MONOPOLISTIC COMPETITION, TRADE,

AND

ENDOGENOUS SPATIAL FLUCTUATIONS*

GIANMARCO I.P. OTTAVIANO**

Dipartimento di Scienze Economiche, Università di Bologna
Bologna, Italy 40125

Centre for Economic Policy Research
London, U.K. W1X 1LB

January 1996

ABSTRACT: This paper investigates the possibility of endogenous fluctuations in the international distribution of economic activities in the presence of increasing returns, monopolistic competition, trade and convex adjustment costs. Differently from the existing literature, it does not allow for any local productive externalities. Using a 2-country dynamic general equilibrium model, it derives necessary and sufficient conditions for the existence of self-reinforcing relocation processes. It shows that the occurrence of multiple equilibria and endogenous fluctuations is associated with a high degree of increasing returns to scale as well as low trade and adjustment costs. Under such circumstances relocation processes are driven by self-fulfilling expectations.

KEYWORDS: monopolistic competition, increasing returns, international trade, spatial fluctuations

Journal of Economic Literature Classification Number: E32, F12, F22, R12

---

* I am indebted to Richard Baldwin, Guido Candela, Larry Karp, Philippe Martin, Diego Puga, Anthony Venables and the seminar participants at IGIER Milan, at GIIS Geneva and at the AUEB/CEPR European Research Workshop in International Trade for helpful comments on an earlier draft. They are not responsible for any mistake.

** MAILING ADDRESS: Università di Bologna, Dipartimento di Scienze Economiche, Strada Maggiore 45, 40125 Bologna. Tel. +39-51-6402611. Fax. +39-51-6402664. E-mail. ottavian@boph01.cineca.it

Introduction

The importance of understanding the link between imperfect competition, multiple equilibria and endogenous fluctuations has been emphasized by several authors in recent years (Galì [6]). In particular, an increasing number of contributions in a wide range of different fields has privileged the assumption of monopolistic competition (Matsuyama [23]).

This paper investigates the possibility of multiple equilibria and endogenous fluctuations in the international distribution of economic activities in the presence of increasing returns, monopolistic competition and trade. It differs from existing works for two main reasons. On one hand, differently from Galì [7], it does not rely on the assumption of "local externalities" to obtain multiple equilibria. On the other hand, differently from Galì [5], it introduces trade between different locations and analizes the influence of trade costs on the international spatial dynamics.

The contribution is closely related to recent work on the international location effects of trade barriers in the wake of Krugman [11,12] and Venables [30]. These authors have shown that the explicit simultaneous consideration of increasing returns and trade costs generates complementarities between the location decisions of different agents and gives rise to cumulative location processes and multiple spatial equilibria. Previously, trade models with increasing returns had been widely used to assess the relative importance of increasing returns versus comparative advantage in causing international specialization (Helpman and Krugman [9]). The recent works have argued that the same models can be used to show that, when production exhibits increasing returns to scale, trade barriers affect the properties of the international distribution of economic activity in terms of both the number of stationary states and their stability. The framework is S-D-S monopolistic competition (Spence [29], Dixit and Stiglitz [2]). Obstacles to trade are shown to affect (non-linearly) the balance between "centripetal forces" that favour agglomeration of economic activities in few alternative sites, and "centrifugal forces" that favour their dispersion in many different places (Henderson [10]). In general, the location effects of different levels of trade barriers are sensitive to the way the economy is designed in terms of returns to scale in different sectors, linkages among sectors and factor mobility (Krugman and Venables [16,17]).

While raising issues that are essentially dynamic, this field of research has mostly confined itself to ad hoc Marshallian dynamics. While stressing the possibility of multiple spatial equilibria, it has dismissed the complex question of equilibrium selection. In the context of industrialization, Marshallian dynamics have been critized both by Krugman [13] and, in detail, by Matsuyama [22] under the name of "Marshallian tatonnement approach". This approach assumes that agents care only about current returns and it could be supported by a dynamic structure with an infinite rate of time preference. However, it also assumes smooth transitional dynamics, which are not generally consistent with a myopic time horizon. The major shortcoming of tatonnement arguments is the identification of initial conditions as the only factor that selects the stationary state which is eventually reached. The importance of path-dependency ("history") is widely accepted (Arthur [1]). Nonetheless the possibility of self-fulfilling expectations remains not only a major logical consequence of complementarities, but also a subtle policy issue when considering intervention on policy parameters such as, in this case, trade barriers. With complementarities, multiple equilibria as well as multiple transition paths are possible and variations in the parameters values affect the global dynamic properties of the system.

Some recent contributions investigate the global dynamic consequences of monopolistic competition, increasing returns and trade. Krugman [11] introduces the dynamic issues listed above. However, he relies on local production externalities to obtain multiple equilibria and, therefore, he cannot say anything about the influence of trade barriers on the global dynamics of the system. The same applies to Galì [7] and, as far as there is no trade, to Galì [5]. Krugman [14] takes one step further but its dynamics are still largely ad hoc, since they are not microfounded on agents' forward looking behaviour. The same is true for Krugman and Venables [18], whose interest is focussed on the enrichment of the geometry of the economic space, as well as for Puga [26]. So, existing models still provide an incomplete framework for understanding the dynamic location effects of trade barriers with monopolistic competition.

This paper develops a two-country dynamic model with forward looking migration decisions. The model represents a simple dynamic extension of the location model proposed by Martin and Ottaviano [20] building on Flam and Helpman [3] and Krugman [12]. The structure of the model is as simple as possible. In a two-country world, there are two sectors: the 'traditional' sector is perfectly competitive and produces a freely-traded homogeneous good with constant returns to scale using unskilled labour as the only input; the 'industrial' sector is monopolistically competitive and produces a horizontally differentiated good with increasing returns to scale using both unskilled and sector-specific skilled labour. International trade in the differentiated good incurs frictional costs. While unskilled workers are internationally immobile, skilled workers can migrate at a cost. Their forward looking migration decisions interact with increasing returns as well as frictional trade costs and can generate the complementarities (agglomeration economies) required to have multiple spatial equilibria. Such a multiplicity stems from the interaction of centripetal and centrifugal forces. The former arise from the imperfect substitutability between different varieties of the industrial good and increasing returns in their production; the latter comes from the existence of a spatially dispersed demand due to immobile unskilled workers. Both types of forces are active only if trade costs in the differentiated good are not null, otherwise industrial location is indifferent.

Agglomeration affects the incentive to relocate through two channels: (i) the whole set of varieties is cheaper in a certain country the bigger the number of domestically produced varieties i.e. the bigger the number of resident skilled workers ('price index effect'); (ii) skilled workers' income in a certain country can be higher or lower the bigger their number, depending on parameters values ('factor price effect'). If skilled workers' income in a certain country grows with their number, then agglomerations in either country are the only stable equilibrium states, because both effects work in the same direction. On the contrary, if income falls with their number, then the price index effect and the factor price effect work in opposite directions. If the former effect dominates, then agglomeration emerges as a stable equilibrium as before. On the contrary, if the latter is stronger, dispersion is the only stable equilibrium. It can also be that, while the former dominates if the initial distribution of skilled workers is very uneven, the latter dominates for a more even distribution. In that case, both agglomeration and dispersion are stable spatial equilibria. The price index effect is stronger the higher the expenditure share of the industrial good, the lower the trade cost and the smaller the degree of substitutability between varieties, which can be interpreted as an inverse index of market power or of the degree of equilibrium returns to scale (Krugman [12]). So, only a relatively small expenditure share of the industrial good, relatively high trade costs and weak returns to scale will grant the uniqueness of the equilibrium point.

The solution of the model leads to a non-linear system of differential equations that is studied by perturbation methods (Nayfeh [24], Naifeh and Mook [25], Guckenheimer and Holmes [8]). The paper shows how detailed pictures of the phase plane can be drawn for all possible values of the parameters. Necessary and sufficient conditions are derived for each alternative picture in terms of the parameters values. Relatively large expenditure share of the industrial good, relatively low trade and relocation costs as well as relatively strong returns to scale are shown to be associated with endogenous spatial fluctuations. Originally devised in pure mathematics and physics, perturbation methods have been introduced to economists by Matsuyama [22]. Essentially, such methods start with observing that, under some restrictive assumptions about the parameters, the dynamic system to be solved happens to be a special system whose solutions are known completely. Then, they study small perturbations of it by relaxing the restrictive assumptions.

The paper is in five parts. The first part develops the model obtaining a system of two differential equations. It also derives necessary and sufficient condition for multiple equilibria to exist. The second shows that, when the rate of time preference is null, the system is Hamiltonian and its solutions are known. The third section perturbes the Hamiltonian system letting the rate of time preference to be different from zero and illustrates the possibility of endogenous spatial fluctuations. The fourth section investigates under which conditions endogenous spatial fluctuations arise. The fifth part concludes.

1. A dynamic spatial model

The economy consists of two countries, A and B, with fixed endowments of unskilled labour L and skilled labour H. Every worker is endowed with one unit of labour. While unskilled labour is internationally immobile and evenly distributed between countries, each hosting L/2 unskilled workers, skilled labour can move and its international distribution will be endogenously determined. Call $h(t)\in[0,1]$ the share of skilled workers in country A at time t. Then, setting H=1 by choice of units, $h(t)$ becomes the number of skilled workers in A and $1-h(t)$ the number of skilled in B.

