ELSEVIER

Journal of Development Economics
Vol. 49 (1996) 137–150

JOURNAL OF
Development
ECONOMICS

Trade policy and the Third World metropolis

Paul Krugman *, Raul Livas Elizondo

Received 1 February 1995; revised 1 October 1995

Abstract

Many of the world's largest cities are now in developing countries. We develop a simple theoretical model, inspired by the case of Mexico, that explains the existence of such giant cities as a consequence of the strong forward and backward linkages that arise when manufacturing tries to serve a small domestic market. The model implies that these linkages are much weaker when the economy is open to international trade; in other words, the giant Third World metropolis is an unintended by-product of import-substitution policies, and will tend to shrink as developing countries liberalize.

JEL classification: F12; F13; F15; R12

Keywords: Urban economics; Cities; Regions; International trade; Protection; Mexico

1. Introduction

Half a century ago, really large cities were found mainly in advanced industrial nations. Today, many of the world's largest cities are in developing countries. Many, perhaps most, observers suspect that the emergence of these huge urban concentrations is unhealthy. Bairoch (1988), for example, has called Third World metropolises "Romes without empires", and suggests that they are parasitic entities that drain the economic vitality from their host economies. Some developing country governments have encouraged decentralization of industry in an effort to curb the growth of their biggest cities, with little effect.

One might have expected that the remarkable phenomenon of the Third World metropolis would be a major preoccupation of development economists, and that

* Corresponding author.

SSDI 0304-3878(95)00055-0

policy analysis in developing countries would routinely focus on the question of how any proposed policy change would affect the geographic concentration of population. But this does not seem to be the case. Admittedly, urban economists, especially urban systems theorists like Henderson (1988), make extensive use of evidence from developing countries. In the development literature proper, however, urbanization in general and the growth of giant cities in particular are addressed obliquely, if at all. In the *Handbook of Development Economics*, for example, the chapter by Williamson (1988) treats rural–urban migration at considerable length, but barely touches on why so many manufacturing jobs are concentrated in huge urban areas in the first place. When economists discuss such issues as trade policy in developing countries, they generally pay little attention to the effects of such policies on the internal economic geography of those countries.

The purpose of this paper is to argue that such neglect is a mistake; that the trade policies of developing countries and their tendency to develop huge metropolitan centers are closely linked. It argues that the rise of giant metropolises in developing countries after World War II may have been due in large part to the rise of import-substituting industrialization policies. Correspondingly, the shift away from such policies may well limit the future growth of huge Third World cities. The inspiration for the paper is the case of Mexico, which contains what is probably the world's most populous city, but which has begun a noticeable process of decentralization as it liberalizes trade. We argue, however, that the case is more general: closed markets promote huge central metropolises, open markets discourage them.

The paper is in five parts. Section 2 presents an intuitive version of the basic argument. Section 3 lays out the assumptions of an illustrative formal model. Section 4 shows how forward and backward linkages can support a large metropolis in a closed economy. Section 5 shows how the existence of such a metropolis depends on the openness of the economy. Finally, Section 6 offers some suggestions for further research.

2. Trade policy and metropolitan concentration

Mexico City is arguably the world's largest urban center. The disadvantages of such a massive population concentration are apparent at first sight and first breath. One might have expected manufacturing to avoid the city's high land rents and relatively high wage rates by Mexican standards, let alone its congestion and pollution. As late as 1980, however, and in spite of the *maguiladora* program designed to encourage export-oriented manufacturing near the US border, Mexico City still accounted for more than 40% of the nation's manufacturing employment, more than half its manufacturing value-added. The proportion has declined substantially since then, and it is indeed this decline that motivated this paper; but as a starting point we need to explain why so much population and industry concentrated in Mexico City in the first place.

A major reason for the concentration of manufacturing in Mexico City was surely the powerful backward and forward linkages the site offered. Firms manufacturing for the Mexican domestic market had an incentive to choose production sites with good access to consumers; the huge and relatively affluent population concentration at Mexico City ensured that sites close to the capital offered the best market access; in Hirschman (1958) terms, the capital offered strong backward linkages. Firms would also want good access to the products of other firms, whether these goods were in the consumption basket of their workers or were intermediate inputs into their own production; the wide variety of goods produced near Mexico City ensured that it offered the best access to such inputs; again in Hirschman's terms, the capital provided forward linkages as well. These backward and forward linkages played a major role in overcoming the disadvantages of high rents, wages, congestion and pollution.

