: Division of labor 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, specialization models lack a canonical representation. This is because they are both new and complex, characterized 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 specialization models, and explaining why one sub-class of these models is particularly difficult to illustrate easily.

JEL classifications: [F12],[O12],[O41]
keywords: specialization, trade and development, location of industry, agglomeration

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. While working on this paper, the authors benefitted 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.

address correspondence to: J.F. Francois, Faculty of Economics H8-23, Erasmus University Rotterdam, postbus 1738, 3000DR Rotterdam, NETHERLANDS. Fax: +31 10 452 5808.
Email: francois@few.eur.nl.

A Geometry of Specialization

NONTECHNICAL SUMMARY

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. These are strategic trade models and division of labor models. The strategic trade models entered the literature with simple graphical representations developed in the industrial organization literature, while the division of labor models have proven to be considerably more resistant to simple representation.

The recent specialization literature leans on special models built around specific functional forms, and often involves numeric simulation. Even so, a set of general results (low level equilibrium traps, catastrophic adjustment, agglomeration effects) does 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 generalized treatment that links this pattern of results to the general properties of models with increasing returns due to specialization. In the process, we demonstrate that important results in the recent literature depend critically on the stability and transformation properties that characterize 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.

We begin with two versions of national production externality (NPE) models. In the first, a closed-economy version of the model, we develop the basic elements of the Ethier-type division of labor model 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, 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 differentiated goods is permitted. The graphical analysis makes the locus of this difficulty clear. In addition, we develop a new graphical apparatus that is directly analogous to the Baldwin envelope for the case of division of labor models. 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.

A Geometry of Specialization

1. Introduction

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. These are strategic trade models and division of labor models. The strategic trade models entered the literature with simple graphical representations developed in the industrial organization literature, while the division of labor models have proven to be considerably more resistant to simple representation.

The recent specialization literature leans on special models built around specific functional forms, and often involves numeric simulation. Even so, a set of general results (low level equilibrium traps, catastrophic adjustment, agglomeration effects) does 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 generalized treatment that links this pattern of results to the general properties of models with increasing returns due to specialization. In the process, we demonstrate that important results in the recent literature depend critically on the stability and transformation properties that characterize 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.

The central role of division of labor 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, 1982a) 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 labor 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 both trade patterns (Ethier 1979, 1982; Markusen 1988, 1989; van Marrewikj *et al*. 1997) and trade policy (Markusen 1990; Francois 1992, 1994; Lovely 1997). One of the most intersting 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; 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; Chapter 6) and has been used extensively in development theory (Rodriguez-Clare 1996 and Rodrik 1996) and regional economics (Holtz-Eakin and Lovely, 1996a and 1996b).

In this paper we proceed as follows. To provide some structure to the exercise, we have divided the general family of specialization models into 5 types, which are 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 the basic elements of the Ethier-type division of labor model 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. In addition, in this section we develop a new graphical apparatus that is directly analogous to the Baldwin envelope for the case of division of labor models. 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.

2. National Production Externalities in Autarky: Model I

2.A 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, labor (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 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 labor 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. The market for components is monopolistically competitive, and bundles production is perfectly competitive.

Ethier's key insight was that the Spence (1976)-Dixit-Stiglitz (1977) model of preference for variety, applied to international trade by Krugman (1979, 1980), when applied to production consititutes the basis of a model of division of labor. The model contains two main elements: (1) A technology reflecting increasing returns to "division of labor"; and (2) Something limiting the division of labor (i.e. "the extent of the market").

The first element is given by a CES function that costlessly aggregates components (x_i) into finished manufactures:[1]

$$M = \left[\sum_{i \in n} x_i^{\phi} \right]^{\frac{1}{\phi}} . \quad \textbf{(1)}$$

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).[2] In particular, note that if $x_i = x \; \forall \; i \in n$, equation (1) reduces to $M = n^{\frac{1}{\phi}} x$. 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,

---

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

$$M = n^{\alpha} \left[\sum_{i \in n} \frac{x_i^{\phi}}{n} \right]^{\frac{1}{\phi}} . \quad \textbf{(1a)}$$

Note that, in both formulations, the elasticity of substitution between varieties is given by $1/(1 - \phi)$--i.e. the higher is ϕ, the more easily can the x_i be substituted for one another in production. In (1a) 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 (1) and (1a) 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 in (1), where returns to division of labor emerge directly from the imperfect subsitutability between varieties of inputs. For an analysis of policy that exploits the distinct effects in (1a) see Holtz-Eakin and Lovely (1996a and 1996b).

[2]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.

i.e.: $\frac{\partial M}{\partial n} = \frac{1}{\phi} n^{\frac{1-\phi}{\phi}} x > 1.$[3] The smaller is ϕ, the stronger are the returns to the division of labor.

Component production is the final essential element of the standard division of labor model -- and the locus of the actual division of labor. Where production of final manufactures is characterized 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 labor (denoted m) according to the relation:

$$m_i = a x_i + b. \qquad \textbf{(2)}$$

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 $x_i = x \; \forall \; i \in n$, then $m = n(ax + b)$. As a result of decreasing costs, no two firms will produce the same type of component. The left panel of Figure 1 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 $\left(\bar{K}, \bar{L}\right)$ this is summarized 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 ψ is a linear relationship, i.e. $\psi(w) = a_{wW} w$. 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

---

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

non-linear function.

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

$$P_m = -B'(m). \tag{3}$$

Now we want to link bundles to final production of manufactures: $M = \theta(m)$.[4] 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 $-B'(m)[ay_j + b]$ and total revenue is just $p_j x_j$, so the condition that marginal revenue equals marginal cost can be rearranged to get an expression for p_j:

$$p_j = -B'(m)\frac{a}{\phi}. \tag{4}$$

In deriving (4) we substitute $\frac{dx_j}{dp_j} = \frac{1}{1-\phi}\frac{x_j}{p_j}$, which can be derived from the demand curve for component j (See Ethier, 1982, equation (4)). The profit of each component producing firm is $\pi_j = p_j x_j + B'(m)[a x_j + b]$ 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 equation (4), we get:

$$x_j = \frac{b\phi}{a(1-\phi)}. \tag{5}$$

Since this is made up entirely of parameters that are constant across component producing firms, equation (5) 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:

---

[4]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.

$$n = \frac{(1-\phi)}{b} m . \qquad \textbf{(6)}$$

Note the implication that *m* and *n* are linearly related. Ethier works with a function *k*(*m*), given by *M* = *km*. We can use (5) and (6), along with the fact that $M = n^{\frac{1}{\phi}} x$, to get an expression for $k = \frac{M}{m}$:[5]

$$k = \left(\left[\frac{(1-\phi)}{b} \right]^{\frac{1}{\phi}-1} \frac{\phi}{a} \right) m^{\frac{1}{\phi}-1} . \qquad \textbf{(7)}$$

Furthermore, since the expression in parentheses is entirely made up of parameters, we can write this as $k = A m^{\frac{1}{\phi}-1}$. Since 0 < ϕ < 1, *k* is clearly an increasing function of *m*, i.e. ϕ′ > 0, and since ϕ″ > 0, θ is a strictly convex function, reflecting strictly increasing returns in the production of *M*. From this, it is easy to define the function θ(*m*) that maps bundles into final outputs as:

$$M = \theta(m) := km \equiv \left[A m^{\frac{1}{\phi}-1} \right] m \equiv A m^{\frac{1}{\phi}} . \qquad \textbf{(8)}$$

It will be useful in the later analysis to have expressions for the first and second derivatives of θ:

$$\theta'(m) = \frac{1}{\phi} A m^{\frac{1}{\phi}-1} > 0 ,$$
$$\theta''(m) = \frac{\frac{1}{\phi}-1}{\phi} A m^{\frac{1}{\phi}-2} > 0 . \qquad \textbf{(9)}$$

Summarizing 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:

[5]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 γ(*M*) in the Herberg-Kemp/Mayer analysis.

$$
\begin{aligned}
w &= B(m) \\
W &= \psi(w) \\
M &= \theta(m) := A\, m^{\frac{1}{\phi}}
\end{aligned}
\qquad \textbf{(10)}
$$

Given the production structure assumed above, and summarized in equations (10), the supply side of the model has a simple graphical representation (Mayer, 1972; figure 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.

Herberg and Kemp (1968) and Mayer (1972) have intensively studied precisely the system described in (10) 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, pg. 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]

$$
\begin{aligned}
\tau' &= \frac{dM}{dW} = \left(\frac{\theta'}{\psi'}\right)\beta', \text{ and} \\
\tau'' &= \frac{d^2M}{dW^2} = \frac{\theta'}{(\psi')^2}\left[\left(\frac{\theta''}{\theta'}\right)(\beta')^2 - \left(\frac{\psi''}{\psi'}\right)\beta' + \beta''\right].
\end{aligned}
\qquad \textbf{(11)}
$$

In the baseline case of HOS structure for bundles production, $\psi' = 1$ and $\psi'' = 0$, so the expressions in (11) are considerably simplified to:

$$
\begin{aligned}
\tau' &= \frac{dM}{dW} = \theta'\beta', \text{ and} \\
\tau'' &= \frac{d^2M}{dW^2} = \theta'\left[\left(\frac{\theta''}{\theta'}\right)(\beta')^2 + \beta''\right].
\end{aligned}
\qquad \textbf{(12)}
$$

[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*.