Preferences are represented as a log transformation of a Cobb-Douglas utility function over a differentiated ("industrial") good D(t) and a homogeneous ("traditional") good Y(t) with shares α and $1-\alpha$ respectively, being $\alpha\in(0,1)$:

$$U_i(t) = \ln\left[\left(\frac{D_i(t)}{\alpha}\right)^{\alpha}\left(\frac{Y_i(t)}{1-\alpha}\right)^{1-\alpha}\right] \quad , \; i = A,B \tag{1}$$

The differentiated good D is a C.E.S. aggregate of $n_i(t)$ symmetric domestic varieties and $n_j(t)$ symmetric foreign varieties:

$$D_i(t) = \left[n_i(t)d_{ii}(t)^{\frac{\sigma-1}{\sigma}} + n_j(t)d_{ji}(t)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} \tag{2}$$

where $\sigma\in(1,+\infty)$ is the elasticity of substitution between any two varieties and the own-price elasticity of demand for each variety (Spence [29], Dixit and Stiglitz [2]).

The homogeneous good Y is produced in a perfectly competitive sector that uses unskilled labour as the only input. Technology exhibits constant returns to scale with unit input coefficient equal to 1. On the contrary, the differentiated good is produced in a monopolistically competitive sector using both skilled and unskilled labour. Entry and exit are free so that profits are zero in equilibrium. The production of each variety exhibits increasing returns to scale and technology is represented by a linear cost function with a fixed component g=1 undertaken in terms of skilled labour and a marginal component k undertaken in terms of unskilled labour. Because of increasing returns, in equilibrium there is a 1:1 relationship between firms and varieties. Moreover, since g=1 there is also a 1:1 relationship between firms and skilled workers. Thus any moment each country will produce as many varieties (have as many firms) as the number of skilled workers it hosts: $n_A(t)=h(t)$ and $n_B(t)=(1-h(t))$.

Trade in the homogeneous good is free while it is costly in the differentiated good. Trade costs are frictional and are modeled as iceberg costs in the sense of Samuelson [28]: if a unit of the differentiated good has to reach a certain country from the other, $\tau\in[1,+\infty)$ units must be shipped.

Since unskilled labour is perfectly mobile between sectors in the same country, free trade in the homogeneous good pins down its salary to unity, even if it is internationally immobile. On the contrary, skilled labour can migrate but its mobility incurs a cost which is supposed to be a quadratic function of the rate of migration $\dot{h}(t)\equiv dh(t)/dt$ with coefficient $1/(2\gamma)$ and $\gamma\in(0,+\infty)$.

The aim is to investigate the long run distribution of firms as the end result of forward-looking migration decisions by atomistic skilled workers who have perfect foresight about the future paths of indirect utilities in the two locations, $\{W_A(t), W_B(t)\}_0^\infty$, that they take as given. The problem of migration can be solved 'as if' there were a single representative skilled worker, holding the whole endowment of skilled labour, that decided to allocate his labour intertemporally between the two countries while taking the future path of indirect utilities as given. Thus the solution to the problem of the equilibrium evolution of the system can be found by maximizing with respect to $\{h(t), \dot{h}(t)\}_0^\infty$, and for initial $h(0)$, the present value of the welfare flow:

$$PV(0) = \int_0^\infty \left\{ W_A(t)h(t) + W_B(t)[1-h(t)] - \frac{1}{2\gamma}(\dot{h}(t))^2 \right\} e^{-\delta t}\, dt \tag{3}$$

where $\delta\in(0,+\infty)$ is the rate of time preference. As it will become clear later on, the case of a zero rate of time preference, which would prevent $PV(0)$ from converging, will prove notheless useful as a benchmark for studying the dynamic properties of the economy by perturbation techniques.

The migration decision is based on a shadow price defined as follows. Let $v_A(t)$ and $v_B(t)$ be the expected discounted sum of future utility minus moving costs of an agent currently in A and in B respectively. Let T be the first time the economy reaches an equilibrium point. T can be null, finite or infinite depending on the initial conditions and the type of equilibrium point the system is heading towards. It will be null if the the economy starts at an equilibrium point; it will be infinite if the economy heads towards a saddle point; it will be finite if the economy is going to rest at a boundary. Then, in general:

$$v_A(t) = \int_t^{T} W_A(s)\, e^{-\delta(s-t)} ds + v_A(T^-)\, e^{-\delta(T-t)} \tag{4a}$$

$$v_B(t) = \int_t^{T} W_B(s)\, e^{-\delta(s-t)} ds + v_B(T^-)\, e^{-\delta(T-t)} \tag{4b}$$

where $v_i(T^-)$ denotes the limit from the left, which must be determined by a terminal condition of the model.

Moreover, since agents in each location have the option to move to the other location by paying the marginal relocation cost, $|\dot{h}(t)|/\gamma$, it must be:

$$v_A(t) \geq v_B(t) - \frac{|\dot{h}(t)|}{\gamma} \quad \textit{with equality if} \quad \dot{h}(t) < 0 \tag{5a}$$

$$v_B(t) \geq v_A(t) - \frac{|\dot{h}(t)|}{\gamma} \quad \textit{with equality if} \quad \dot{h}(t) > 0. \tag{5b}$$

The shadow price is then defined as:

$$v(t) = v_A(t) - v_B(t). \tag{6}$$

This shadow price represents the difference in "private" value between having a unit of skilled labour in location A rather than in location B. It represents the incentive to migrate.

Eq.s (5) and (4) can be used to derive the economy laws of motion. They imply respectively:

$$\dot{h}(t) = \gamma v(t) \tag{7a}$$

$$\dot{v}(t) = \delta v(t) - [W_A(t) - W_B(t)] \tag{7b}$$

For any internal point $(h(t),v(t))$, such that $h(t)\in(0,1)$, there is no incentive to migrate if and only if $v(t)=0$. As Fukao and Benabou [4, Proposition 2] prove, this must be true also for points at the boundary, such that $h(t)=0$ or $h(t)=1$. The first part of the proof shows that, if the system hits a boundary, it will happen in finite time T. The second part of the proof shows that, since T is finite, $v(T)$ has to be zero. While the former is purely mathematical, the latter has an intuitive economic appeal. The intuition behind the proof is by contrast. Suppose the economy is in the neibourhood of the terminal point moving along an equilibrium trajectory. Suppose also that it will reach the terminal point in finite time T. It will then stay there forever. In such a situation, if $v(T)$ is not null, a skilled worker who has not migrated yet, will wait until T and gain: by waiting until the very last moment, he will face zero costs of migration and gain the absolute value of $v(T)/\delta$. However, if such a remunerative deviation from the original trajectory is possible, that trajectory cannot be an equilibrium trajectory which contradicts the hypothesis. So, a zero value of $v(T)$ has to be imposed at the boundary to avoid arbitrage behaviour by migrants. The boundary conditions (0,0) and (1,0) on $(h(T),v(T))$ complete the specification of system (7).

In order to close the model, the instantaneous utility differential between A and B is needed. The instantaneous flow indirect utility of a typical skilled worker in country i earning a wage r_i is:

$$W_i(t) = \ln\left[\frac{r_i(t)}{\left(q_i(t)\right)^{\alpha}}\right] \tag{8}$$

where q_i is the C.E.S. price index associated with the quantity index (2):

$$q_i(t) = \left[n_i(t)p_i(t)^{1-\sigma} + n_j(t)p_j(t)^{1-\sigma}\tau^{1-\sigma}\right]^{\frac{1}{1-\sigma}} \tag{9}$$

It follows that q_i^{α} is the exact price index and W_i is the real wage of a typical skilled worker in country i. p_i and p_j are the local profit-maximizing prices of domestic producers so that $p_j\tau$ is the consumer price in country i of a typical variety made in country j. The choice of p_i that maximizes profits obeys the standard rule in monopolistic competition: $p_i = k\sigma/(\sigma-1)$. With the appropriate choice of k, the price of all differentiated goods is set to 1. In equilibrium, $\sigma/(\sigma-1)$ is the ratio of average cost to marginal cost. Therefore, as in Krugman [12], σ can be interpreted as an inverse index of equilibrium returns to scale. Due to free entry, skilled workers' wages equal the difference between revenues and unskilled labour costs: $r_i = x_i - kx_i = x_i/\sigma$ where x_i is production of each differentiated good. Thus, instantaneous indirect utility can be written as:

$$\begin{aligned} W_A(t) &= \ln\left\{\frac{x_A}{\sigma}\left[h(t) + (1-h(t))\tau^{1-\sigma}\right]^{\frac{\alpha}{\sigma-1}}\right\} \\ W_B(t) &= \ln\left\{\frac{x_B}{\sigma}\left[h(t)\tau^{1-\sigma} + (1-h(t))\right]^{\frac{\alpha}{\sigma-1}}\right\} \end{aligned} \tag{10}$$

Finally, total production of firms in A and B can be determined by solving the system requiring that demands (inclusive of trade costs) must equal supplies:

$$x_A(t) = \frac{\alpha\left[\frac{L}{2} + \frac{x_A(t)}{\sigma}h(t)\right]}{h(t) + \rho(1-h(t))} + \frac{\rho\alpha\left[\frac{L}{2} + \frac{x_B(t)}{\sigma}(1-h(t))\right]}{\rho h(t) + (1-h(t))} \tag{11a}$$

$$x_B(t) = \frac{\rho\alpha\left[\frac{L}{2} + \frac{x_A(t)}{\sigma}h(t)\right]}{h(t) + \rho(1-h(t))} + \frac{\alpha\left[\frac{L}{2} + \frac{x_B(t)}{\sigma}(1-h(t))\right]}{\rho h(t) + (1-h(t))} \tag{11b}$$

where $\rho \equiv \tau^{1-\sigma}$ is the ratio of total demand by domestic residents for each foreign variety to demand for each domestic variety.

