PART I: A STANDARD ANALYSIS OF FACTOR MOBILITY

CHAPTER 1: BENEFICIAL vs. DISTORTIONARY MOBILITY OF FACTORS OF PRODUCTION

Introduction

Classical economic setups suggest that factors of production move, if not constrained, from locations where their marginal product is low to other locations where their marginal product is high. In these setups, perfect competition with complete information prevails and there are no distortions (created by taxation, externalities, etc.), so that the private returns to the factor owners coincide with the social returns. Accordingly, factor mobility induced by private factor return differentials is beneficial both for the owners of the factors that actually move from one location to another and to the source and destination economies.

Factor mobility may wear two guises. First, there is the mobility of the factor of production itself, without the owner changing his national residence. Second, one can look at the mobility of the owner with his factor of production. The first kind of mobility is typical for capital. The phenomenon of guest workers can also be viewed as a factor mobility of the first kind. Guest workers are typically not eligible for all the amenities (especially in the area of social insurance) of the host country. In our new age of information technology, many professionals can provide their services via the internet, and other electronic means, without physically moving to the location where their services are received - the so-called "weightless trade".[1] The second type of mobility typically characterizes labor and is usually termed migration. It raises a host of issues and considerations associated with the welfare state which are not relevant for factor mobility of the first kind: unemployment insurance, pensions, health insurance and care, education, etc.

The possibility of separation between the mobility of the factor of production and the mobility of his owner underscores also the distinction between the Gross Domestic Product (GDP) of a country and its Gross National Product (GNP). The first term includes all the value added (or income) which is produced in the country in question. The second term subtracts from the GDP of a country the value added produced in this country by factors of production owned by foreigners and adds to it the value added produced abroad by factors of production owned by residents of this country. Factor mobility of the first kind affects primarily the GDP of the host and the source countries, while factor mobility of the second type affects both their GDP and GNP. When analyzing the welfare consequences of factor mobility, it is the GNP which is relevant as it can serve as a macro-economic proxy for welfare. For instance, it may be argued that the heavy subsidization of FDI in Ireland in the last two decades resulted in impressive GDP growth rates but with much less pronounced effect on the well-being of Irish residents, as proxied by the GNP growth rates.[2]

In this part we focus our attention on factor mobility without factor owner mobility. (Successive parts of the book deal with factor mobility of the second kind.)

One-Good Case

Suppose first that there is only one final good. Therefore, in this case there is no scope for trade in goods, (i.e. of one good for another) as in standard trade models. But a person residing in one country who provides the services of the factor of production that she owns in the other country, can still retract the remuneration accruing to that factor in the other country to her own country.

The welfare impact of factor mobility can be neatly presented with the aid of the familiar scissors diagram (Figure 1.1) in which the marginal product of a mobile factor (say, capital), for two countries (home and foreign) that comprise the world economy, are depicted originating at opposite ends. Following MacDougall (1960), suppose that originally the world allocation of capital is at A, with the home country having a higher marginal product of capital than the foreign country. If capital flows from the foreign country to the home country up until the point at which the marginal product of capital is the same in the two countries, bringing the world allocation of capital to point E, then the world output is at a maximum.

In a laissez-faire, competitive environment with complete information and no barriers to factor mobility, an amount of AE units of capital will indeed flow from the foreign country to the home country. This is because in the aforementioned classical setup the market return to capital is equal to its marginal product, so that it will pay the owners of capital in the foreign country to invest the amount of AE units of it in the home country. Furthermore, not only world output (namely, the sum of the home and foreign GNP) rises, but the GNP of each country also rises as well: The GNP of the home country rises from $O_{\mathsf{H}}MKA$ to $O_{\mathsf{H}}MRQA$[3] and the GNP of the foreign country rises from $O_{\mathsf{F}}NSA$ to $O_{\mathsf{F}}NRQA$,[4] so that world output rises by KSR.

Trade in Goods

In this section we extend the preceding analysis to the standard trade models in which some of the goods are traded for other goods. Specifically, there are traded and nontraded goods. This setup guarantees that in the absence of factor mobility, primary inputs will be differently priced internationally (not only in the absence of trade in goods which is obvious), but also when trade in goods takes place (for the same reasons that are advanced in the standard trade models; see, for example, Jones (1967) and the next chapter).