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 Figure 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 $\frac{\theta''}{\theta'} \to \infty$ as $m \to 0$, then the transformation function must be convex in the neighborhood of zero manufacturing output.[7] Given the derived expressions in equation (9), it is easy to see that

$$\frac{\theta''}{\theta'} = \frac{\frac{1}{\phi} - 1}{m} > 0\,, \tag{13}$$

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 (i.e. almost infinitely) tightly curved in the neighborhood 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 convex portion of $T(M)$ in the neighborhood of zero W output. In particular, it is easy to see that (13) and β' both get smaller as the output of M increases, implying that the first term in the square brackets in (12) gets 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 neighborhood of zero M output*. This is an important point. The frontier may, in general, be characterized 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.

2.B 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 1982b; Gomory, 1994) labor 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.

[7] This is a result of Mayer's (1972, pp. 106-109) which refines a result originally presented in Herberg and Kemp (1968).

(Figure 2a). Wheat is produced with a constant returns to scale production function, components are produced with a fixed and variable component (now paid entirely in labor) and manufactures are produced from components according to equation (1). We now have that the "bundles" transformation function (i.e. the labor constraint) is characterized by $B' = -1$ and $B'' = 0$. Since ψ is still a linear map with a slope of 1, we find from equation (10) that $T' = -\theta'$ 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 Figure 2a since the θ function is the only source of curvature, while both B and ψ have unit slopes.

The Ricardo-Viner structure (Figure 2b) has been extensively used in an important series of papers by Markusen (1988, 1989, 1990a, 1990b, 1991). Consider the simplest version of this model: wheat is produced with mobile labor and specific capital; components are produced with labor 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 labor 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 labor (i.e. $\psi' > 0$ and $\psi'' < 0$). As a result, we cannot use the expressions for τ' and τ'' in (11) but must use those in (12). On the other hand, the Ricardian resource constraint still permits us to take $B' = \beta' = -1$ and $B'' = \beta'' = 0$, so we can write:

$$\tau' = \frac{dM}{dW} = -\left(\frac{\theta'}{\psi'}\right), \text{ and}$$
$$\tau'' = \frac{d^2M}{dW^2} = \frac{\theta'}{(\psi')^2}\left[\left(\frac{\theta''}{\theta'}\right) + \left(\frac{\psi''}{\psi'}\right)\right]. \qquad \textbf{(14)}$$

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 equation (13), goes to positive infinity in the neighborhood 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 neighborhood of zero manufacturing output*.[8] This is an important source of multiple equilibria in the literature.

2.C *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 Figure 1b. 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 specialization 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 Figure 1b, domestic prices will not be tangent to the $T(\cdot)$ frontier at this point. (See Markusen 1990). 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), the standard Rybczynski and Stolper-Samuelson results hold (in terms of wheat and bundles). However, the welfare calculus is complicated by variety effects.

2.D *Economic Growth*

In addition to the implications of returns to specialization 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

---

[8]Following Markusen (1989) we can get some additional leverage by considering specific functional forms. For example, in the Cobb-Douglas case we have $\frac{\psi''}{\psi'} = \frac{\alpha - 1}{L_W} < 0$. This expression goes to negative infinity as output of wheat goes to zero. Mayer (1972), again expositing a result due to Herberg and Kemp (1969), shows that in this case *T* must be strictly concave in the relevant neighborhood. 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).

[9]See Bhagwati, Panagariya, and Srinivasan (forthcoming) for a concise discussion of the optimal variety issue.

model, there will be an expansion of the production possibility frontier (the $T(\cdot)$ frontier), with a bias toward the capital intensive sector. With labor 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/labor ratio and zero growth. This process can be fundamentally altered, however, by the simple addition of returns to specialization. Because the θ function is no longer linear, the decline in the return to capital is moderated by returns to specialization. (Grossman and Helpman, 1991, Chapter 4). If returns to specialization 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 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. Specialization economies can deliver the required externalities. It is the curvature of the θ function that proves critical to determining whether specialization 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)

3. National Production Externalities 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 (i.e. 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 equation (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, figures 1 - 2 continue to characterize 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).

3.B Stability, Low-level Development Traps, and the "Big Push"

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 specialize 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, 1982b).[10]

Many of the insights of the recent literature on forward linkages, development, and specialization (Rodrik 1996; Rodriguez-Clare 1996; Rivera-Batiz and Rivera-Batiz 1991; Venables 1996) follow directly from this property of local agglomeration models. Basically, because specialized primary/wheat production involves a stable equilibrium, and because more developed economies, by definition, have cost advantages related to larger and more specialized upstream industries, there is a tendency for underdeveloped countries to stay that way. Figure 3 illustrates this point, with a national $T(\cdot)$ frontier that

[10] Where Panagariya and Ethier adopt a Ricardian model, Markusen and Melvin (1981, proposition 1) and Ide and Takayama (1993a, 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 characterize 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 (1982b) for the Ricardian Case; Panagariya (1986) for the Ricardo-Viner case, and Ide and Takayama (1991, 1993a) for the HOS case.

has two inflection points. Internal equilibria in the concave regions will be Marshallian stable, while equilibria in the convex regions will be unstable.[11] If we assume that an underdeveloped country opens up to trade, it will start with an equilibrium like E_A. The stability properties of the model mean that there will be a catastrophic collapse of the manufacturing sector, with the economy settling at point E_T.

Local agglomeration effects also mean that there is scope for beggar-thy-neighbor commercial policy. Looking again at Figure 3, a successful temporary intervention that pushes an economy past a critical minimum level of industrialization (a "Big Push" past the first unstable equilibrium) will lead the economy onto a self-sustained march up the $T(\cdot)$ frontier to the good equilibrium, represented by E_G. In a large country context, it is quite possible that this will also involve pushing other countries down to their own specialized equilibria, with a consequent collapse of the trading partner's industry. In such a world, catastrophic agglomeration may accompany catastrophic collapse.

The beggar-thy-neighbor implications of the model follow from the importance of regional agglomeration effects relative to global agglomeration effects. In the open economy NPE model (the Type II model), we have the extreme case of only local effects. The next sections deal with the IPE class of models (Types III & IV), where agglomeration effects are more global (at least in the absence of trading costs). Not surprisingly, we will see that the importance of location is less important (even trivial) when we have purely global agglomeration effects. Even in this class of models, however, the introduction of trading costs then pushes us back toward the realm where relative country size and first mover advantage matters.

3.C More on Growth

The growth properties of Type I models covered in Section 2 extend to Type II models as well, where we now have the addition of terms-of-trade effects. Hence, just as we have good and bad static equilibria, the same labels can also be applied to dynamic paths. For example, following Rodriguez-Clare, assume that

---

[11]This isn't quite a general property of the model as developed to this point. As Ide and Takayama (1991) make clear, reflecting a criticism due to Panagariya (1980), this rests on Jones' (1968) 3 additional production assumptions. As this paper is about the geometric properties of specialization models, a fair question (though well beyond the scope of the present paper) is the formal relationship of stability to the geometric properties developed here. We limit ourselves here to noting the following. It can be demonstrated that if we have stable regions on the MW frontier in Type I/II models (meaning that sufficient assumptions for stability are met), then those stable regions will be loosely (though not always) associated with concave regions of the MW frontier in the upper right quadrant. This is because, for both stability and concavity, we require that the transformation constraints represented by the bundles frontier in the lower left quadrant be sufficiently stronger than the scale effects represented in the upper left quadrant.

we have a country specialized in W production, and assume that the return to capital will be relatively low in this equilibrium.[12] In a dynamic context where capital accumulation is sensitive to the rate of return, this means that the long-run path of the capital stock will be locked in at low levels. Similarly, foreign investment may stay away for the same reason. If this can be overcome (and again, a temporary Big Push may be sufficient), then the economy may be moved onto a long run growth path that involves more manufacturing and more capital.

4. International Production Externalities : Models III and IV

While, as we have seen, Ethier's model of the division of labor 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 specialization 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 formalization 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 labor 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 visualize 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.

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 (i.e. $\psi(w)$). In all countries, all markets are taken to be perfectly competitive, except the market for components which is monopolistically competitive. A given country, $j \in R$, assembles components into final manufactures according to the following aggregator function:

$$M^j = \left[\sum_{r \in R} \sum_{i \in n_r} (x_i^r)^\phi \right]^{\frac{1}{\phi}} . \qquad \textbf{(15)}$$

---

[12] In Rodriguez-Clare (1996), there are actually 2 sectors that draw on specialized intermediate inputs, with one sector being more intensive in the specialized inputs, and hence benefitting most from specialization economies. In the present context, this involves the M secor being much more intensive in specialized intermediates.