Solving system (11) yields:

$$\frac{x_A(t)}{x_B(t)} = \frac{\rho h(t) + \psi(1-h(t))}{\psi h(t) + \rho(1-h(t))} \tag{12}$$

where the constant ψ is defined as follows:

$$\psi \equiv \frac{1}{2}\left[1 + \rho^2 - (1-\rho^2)\frac{\alpha}{\sigma}\right]$$

Eq. (12) shows how the international distribution of skilled workers affects the international differences between their wage rates (*factor price effect*). This effect cannot be signed unambiguously because it depends on the interaction between two types of components, whose relative strength differs for different values of the parameters α, σ, and τ . On one hand there are the 'centripetal' components $\rho h(t)$ and $\rho(1\text{-}h(t))$, which tend to yield higher wages in the skilled labour abundant country. On the other hand, there are the 'centrifugal' components $\psi(1\text{-}h(t))$ and $\psi h(t)$, which work in opposite direction making wages higher in the country with fewer skilled workers. Centripetal forces arise from the incentive to concentrate production in a single location when trade costs are low. Centrifugal forces stem from the localized demand by immobile unskilled workers. Localized demand linkages are stronger the lower the share of expenditure α devoted to the differentiated good and the higher the elasticity of substitution σ. By simple algebra, it can be seen that x_A/x_B is increasing (decreasing) in h if $\rho>(<)[(\sigma-\alpha)/(\sigma+\alpha)]$ i.e.

$$\tau < (>) \left[\frac{\sigma+\alpha}{\sigma-\alpha}\right]^{\frac{1}{\sigma-1}} \qquad \textbf{(13)}$$

Thus, immigration has a positive effect on local skilled workers wages if, *ceteris paribus*, trade costs are low, returns to scale are strong and the industrial sector is important in terms of its share of aggregate expenditures. A positive wage effect is sufficient to yield agglomeration.

The instantaneous indirect utility differential between skilled workers in A and B can be evaluated as:

$$\begin{aligned} W_A(t) - W_B(t) &= \ln\big(\rho h(t) + \psi(1-h(t))\big) - \ln\big(\psi h(t) + \rho(1-h(t))\big) + \\ &+ \frac{\alpha}{\sigma-1}\left[\ln\big(h(t) + \rho(1-h(t))\big) - \ln\big(\rho h(t) + (1-h(t))\big)\right] \end{aligned} \qquad \textbf{(14)}$$

Along with (7a,b), Eq. (14) completes the characterization of the optimal evolution of the economy. To simplify the notation, call $f(h)\equiv W_A(t)\text{-}W_B(t)$ where the dependence of h upon time is left implicit.

Eq. (14) shows how the international distribution of skilled workers affects the international distribution of their welfare (real wage). It consists of two parts. The first two logarithmic terms measure the welfare value of the wage rate differential. The other terms measure the welfare value of the different levels of prices and takes account of the relation between the distribution of skilled workers and the distribution of the exact price index (*price index effect*). Differently from the wage effect, the price index effect can be signed unambiguously. The difference between the two terms inside the squared brackets in (14) is a monotone increasing function of h: the price index is always lower where there are more skilled workers and therefore more varieties are produced domestically.

The qualitative features of the instantaneous utility differential are readily assessed, even if it is hard to solve for its zeroes analytically. In fact, one can readily spot one of its zeroes and assess the number of the others. Due to the symmetry of the model, $f(h)=0$ for $h=0.5$. As it can be easily checked, this is always true for any value of the parameters. Therefore, $f(h)$ always has at least one zero where the international distribution of industry is even. Moreover, since the sign of $f'(h)$ depends on the sign of its quadratic numerator, $f(h)$ will change slope at most twice. That is, either $h=0.5$ is the only zero of $f(h)$ or there are (no more and no less than) two other zeroes, say h_L and h_H. Due to symmetry, if they exist, h_L and h_H have to be symmetric with respect to $h=0.5$. Figure 1 shows the four alternative shapes of $f(h)$. In cases (a) and (c), $f(h)$ is decreasing for $h=0.5$, while in (b) and (d), it is increasing. In the first case, $f'(0.5)$ is negative, while it is positive in the second case. By simple algebra, one gets to:

$$\varphi \equiv f'(0.5) = \rho - \frac{\sigma-\alpha}{\sigma+\alpha}\frac{\sigma-\alpha-1}{\sigma+\alpha-1} \qquad \textbf{(15)}$$

Therefore, $f(h)$ is decreasing (increasing) in $h=0.5$ if $\varphi<(>)0$ and this happens for large (small) τ, large (small) σ and small (large) α. As already argued, that is the case when centrifugal (centripetal) forces dominate. The above conditions on φ help to discriminate between situations like (a) as well as (c) on one side and (b) as well as (d) on the other. However, in order to distinguish between (a) and (c) or between (b) and (d), more information is needed. Additional information on the shape of $f(h)$ comes from the evaluation of its behaviour at the boundaries $h=0$ and $h=1$. In cases (a) and (d), $f(0)>0$ and $f(1)<0$. The opposite is true in (b) and (c). Then, the sign of $f(0)$ and $f(1)$ can be used to distinguish between (a) and (d) on one side and (b) and (c) on the other. It is readily assessed that $f(0)= -f(1)$. In particular:

$$\Phi \equiv f(1) = -f(0) = \rho - \frac{\rho^{\frac{\alpha}{\sigma-1}}}{2}\left[1+\rho^2-(1-\rho^2)\frac{\alpha}{\sigma}\right] \qquad \textbf{(16)}$$

Therefore $f(h)$ will behave as in (a) and (d) [(b) and (c)] if $\Phi<(>)0$.

Since $f(0.5)$ is always null and $f'(h)$ changes sign at most twice, by crossing the conditions on Φ and φ, one gets necessary and sufficient conditions for each picture to describe the actual behaviour of $f(h)$. These conditions are summarized in Table I.

Case (d) has to be ruled out. The reason why is the following. As it can be easily checked, necessary condition for Φ to be negative is $\alpha>\sigma-1$. But this requirement is necessary and sufficient to have φ to be negative too. Therefore $\Phi<0$ implies $\varphi<0$, while the reverse is not true: if $\alpha>\sigma-1$ and therefore $f(1)$ is positive, $h=0.5$ is the only zero of $f(h)$.

So, from (15) and (16) one can conclude that, *ceteris paribus*, case (a) will prevail for high trade costs, weak returns to scale and a small industrial expenditure share. Case (b) will prevail for relatively low trade costs, strong returns to scale and a large industrial expenditure share. Finally, case (c) will be characterized by intermediate values of all the parameters.

In most of recent contributions on location and trade with S-D-S monopolistic competition, the information captured by pictures like Figure 1 is the only one used to draw dynamic insights from an otherwise static setting. This is the so-called "Marshallian tatonnement approach" (Matsuyama [22]), which is equivalent to assuming that agents only care about current returns i.e. the rate of time preference δ is infinite. In the present model, under that assumption, only the instantaneous indirect utility differential $f(h)$ matters for relocation decisions. In particular, for any initial h, skilled workers will migrate from country B to country A if $f(h)$ is positive, while they will move in the opposite direction if $f(h)$ is negative. They will settle down only if $f(h)$ is null for any $h \in (0,1)$, and at endpoints h=0 or h=1 if, respectively, $f(0)<0$ or $f(1)>0$. Therefore, in case (a) there is a unique equilibrium at h=0.5. In case (b) there are three equilibria at h equal to 0, 0.5 and 1. In case (c) there are five equilibria at 0, h_L, 0.5, h_H, 1. Which equilibrium will eventually be selected depends on the initial value of h. Unless the economy starts at an equilibrium, in which case there is no migration, the flows of skilled workers will de directed towards the country with the higher real wage. So, in case (a) migration will lead to the evenly dispersed equilibrium h=0.5, which is therefore stable. On the contrary, in case (b), it will lead to agglomeration in either country. The economy will eventually reach either h=0 or h=1, which are both stable equilibria, while h=0.5 is unstable. Finally, in case (c) both even dispersion and agglomeration are stable outcomes. The other two unevenly dispersed equilibria are unstable. In any case, as Matsuyama [22] points out, the transition is usually assumed to happen *gradually*, which, in itself, is in conflict with the other assumption that agents only care about current returns. Marshallian dynamics can be summarized by the following first order differential equation:

$$\dot{h} = \eta f(h) \qquad \eta > 0 \tag{17}$$

This *ad hoc* formulation deprives the dynamic model of much of its richness, since it reduces system (7) to a single first-order linear equation.

As it stands, the tatonnement argument is clearly nongeneric, since it rests on the nongeneric hypothesis of an infinite rate of time preference δ, and hardly consistent, since it usually claims for a gradual transition. Using the dynamic model developed so far, in which migration costs do impose smooth adjustment, next sections perform a more general stability analysis allowing for a finite rate of time preference. Conditions are assessed under which the tatonnement approach can be used as a heuristic device for qualitative analysis.

2. An integrable system: the Hamiltonian

By substitution between (7a) and (7b), the dynamics of the economy are equivalently represented by the following equation:

$$\frac{1}{\gamma}\ddot{h} + f(h) = \frac{\delta}{\gamma}\dot{h} \qquad \textbf{(18)}$$

This is a second order differential equation that contrasts with the *ad hoc* first order equation of the tatonnement approach.