Following Helpman and Razin (1983), we present in this section an analysis of welfare gains from factor movements for a small competitive economy with a constant returns-to-scale technology. For simplicity, we aggregate all traded goods into a single commodity Y and choose $p_{\mathsf{Y}} = 1$ as its price. The aggregation is based on the assumption that relative prices of traded goods do not change as a result of factor movements (the small country assumption in commodity markets), so that we can abstract from welfare changes that result from adjustments in the terms of trade of goods. We also assume that there is a single nontraded good X whose price in terms of Y is p.

Assuming the existence of a representative consumer, or a social welfare function which is maximized with costless income redistribution, our country's welfare level can be represented by an indirect utility function $v(p, GNP)$, where GNP stands for gross national product (or income) measured in units of Y. Assuming that all foreign source income stems from international mobility of capital, GNP equals GDP minus rental payments on domestically employed foreign capital. Hence:

$$GNP = GDP(p, L, K + \Delta) - \rho\Delta, \tag{1.1}$$

where $GDP(\cdot)$ stands for the gross domestic product function, L and K stand for domestically owned labor and capital (assumed to be inelastically supplied), Δ stands for foreign capital employed in the home country when $\Delta > 0$, and domestic capital employed abroad when $\Delta < 0$. Finally, ρ represents the rental rate on Δ. Note that GDP depends both on the exogenously given stocks of primary factors (L and $K + \Delta$) and the relative price p of nontraded goods which guides the intersectoral allocation of K and L. In fact, the function $GDP(\cdot)$ is a restricted revenue function in the sense that the total inputs of labor (L) and capital ($K + \Delta$) to the two industries are exogenously given.

Employing the envelope theorem, one can derive the partial derivatives of the restricted profit function $GDP(\cdot)$:

$$\frac{\partial GDP(p, L, K + \Delta)}{\partial p} = X \tag{1.2a}$$

$$\frac{\partial GDP(p, L, K + \Delta)}{\partial L} = w \tag{1.2b}$$

$$\frac{\partial GDP(p, L, K + \Delta)}{\partial (K + \Delta)} = r \tag{1.2c}$$

where w and r are the domestic wage rate and the domestic rental price of capital.

When foreign capital is employed in the home country, then $\rho = r$; the assumption is that foreign-owned capital commands the same rental rate as domestic-owned capital. On the other hand, when domestic capital is employed abroad its rental rate in the foreign country is ρ, which may or may not be a function of the size of investment abroad.

Choosing a transformation of the utility function such that in equilibrium the marginal utility of income (i.e., $\partial v / \partial GNP$) equals one, differentiation of $u = v(\cdot)$, using (1.1) and the properties of the indirect utility and GDP functions, yields:

$$dU = \frac{\partial v}{\partial p} dp + \frac{\partial v}{\partial (GNP)} \left[\frac{\partial (GDP)}{\partial p} dp + \frac{\partial (GDP)}{\partial (K + \Delta)} d\Delta - \rho d\Delta - \Delta d\rho \right]. \tag{1.3}$$

Employing Roy's identity, we find that

$$\frac{\partial v}{\partial p} = - \frac{\partial v}{\partial (GNP)} D_{\mathsf{X}}, \tag{1.4}$$

where D_{X} is the consumption of good X. Substituting (1.4), (1.2a) and (1.2c) into equation (1.3) yields:

$$dU = \quad (r-\rho)d\Delta + (X - D_{\mathsf{X}})dp - \Delta d\rho. \tag{1.5}$$

.

(Recall that $\partial v/\partial(GNP) = 1$.)

Since X is not traded, in equilibrium $X = D_{\mathsf{X}}$, and equation 1.5 reduces to:

$$dU = (r-\rho)d\Delta - \Delta d\rho. \tag{1.6}$$

Suppose that r is smaller than the rental rate that domestic capital can obtain abroad. Then owners of domestic capital will shift part of it into foreign operations, thereby increasing domestic welfare due to the first term on the right-hand-side of (1.6) (since $r < \rho$ and $d\Delta < 0$). If the foreign rental rate is unaffected by the home country's investment abroad, the second term on the right-hand-side of (1.6) equals zero (because $d\rho = 0$). In this case domestic welfare unambiguously rises ($dU > 0$). If, on the other hand, the foreign rental rate on domestic capital invested broad declines with the size of the investment and we start with a positive investment level ($\Delta < 0$), the second term generates a negative welfare effect, but this negative welfare effect is negligible for small investment levels. In the case under discussion, dU evaluated at $\Delta = 0$ is positive, so that it pays to invest abroad, at least a little. (The negative welfare effect (which does not exist at $\Delta = 0$) stems from possible market power in foreign investment.)