This is, of course, a circular argument, which in economic geography and development economics is a virtue, not a vice! Manufacturers choose to produce in Mexico City because of the concentration of demand and inputs there, but there is a concentration of demand and inputs in Mexico City in large part precisely because so many producers have chosen that site. So the size of the national metropolis is the result of a self-reinforcing process of agglomeration. One needs to address the specifics of Mexican history to ask why Mexico City rather than some other site was the place on which the circular causation converged, but the important thing from the economist's point of view is that there was a logic that mandated concentration somewhere. And of course this same logic applies to countries other than Mexico.

This much is intuitive, even obvious. What may be less obvious is that the argument relies critically on two somewhat hidden assumptions: significant economies of scale and industrialization oriented primarily toward the domestic market.

The role of economies of scale may be seen by noticing that our description of the locational choices of firms implicitly assumes that they must choose only one or at most a few sites to serve the domestic market. Given this constraint, it makes sense to choose Mexico City, and serve the rest of the market from there. But what if it were possible to build a number of small factories at little cost in efficiency? Then one could build a factory to serve each local market. Given the high land and wage costs of Mexico City, there would be no point in exporting manufactures from there; indeed, if anything one might prefer to supply the capital at least in part from lower-cost sites elsewhere. Without the incentive to produce from a single central site, however, the logic of cumulative agglomeration would break down. As development economists have long understood, backward and forward linkages only become economically meaningful in the presence of sufficiently strong scale economies; the same must be true of a story of urban concentration that rests upon these linkages.

The importance of reliance on the domestic market can be seen by asking what would happen if the typical manufacturer sold primarily to export markets and relied primarily on imported inputs. Then there would be little advantage to a location near a country's metropolitan center - little backward linkage, because most output is sold abroad, little forward linkage, since most inputs come from abroad. Meanwhile, the disadvantages of expensive land and labor would loom just as large. So our story about the metropolis depends on the assumption that industrialization is inward-looking. [1]

These observations do not pose a problem for our story about Mexico City, because in 1980 Mexican manufacturing was primarily oriented toward the domestic market, and this market was sufficiently small that scale economies were of major importance. This inward orientation was not, however, a fact of nature: it was a result of policy. In other words, our story suggests that the extraordinary concentration of population and production in Mexico City, and by extension in other Third World metropolises, was an unintended by-product of import-substituting industrialization.

The rough outline of Mexican economic history supports this view. Recent work by Hanson (1992) and Livas Elizondo (1992) shows that before the beginnings of import substitution Mexico City was far less dominant in Mexico's economy and manufacturing sector than it was later to become, and that since liberalization began in the 1980s there has been a dramatic shift of manufacturing away from Mexico City, especially to the northern states. Admittedly, the Mexican experiment is not as pure as we would like: the northern states are not only less congested than Mexico City, they are also closer to the US border. Our informal argument suggests, however, that much the same history would have unfolded even if there were no special locational advantage to northern production, and that trade liberalization will shrink metropolises in other Third World countries as well.

It is not, however, enough simply to make a plausible informal argument. To solidify our story, we must embed it in a fully worked out model. So we turn next to such a model.

3. A formal model

Any interesting model of economic geography must involve a tension between the "centripetal" forces that tend to pull population and production into agglomerations and the "centrifugal" forces that tend to break such agglomerations up.

[1] We might also note, somewhat parenthetically, that economies of scale are in practice more likely to be significant when industrialization is oriented toward the domestic market. Mexico, which is a big economy for the developing world, has a market only about 3% as large as that of the US, so that presumably there are many more sectors in which minimum efficient scale is large relative to sales than there would be if Mexico were selling to an integrated North American market. Unfortunately, the special assumptions made below in order to keep the formal model tractable tend to obscure this point.

