The effect of taxes and subsidies on the location of economic activity and welfare

Second Draft

Egle Tafenau*

1 Introduction

Regional science has been a marginal field in economics. Only since the beginning of 1990s - since the seminal papers of Paul Krugman (1991a, 1991b) - the interest in the theory of location of economic activities has risen. The lack of interest was largely caused by the impossibility to deal with the locational issues analytically, as circular causality - firms move to agglomerations as this enables more efficient production; as the number of firms in the agglomeration increases, additional motivation for the rest of the firms to move there is created - causes lots of complications. In explaining why cities emerge and economic activity tends to agglomerate, the existence of increasing returns has an important role. Until the model of imperfectly competitive economy by Dixit & Stiglitz (1977), it was not possible to include this aspect into the economic models explicitly.

New economic geography (NEG) follows the lines of the new trade theories and new growth theories in assuming imperfect competition, increasing returns and applying usually the structure of economy proposed by Dixit & Stiglitz (1977). The modeling framework is general equilibrium, based on the optimization decisions of individual agents. The most important outcome of the NEG models is that even regions that are initially identical in their factor endowments may end up having very different production structures.

The first NEG models were not analytically solvable (e.g. Krugman (1991a,b); also the models presented in the book by Fujita et al. (1999) who show the possibilities of using the NEG approach for discussing various issues in regional, international and urban economics) - in their conclusions and results they relied on numerical examples.

*Ph.D. student at the University of Tartu (Estonia) and Christian-Albrechts-Universität zu Kiel (Germany). Contact: Egle.Tafenau@ut.ee. Financial support from the Estonian Science Foundation (research grant Nr. 6475) is gratefully acknowledged.
It was not possible to show explicitly how the regional distribution of economic activity depends on e.g. trade costs, initial distribution of economic activity or the endowments of production factors and therefore it was also not possible to carry out an explicit policy analysis. In the second half of the 1990s also analytically solvable models like the so-called footloose capital model, footloose entrepreneur model and linear models were developed (for a detailed overview see e.g. Baldwin et al. (2003)). Based on these models, policy implications have been analyzed. Neary (2001) has suggested that the NEG can be most useful exactly for analyzing policy implications. Nevertheless, so far the analysis in this respect has been relatively moderate.

The purpose of the paper is to show how different taxing schemes might influence the location of economic activity and welfare. We do it in the context of the footloose capital model (Martin & Rogers 1995). We compare the effects of income and consumption taxes, with possibly asymmetric tax rates in the two regions or at different income sources.

First we give a short overview of the ideas behind the NEG models. After that the possibilities of policy analysis will be introduced. Then we present the effects on the distribution of economic activity and income inequality of introducing different tax schemes into the footloose capital model. Thereafter we compare the welfare effects of the different tax systems. The final section concludes.

2 Basic models of NEG

NEG models are not the only possible way to discuss the regional development and policy issues, nevertheless, they are the only models aiming directly to address regional economic issues relying explicitly on microfoundations. Regional issues have been analyzed also in the context of models that are developed for national economies (e.g. the Keynesian multiplier model of incomes and expenditures, Leontief-type input-output models, growth models; for an overview see Armstrong & Taylor (2000)). Nevertheless, during the last 15 years the NEG type models have become the basis for the analysis of the location of economic activity and the impact on regional development goals (see e.g. Baldwin et al. 2003).

The NEG literature started with Krugman’s (1991b) paper. The NEG models use similar modeling techniques as the new growth theory or new trade theory, relying on the Dixit & Stiglitz (1977) type imperfectly competitive economy with increasing returns to scale in production. In the heart of the NEG models stands the interaction between centrifugal and centripetal forces, which creates the circular agglomeration process.

In the NEG models it is usually assumed that there are two sectors in the economy, one of which has increasing returns to scale due to some fixed input requirement (so-called manufacturing or modern sector) and the other with constant returns to scale.
(producing agricultural or traditional goods). There are also two production factors, these could be different types of labour, one of which is mobile between the regions and needed in the modern sector and the other is immobile and can be employed only in the traditional sector (e.g. in Krugman 1991b); labour and capital as in Martin & Rogers (1995) (so-called footloose capital model), in which case labour is mobile between the sectors, but not between regions, the inter-regionally mobile factor is physical capital, but it moves without its owner; or labour and human capital (Forslid & Ottaviano 2003, so-called footloose entrepreneur model) with human capital being the factor mobile between the regions and labour having to stay in the region of origin, but being able to move from one sector to the other.