Roman subscripts and superscripts are country identifiers, greek superscripts are numbers (i.e. powers). In the two-country case the Home country will have no superscript and Foreign magnitudes will be starred. I.e., 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 (given by equation (5)) into equation (16) 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.[13] As a result, since $\sum_{i \in n_r} (x_i^{\,r})^\phi = n_r x_r^\phi$, and letting $n^G = \sum_{r \in R} n^r$, we can write (15) as:

$$M^j = \left[\sum_{r \in R} n^r \left(x^r \right)^\phi \right]^{\frac{1}{\phi}} = \left(n^G \right)^{\frac{1}{\phi}} x \,. \qquad \textbf{(16)}$$

Furthermore, since all component producers produce the same quantity (given by equation (5)) 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:

$$\delta_j = \frac{m^j}{m^G} \,. \qquad \textbf{(17)}$$

What we are really interested in is an expression for $\theta(m)$ incorporating the possibility of imported intermediate components. The aggregator in (16) is essentially the same as that in (1), so for national component producers, the equilibrium marginal conditions are still given by (4), the volume of output of each component producing firm is still a constant determined by parameters as in (5), and the number of component producers (and, thus, varieties of components) in country r (n_r) is still a linear function of the

[13]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.

volume of bundle production (6). 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 equation (7):

$$\begin{aligned} k^j &= \left(\left[\frac{(1-\phi)}{b} \right]^{\frac{1}{\phi}-1} \frac{\phi}{a} \right) \left(m^G\right)^{\frac{1}{\phi}-1} \\ &= A\left(m^G\right)^{\frac{1}{\phi}-1}. \end{aligned} \tag{18}$$

Note that this, in fact, is precisely the same expression that we derived in (7). Thus, we can still write $M^j = \theta(m^j, \boldsymbol{m}^{-j}) = km^j$:

$$M^j = \theta^j(m^j, \boldsymbol{m}^{-j}) := k m^j \equiv A\left(\sum_{r \in R} m_r\right)^{\frac{1}{\phi}-1} m^j, \tag{19}$$

where k and θ are now functions of the *global* level of component production. It is also useful to note that, since k is the same for all countries in equilibrium, we can rewrite (17) as:

$$\delta_j = \frac{m^j}{m^G} = \frac{k m^j}{k m^G} = \frac{M^j}{M^G}. \tag{17′}$$

Before considering the two-country case (as an approach to the R country case), we briefly note the analytical simplification purchased by assuming either 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, rest-of-world (or large country) production completely determines the magnitude of the term in parentheses on the RHS 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.

4.A *Allocation Curves and the Baldwin Envelope*

Now suppose that there are two countries (Home and Foreign), both large. Ethier's allocation curves, graphed below the SW quadrant in Figure 4, are used to identify the equilibrium quantities m and

m^*.[14] At this equilibrium, m and m^* determine the equilibrium value of $k = \bar{k}m$. From this we have $\theta = km$. That is, θ is a linear function (shown in the NW quadrant of Figure 4). 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 (i.e. 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 figure 4, this will be a tangency between an indifference curve and the national income line. Using the equilibrium $\bar{k}m$ 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, adjusted by the fact that $kP = p$. 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 characterization 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 k and θ in the latter case. Since θ is not a technological fact, we cannot draw a

[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.

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 *k* as a constant value to draw Figure 4. As a result, there cannot be offer curves or excess supply curves of the usual sort. This, of course, is why Ethier developed the allocation curve technique. As an aid to visualizing trade policy, we now construct an *experiment dependent set of production and consumption frontiers.* The consumption frontier is an analogue to the Baldwin envelope. 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. Every level of Home *m* output is associated with a new policy dependent equilibrium characterized by a new policy-dependent value of *k*, 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 by the same *k*. Now we define two experiment dependent functions relating *m* to *M*. The first, denoted Θ_P, is the locus of all equilibrium points on the linear θ functions in *mM* space. The Θ_P function can now be used to trace out the experiment dependent production frontier $T(\cdot)$ in Figure 5, which we will refer to as the realized product transformation (RPT) frontier. The second step involves finding the locus of all points identified by consumption of final goods at the experiment equilibria along the RPT. It is this envelope, designated *FN* in Figure 5, that traces out the analogue to the Baldwin envelope in the sense that the tangency between the highest indifference curve and this envelope is welfare maximizing for the Home country.

4.B 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.[15] Their existence will depend on the relative curvature of home and foreign $B(\cdot)$ frontiers and the relative importance of returns to specialization.

Recall that we have drawn an experiment-dependent $T(\cdot)$ frontier in Figure 5, where we add

[15] 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.

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 subdisy 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 can not 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 specialization economies, and the underlying preference structure. When we have a foreign region made up of many countries, the non-linearity of n^* will be even more evident. It is, in fact, the non-linearity 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 equation (19).

4.C Development and the location of industry

One insight that carries over from the discussion of NPE models to IPE models is that, again, there are likely to be good and bad equilibria. This is again illustrated by Figure 5. In the figure, we assume an initial equilibrium with income defined by the line P_0, and with consumption at point C_0. Government intervention may be used to move the economy from P_0 to P_1, with an increase in consumption (and welfare) from C_0 to C_1. Depending on the stability properties of the model in the region of these two equilibria, it may be sufficient to again engage in a Big Push, with the economy sustaining the new equilibrium on its own.

It is useful to note that, for the example developed in Figure 5, 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 re-locate 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).

We illustrate one last point with Figure 5. Like the class of Type II models, countries may be trapped in a low level equilibrium trap. We have illustrated one in the figure for a less-developed country with production on the horizontal axis, and with consumption at point *N*.

5. The IPE Model And Trading Costs

5.A The Ricardian IPE model and trading costs

An important area of research with Type IV models involves the implications of trading costs. (Venables 1996a,b; Krugman and Venables 1995, 1996). 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 Figure 6a. The autarky transformation frontier $T(\cdot)$ is represented in the upper right quadrant by the curve *142*. Consider next the integrated 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.[16] 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, where *k* remains constant as long as we are reproducing the integrated equilibrium. 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-

[16] Formally, the expression for k' in the region of an internal equilibrium, with identical Ricardian technologies, collapses to zero for the Type III/IV model.

tangency condition first discussed in Section 2, which will also characterize the integrated equilibrium. In the absence of terms-of-trade effects (which we have sterilized 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.[17]

Next, consider the effect of transport costs. This is represented in Figure 6b. We will have an inward shift of the $T(\cdot)$ frontier in regions where both countries produce *m*.[18] Whether or not either country specializes will depend on relative trading costs for intermediate and final goods. If the foreign country does specialize 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 specialization 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 *k* function (see the upper left quadrant), with the Θ mapping also rejoining the autarky mapping at the point where the foreign country specializes.

The appearance of the $T(\cdot)$ frontier, with trading costs, resembles closely the frontier for the Type II class of models. In particular, in the present example of the Ricardian model, we have the type of

---

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

[18] 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 (See 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, specialization economies (and related convexities) will carry downstream to local *M* production.

frontier examined by Kemp (1964, 1970).[19] 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 generalization, 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 specialization will not necessarily be continuous (hence also for k and the Θ mapping).

5.B The Heckscher-Ohlin IPE model and trading costs

We turn next to the characterization of the $T(\cdot)$ frontier for the Type IV version of the economy developed in Section 2. Recall that the economy is characterized by a Heckshcer-Ohlin structure underlying the $B(\cdot)$ frontier. In autarky, the structure of the economy can be represented as in Figure 7a. The Bundles frontier $B(\cdot)$ is strictly concave to the origin, while the existence of specialization economies, through k, imply increasing returns in production of M, and hence we have a realized product transformation frontier $T(\cdot)$ in the upper right quadrant that is characterized 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, the above diagram 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, Chapter 7) that for a global set of factor endowments $\boldsymbol{F}$ within the factor price equalization 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.[20] Hence, within $\boldsymbol{F}$, we know that the production of m and w will be that necesary to allow reproduction of the integrated equilibrium level of output. This defines one equilibrium point on the $B(\cdot)$

[19] 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 the previous note.

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

frontier, and also on the RPT curve, as represented in Figure 7a. 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 specialization of the foreign country, either in m^* or in W^*. In the case of full foreign specialization in W, the value of k is then 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 specialization 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 specialization in m^*, the contribution of m^* to the value of k becomes a constant, and we are in a situation of strictly national returns to scale (where such returns are greater that 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 k will be strictly lower (unless the foreign region is specialized in W) than without trading costs, and hence that Θ in Figure 7b will shift in. Again, for points beyond where the foreign region specializes 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 Figure 7b, 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. To tackle this problem, we offer a slightly different graphical representation of trade policy and welfare in Figure 8. Again, we have produced an experiment-driven Θ function, this time designated Θ_T. The difference from Figure 7 is that the derivation of this transformation function involves variations in an import tax/subsidy on imported intermediates. We know that one point on the Θ_T function, and hence the corresponding point on the $T(\cdot)$ frontier, will correspond to the free trade point on the frontier described in Figure 7. Other tax/subsidy values will correspond to variations in *m* production. Again, each point on the $T(\cdot)$ frontier represents an equilibrium level of realized production for some arbitrary value assigned to the import tax/subsidy on intermediates. From each point, there will be a corresponding budget line, and a corresponding consumption point. The envelope of such points is designated by *FN*. As in Figure 4, welfare maximization, given that our available policy instrument, will involve the production point corresponding to the highest possible social indifference curve along the *FN* frontier.[21] In general, welfare maximization through a combination of policy instruments will involve the highest social indifference curve within the relevant set of equilibrium consumption values, contained by the consumption envelope.

6. Summary

Division of labor 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, specialization models so far lack a canonical representation. This is because they are both new, and also highly complex. Typically, they are characterized by multiple equilibria, instability, and emergent structural properties under parameter transformation.

Given the prominence of specialization 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 characterize the generic model. We have also highlighted why one sub-class of these models is particularly difficult to illustrate easily.

[21] Depending on the instrument being varied, this may or may not be tangent to the income/budget line at that point on the consumption envelope.

References

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

Chao, C.C. and A. Takayama (1990). "Monopolistic Competition, Non-Homotheticity, and the Stability of the Chamberlinian Tangency Solution". *International Economic Review*; V.31-#1, pp. 73-86.

Chipman, J. (1970). "External Economies of Scale and Competitive Equilibrium". *Quarterly Journal of Economics*; V.84-#3, pp. 347-385.