Centripetal forces can include both pure external economies and a variety of market size effects, such as the forward and backward linkages described above. Centrifugal forces can include pure external diseconomies such as congestion and pollution, urban land rents, and the attraction of moving away from highly competitive urban locations to less competitive rural ones.

In this model we choose to include only the centripetal forces that arise from the interaction among economies of scale, market size and transportation costs, i.e., backward and forward linkages. There are undoubtedly other external economies at work in real urban areas, but they are omitted in the interest of keeping the model as simple as possible and of keeping a reasonable distance between assumptions and conclusions.

For similar reasons, the only centrifugal force allowed is commuting cost/land rent. In several recent papers (Krugman, 1991, Krugman, 1993a, Krugman, 1993b) one of us has adopted instead a modeling approach in which the centrifugal force is the pull of a dispersed rural market; while that approach has some important virtues, it seems both less to the point and less realistic than a focus on land rent in the current context.

We imagine, then, an economy consisting of three locations 0, 1, and 2. Location 0 is the "rest of the world", while 1 and 2 are two domestic locations (e.g., Mexico City and Monterrey). There is only one factor of production, labor. A fixed domestic supply of labor (L) is mobile between locations 1 and 2, but there is no international labor mobility.

It will be assumed that in each location production must take place at a single central point. [2] Workers, however, require land to live on. To make matters simple, we make several special assumptions. First, we assume that each worker needs a fixed living space, say one unit of land. Second, we assume that the cities are "long and narrow", so that workers are effectively spread along a line. This has the implication that the commuting distance of the last worker in location j is

$$d_j = L_j/2 \tag{1}$$

Finally, we assume that commuting costs are incurred in labor: a worker is endowed with one unit of labor, but if he must commute a distance d, he arrives with a net amount of labor to sell of only

$$S = 1 - 2\gamma d \tag{2}$$

These assumptions immediately allow us to describe the determination of land rent given the labor force at a location. Let w_j be the wage rate paid at the city center per unit of labor. Workers who live at the outskirts of the town will pay no land rent, but will receive a net wage of only $(1 - \gamma L_j)w_j$ because of the time

[2] Ideally the need for a central business district would itself be derived from the model, but this is left for later research.

spent in commuting. Workers who live closer to the city center will receive more money, but must pay an offsetting land rent. The wage net of commuting costs declines as one moves away from the city center, but land rents always exactly offset the differential. Thus the wage net of both commuting and land rents is $(1-\gamma L_j)w_j$ for all workers. Total land rents are equal to the area of the triangle above that net wage.

We may also note, for future reference, that the total labor input of a location, net of commuting costs, is

$$Z_j = L_j(1-0.5\gamma L_j) \tag{3}$$

and that the location's total income, including the income of landowners, is

$$Y_j = w_j Z_j \tag{4}$$

Commuting costs and the resulting land rent are obviously diseconomies of city size. To explain agglomeration, we must introduce compensating advantages of concentration. These must arise from economies of scale. Unless economies of scale are purely external to firms, however, an approach we have rejected, they must lead to imperfect competition. So in introducing scale economies we must do so in a way that allows a tractable model of imperfect competition.

Not surprisingly, the easiest way to do this is with the familiar tricks of the Dixit–Stiglitz monopolistic competition model (Dixit and Stiglitz, 1977). We suppose that there are a large number of symmetric potential products, not all actually produced. Each producer acts as a profit-maximizing monopolist, but free entry drives profits to zero.

Specifically, we assume that everyone in the economy shares the CES utility function

$$U = \left[\Sigma_i C_i^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} \tag{5}$$