Now suppose that r exceeds the rental rate that foreign capital receives abroad. Then foreigners will invest in the home country, earning the domestic rental rate of r. Thus, $r \equiv \rho$ in this case, and (1.6) reduces to $dU = -\Delta dr$. However, due to diminishing marginal product of capital, the rental rate on capital declines with capital inflows so that for positive investment levels ($\Delta > 0$) welfare increases.

This analysis illustrates again the two points made in the one-good case. First, capital mobility can raise welfare. Second, in a competitive, distortion-free environment (absent terms-of-trade effects) private considerations about the location of capital guided by the differential between r and ρ coincide with social considerations in the sense that social welfare increases as a result of private decisions to shift capital from the low return to the high return location.

Factor Mobility in the Presence of Distortions

It should be emphasized that the social benefit generated by laissez-faire capital mobility which was demonstrated in the preceding two sections holds in a classical setup, where perfect competition with complete information and no other distortions (such as taxes) prevail. As a matter of fact, with imperfectly competitive markets and/or tax distortions, factor mobility may be harmful. Such deviations from the classical setups are more likely to occur in the factor markets rather than in the goods markets. Labor markets are particularly notorious for their imperfections, due to unionism, state regulation (e.g. minimum wage laws), incomplete information about job availability and workers' characteristics, and relatively heavy payroll taxes. Similarly, capital markets are also plagued by imperfect information phenomena manifested by severe moral hazards, adverse selection, debt and bank runs, herd behaviour, etc. On the other hand, goods are typically homogenous and information about them is transparent; trade enhances competition in their markets and indirect taxes tend to be uniform (e.g. VAT), thereby avoiding inter-commodity distortions.[5]

It is quite straightforward to show that laissez-faire factor mobility can be harmful in distortive environments. We shall consider two deviations from the classical setup: taxes and noncompetitive markets.

Taxes

It is most convenient to employ the one-good case depicted in Figure 1.2 (which is drawn on Figure 1.1) in order to analyze the effects of taxes.

Recall that in the aforementioned case, social welfare increases when a well specified quantity of capital (i.e. AE) moves from the foreign to the domestic economy. However, in a distorted environment, social welfare can fall either when too much capital flows in, or when the flow of capital is reversed. For instance, suppose that the home and the foreign countries levy source-based taxes at rates τ and τ^*, respectively, on the income from capital that accrues in their jurisdictions.[6]

For instance, suppose that the foreign country levies a source-based tax at the rate of τ^* which is equal to VG/VF, while the home country levies no tax (i.e. $\tau = 0$). In this case the schedule of after-tax return to capital in the foreign country falls from NP to $N'P'$ and a quantity of AF units of capital flows from the foreign country to the home country. World output changes from $O_\mathsf{H}MKSNO_\mathsf{F}$ to $O_\mathsf{H}MGVNO_\mathsf{F}$, amounting to a decline of RVG, minus KSR in world output. (Of course, a similar allocation can be achieved with both countries levying a tax, but with the foreign country levying a higher rate (that is, $\tau^* > \tau > 0$).) Similarly, if the home country levies a much higher tax rate than the foreign country, then the direction of capital flows may be reversed, causing a decline in world output.

Increasing Returns

Another possibility of welfare-reducing factor mobility can occur when there are increasing returns, forcing prices to deviate from marginal costs. To illustrate this, we revert to the two-good model described earlier in this chapter, while introducing to the model increasing returns-to-scale. We continue to assume that the traded good (Y) is produced with a constant-returns-to-scale technology. However, the nontraded good (X) is now a composite good of symmetric differential products, produced by N firms, each possessing increasing returns-to-scale technology. All of the N firms which produce X possess identical technologies. Thus, they charge the same price, and due to free entry that entails zero profits, they engage in average cost pricing. The free entry assumption pins down analytically the number of firms (N) in the industry, as always. Good Y is produced with constant returns-to-scale technology in which the unit (average) cost, denoted by $c_{\mathsf{Y}}(w, r)$ is constant, unrelated to the scale of production. Therefore, at equilibrium in the market for Y :

$$1 = c_{\mathsf{Y}}(w, r) \tag{1.7}$$

(recalling that Y is the numeraire whose price is set to unity). However, the average cost of production of X decreases with the scale of production. Denoting by x the scale of production of each firm in the X industry, we conclude that at equilibrium in the market for X :

$$p = c_{\mathsf{X}}(w, r; x) \tag{1.8}$$

where $c_{\mathsf{X}}(w, r; x)$ is the average cost of producing x units of X.