Most of the NEG models rely on the Dixit & Stiglitz (1977) model of monopolistically competitive economy, assume iceberg type trade costs and CES-type utility functions (e.g. Krugman 1991b, Martin & Rogers 1995, Venables 1996, Krugman & Venables 1995, Puga 1999, Forslid & Ottaviano 2003) and this is one reason why NEG models have been strongly criticized - they seem to rely on very specific assumptions. Nevertheless, Ottaviano et al. (2002) have shown that when using quasi-linear quadratic utility functions and that transport of a manufactured good needs to be covered with the good of the CRS sector, gives qualitatively same results.

The main result of the NEG models is that regions which are originally identical might develop to have completely different industrial structure. The main factor of interest is the share of modern firms in each region. Most of the models come to the result that for very high trade costs the symmetric outcome is the only stable equilibrium, but if the trade costs decrease, agglomeration in one of the regions will be the outcome of the market forces. Which region gets the so-called industrial core depends on “accidents” or expectations. The outcomes of the NEG models are discussed in the context of increasing integration of the regions (or decreasing trade costs between the regions).

Baldwin et al. (2003, pp. 34-36) distinguish seven key features of the core-periphery model.

- Home market effect and home market magnification: in answer to an exogenous change in the location of demand, the industry relocates more than proportionally to the enlarged region. The home market magnification means that the home market effect is the stronger the freer is trade between the regions.

- Circular causality: agglomeration forces are self-reinforcing - relocation of some industry motivates also other firms to relocate to the same region.

- Endogenous asymmetry: if the trade costs decrease progressively, the initially even distribution of economic activity (symmetric regions) will change to asymmetric distribution.
• Catastrophic agglomeration: there is a critical level of trade costs at which a very small reduction of them leads the symmetric regions to reorganize into core-periphery pattern, if a shock disturbs the symmetric equilibrium. At high trade costs such a small decrease would have no impact on the spatial distribution of industry.

• Locational hysteresis: for the intermediate trade costs where both the symmetric and agglomerated equilibriums are sustainable, if a shock induces a change in the spatial pattern of economic activity, the economy does not return to its initial equilibrium if the shock is removed - a temporary shock has permanent consequences.

• Hump-shaped agglomeration rents: if the economy is organized as a core and a periphery, the mobile factor is usually not indifferent to location, as it would lose income by moving from the core to the periphery. These agglomeration rents first rise and then fall in answer to the proceeding reduction of trade costs, in the range of the trade costs where the full agglomeration outcome is sustainable.

• The overlap and self-fulfilling expectations: there exists a range of trade costs where both the symmetric and core-periphery outcomes are locally stable long-run equilibria. If there is a change in expectations about which equilibrium will be the outcome in the future, a jump between the symmetric outcome and a full agglomeration outcome is possible.

3 Policy in the NEG models

Ottaviano (2003), when discussing the key policy implications of the NEG models, mentions that policies not directly aimed to influence regional economic activity patterns might nevertheless have an effect on these. There have been several trials to introduce policy measures and analyze their effects in the NEG models (e.g. see Baldwin et al. 2003), with attention on different issues like for example tax competition, political economy, infrastructure policies or regional subsidies.

The current paper is most tightly in the lines of Dupont & Martin (2006). They discuss the location and welfare effects of capital and employment subsidies, financed by a local or global income tax. In the current paper we analyze analogously the subsidy to profits, but assume another kinds of taxes. First the case of different tax rates on labour and capital incomes are discussed, after that we introduce value added taxes. The analysis is based on the footloose capital model.
3.1 The footloose capital model

The footloose capital model uses the following assumptions.\(^1\)

The economy consists of two regions, region \(A\) and region \(B\). There are two sectors in the economy, the traditional sector (indexed by \(T\)) producing a homogeneous good and the modern (or manufacturing) sector (index \(M\)), producing \(n^w\) varieties. There are also two production factors, labour and capital. Labour is immobile between regions, whereas capital can flow freely between them. It is assumed that the capital moves without its owners, therefore the capital income is repatriated.

The consumers consume both modern and traditional goods and their utility is given by

\[
U = C_M^\mu C_T^{1-\mu}, \quad C_M = \left(\int_{i=1}^{n^w} c_i^{1-1/\sigma} di\right)^{1/(1-1/\sigma)}, \quad 0 < \mu < 1 < \sigma
\]

where \(C_T\) is the consumption of the traditional goods and \(C_M\) is the CES index of the modern varieties, with \(c_i\) being the demand for the variety \(i\) and \(\sigma\) the elasticity of substitution. \(\mu\) is the share of expenditures spent on the modern goods.