Upon integrating, one obtains:

$$\frac{1}{2\gamma}\dot{h}^2 = u + \frac{\delta}{2\gamma}h^2 - F(h) \qquad \textbf{(19)}$$

where $F(h)=\int f dh$ and u is a constant of integration. In mechanics, the different components of equations like (18) and (19) have evocative names. The function $f(h)$ is called the (non-linear) restoring force, $F(h)$ is the potential (stored) energy, $\frac{1}{2\gamma}\dot{h}^2$ is the kinetic energy, $\frac{\delta}{\gamma}\dot{h}$ is the (linear) damping mechanism, and u is the total energy level per unit mass, which is determined by the initial conditions. For example, with $f(h)=\sin h$ the above equations would describe a simple pedulum consisting of a particle of unit mass fixed to one hand to a rod of length $1/\gamma$, the other hand of the rod being fixed to a pivot. In that case, h would measure (in radians) the angular displacement from the downwards vertical through the point where the other hand of the rod is fixed. Then the damping mechanism would represent the air resistance and the friction at the pivot, were they directly proportional to the angular velocity.

However, if one wanted to find a mechanical analogue to the spatial system under consideration, a simple pendulum would be just “too simple”. An alluring metaphore (not an exact one) is that of a pendulum which is subject to two opposing forces: one would lead the free hand of the pendulum to a rest in the downward position, the other would lead it to the upward position. The former would be ‘centripetal’ while the latter would be ‘centrifugal’ in the economic sense introduced in the previous section. The trade-off between the two forces would be measured by the potential energy $F(h)$.

From (19) it follows that motion (i.e. non-zero kinetic energy) is possible only for those initial conditions where the potential energy $F(h)$ is less than the total energy u once the damping has been taken into account. Then, by definition, in an *equilibrium point* of the system the velocity $\dot{h}=\gamma v$ and the accelleration $\ddot{h}$ are null so that the kinetic energy is zero. This happens for internal values of h such that $f(h)=0$ and for values $h=0$ and $h=1$ because of the boundary condition $v=\dot{h}/\gamma=0$. Therefore, equilibrium points do not depend on the values of δ and γ.

Solving the above equations is not an easy task. The reason why is that the restoring force $f(h)$ is nonlinear and therefore precludes an exact analytical solution. This is a general feature of recent models of location and trade with monopolistic competition in the spirit of Krugman [12] and Venables [30]. In order to obtain information about solutions, few methods are available: approximations, numerical solutions, or combinations of both. So far, in static or *ad hoc* dynamic models, numerical solutions have been widely preferred (Krugman [12], Krugman and Venables [16,17], Puga and Venables [27]). Only Krugman and Venables [18] have used a combination of numerical and analytical techniques. This paper takes the direction of approximations and uses parameter perturbation (asymptotic) methods (Nayfeh [24]). "Generally, in perturbation methods one starts with an (integrable) system whose solutions are known completely and studies small perturbations of it. Since the unperturbed and perturbed vector fields are close, one might expect that solutions will also be close, but, as we shall see, this is not generally the case, in that the unperturbed systems are often structurally unstable [:] arbitrarily small perturbations of such systems can cause radical qualitative changes in the structure of solutions. However, these changes are generally associated with limiting, asymptotic behavior and one does usually find that unperturbed and perturbed solutions remain close for *finite* times" (Guckenheimer and Holmes [8]). In our case, the integrable system to start with is a Hamiltonian system. In fact, when the parameter of the rate of time preference δ is null, the system is 'conservative' and the properties of its solution can be studied through a *Hamiltonian representation.* Then, the properties for δ different from zero can be understood perturbing the Hamiltonian representation by setting δ in a neighbourhood of zero. Therefore, while $\delta=0$ is not acceptable if $PV(0)$ in Eq. (3) has to converge, it is nonetheless useful as a benchmark for studying the dynamics of this economy characterized by $\delta>0$.

When $\delta=0$ there is no damping and the system is said to be *conservative* or *Hamiltonian.* Writing the total energy level as a function of h and $\dot{h}$, $u=\frac{1}{2\gamma}\dot{h}^2+F(h)$, it follows that $du/dt=0$ implying that the total energy level is constant along a solution curve of (19). In different words, given an initial level of total energy, the system moves along paths that preserve that initial energy level. This property can be used to find the solution trajectories of the system. By (7a) the total energy level can be rewritten as a function H of h and v:

$$H(h,v)=\frac{\gamma}{2}v^2+F(h) \qquad \textbf{(20)}$$

where, by definition,

$$F(h)=\int f(h)dh=\frac{(\rho-\psi)h+\psi}{\rho-\psi}\{\ln[(\rho-\psi)h+\psi]-1\}-\frac{\psi h+\rho(1-h)}{\psi}\{\ln[\psi h+\rho(1-h)]-1\}+$$
$$+\frac{\alpha}{\sigma-1}\left\{\frac{(1-\rho)h+\rho}{1-\rho}[\ln[(1-\rho)h+\rho]-1]-\frac{1-(1-\rho)h}{1-\rho}[\ln[1-(1-\rho)h]-1]\right\}$$

$H(h,v)$ is called the Hamiltonian function or simply the *Hamiltonian.* Since, by definition, when $\delta=0$, $dH/dt=(\gamma v)\dot{v}+f(h)\dot{h}=\dot{h}\,\dot{v}-\dot{v}\,\dot{h}=0$, its level curves are the solution curves of the system, the level being set by the total level of energy corresponding to the initial conditions. These conditions are the initial division of firms between countries, $h(0)$, and the initial discounted value of the expected future stream of the indirect utility differential, $v(0)$. The evaluation of such a stream depends on the equilibrium point which is expected to be reached. If centripetal forces are strong enough to give rise to agglomeration economies, then, for a given $h(0)$, multiple equilibria could be perfectly forecast. Then each forecast will determine a value for $v(0)$ and consequently the initial total energy level. This level will be conserved along the trajectory to the forecast equilibrium point, which will be eventually reached just because it is expected (*self-fulfilling expectation*). So, with multiple equilibria, the initial conditions (*history*) cannot always determine the final oucome of the relocation process.

A typical property of Hamiltonian systems is that all internal equilibrium points are either centres or saddle points. This comes from the fact that the trace of the Jacobian matrix of the associated system of differential equations is zero at an internal equilibrium point. Equilibrium points correspond to extrema of the Hamiltonian. In correspondence of a *centre*, the Hamiltonian has a maximum or a minimum. Therefore it is a degenerate level curve of the Hamiltonian. In the neibourhood of a centre level curves are non-overlapping closed trajectories or orbits (Jordan curves). A centre is stable in the sense of Liapunov (1966) because a small disturbance will result in a closed orbit around the centre along which the state of the system stays close to it. The motions corresponding to the closed orbits are periodic. Since $F(h)$ is nonlinear, the period is a function of the amplitude of the motion and thus of the energy level. In general, concentric orbits do not extend the same distances to the right and the left of the centre so that the midpoint of the motion shifts away from the static centre as the amplitude increases. This shift is often called *drift* or *steady-streaming* and characterizes non-linear systems with respect to linear ones. Consequently, the properties of the trajectories, which are assessed in a neighbourhood of the centre, are not generally valid away from the centre. That is why, in the case of nonlinear systems, the local stability analysis has to be supplemented by a global analysis. Differently from the case of a centre, in correspondence of a *saddle point* there is the meeting of four branches of a level curve. The branches (trajectories) passing through a saddle point are called *separatrices.* None of the other trajectories passes through a saddle and the separatrices are asymptotes to all other trajectories. An infinite amount of time is required to pass along a separatrix from any point in the neighbourhood of the saddle point to the saddle point itself. A saddle point is an unstable equilibrium point because any small disturbance will result in a trajectory along which the state of the system deviates more and more from the saddle point as time passes. So, even if centres and saddle points are both equilibrium points, in the neibourhood of these points the motions produced by small disturbances are very different.

Figure 2 shows the three possible shapes of the Hamiltonian corresponding to the three possible shapes of $f(h)$ and the subtitles (a), (b) and (c) are given accordingly. Figures 3 to 5 depict the qualitative properties of the system as they can be read through the Hamiltonian. In the upper portion of each picture, the undulating line represents the potential energy, that is the section of the Hamiltonian at v=0. Since the sign of the Hessian of $H(h,v)$ is sign$[(\partial^2 H/\partial v^2)(\partial^2 F(h)/\partial h^2)]$ =sign$[\gamma f'(h)]$=sign$[f'(h)]$, this section summarizes all the relevant information on the way all the parameters (other than γ) influence the shape of $H(h,v)$. The straight dotted horizontal lines represent total energy levels. Each total energy level corresponds to a different motion, and the vertical distance between a given horizontal line and the undulating line represents the kinetic energy for that motion. Thus motion is possible only in those regions where the potential energy lies below the total energy level. The lower portion of each figure shows the variation $v=\dot{h}/\gamma$ along with h. Such a graph is called a *phase plane*. The dark solid curves represent some of the level curves of the Hamiltonian (i.e. the integral curves of the system), while the light arrows show the direction of the field. For a given set of initial conditions (i.e., for a given total energy level), the response of the system can be viewed as the motion of a point along a one-parameter (time) curve, which is called *trajectory*.

In Figure 3, where trade costs are high, returns to scale are weak and the industrial good expenditure share is small, $F(h)$ has a maximum at $h^*=0.5$. When the energy level u is less than $F(h^*)$, each level curve consists of two branches, which intersect the h-axis and are similar in shape to branches of hyperbolas, one opening to the right and the other opening to the left. When $u>F(h^*)$, each curve consists also of two branches, but in this case they do not intersect the h-axis. When $u=F(h^*)$, the level curves consist of four branches that meet at the point $(h^*,0)$, which is a singular point and, in particular, it is called a *saddle point* (or *col*). The branches passing through the saddle point are the separatrices. The saddle point is the only singular point. Agents know that and will form their expectations accordingly. The system will then jump on a separatrix and will move towards the saddle point without reaching it in finite time. With high trade costs, low returns to scale and a small industrial good expenditure share, the economy tends towards a spatial distribution of firms which is *dispersed*. Centripetal forces are weaker than centrifugal ones. The metaphorical pendulum rests in the upward position.