Similarly, the unit capital and labor requirements in industry Y (a_{KY} and a_{LY}, respectively) are independent of the scale of production; whereas the unit capital and labor requirements in industry X (a_{KX} and a_{LX}, respectively) are assumed to be declining in the scale of production, x, due to increasing returns-to-scale. Thus, equilibrium in the factor markets requires that:

$$a_{\mathsf{LY}}(w, r)Y + a_{\mathsf{LX}}(w, r; x)X = L, \tag{1.9}$$

$$a_{\mathsf{KY}}(w, r)Y + a_{\mathsf{KX}}(w, r; x)X = K + \Delta, \tag{1.10}$$

where $X = Nx$.

Given x, the restricted revenue function which depends, as above, also on p, L and $K + \Delta$, is essentially equal to $GDP(p, L, K + \Delta; x)$. The partial derivatives of this function can be found as before, by the envelope theorem:

$$\partial GDP(p, L, K + \Delta; x)/\partial p = X = Nx, \tag{1.11a}$$

$$\partial GDP(p, L, K + \Delta; x)/\partial L = w, \tag{1.11b}$$

$$\partial GDP(p, L, K + \Delta; x)/\partial (K + \Delta) = r. \tag{1.11c}$$

As was already mentioned, the difference between this GDP function and that used earlier in this chapter is the dependence of the present one on x, the individual firm's output level (or, alternatively, its scale of production).

One can observe from the set of equilibrium equations (1.7)-(1.10) that our non-constant returns-to-scale economy is formally similar to a constant returns-to-scale economy with a technical progress coefficient. An increase in x reduces average costs $c_{\mathsf{X}}(\cdot)$, because the elasticity of $c_{\mathsf{X}}(\cdot)$ with respect to x is negative:

$$b'(w, r, x) \equiv \frac{\partial c_{\mathsf{X}}}{\partial x}\frac{x}{c_{\mathsf{X}}} = -1 + \varphi(w, r; x) < 0, \tag{1.12}$$

where $\varphi(w, r; x)$ is the elasticity of total cost (namely, xc_{X}) with the respect to the firm's output (namely, x).

Due to the increasing returns-to-scale, φ must be smaller than one, so that $b' < 0$. The absolute value of b', denoted by b, is

$$b(w, r; x) = -\varphi(w, r; x) + 1 > 0. \tag{1.13}$$

We can now adopt the familiar analysis of technical progress in the standard model of two factors, two goods and constant returns-to-scale technologies developed by Jones (1965). He showed that

$$b = \theta_{LX} b_L + \theta_{KX} b_K, \tag{1.14}$$

where b_L is the absolute value of the elasticity of $a_{LX}(\cdot)$ with respect to x, b_K is the absolute value of the elasticity of $a_{KX}(\cdot)$ with respect to x, and θ_{jX} is the share of factor j in costs of production; $j = L, K$. As Jones (1965) shows, a one percentage point increase in x has the same effect on output levels as a b percent increase in the price p, plus a $\lambda_{LX} b_L$ percent increase in the labor force, plus a $\lambda_{KX} b_K$ percent increase in the capital stock, where λ_{LX} is the share of labor employed in the production of X and λ_{KX} is the share of the capital stock employed in the production of X. This can be explained as follows. Suppose x is increased by a one percentage point and the number of firms N is reduced by a one percentage point, so that aggregate output in sector X (namely, Nx) does not change. As a result of the increase in x, each firm will increase its employment of labor by ε_{Lx} percent, where ε_{Lx} is its elasticity of labor demand with respect to output, so that the sector's demand for labor will increase by ε_{Lx} percent. On the other hand, due to the decline in the number of firms in the industry, the industry's labor demand will fall by 1 percent, so that $b_L \equiv 1 - \varepsilon_{Lx}$ is the proportion of the industry's labor force that is being released as a result of these changes. Since the industry employs the proportion λ_{LX} of the total labor force, $\lambda_{LX}b_L$ is the industry's saving of labor as a proportion of the total labor force. Similarly, $\lambda_{KX}b_K$ is the proportion of total capital saved by industry X as a result of a one percent increase in x, holding aggregate output of good X constant (with the adjustment being made by means of a decline in the number of firms in the industry). In addition to these factor supply effects, a one percentage point increase in x reduces unit production costs by b percent.