To produce any good i at location j involves a fixed as well as a variable cost:

$$Z_{ij} = \alpha + \beta Q_{ij} \tag{6}$$

The properties of this model are by now very familiar. As long as many goods are produced, and as long as we make appropriate assumptions on transportation costs (see below), each producer faces an elasticity of demand equal to the elasticity of substitution, and will therefore charge a price that is a constant markup over marginal cost:

$$P_j = \frac{\sigma}{\sigma-1}\beta w_j \tag{7}$$

Given this pricing rule and the assumption that free entry will drive profits to zero, there is a unique zero-profit output of each product:

$$Q = \frac{\alpha}{\beta}(\sigma - 1) \tag{8}$$

And the constancy of output of each product implies that the number of goods produced at each location is simply proportional to its net labor input after commuting:

$$n_j = \frac{Z_j}{\alpha\sigma} \tag{9}$$

It is worth dwelling for a moment on Eq. (9). Increasing returns at the level of the firm are an essential feature of the story in this paper. Yet they will seem to be almost invisible from this point on. Where did they go? The answer is that they are embedded in Eq. (9): the fact that a location with large net labor input produces a greater variety of goods than one with smaller labor input drives all of the results.

It will save notation later if we make two useful choices of units. First, let us choose units so as to make the f.o.b. price of goods produced at any given location equal to the wage rate at the region's city center. Thus we have

$$P_j = W_j \tag{10}$$

Second, notice that there is no reason why we need to count goods one at a time. We can equally well count them in ''batches'', say of a dozen each. So we can play with the batch size; and to save notation, we let the batch size be such that

$$n_j = Z_j \tag{11}$$

Next, we introduce costs of transacting between locations. In order to preserve the constant elasticity of demand facing firms, these must take Samuelson's ''iceberg'' form in which transport costs are incurred in the goods shipped. Thus we assume that when a unit of any good is shipped between location 1 and location 2, only $1/\tau$ units actually arrive; thus the c.i.f. price of a good shipped from either domestic location to the other is τ times its f.o.b. price. Only a fraction $1/\rho$ of a good imported from location 0 is assumed to arrive in either location 1 or 2. For simplicity, exports are assumed to take place with zero transportation costs. [3]

We take τ to represent ''natural'' transportation costs between locations. The parameter ρ, however, is meant to be interpreted as combining natural transport

[3] Even though we make exports costless, an increase in ρ, which reduces imports, must necessarily decrease exports as well. The mechanism through which this happens is through a rise in the prices of domestic relative to foreign output, in effect through a real overvaluation that prices domestic goods out of world markets.

costs with artificial trade barriers. It would be straightforward (and would yield similar results) in this model to introduce an explicit ad valorem tariff whose proceeds are redistributed, but here we simply imagine that any potential revenue is somehow dissipated in waste of real resources - not too unrealistic a view, if the rent-seeking story is to be believed.

Given these transportation costs and the utility function, we may define true consumer price indices for manufactured goods in each location. First, let us define the shares of the three locations in the total number of products produced, which are equal to their shares of net labor input:

$$\lambda_j = \frac{n_j}{\Sigma_k n_k} = \frac{Z_j}{\Sigma_k Z_k} \tag{12}$$

Let the wage rate in location 0 be the numeraire; then the true price indices are

$$T_0 = K\left[\lambda_0 + \lambda_1 w_1^{1-\sigma} + \lambda_2 w_2^{1-\sigma}\right]^{\frac{1}{1-\sigma}} \tag{13}$$

$$T_1 = K\left[\lambda_0 \rho^{1-\sigma} + \lambda_1 w_1^{1-\sigma} + \lambda_2 (w_2 \tau)^{1-\sigma}\right]^{\frac{1}{1-\sigma}} \tag{14}$$

$$T_2 = K\left[\lambda_0 \rho^{1-\sigma} + \lambda_1 (w_1 \tau)^{1-\sigma} + \lambda_2 w_2^{1-\sigma}\right]^{\frac{1}{1-\sigma}} \tag{15}$$

where

$$K = (n_0 + n_1 + n_2)^{\frac{1}{1-\sigma}} \tag{16}$$

We will take Z_0 as given. Suppose that we know the allocation of labor between locations 1 and 2. Then this will allow us to determine Z_1 and Z_2. As we will see later, we can then solve the model for equilibrium wage rates w_j. Labor is, however, mobile, and we will only have a full equilibrium if all domestic workers receive the same net real wage. This net real wage in location j can be defined as

$$\omega_j = w_j(1 - \gamma L_j)/T_j \tag{17}$$

A situation in which real wages are equal in the two domestic locations is an equilibrium. Such an equilibrium may, however, be unstable under any plausible adjustment story. To get some rudimentary dynamics, we impose a simple Marshallian adjustment mechanism,

$$\mathrm{d}L_1/\mathrm{d}t = -\mathrm{d}L_2/\mathrm{d}t - \delta(\omega_1 - \omega_2) \tag{18}$$

We could try to justify this mechanism in terms of explicit moving costs, and take account of forward-looking behavior, but that would go beyond the scope of the present paper.