Dixit, A. and V. Norman, (1979). *Theory of International Trade*, Cambridge University Press.

Dixit, A. and J. Stiglitz (1977). "Monopolistic Competition and Optimum Product Diversity". *American Economic Review*, V.67-#3, pp. 297-308.

Eaton, J. and A. Panagariya (1979). "Gains from Trade under Variable Returns to Scale, Commodity Taxation, Tariffs and Factor Market Distortions". *JIE*; V.9-#4, pp. 481-501.

Ethier, W. (1979). "Internationally Decreasing Costs and World Trade". *Journal of International Economics*; V.9-#1, pp. 1-24.

Ethier, W. (1982a). "National and International Returns to Scale in the Modern Theory of International Trade". *American Economic Review*; V.72-#3, pp. 388-405.

Ethier, W. (1982b). "Decreasing Costs in International Trade and Frank Graham's Argument for Protection". *Etrica*; V.50-#5, pp. 1243-1268.

Ethier, W. and H. Horn (1991). "Services in International Trade". in E. Helpman and A. Razin, eds. *International Trade and Trade Policy*. Cambridge: MIT Press, pp. 223-244.

Francois, J. (1992). "Optimal Commercial Policy with International Returns to Scale". *Canadian Journal of Economics*; V.25-#1, pp. 184-195.

Francois, J. (1994). "Global Production and Trade: Factor Migration and Commercial Policy with International Scale Economies". *International Economic Review*; V.35-#3, pp. 565-581.

Francois, J. and D.W. Roland-Holst (1997), "Scale Economies and Imperfect Competition." in J.F. Francois and K.A. Reinert, eds., *Applied Methods for Trade Policy Analysis: A Handbook*, Cambridge University Press: Cambridge.

Gomory, R. (1994). "A Ricardo Model with Economies of Scale". *Journal of Economic Theory*; V.62-#2, pp. 394-419.

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

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. 325-365.

Helpman, E. (1985). "International Trade in Differentiated Middle Products". in K. Jungenfelt and D. Hague, eds. *Structural Adjustment in Developed Open Economies*. New York: St. Martins, pp. 3-34.

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

Herberg, H. and M. Kemp (1968). "Some Implications of Variable Returns to Scale". *Canadian Journal of Economics*; V.2-#3, pp. 403-415.

Herberg, H. and M. Kemp (1991). "Some Implications of Variable Returns to Scale: the Case of Industry Specific Factors". *Canadian Journal of Economics*; V.24-#3, pp. 703-704.

Holtz-Eakin, D. and M. Lovely (1996a). "Scale Economies, Returns to Variety, and the Productivity of Public Infrastructure". *Regional Science and Urban Economics*; V.26-#2, pp. 105-23.

Holtz-Eakin, D. and M. Lovely (1996b). "Technological Linkages, Market Structure, and Production Policies". *Journal of Public Economics*; V.61-#1, pp. 73-86.

Hunter, L., J. Markusen and T. Rutherford (1994). "Trade Liberalization in a Multinational Dominated Industry: A Theoretical and Applied General Equilibrium Analysis". in J. Francois and C. Shiells, eds. New York: CUP.

Ide, T. and A. Takayama (1991). "Variable Returns to Scale, Paradoxes and Global Correspondences in the Theory of International Trade". in A. Takayama, et al. eds. *Trade, Policy, and International Adjustments*. San Diego: Academic Press, pp. 108-154.

Ide, T. and A. Takayama (1993a). "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, pp. 67-101.

Ide, T. and A. Takayama (1993b). "Variable Returns to Scale and Dynamic Adjustments: the Marshall-Lerner Condition Reconsidered". in R. Becker, M. Boldrin, R. Jones and W. Thomson, eds. *General Equilibrium, Growth, and Trade II*. San Diego: Academic Press, pp. 507-540.

Jones, R. (1968). "Variable Returns to Scale in General Equilibrium Theory". *International Economic Review*; V.9-#3, pp. 261-272.

Jones, R. and H. Kierzkowski (1990). "The Role of Services in Production and International Trade". in R. Jones and A. Krueger, eds. *The Political Economy of International Trade*. Oxford: Basil Blackwell; pp. 30-48.

Kemp, M.C. (1964). *The Pure Theory of International Trade*. Englewood Cliffs, N.J.: Prentice-Hall. Chapter 8.

Kemp, M.C. and T. Negishi (1970). "Variable Returns to Scale, Commodity Taxes, Factor Market

Distortions and Their Implications for Trade Gains." *Swedish Journal of Economics* 72 (January), pp. 1-11.

Krugman, P. (1979). "Increasing Returns, Monopolistic Competition, and International Trade". *Journal of International Economics*; V.9-#4, pp. 469-479.

Krugman, P. (1980). "Scale Economies, Product Differentiation, and the Pattern of Trade". *American Economic Review*; V.70-#5, pp. 950-959.

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: Univeristy of Michigan Press, pp. 23-46. [comments by Dixit and Matsuyama follow]

Krugman, P. and A. Venables (1995). "Globalisation and the Inequality of Nations". *Quarterly Journal of Economics*; V.110-#4, pp. 857-880.

Lopes-de-Silanes, F., J. Markusen and T. Rutherford (1994). "Complementarity and Increasing Returns in Intermediate Inputs". *Journal of Development Economics*; V.45-#1, pp. 101-119.

Lovely, M. (1997). "Playing by the New Subsidy Rules: Capital Subsidies as Substitutes for Sectoral Subsidies". *Journal of International Economics*; V.43-#3/4, pp.463-482.

Markusen, J. (1988). "Production, Trade, and Migration with Differentiated, Skilled Workers". *Canadian Journal of Economics*; V.21-#3, pp. 492-506.

Markusen, J. (1989). "Trade In Producer Services and in Other Specialized Inputs". *American Economic Review*; V.79-#1, pp. 85-95.

Markusen, J. (1990). "Micro-foundations of External Economies". *Canadian Journal of Economics*; V.23-#3, pp. 495-508.

Markusen, J. (1990). "Derationalizing Tariffs with Specialized Intermediate Inputs and Differentiated Final Goods". *Journal of International Economics*; V.28-#3/4, pp. 375-383.

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: MIT Press, pp. 245-269.

Markusen, J. and J. Melvin (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. (1995). "Complementarities and Cumulative Processes in Models of Monopolistic Competition". *Journal of Economic Literature*; V.33-#2, pp. 701-729.

Matsuyama, K. (1996). "Why are There Rich and Poor Countries? Symmetry-Breaking in the World Economy". *Journal of the Japanese and International Economies*; V.10-#4, pp. 419-439.

Matusz, S. (1996). “International Trade, the Division of Labor, and Unemployment”. *International Economic Review*; V.37-#1, pp. 71-84.

Mayer, W. (1972). "Homothetic Production Functions and the Shape of the Production Possibility Locus". *Journal of Economic Theory*; V.8-#2, pp. 101-110.

Panagariya, A. (1980). “Variable Returns to Scale in General Equilibrium Once Again”. *Journal of International Economics*; V.10-#4, pp. 221-230.

Panagariya, A. (1981). "Variable Returns to Scale in Production and Patterns of Specialization". *American Economic Review*; V.71-#1, pp. 221-230.

Panagariya, A. (1986). “Increasing Returns, Dynamic Stability, and International Trade”. *Journal of International Economics*; V.20-#1/2, pp. 43-63.

Puga, D. and A. Venables (1996). “The Spread of Industry: Spatial Agglomeration in Economic Development”. *Journal of the Japanese and International Economies*; V.10-#4, pp. 440-464.

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*; V.34-#1/2, pp. 287-307.

Rodríguez-Clare, A. (1996). “The Division of Labor and Economic Development”. *Journal of Development Economics*; V.49-#1, 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*; V.40-#1/2, pp. 1-22.

Spence, A.M. (1976). "Product Selection, Fixed Cost and Monopolistic Competition". *Review of Economic Studies*; V.43-#2, pp. 217-235.

van Marrewikj, C., J. Stibora, A. de Vaal, and J.M. Viaene (1997). “Producer Services, Comparative Advantage, and International Trade Patterns”. *Journal of International Economics*; V.42-#1/2, pp. 195-220.

Venables, A. (1995). “Economic Integration and the Location of Firms”. *American Economic Review*; V.85-#2, pp. 296-300.

Venables, A. (1996). “Trade Policy, Cumulative Causation, and Industrial Development”. *Journal of Development Economics*; V.49-#1, pp. 179-197.

Venables, A. (1996). “Equilibrium Location of Vertically Linked Industries”. *International Economic Review*; V.37-#2, pp. 341-358.

Venables, A. and P. Krugman (1996). “Integration, Specialization, and Adjustment”. *European Economic Review*; V.40-#3/5, pp. 959-967.

Wong, K.Y. (1996). “A Comment on ‘Some Implications of Variable Returns to Scale: The Case of

Industry-Specific Factors'". *Canadian Journal of Economics*; V.29-#1, pp. 240-244.