The homogeneous good is produced with a constant returns to scale technology, using labour as the only input. Its units are chosen such that the amount of output is equal to the labour input. As the result of this assumption, the wages are equal to the price of the traditional good, which is chosen to be the numeraire: \(p_T = w_T = 1\). The homogeneous good is traded without costs across the regions and therefore also the prices of the homogeneous good and the wages have to be equal in the two regions.\(^2\)

Production in the modern sector incurs increasing returns to scale: there is some fixed cost in producing each variety. This has to be covered with capital input, which requires return \(\pi\). We choose units of capital such that the capital input needed for producing a variety is equal to unity. This implies that the total number of varieties is equal to the world stock of capital: \(n^w = K^w\). The variable cost is associated with labour input, specifically it is needed \(a_M\) units of labour per a unit of output. As the labour can move freely between the sectors, the wages in both sectors equalize. Using these assumptions, the total cost function of a manufacturing firm is

\[
TC = \pi + a_M x,
\]

where \(x\) is the output of a typical modern firm.

Due to the scale effects each variety is produced exactly by one firm and it is assumed that each firm produces only one variety. The trade of modern goods from one region to the other is costly. It is assumed that the trade costs are so-called iceberg-type: in

---

\(^1\)We follow here the notation of Dupont & Martin (2006).

\(^2\)We also have to assume that none of the regions is large enough to satisfy alone the aggregate demand for the traditional good.
order to supply \( x \) units of the good in the other region, \( \tau x \) units have to be shipped, with \( \tau > 1 \).

In the equilibrium the region \( A \)’s aggregate demand for the homogeneous good is \( C_T = (1 - \mu)E \), where \( E \) is the disposable income in region \( A \). The region \( A \)’s demand for each manufacturing variety is given by

\[
c_j = \frac{p_j^{-\sigma} \mu E}{\int_{i=0}^{nu} pi^{1-\sigma} di}, \quad E = \pi K + L. \tag{3}
\]

Here there is no taxation assumed. If income taxes are added into the model, the equation for expenditure changes and the expenditure entering other equations is disposable income. In case of a value added tax, the prices should be interpreted as after-tax prices (consumer prices).

Under Dixit-Stiglitz monopolistic competition and the assumed utility function the firms set the prices of modern goods at a constant markup over the marginal cost: \( p = a_M/(1 - 1/\sigma) \). The foreign prices have to cover also transportation costs, therefore \( p^* = \tau p \). Moreover, the Dixit-Stiglitz monopolistic competition implies that the operating profit of a manufacturing firm is the value of sales divided by the elasticity of substitution: \( \pi = px/\sigma \). Using the demand function and the monopolistic prices, the equilibrium profits can be written as

\[
\begin{align*}
\pi &= b \frac{E^w}{K^w} \left[ \frac{s_E}{s_n + \phi (1 - s_n)} + \frac{\phi (1 - s_E)}{s_n \phi + 1 - s_n} \right]; \quad b \equiv \frac{\mu}{\sigma} < 1; \quad \phi \equiv \tau^{1-\sigma} \\
\pi^* &= b \frac{E^w}{K^w} \left[ \frac{\phi s_E}{s_n + \phi (1 - s_n)} + \frac{1 - s_E}{s_n \phi + 1 - s_n} \right],
\end{align*}
\tag{4}
\]

where \( s_n \) is the region \( A \)’s share of industry and \( s_E \) the region \( A \)’s share of expenditure. \( E^w \) and \( K^w \) are the national expenditure and national capital, respectively.\(^3\)

For the general equilibrium the following conditions have to be fulfilled in case of the footloose capital model without policy:

\[
s_E = (1 - b)s_L + b s_K, \tag{5}
\]

\[
\pi = \pi^* = b \frac{E^w}{K^w}, \tag{6}
\]

\[
s_n = \frac{1}{2} + \left( \frac{1 + \phi}{1 - \phi} \right) \left( s_E - \frac{1}{2} \right). \tag{7}
\]

It is assumed that the location of firms is given in the short run, thus, the equation for the equality of profits in both regions (equation 6) and the equation for the share of firms in region \( A \) (equation 7), determined form the condition for the equality of profits in the two regions, have to be fulfilled only in the long run. The equation for

\(^3\)Usually in the NEG literature the subscript \( w \) refers to the world, but in the current context we interpret the world as consisting of one country.
the region A’s share of expenditures (5)—derived from the definition of the share of expenditure—has to be valid always, both in the short and the long run.

The firms (capital) move from one region to the other whenever the nominal profits in the other region are higher than in the initial region. The price level does not matter as the capital income is repatriated.

3.2 Introducing policy into the footloose capital model

In order to introduce policy into the footloose capital model, some equilibrium conditions have to be added. Policy measures mean redistribution of money—money is collected from some economic agents and given to some others. This introduces distortions into the economy. We assume first that the government budget must be balanced: the whole tax revenue is paid out as subsidies. Second, it has to be guaranteed that factor markets stay in equilibrium after introducing policy. This is no problem for the case of capital as the assumption that the number of firms depends on the amount of available capital is set already in the basic model. For the labour market the equilibrium condition is given by equation (8):

\[ L^w = E^w(1 - b) \text{ with } E^w = L^w + \pi K^w, \]  

where the condition that the profits (including subsidies) have to be equal in the two regions has been employed. If income taxes are introduced, the national expenditure has to be replaced by the country’s disposable income.