In Figure 4, trade costs are low, returns to scale are strong and a large share of expenditures is devoted to the differentiated good. $F(h)$ has a minimum in $h^*=0.5$. When $u=F(h^*)$, the level curve degenerates into a single equilibrium point, which is a *centre*. When $u<F(h^*)$, there is no real solution, while, when $u>F(h^*)$, each level curve consists of a single closed trajectory (Jordan curve) which need not be an ellipse surrounding the centre $(h,0)$. In that case, firms keep on moving to and fro between any two points $(h,0)$ that are symmmetric with respect to $h^*=0.5$. Expectations will choose the two symmetric points. There can be endogenous "spatial cycles" of any period and any amplitude. The metaphorical pendulum oscillates between any two symmetric position with respect to the horizontal line.

In Figure 5, trade costs, returns to scale and the industrial good share of expenditures are intermediate. As a result the dynamics are richer. $F(h)$ has a maximum in $h^*=0.5$ and two minima in h_L and h_H, which are symmetric with respect to h^*. As before, in correspondence of the maximum there is a saddle point (h*,0), while, in correspondence of the minima, there are two centres, $(h_L,0)$ and $(h_H,0)$. When $u=F(h_L)=F(h_H)$, the level curve consists of the two centres, while, when $u=F(h^*)$, the level curve consists of two trajectories (separatrices) meeting at the saddle point. When $F(h_L)=F(h_H)<u< F(h^*)$, each level curve consists of two closed trajectories, one surrounding the centre $(h_L,0)$ and the other surrounding the centre $(h_H,0)$. When $h\geq F(h^*)$, each level curve consists of a single closed trajectory that surrounds the two centres as well as the saddle point. Again, expectation will choose the evolution of the economy among all these alternatives. Accordingly, one could observe an overall tendency towards the dispersed equilibrium point $(h^*,0)$, oscillations between (very uneven) spatial distributions of firms as in Figure 4, oscillations around either $(h_L,0)$ or $(h_H,0)$ so that the country who has more firms in the beginning will keep its lead forever. Figure 5 clearly shows that, as argued before, in a nonlinear system, the closed trajectories for large amplitudes are not merely scaled-up versions of those for very small amplitudes.

These pictures illustrates the strong dependence of the equilibrium points of the system on the values of parameters. Now, we show that dependence to be the same as in the tatonnement approch.

Essentially, in order to draw the Hamiltonian, two major pieces of information are required: information about saddle points and extrema; information about the behaviour at the boundaries $h=0$ and $h=1$. The two issues are tackled sequentially.

For every internal point (h,v) such that $h\in(0,1)$, the vector of the first derivatives of the Hamiltonian is well defined:

$$\left[\frac{\partial H}{\partial h},\frac{\partial H}{\partial v}\right]=\left[f(h),\gamma v\right] \tag{21}$$

Then, for every saddle point or extremum, each component of the vector has to be zero. This requirement imposes the following necessary conditions:

$$f(h) = 0 \quad \textbf{(22a)}$$

$$v = 0 \quad \textbf{(22b)}$$

The first condition requires the restoring force $f(h)$ to be null. In economic terms, it requires the centrifugal and centripetal forces to offset each other at any instant. As already seen, such a condition is also met by equilibrium points in the tatonnement approach. Therefore, equilibrium points are the same as before. The second condition requires skilled workers to be indifferent between staying where they are or moving. It implies that equilibrium points have to be looked for on the section of the Hamiltonian cut along the h-axis i.e. along the potential energy function $F(h)$. Then, the two conditions together say that saddle points and extrema of $H(h,v)$ are extrema of $F(h)$, points where $F'(h) \equiv \mathrm{d}F(h)/\mathrm{d}h \equiv f(h)=0$.

To determine whether a zero of $f(h)$ is a saddle point or an extremum, sufficient conditions have to be taken account of. The Hessian Matrix of the Hamiltonian is:

$$\begin{bmatrix} \dfrac{\partial^2 H}{\partial h^2} & \dfrac{\partial^2 H}{\partial h \partial v} \\ \dfrac{\partial^2 H}{\partial v \partial h} & \dfrac{\partial^2 H}{\partial h^2} \end{bmatrix} = \begin{bmatrix} f'(h) & 0 \\ 0 & \gamma \end{bmatrix} \quad \textbf{(23)}$$

Since $\gamma>0$, the Hessian determinant is positive if $f'(h)>0$, negative if $f'(h)<0$. If $f'(h)$ were null for some h, $f(h)$ would be constant for that values of h. This never happens unless $\tau=1$, in which case $f(h)$ is always constant since location is irrelevant when there is no obstacle to trade.

Due to γ being positive, the Hamiltonian cannot have any maximum. It will have (local) minima for those h such that $f(h)=0$ and $f'(h)>0$ and saddle points for those h such that $f(h)=0$ and $f'(h)<0$. Thus, minima (maxima) of $F(h)$ are minima (saddle points) of the Hamiltonian. A positive sign $f'(h)$ for every h is sufficient for the Hamiltonian to exhibit at most one extrema and this is a unique minimum.

So, the study of the relevant features of the shape of $H(h,v)$ can be reduced to the study of $f(h)$. Therefore, the same conditions that support alternatives (a), (b) or (c) in Figure 1, will support alternatives (a), (b) or (c) in Figure 2. The necessary and sufficient conditions for each alternative are summarized in Table I.

As discussed in part 1, Table I assesses the relation between the values of the parameters and the number of equilibrium points. In particular, it shows that multiple equilibria arise when trade costs are low, increasing returns are strong and the industrial good expenditure share is large. However, once the dynamics have been explicitly spelled out, initial conditions will not generally be enough to select the final equilibrium. Differently from the tatonnement case, expectations matter. Therefore, in Figure 4 the industrial sector could agglomerate in either country no matter what the initial conditions are. Along the same line of reasoning, in Figure 5, initial conditions cannot tell along which integral curve the system will move. So, the tatonnement argument misses the importance of expectations. In particular, it misses the fact that self-fulfilling expectations and agglomeration forces are twinned, whenever the future is not discounted too heavily. Section 4 will come back to this point.

3. Perturbation of the Hamiltonian

The previous section has shown that, when $\delta=0$ and agglomeration forces are strong enough, the system exhibits endogenous spatial cycles. As already mentioned, the Hamiltonian case in nongeneric and thus uninteresting in itself. Nonetheless it is a useful starting point to study the qualitative features of a nonzero rate of time preference. In fact, the case of $\delta\neq0$ can be understood by perturbing the Hamiltonian system.

When $\delta\neq0$ Eq. (18) exhibits the damping mechanism $\frac{\delta}{\gamma}\dot{h}$. If $\delta>0$ the damping causes the amplitude of the motion to increase (*negative damping*). The opposite would happen if $\delta<0$ (*positive damping*). In both cases, the total energy level is not constant and the system is *nonconservative*. In fact, the presence of damping alters one of the typical properties of Hamiltonian systems in which equilibrium points are either saddle points or centres. The existence of centres is due to the null divergence of the vector field corresponding to a conservative system. The divergence of the vector field is a measure of the expantion of flows defined as:

$$\frac{\partial\dot{h}}{\partial h}+\frac{\partial\dot{v}}{\partial v}=\delta \qquad \textbf{(24)}$$

This follows from (7a,b). When $\delta=0$ divergence is null and flows point neither inward nor outward so that centres and closed orbits around them are possible. On the contrary, when $\delta>0$ ($\delta<0$) divergence is positive (negative) and flows point outward (inward). Therefore equilibrium points cannot be centres and are either saddle points or *focal points*. These are either sources or sinks. In a neighbourhood of a *source* all trajectories lead away from it and consequently they are not generally observed as equilibria. On the contrary, in a neighbourhood of a *sink*, all trajectories lead to it. However, changing the value of δ from zero, has also another major effect on the topological property of the dynamics. If perturbation makes δ positive so that a centre becomes a source, all trajectories from it are isolated. As Matsuyama [22] points out, even if equilibria were multiple, one could still perform a local comparative static exercise by (*ad hoc*) supposing that the economy jumps to the nearby trajectory. On the contrary, if perturbation makes δ negative so that a centre becomes a sink, trajectories are no longer isolated since there is a continuous family of trajectories that lead to a source. In that case, no comparative static exercise can be performed. However, since convergence of *PV*(0) in Eq. (3) requires a positive δ, in our economy the case of a negative δ must be discarded.

The effects of turning from $\delta=0$ to $\delta>0$ on the phase plane are shown in Figure 6. The dark solid curves represent the integral curves of the system. There are no exact solutions for such curves. However they can be approximated at whatever degree of precision (Nayfeh [24], Nayfeh and Mook [25]). The light arrows show the direction of the flow. Cases (a), (b) and (c) correspond respectively to Figure 3, 4 and 5. Cases (a), in which centrifugal forces dominate on centripetal ones, exhibits the same qualitative properties as Figure 3. There is a unique equilibrium point with dispersion and it is a saddle point. Perturbation makes no difference since even in the Hamiltonian case there are no closed orbits. More interesting is case (b) where centripetal forces dominate. In that case, perturbation breaks the closed orbits around the singular point (h^*,0) that becomes a focus. In particular, since flows point outward, it becomes a *source*. When $\delta>0$, the system will eventually reach an endpoint, either (0,0) or (1,0), in finite time. Initial values for h can be divided into two groups. In correspondence of some of them, relatively close to h^*, there are two trajectories each leading towards agglomeration in a different country. For such initial values, expectations will choose along which trajectory the economy will move. The set of those values is what Krugman [13] calls the 'overlap'. On the contrary, for other initial values of h, relatively close either to 0 or 1, there is a unique trajectory. Therefore expectations have no role to play, and the initial value of h alone determines the final outcome. If initially a country has a much lower share of skilled labour, it will eventually lose all its industry because of agglomeration in the other country.