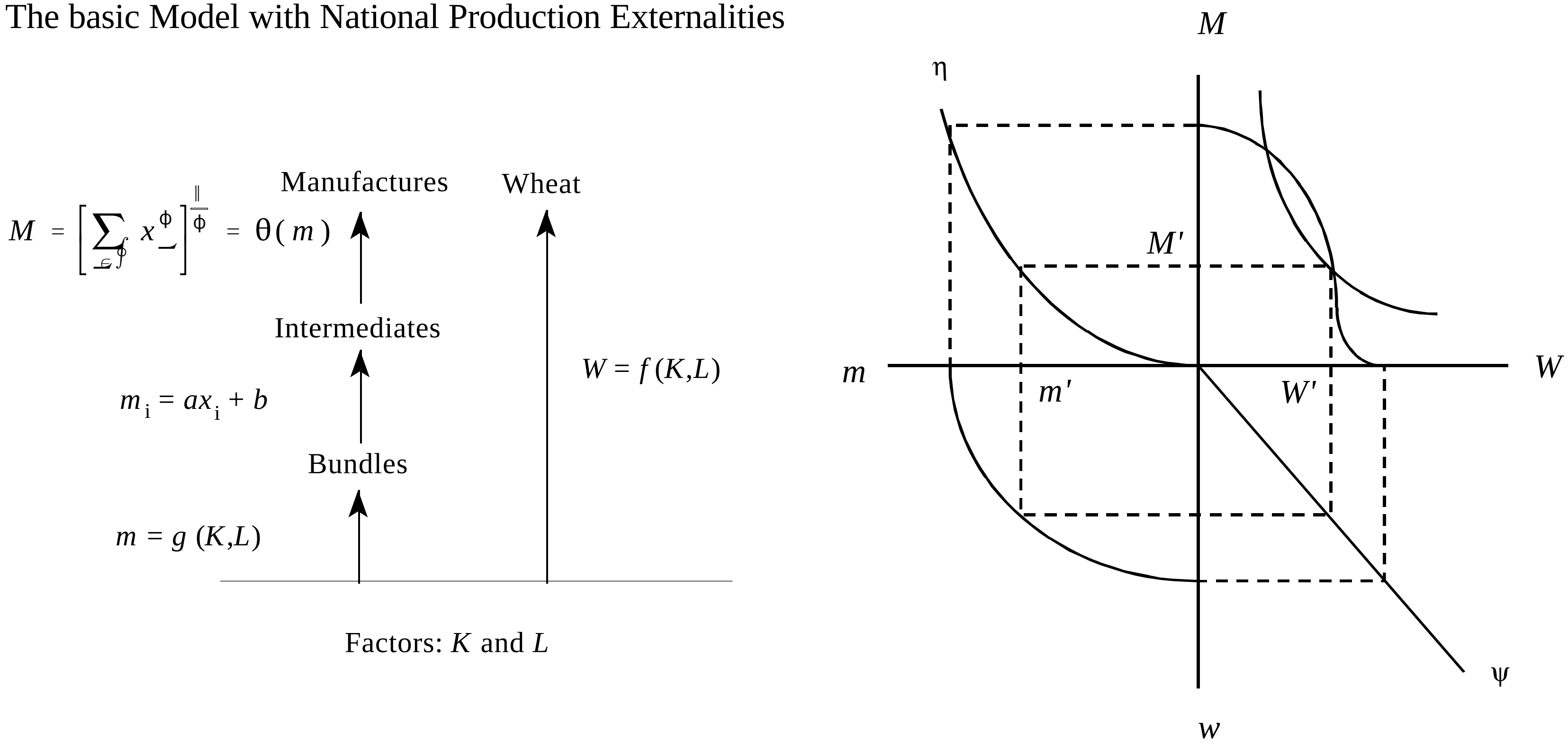
Young, A. (1928). "Increasing Returns and Economic Progress". *Economic Journal*; V.38-#152, pp. 527-540.

Table 1:
A Classification of Specialization Models

		Trade Structure	Description
National Production Externality Models	I	Closed economy	Spence-Dixit-Stiglitz, and closed economy Ethier (1982) interpretations are identical, in reduced form. The properties of this type of model are those of an external scale economy model (Markusen 1990), where production externalities are derived from variety-scaled output.
	II	Open Economy (Trade in final goods only)	This version of specialization has been explored by Markusen (1986,1990), among others. Final goods production in each region will exhibit increasing returns due to specialization. However, without direct trade in intermediates, trade has no effect on the production structure of the economy. In a multi-sector context, as long as sectors are distinct (i.e. each upstream sector serves one downstream sector), each IRTS sector will behave like an external scale economy sector.
International Production Externality Models	III	Trade in intermediates only	This is the Krugman and/or Ethier interpretation, where trade is based on the desire for varieties. In multi-sector models, such trade may be supplemented by trade based on differences in endowments. θ is affected by trade/trade policy and thus differs from national production externality models.
	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.
	V	Open Economy: Trade in intermediates and final goods with intersectoral spillovers	This is a further extension of the Ethier model, with intermediate linkages between sectors leading to inter-sectoral spillovers.

Figure 1

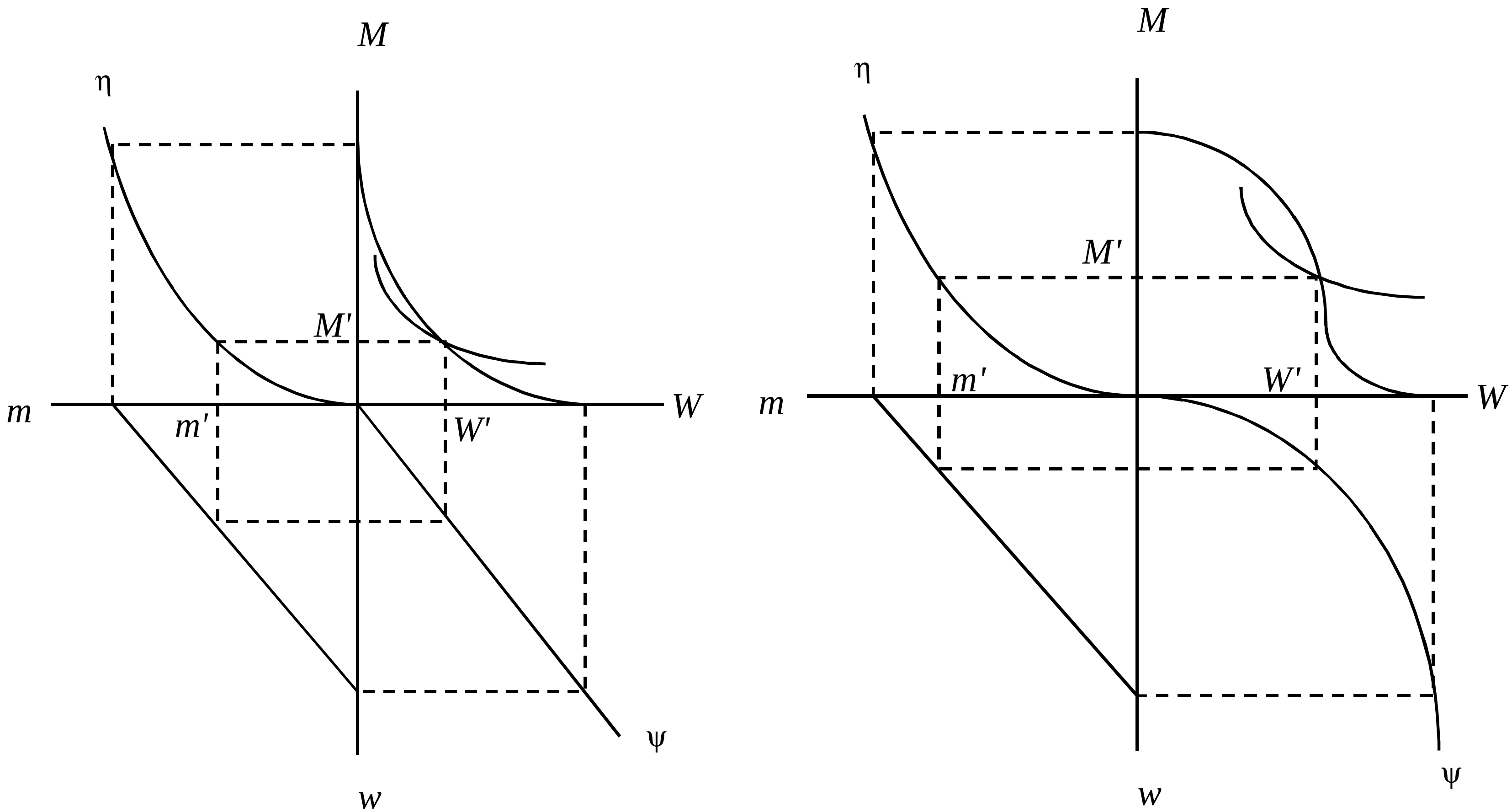
The basic Model with National Production Externalities