We have now laid out a complete formal model. We will turn to the full solution of that model in Section 5. First, however, we consider a special case as a way to highlight the nature of the centripetal and centrifugal forces in the model.

4. Centripetal and centrifugal forces

To understand how this model works, it is useful to consider what would happen if there were no foreign trade, and within that special case to ask only a limited question: under what conditions is concentration of all population in either location 1 or 2 an equilibrium? Once we have seen this case, it will be easier to understand the results we get once the model is "opened up".

Consider, then, a situation in which ρ is very high, so that we can ignore the role of the rest of the world. And furthermore, let us consider the determination of relative real wages when almost all domestic labor is in region 1. If $\omega 2 < \omega 1$ in this case, then concentration of all labor in region 1 is an equilibrium; otherwise it is not.

We first note that the nominal wage paid at the center of city 2 must be less than that at the center of city 1. The reason is that almost all output from a firm in 2 must be sold in 1, and must therefore incur transportation cost. At the same time, the zero-profit output for firms is the same in each location. So goods produced at location 2 must have sufficiently lower f.o.b. prices to sell as much in 1's market as goods produced at 1 (note, however, that these sales include goods used up in transport; final sales to the consumer need be only $1/\tau$ times as large). But the f.o.b. price of goods is simply proportional to the local wage rate, so we must have

$$\frac{w_2}{w_1} = \tau^{(1-\sigma)/\sigma} \tag{19}$$

This wage premium at location 1, which results from its dominant role as a market, corresponds to our concept of backward linkage.

Next we notice that if almost all labor is in location 1, almost all goods consumed in 2 must be imported, implying a higher price of these goods:

$$\frac{T_2}{T_1} = \tau \tag{20}$$

If the wage rate is higher in 1 and the price of consumer goods lower, does this not mean that real wages must be higher in 1? No, against this we must set higher land rent and/or commuting cost. With almost all of the labor force L concentrated in 1, the most remote workers in 1 must commute a distance $L/2$, and all workers who live closer to the center must pay a land rent that absorbs any saving in commuting cost. Meanwhile the small number of workers in 2 pay almost no land rent and have essentially no commuting distance. So the real wage difference turns out to be

$$\frac{\omega_1}{\omega_2} = \tau^{(2\sigma-1)/\sigma}(1-\gamma L) \tag{21}$$

In this expression the first term represents the ''centripetal'' forces, the backward and forward linkages described in Eqs. (19) and (20), while the second term represents the ''centrifugal'' force of commuting cost/land rent.

5. Trade policy and population concentration

To solve the general model, we need to show how to determine equilibrium real wages for any given allocation of domestic labor between locations 1 and 2. Given these equilibrium real wages, we can then ask which allocations are stable. Finally, we ask how the possible equilibria depend on the openness of the economy, as measured by ρ.

As a first step, we ask how consumers in each location spend their income. For example, consider consumers in location 0. Let $p_{1,0}$ be the price of location 1 goods at location 0, etc.. Also, let $c_{1,0}$ be consumption of a typical good from 1 at 0. Then we must have

$$Y_0 = n_0 p_{0,0} c_{0,0} + n_1 p_{1,0} c_{1,0} + n_2 p_{2,0} c_{2,0} \tag{22}$$

where Y_0 is the location's complete income. But we also know that

$$c_{0,0} = c_{1,0}(p_{0,0}/p_{1,0})^{-\sigma} \tag{23}$$

and that

$$c_{2,0} = c_{1,0}(p_{2,0}/p_{1,0})^{-\sigma} \tag{24}$$

Putting these together, and substituting the definition of the true price index at location 0, we find