In case (c), agglomeration and dispersion characterize alternative stable equilibrium points. If countries have initially relatively similar endowments of skilled workers, the economy will tend to an even distribution of firms. On the contrary, if initially they are very different, the country which is initially well endowed with skilled workers, will gain all firms. In case (c) there are two "overlaps" around (h_L,0) and (h_H,0). For values of h inside the "overlaps", expectations will decide whether there is going to be dispersion or agglomeration.

4. History-dependent versus expectation-driven trajectories

"Overlaps" exist as in Figure 6 only if the solution of the system yields multiple whirling trajectories, in which case the economy exhibits expectation-driven endogenous spatial fluctuations. This section investigates the generality of this feature.

Since the demarcation curves, $\dot{h} = 0$ and $\dot{v} = 0$, are everywhere differentiable, expectations matter if there are sources and whirling trajectories flowing cyclically away from them. This happens whenever there exist complex roots in the characteristic equation of the reduced linearization of the system around an equilibrium point. For a given equilibrium point $(h,0)$, the Jacobian matrix associated to system (7) is:

$$J_h \equiv \begin{bmatrix} 0 & \gamma \\ -f'(h) & \delta \end{bmatrix} \quad \textbf{(25)}$$

The characteristic equation has complex roots if $(\text{tr}J_h)^2 < 4|J_h|$, where $\text{tr}J_h$ and $|J_h|$ are respectively the trace and the determinant of the Jacobian matrix. The condition for the existence of two complex roots can then be written as:

$$\delta^2 - 4\gamma f'(h) < 0 \quad \textbf{(26)}$$

Necessary condition for (26) to hold is $f'(h)>0$ i.e. marginal immigration in one country improves the indirect utility of local skilled workers with respect to the other country. In that case, relatively low values of δ and relatively high values of γ are sufficient to yield complex roots. That should be expected. If the rate of time preference and the cost of adjustment are low, future returns are important and expectations can drive the evolution of the system. To get further insight, consider the equilibrium point (0.5,0). $f'(0.5)$ takes the value φ as defined in equation (15). As already pointed out, $\varphi>0$ implies the existence of agglomeration economies and the perturbed system will rest only when complete concentration has taken place (h=0 or h=1). Nonetheless, a positive φ is not sufficient to make (26) negative: it also has to be relatively large. So, while the existence of agglomeration economies is necessary for expectations to have a role, it will be also sufficient only if such economies are relatively strong.

Whenever this is not the case, the existence of only real roots rules out the possibility of whirling trajectories and initial conditions will therefore select the final equilibrium point. So, only when (26) does not hold for *any* equilibrium point, the tatonnemet argument is good enough to describe the direction of the migration flows. When agglomeration economies are not strong enough, δ is large and γ is small, multiple equilibria do not imply multiple trajectories. In particular, they do not imply endogenous spatial fluctuations.

5. Final remarks

With S-D-S monopolistic competition and increasing returns in production, obstacles to trade give rise to complementarities between agents location decisions. These complementarities tend to favour the spatial agglomeration of economic activities. Such a tendency is opposed by the presence of localized inputs and demands that tend to favour spatial dispersion. The nonlinear interaction between those opposing forces accounts for the complexity of the economic landscape. Spatial equilibria are multiple and often path-dependent. However, complementarity in itself raises also the issue of self-fulfilling expectations. When the complementarities among agents' location decisions are strong, an equilibrium can emerge just because everyone expects it to emerge. These issues point to a deeper understanding of the link between monopolistic competition, increasing returns, trade barriers and endogenous spatial fluctuations.

In a two-country, two-factor model with two sectors, one perfectly competitive and the other monopolistically competitive, this paper has shown that, when agents do not discount the future too heavily, spatial fluctuations occur if the monopolistically competitive sector absorbes a relatively large share of aggregate expenditures, if the market power (i.e. returns to scale) of firms in that sector is strong and if obstacles to trade and relocation are small.

Even if the model presented is highly stylized, it represents an interesting and novel example of how recent developments in trade and location theory can add to the more general understanding of the link between monopolistic competition, multiple equilibria, and endogenous fluctuations. From the more narrow point of view of trade theory, it suggests that, with imperfect competition, even very small changes in the policy parameters can have dramatic consequences on the global stability properties of the spatial distribution of economic activities. Accordingly, we believe that more effort should be devoted to a deeper understanding of those consequences.

References

1. B. Arthur, Self-reinforcing mechanisms in economics, in P.W.Anderson, K.J.Arrow and D.Pines, eds., "The Economy as an Evolving Complex System", Santa Fe Institute Studies in the Sciences of Complexity, vol.5, Addison-Wesley, Redwood City, 1988.

2. A.K. Dixit, and J. Stiglitz, Monopolistic competition and optimum product diversity, *Amer. Econ. Rev.* **67** (1977), 297-308.

3. H. Flam and E. Helpman, Industrial Policy under Monopolistic competition", *J. Int. Econ.* **22** (1987), 79-102.

4. K. Fukao, and R. Benabou, History versus expectations: A comment, *Quart. J. Econ.* **108** (1993), 535-542.

5. J. Galì, Expectations-driven spatial fluctuations, in "The Location of Economic Activity: New Theories and Evidence", Centre for Economic Policy Research and Cambridge University Press, Cambridge, 1994.

6. J. Galì, Monopolistic competition, business cycles, and the composition of aggregate demand, *J. Econ. Theory* **63** (1994), 73-96.

7. J. Galì, Local externalities, Convex adjustment costs, and sunspot equilibria, *J. Econ. Theory* **64** (1994), 242-252.

8. J. Guckenheimer and P. Holmes, "Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields", corrected third printing, Springer-Verlag, New York, 1990.

9. E. Helpman and P. Krugman, "Market Structure and Foreign Trade", MIT Press, Cambridge, Massachusets, 1985.

10. J.V. Henderson, "Economic theory and the cities", Academic Press, New York, 1977.

11. P. Krugman, "Geography and Trade", MIT Press, Cambridge, Massachusetts, 1991.

12. P Krugman, Increasing returns and economic geography, *J. Polit. Econ.* **99** (1991), 483-499.

13. P. Krugman, History versus expectations, *Quart. J. Econ.* **106** (1991), 651-667.

14. P. Krugman, A dynamic spatial model, Working Paper No. 4219, National Bureau of Economic Research, Cambridge, Massachusetts, 1992.

15. P. Krugman, Complex landscapes in economic geography, *Amer. Econ. Rev.* **84** (1994), 412-416.

16. P. Krugman and A.Venables, Integration, specialization and adjustment, Discussion Paper No. 886, Centre for Economic Policy Research, London, 1993.

17. P. Krugman and A.Venables, Globalization and the inequality of nations, Discussion Paper No. 165, Centre for Economic Performance, London School of Economics, 1994.

18. P. Krugman and A. Venables, The seamless world: A spatial model of international specialization, Discussion Paper No. 1230, Centre for Economic Policy Research, London, 1995.

19. A.M. Liapunov, "Stability of Motion", Academic Press, New York, 1966.

20. P. Martin and G.I.P. Ottaviano, The geography of multi-speed Europe, Discussion Paper, Centre for Economic Policy Research, London, 1995 (forthcoming).

21. P. Martin and C.A. Rogers, Industrial Location and Public Infrastructure, *J. Int. Econ.* (1995), forthcoming.

22. K. Matsuyama, Increasing returns, industrialization, and indeterminacy of equilibrium, *Quart. J. Econ.* **106** (1191), 617-650.

23. K. Matsuyama, "Complementarities and cumulative processes in models of monopolistic competition", *J. Econ. Lit.* **33** (1995), 701-729.

24. A.H. Nayfeh, "Perturbation Methods", Wiley, New York, 1973.

25. A.H. Nayfeh and D.T. Mook, "Nonlinear Oscillations", Wiley, New York, 1979.

26. D. Puga, Regional integration and the dynamics of urbanization, Discussion Paper No. 212, Centre for Economic Performance, London School of Economics, 1995.

27. D. Puga and A. Venables, Preferential trading arrangements and industrial location, Discussion Paper No. 267, Centre for Economic Performance, London School of Economics, 1995.

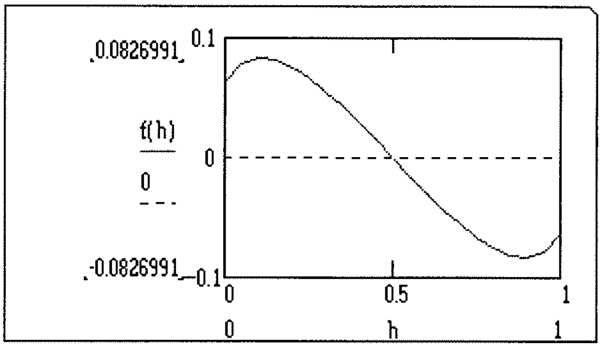
28. P. Samuelson, Spatial price equilibrium and linear programming, *Amer. Econ. Rev.* **42** (1952), 283-303.

29. M.A. Spence (1976), Product selection, fixed costs, and monopolistic competition, *Rev. Econ. Stud.* **43** (1976), 217-235.