panel a: A Schematic Representation of Production Structure

panel b: The NPE Model with HOS Technology for Bundles

Figure 2
Ricardian Variations



Panel a: Ricardian Technology for Bundles Production

Panel b. Ricardo-Viner Technology for Bundles Production

Figure 3

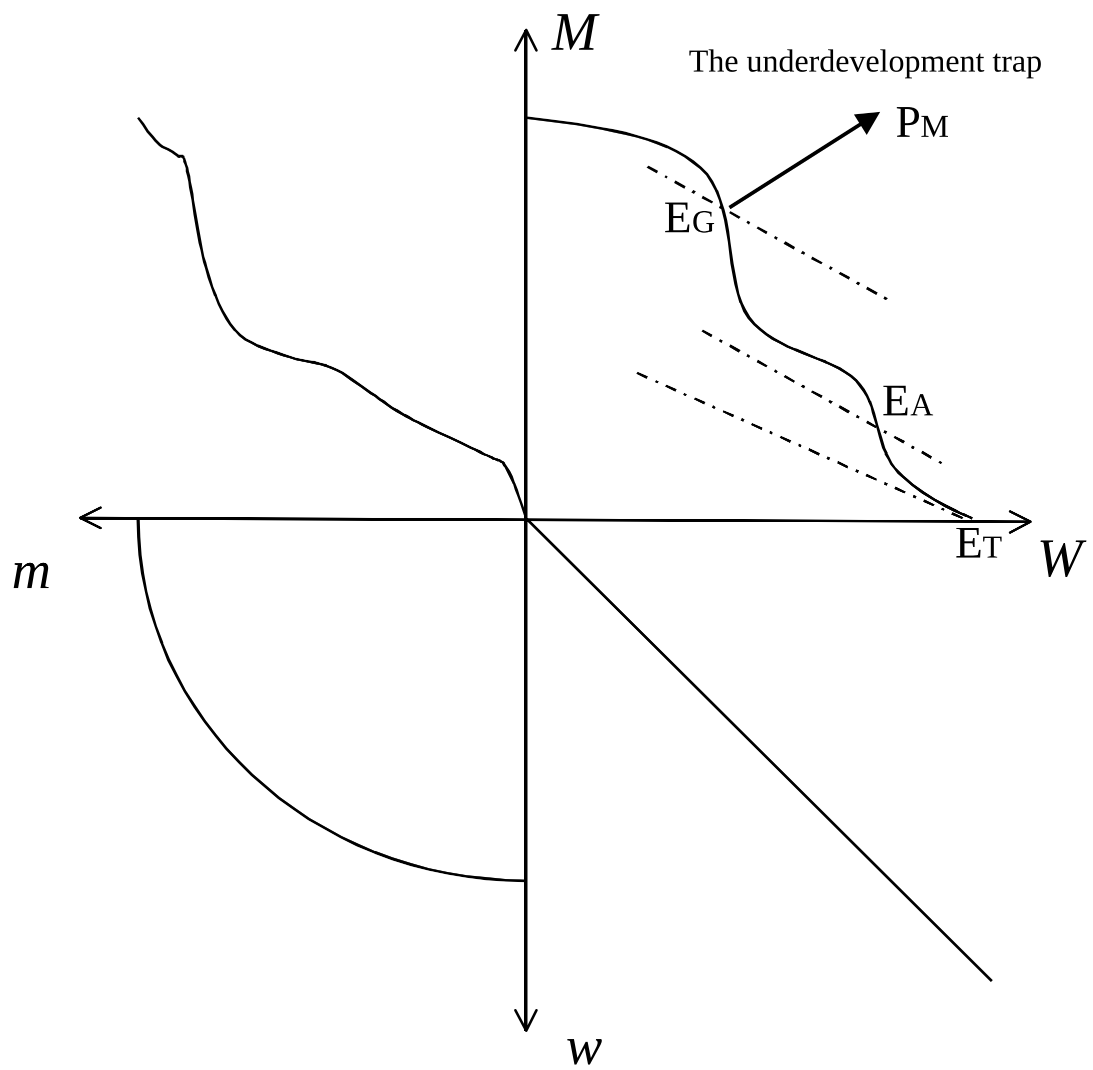


Figure 4

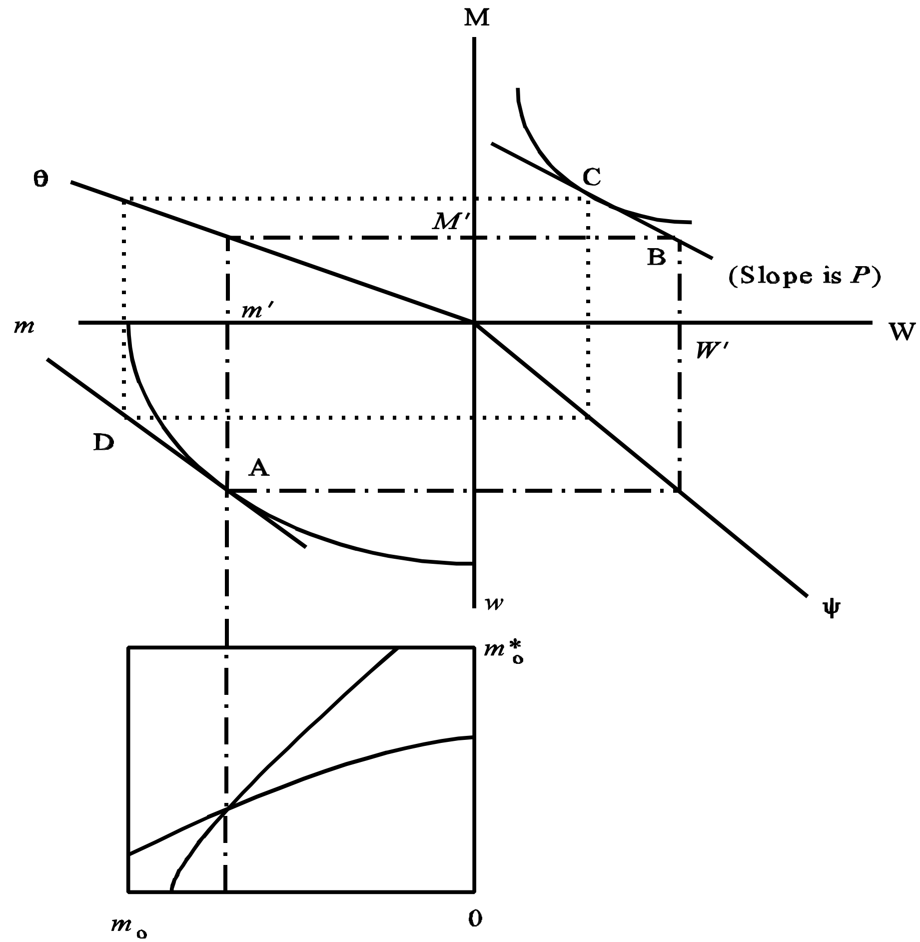


Figure 5

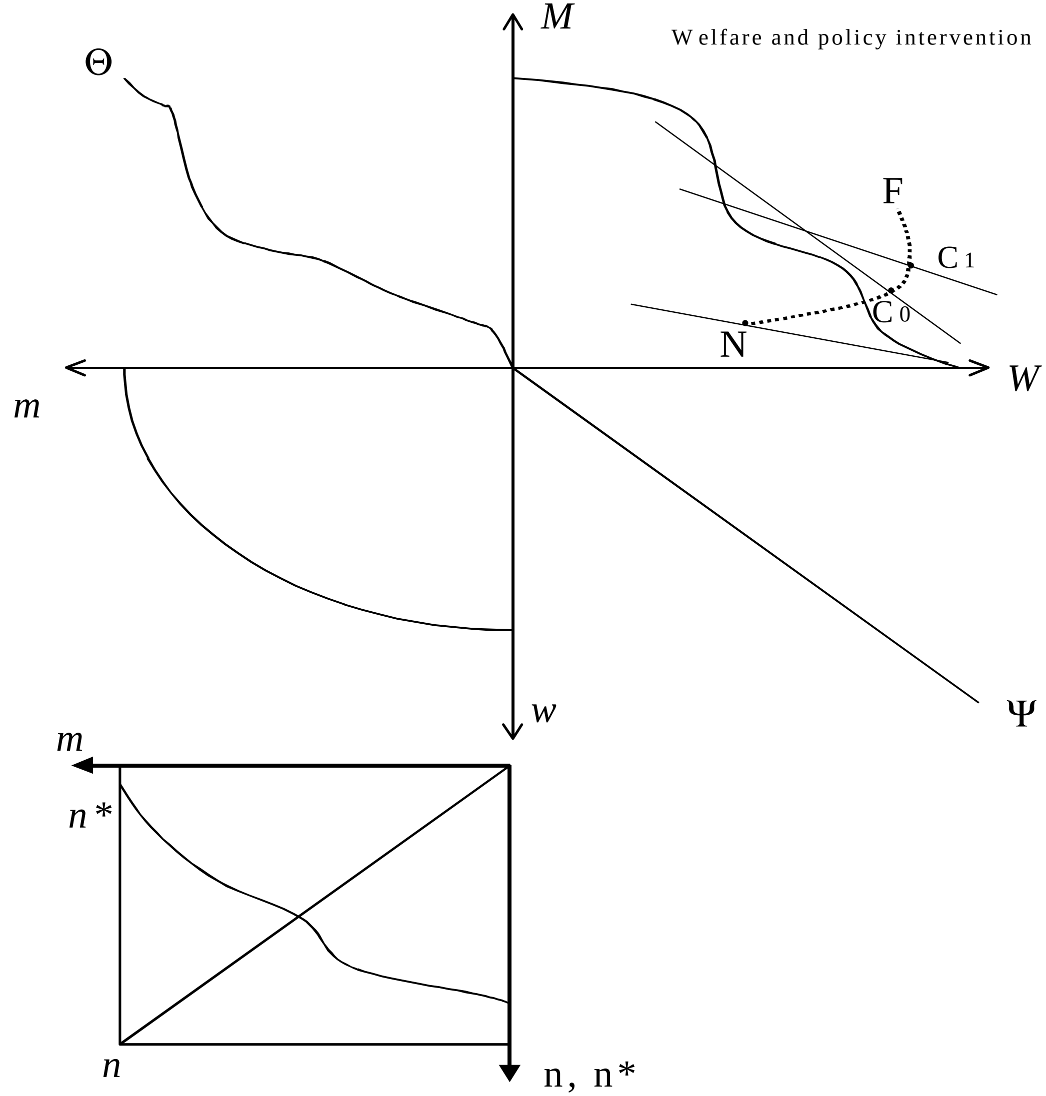
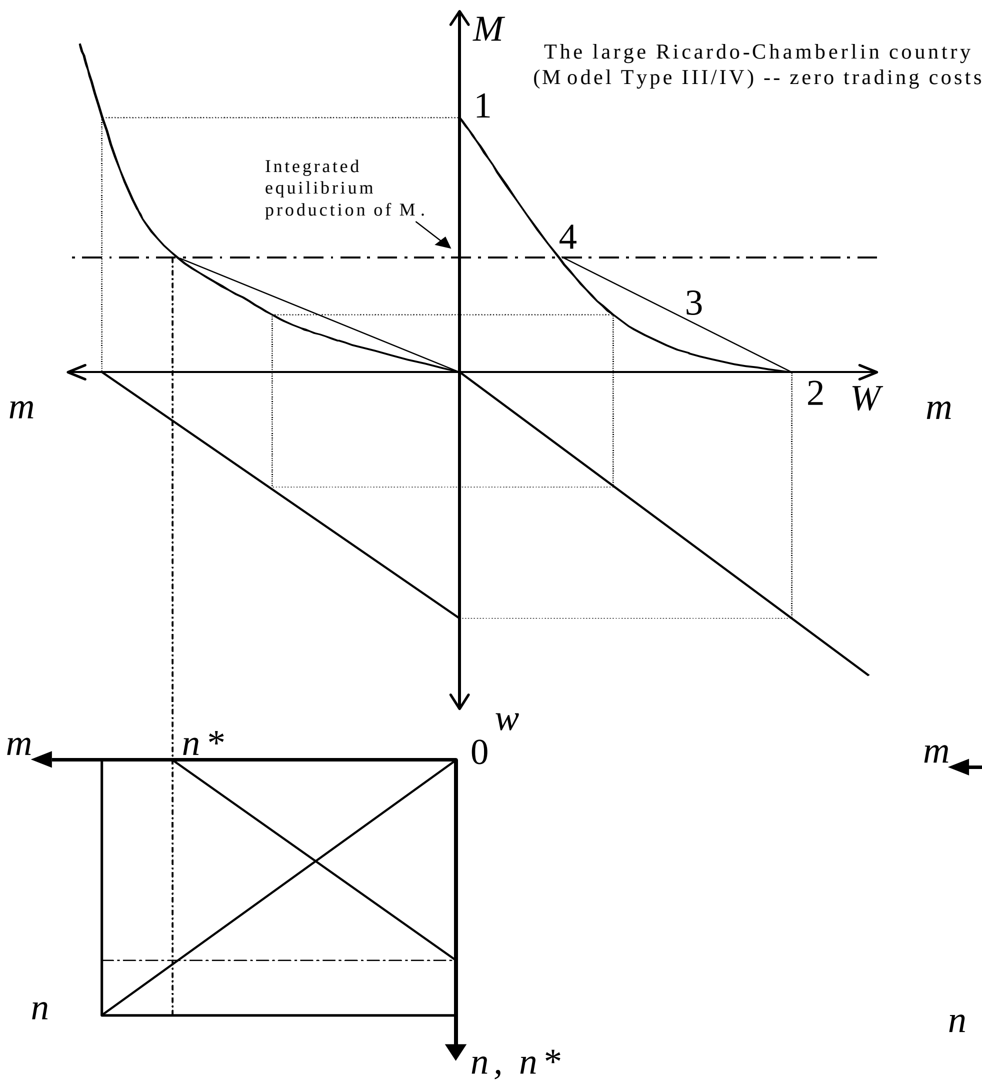
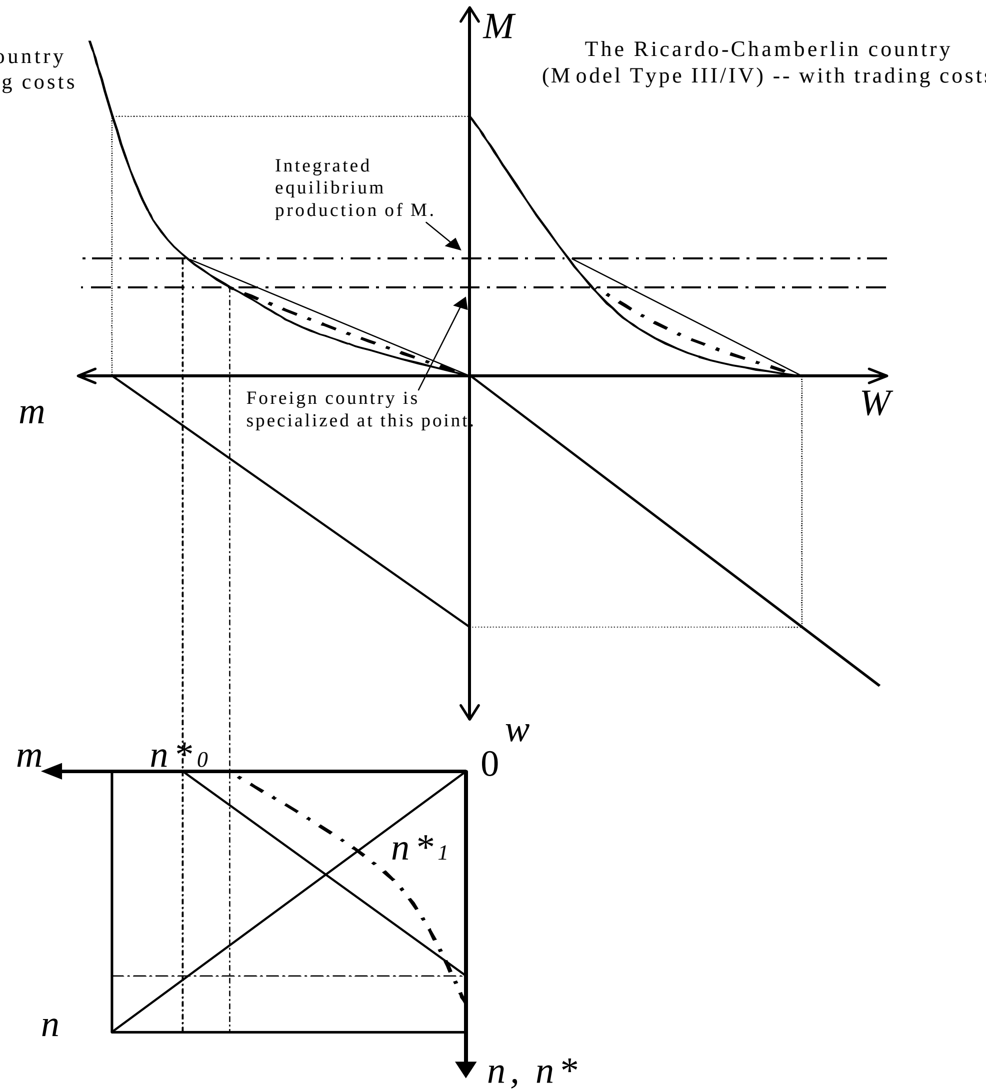