$$p_{1,0} c_{1,0} = Y_0 \left[\frac{p_{1,0}}{T_0}\right]^{1-\sigma} \tag{25}$$

Eq. (25) tells us the total expenditure of consumers at 0 on a typical good from 1. We can derive similar expressions for consumers at each of the other locations. But the total income of location 1 is simply the global expenditure on goods produced there:

$$w_1 Z_1 = n_1 \left[Y_0 \left(\frac{w_1}{T_0}\right)^{1-\sigma} + Y_1 \left(\frac{w_1}{T_1}\right)^{1-\sigma} + Y_2 \left(\frac{w_1 \tau}{T_2}\right)^{1-\sigma} \right] \tag{26}$$

or, substituting once again,

$$w_1 = \left[Y_0 T_0^{\sigma-1} + Y_1 T_1^{\sigma-1} + Y_2 (T_2/\tau)^{\sigma-1} \right]^{1/\sigma} \tag{27}$$

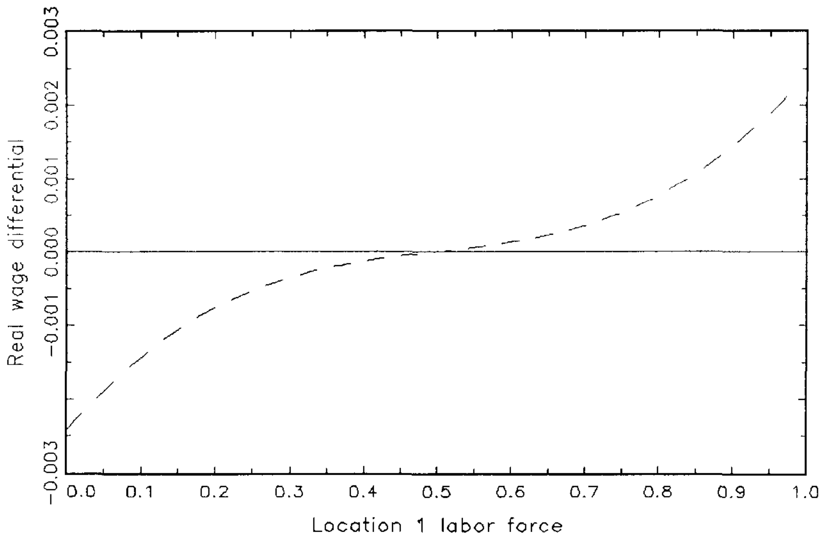

Fig. 1. Real wage differential $\omega_1 - \omega_2$ against the labor force in location 1, with rho = 1.83.

By similar reasoning, we also find that

$$w_2 = \left[Y_0 T_0^{\sigma-1} + Y_1 (T_1/\tau)^{\sigma-1} + Y_2 T_2^{\sigma-1} \right]^{1/\sigma} \tag{28}$$

We now have a system of equations that can be solved for any given allocation of labor between 1 and 2. Given such an allocation, we can determine z_j and hence n_j for each region. We can then simultaneously solve for income using (4), for the true price indices using (13)–(15), and for the wage rates in terms of the numeraire using (27) and (28). We can then use the true price indices to solve for real wage rates.

Unfortunately, even though the logic of this model is quite simple and the results we get will make intuitive sense, the model is too complicated to solve analytically. So at this point we are driven to numerical examples.

Several numerical examples are shown in Figs. 1–3. In each case, we plot the real wage differential $\omega_1 - \omega_2$ against the labor force in location 1. Any point where the wage differential is 0 is an equilibrium; such an equilibrium is stable if the schedule is downward-sloping, unstable if it is upward-sloping. There may also be corner equilibria: if all labor is concentrated in location 1, it will stay there if $\omega_1 > \omega_2$, and conversely.

In all three figures we assume $L = 1$, $\sigma = 4$, $\tau = 1.4$, $\gamma = 0.2$, $z_0 = 10$. What we vary is the "protection" parameter ρ. Our informal analysis suggests that a closed economy should be more likely to have population concentrated in one metropolis, so we consider what happens when ρ is gradually reduced through the critical range at which the qualitative behavior of the economy changes.