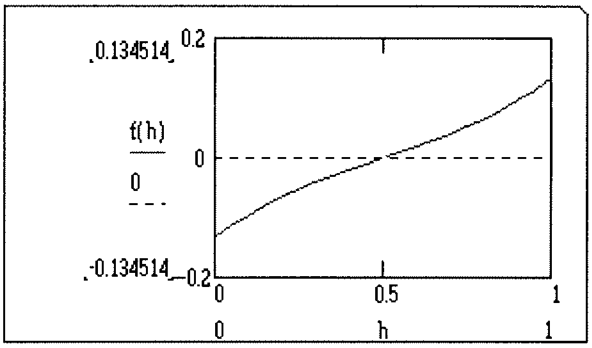
30. A.J. Venables, Equilibrium locations of vertically linked industries, Discussion Paper No. 802, Centre for Economic Policy Research, London, 1993.

	$\Phi<0$	$\Phi>0$
$\varphi<0$	Fig. 5 (a)	Fig. 5 (c)
$\varphi>0$	Fig. 5 (d)	Fig. 5 (b)

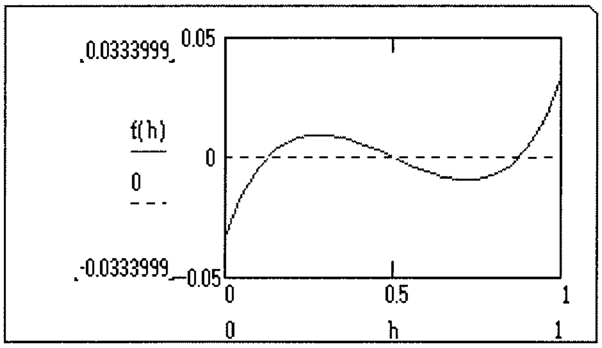
Table 1 - Necessary and sufficient conditions for Figure 1



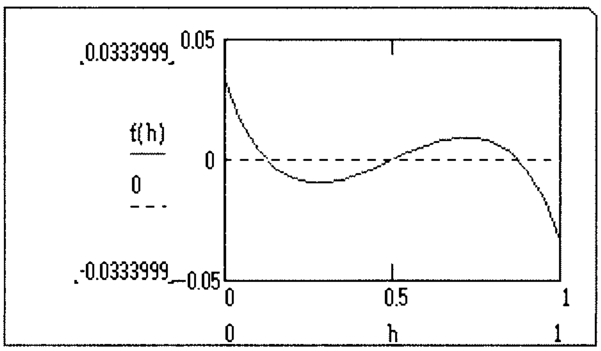
(a)



(b)

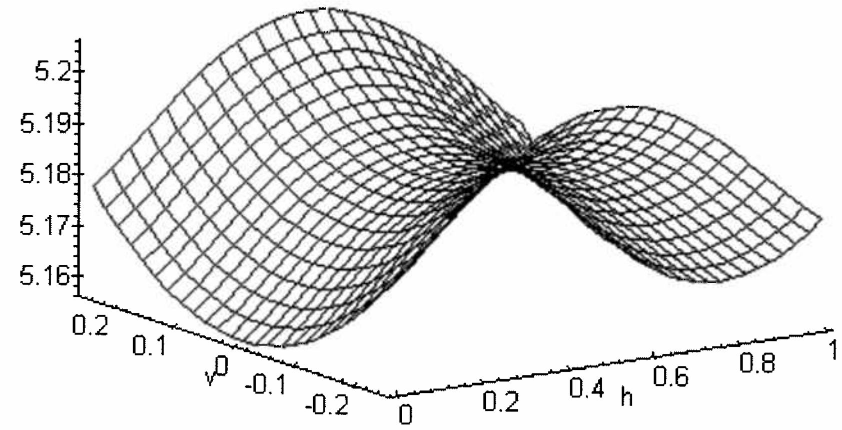


(c)

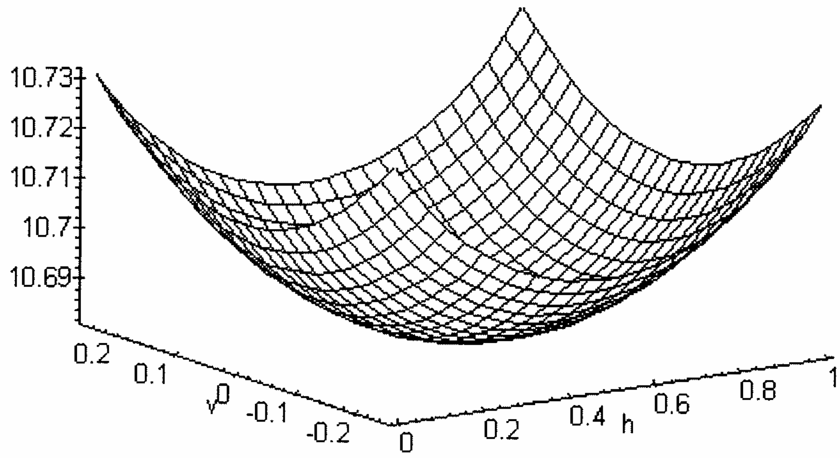


(d)

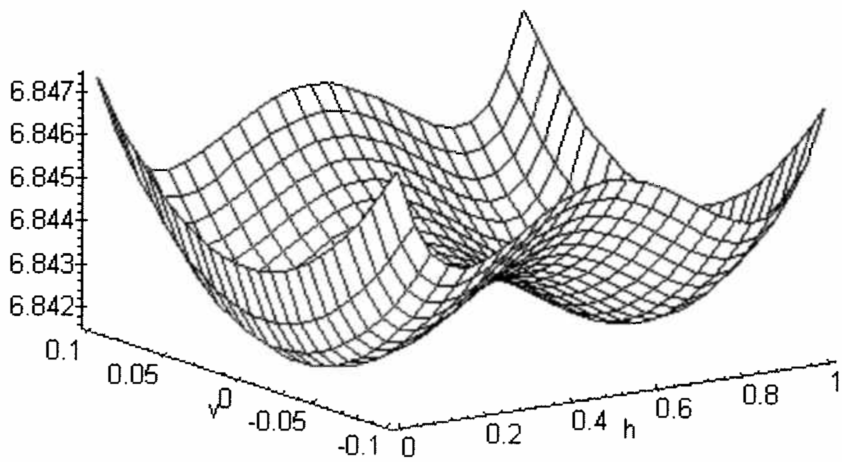
Figure 1 - The indirect utility differential



(a)

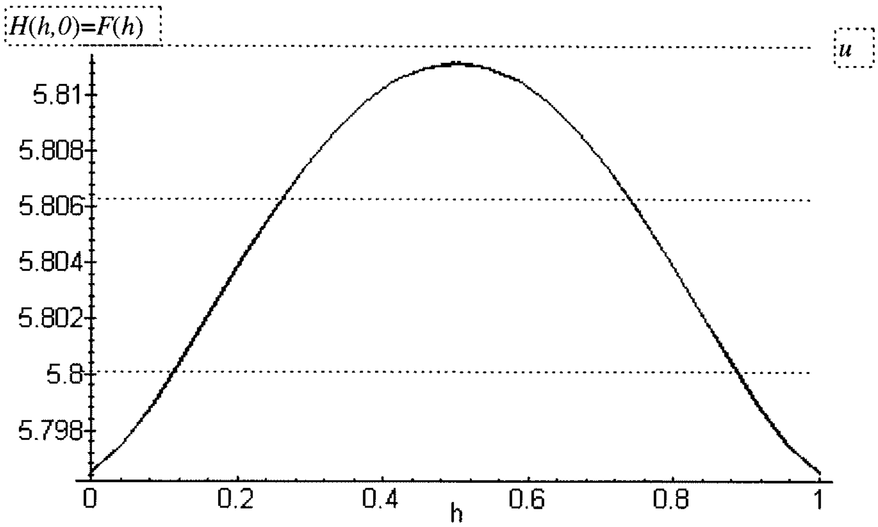


(b)

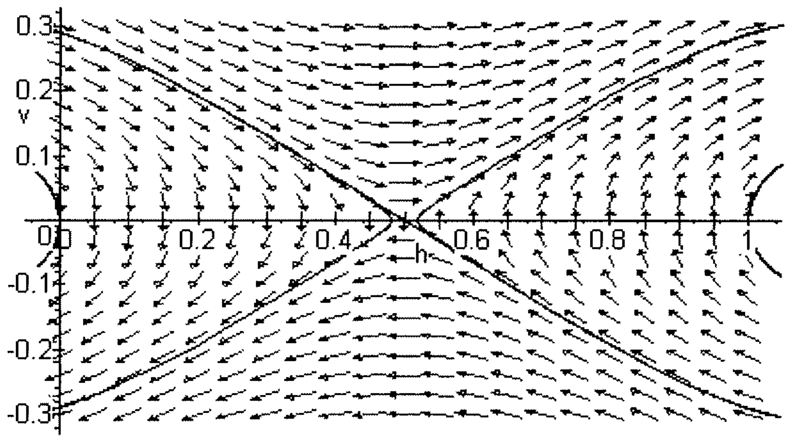


(c)