Figure 6

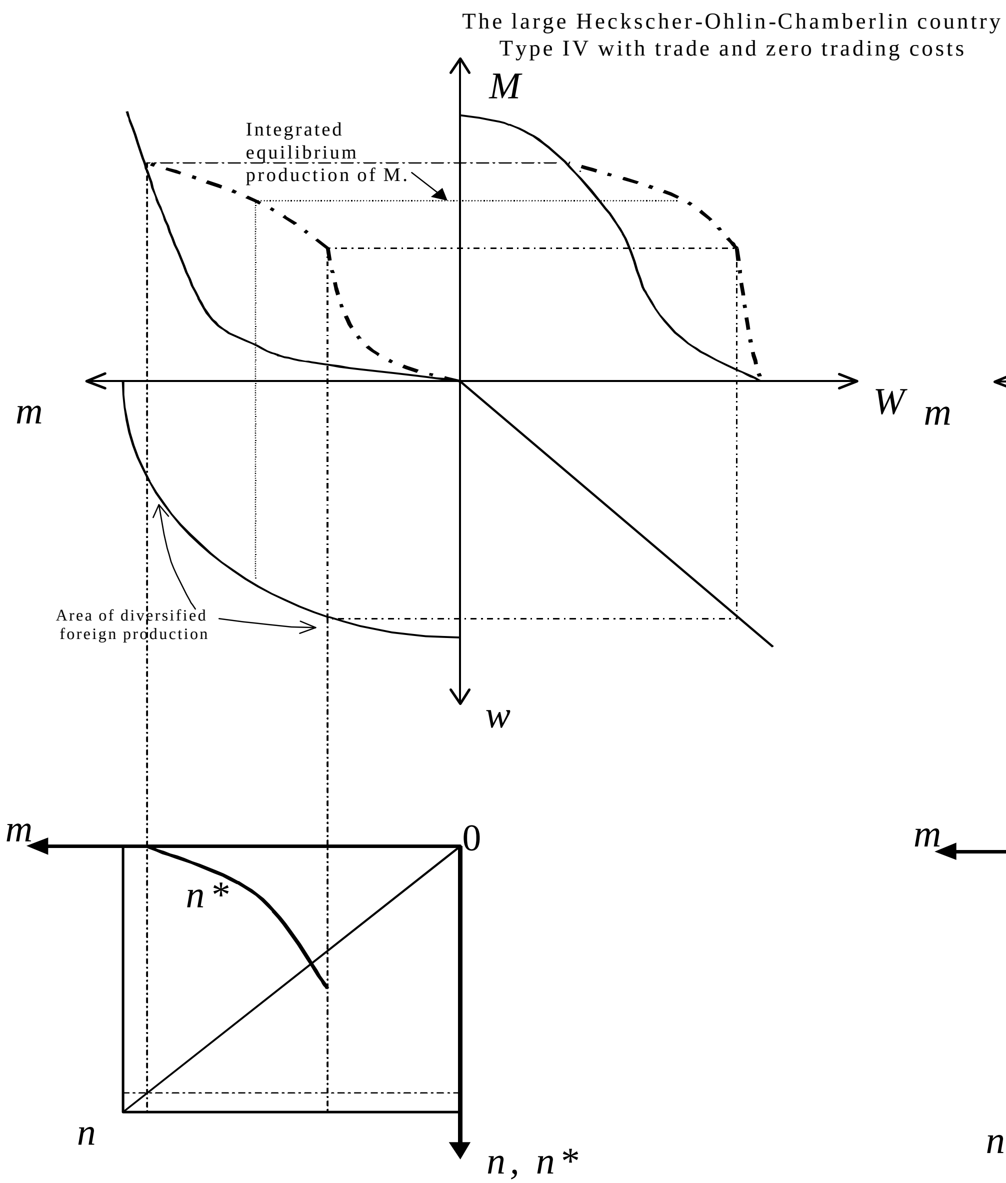


Panela a. The Ricardian model with international production externalities.

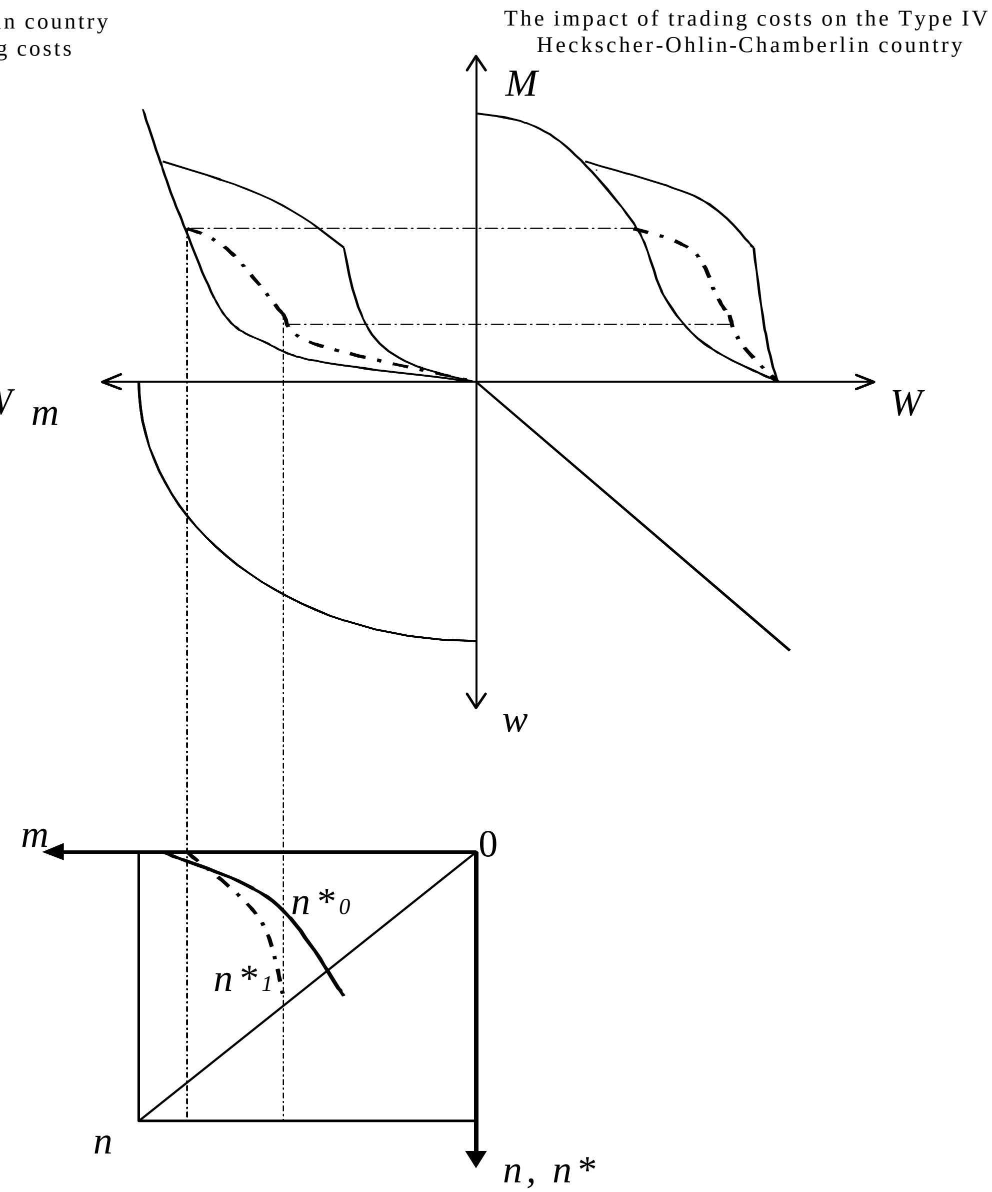


Panel b. Trading Costs in the Ricardian Model with international production externalities

Figure 7

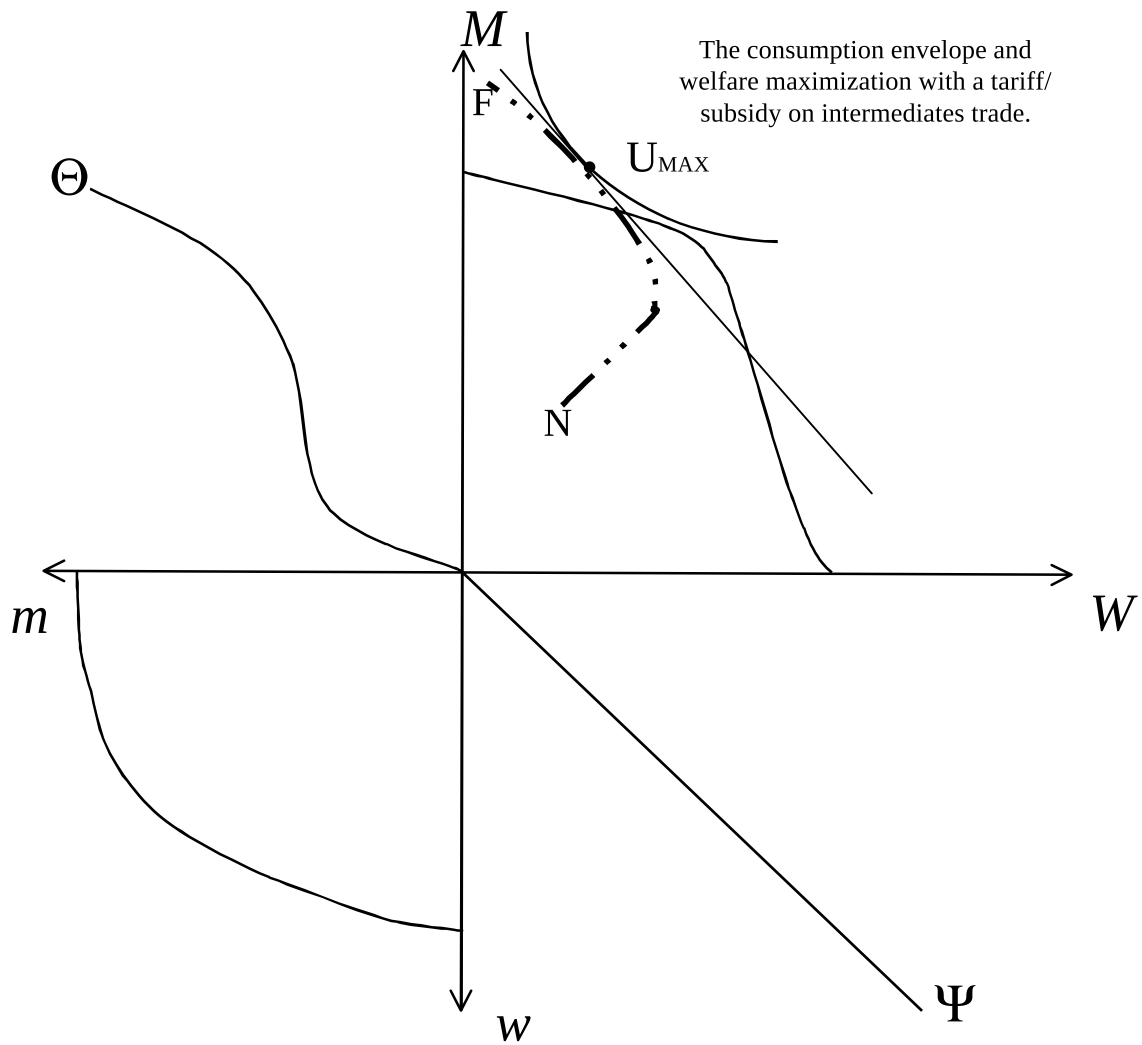


Panel a.



Panel b.

Figure 8


M
The consumption envelope and
welfare maximization with a tariff/
subsidy on intermediates trade.
F
UMAX
Θ
N
m
W
w
Ψ