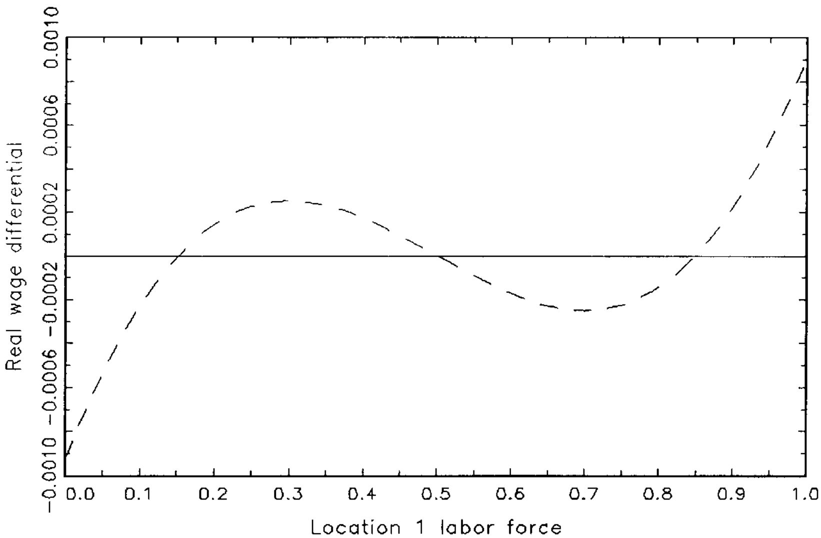

Fig. 2. As Fig. 1, but with rho = 1.81.

In Fig. 1, we have $\rho = 1.83$. The equilibrium in which population is evenly divided between the two locations is unstable, with the only stable allocations being concentration in one or the other location.

In Fig. 2, we show what happens when the economy is opened slightly,

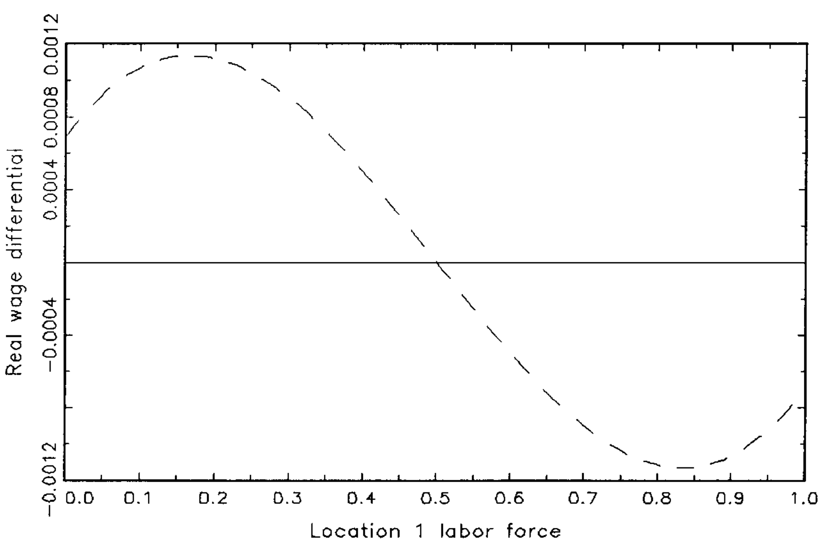

Fig. 3. As Fig. 1, but with rho = 1.79.

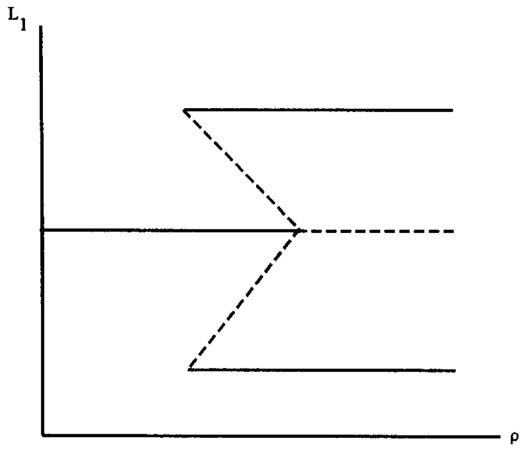

Fig. 4. A schematic summary of the way that the set of equilibria depends on the rate of protection. Plotted is the rate of protection against region 1's share of the labor force.