Figure 2 - The Hamiltonian

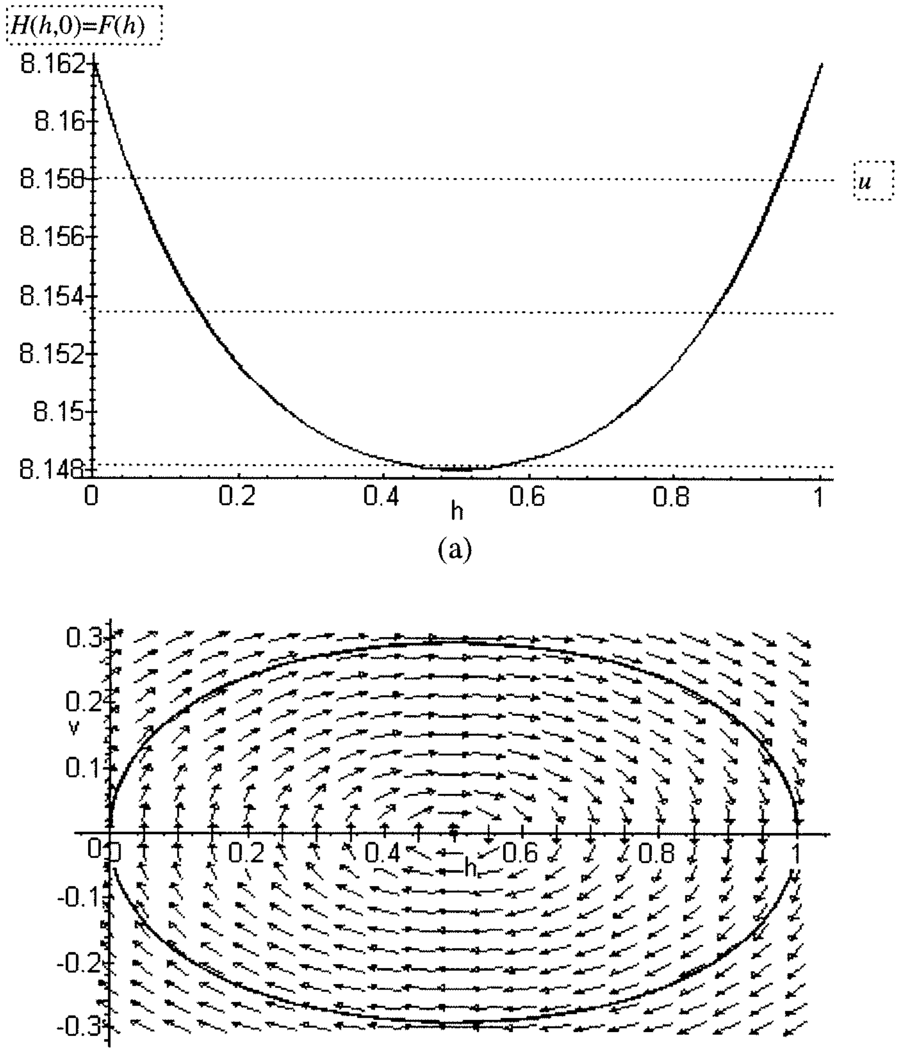


(a)



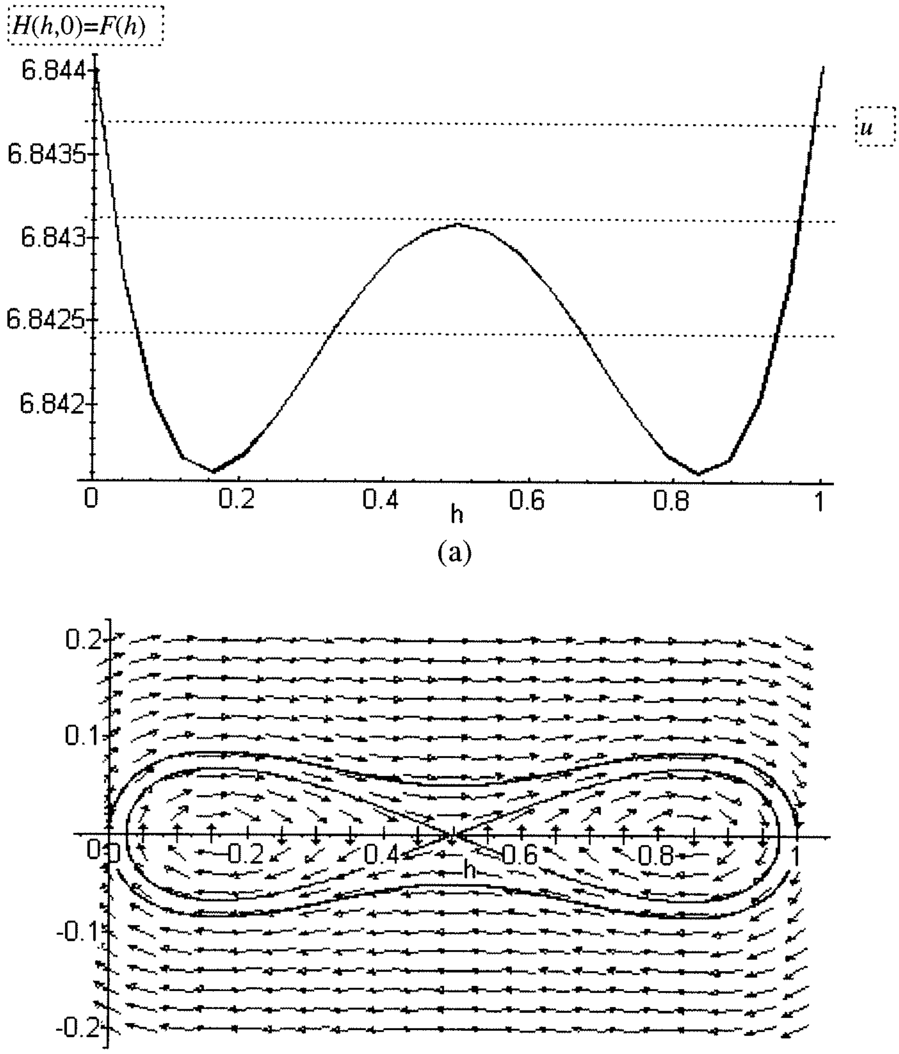
(b)

Figure 3 - The Hamiltonian: a unique saddle point



(b)

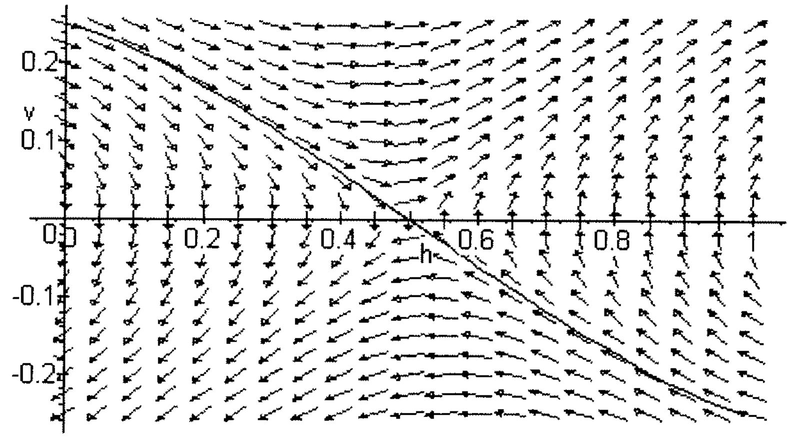
Figure 4 - The Hamiltonian: a unique centre



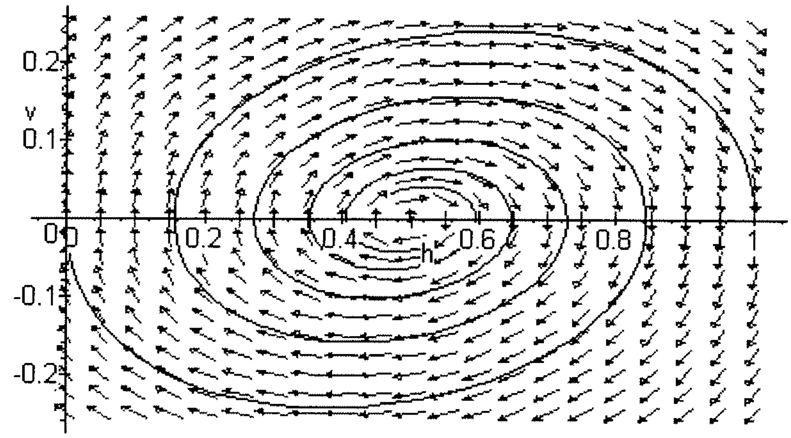
(a)

(b)

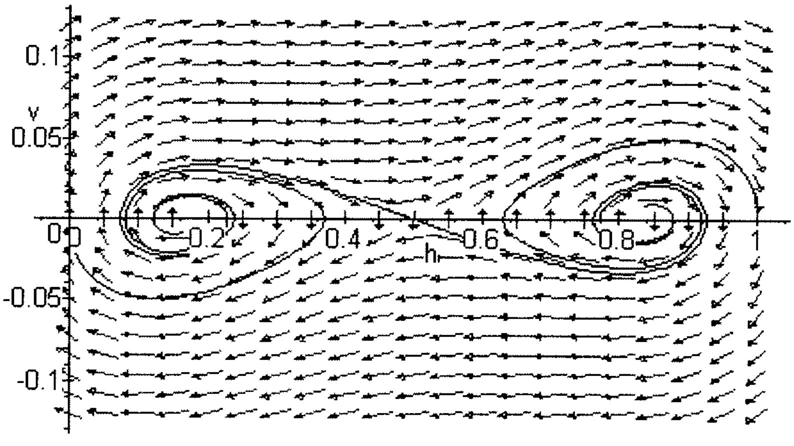
Figure 5 - The Hamiltonian: one saddle and two centres



(a)



(b)



(c)

Figure 6 - Perturbation of the Hamiltonian