Using the above-described elasticity relationship between the effects on output levels of a one percentage point increase in x, and b percent increase in the price of p, plus $\lambda_{jX}b_j$ $(j = L, K)$, percent increases in the supply of factors of production, we can calculate the change in GDP as a result of a one percentage point increase in x as follows:

$$\begin{aligned} \frac{\partial GDP}{\partial x}x &= \left(p\frac{\partial X}{\partial p} + \frac{\partial Y}{\partial p}\right) pb + \left(p\frac{\partial X}{\partial L} + \frac{\partial Y}{\partial L}\right) L\lambda_{LX}b_L + \left(p\frac{\partial X}{\partial K} + \frac{\partial Y}{\partial K}\right) K\lambda_{KX}b_K \\ &= w_L\lambda_{LX}b_L + r_K\lambda_{KX}b_X, \end{aligned}$$

due to the envelope theorem (recall that $GDP = pX + Y$).[7] Hence, using the definition of λ_{jX} $(j = L, K)$, we obtain:

$$\frac{\partial GDP}{\partial x}x = wa_{\mathsf{LX}}Xb_{\mathsf{L}} + ra_{\mathsf{KX}}Xb_{\mathsf{K}} = pX(\theta_{\mathsf{LX}}b_{\mathsf{L}} + \theta_{\mathsf{KX}}b_{\mathsf{K}}) = pXb$$

and

$$\frac{\partial}{\partial x}GDP(p, L, K + \Delta; x) = pN(1 - \varphi), \tag{1.15}$$

where use has been made of the relationships $X = Nx$ and $b = (1 - \varphi)$.

Now define $\bar{r}$ as the increase in GDP that results from an increase in Δ, holding p constant. In the competitive case with constant returns-to-scale technologies analyzed earlier in this chapter, this was shown to equal r - the market rental rate on capital. In the case considered, however, $\bar{r}$ is given by:

$$\bar{r} = \frac{\partial}{\partial(K + \Delta)}GDP(\cdot) + \frac{\partial}{\partial x}GDP(\cdot)\frac{dx}{d\Delta}.$$

Using (1.15) this can be written as:

$$\bar{r} = r + pN(1 - \varphi)\frac{dx}{d\Delta}. \tag{1.16}$$

Since $\varphi < 1$ (due to economies of scale at the firm level), equation (1.16) tells us that an inflow of one unit of capital will increase GDP by more than the market rental rate on capital if it brings about an expansion of every firm's output level in sector X. An inflow of the same amount of capital will increase GDP by less than the market rental rate on capital, or even reduce GDP, if it brings about a contraction of every firm's output level in sector X.

In more general terms, this means that the private sector may undervalue or overvalue the marginal productivity of capital (and that of labor) as far as GDP valuation is concerned, depending on the marginal effect of capital inflows on the size of operation of firms in the sector with economies of scale:

$$\bar{r} \gtreqless r \text{ as } \frac{dx}{d\Delta} \gtreqless 0. \tag{1.17}$$

That is, the private sector undervalues (respectively, overvalues) the effect of capital inflows on GDP when such inflows increase (respectively, decrease) the scale of production. Therefore, such a biased evaluation by the market of the effect of capital inflows on GNP and welfare can occur as well.[8]

For a complete welfare analysis of the effect of a capital inflow by the amount Δ, one has to fill the above model with a formal specification of consumer preferences and the market organization of the increasing returns (nontraded) industry (e.g., imperfect competition with differentiated production). Such specification enables an explicit solution for the scale of production (x and, consequently, N) and the relative price p of the nontraded good, as described, for intance, in Helpman and Razin (1983).

Conclusion

We show that in a classical setup with perfect competition everywhere, no distortion and full information, the private return to a factor of production coincides with the social return. In such a setup, factor mobility induced by international differentials in factor returns is beneficial both for the factor owner and the source and destination countries, provided that distortion-free redistribution is made within each country.

We also show that once we deviate from the classical setup - for instance, through distortive taxes, market imperfections, etc. - private and social returns to factor of production may diverge and factor mobility may reduce welfare. Furthermore, when countries experience some market power, factor mobility may change the terms-of-trade. In this case, even when perfect competition and distortion-free environments prevail in each country, factor mobility can be detrimental to welfare. In the familiar Dornbusch-Fischer-Samuelson (1977) Ricardian model, a country receiving migrants may find its terms-of-trade deteriorating to such an extent so as to immeserise all the native-born. Naturally, these adverse terms-of-trade effects never occur to a small country.