$\rho = 1.81$. An equal-division allocation is now stable. Concentration of population in either location is, however, stable as well. Between the stable equilibria lie unstable equilibria.

Finally, in Fig. 3 we show what happens when ρ is reduced to 1.79. We now have a unique, stable equilibrium in which population is evenly divided between the two locations.

These results confirm our intuition. In a relatively closed economy, the forward and backward linkages are strong enough to create and support a single large metropolis. As the economy is opened, these forces are weakened and the offsetting centrifugal forces make a less concentrated urban system first possible and then necessary.

Fig. 4 offers a schematic summary of the way that the set of equilibria depends on the rate of protection. On the horizontal axis is the rate of protection, on the vertical axis region 1's share of the labor force. Solid lines represent stable equilibria, dotted lines unstable equilibria. When protection is low, the only stable equilibrium is with dispersed production; when it is high, the only equilibrium is with all production concentrated in one or the other region. There is a range (which our numerical example suggests may be pretty narrow) in which both kinds of internal geography are possible.

How would the regional structure of the economy change as trade policy changes? A hypothetical sequence may help illustrate the principles, as well as providing a very stylized history of Mexico. Imagine that the economy initially has low protection, and that it gradually turns inward. At first the economy remains characterized by an equal division of the labor force between regions. Eventually, however, the circular logic of concentration takes over. Whichever region has a head start or small advantage snowballs in size, leading (say) to a situation in which everything is concentrated in region 1.

Now run it in reverse. Starting with a concentrated population, we imagine a process of liberalization. Initially this does not break up the concentration, but eventually there is no longer enough reliance on the domestic market to make the backward and forward linkages strong enough to support the concentration of production, and a cumulative unravelling process takes place.

This is just a particular numerical example, but it does confirm our intutitive argument. We see that a trade policy that closes off the domestic market can lead to the emergence of a central metropolis, while a policy of opening can lead that metropolis to lose its dominant position.

6. Conclusions

The trade policy of developing countries has been the subject of a huge theoretical and empirical literature. Urbanization, though hardly ignored, has not generated a comparable outpouring. In this paper we suggest not only that Third World urbanization is an important subject, but that there is a surprise linkage between trade policy and urban development: closed domestic markets have been a key factor in the emergence of the huge metropolises that dot the developing world.

This paper is only a theoretical exercise, although recent work by Ades and Glaeser (1994) offers at least mild support for its conclusions. Beyond inspiring empirical work, we hope that the paper will help to alert economists to the point that international trade theory and urban economics cannot, ultimately, be regarded as wholly separate disciplines.

References

Ades, A. and Glaeser, E., 1994, Trade and circuses: explaining urban giants, NBER Working Paper #4715.

Bairoch, P., 1988, Cities and Economic Development, Chicago: University of Chicago Press, 1988.

Dixit, A. and Stiglitz, J., 1977, Monopolistic competitions and optimum product diversity, American Economic Review, June.

Hanson, G., 1992, Industry Agglomeration and Trade in Mexico, MIT Ph.D. thesis.

Henderson, J.V., 1988, Urban Development, Oxford: Oxford University Press.

Hirschman, A., 1958, The Strategy of Economic Development, New Haven: Yale University Press.

Krugman, P., 1991, Increasing returns and economic geography, Journal of Political Economy, June.

Krugman, P., 1993a, First nature, second nature, and metropolitan location, Journal of Regional Science.

Krugman, P., 1993b, On the number and location of cities, European Economic Review, May.

Livas Eilzondo, R., 1992, Regional implications of international trade in Mexico, mimeo, MIT.

Williamson, J., 1988, Migration and urbanization, in Chenery, H. and Srinivasan, T.N., eds., Handbook of Development Economics (North Holland, New York, NY).