4 Policy analysis

4.1 Subsidizing profits with revenues from taxes on labour and capital income

If the profits in region B are subsidized—also like in Dupont & Martin (2006)—it has to be that \( \pi = (1 + z^*)\pi^* \). Using this condition, equations (4), and taking the region A’s expenditures as given, it is possible to solve for the share of firms operating in the region A:

\[ s_n = \frac{s_E(1 - \phi^2) - \phi(1 + z^* - \phi)}{(1 - \phi)[1 + z^* - \phi - z^*s_E(1 + \phi)]}. \]  

(9)

On the taxation side, we assume that the labour and capital incomes are possibly taxed at different rates, \( t_L = \rho t_K \), where \( t_L \) and \( t_K \) are the labour and capital income tax rates, respectively, and \( \rho \geq 0 \) is the factor determining their ratio. From this follows that the economy-wide disposable income is

\[ E^w_d = (1 - \rho t_K)L^w + (1 - t_K)\pi K^w. \]  

(10)
This means that the government’s budget constraint is

\[ \rho t_K L^w + t_K \pi K^w = z^* (1 - s_n) \pi^* K^w = \frac{z^*}{1 + z^*} (1 - s_n) \pi K^w. \quad (11) \]

The expenditure share of region A is now

\[ s_E = \frac{(1 - \rho t_K) s_L L^w + (1 - t_K) \pi s_K K^w}{(1 - \rho t_K) L^w + (1 - t_K) \pi K^w}, \quad (12) \]

where \( s_L \) and \( s_K \) are the region A’s share of labour and capital owners, respectively.

The labour market equilibrium condition is

\[ L^w = ((1 - \rho t_K) L^w + (1 - t_K) \pi K^w) (1 - b). \quad (13) \]

This equation says that if the income is taxed, then the pre-tax profits have to rise in order to maintain the labour market equilibrium.

The equilibrium profit stays as in Dupont & Martin (2006):

\[ \pi = (1 + z^*) \pi^* = \frac{b L^w \left[ 1 + z^* - \phi - s_E z^* (1 + \phi) \right] (1 + \phi) (1 - \phi)}{\left[ 1 - \phi \left( 1 + z^* \right) \right] \left( 1 + z^* - \phi \right)}, \quad (14) \]

given region A’s expenditure share \( s_E \).

We assume that the rate of subsidy \( z^* \) is predetermined, as well as the ratio of labour and capital income tax (\( \rho \)). Thus, we have to solve for the capital tax rate from the resource constraint (equation (13)), given the after-subsidy profits:

\[ t_K = \frac{b}{b - 1} L^w + \frac{\pi K^w}{\rho L^w + \pi K^w}. \quad (15) \]

From the government budget constraint, equation (11), the equilibrium level of after-subsidy profits is found:

\[ \pi = \frac{b}{1 - b} \frac{L^w (1 + z^*)}{K^w (1 + s_n z^*)}. \quad (16) \]

The equilibrium capital income tax rate is therefore

\[ t_K = \frac{b (1 - s_n) z^*}{b(1 + z^*) + (1 - b)(1 + s_n z^*) \rho}. \quad (17) \]

Using the equations (17) and (16), and substituting these into the definition of the region A’s share of expenditure, equation (12), we get its expenditure share under such a policy as

\[ s_E - \frac{1}{2} = \frac{1}{2} \frac{b (2 s_K - 1)(1 + z^*) (\rho (1 - b) + b)}{b(1 + z^*) + (1 - b)(1 + s_n z^*) \rho}. \quad (18) \]

Combining this equation with equation (9) gives the equilibrium geography (the shares of the number of modern firms) of the economy and the expenditure shares of the regions.
Figure 1: The share of firms in region A (the tomahawk diagram) in case of different taxes on capital and labour income, based on equation (9). Assumptions: $s_E = 0.55$, blue line: $z^* = 0$, green line: $z^* = 0.2$, red line: $z^* = 1$. Phi: trade freeness.

The effects of subsidies are given at Figure 1. The share of firms in region A does not depend on the taxes, given the share of expenditures. Using the profit subsidies can be effective in relocating the increasing returns to scale industry from the larger region to the smaller one and the effect gets stronger if trade costs decrease.

The results are in principle the same as in Dupont & Martin (2006): the subsidies enable the government to influence the location of the industrial firms. Differences are in the welfare effects of the different groups of economic agents. If the capital owners are taxed more heavily than the workers, the tax burden is carried rather by the capital owners, so that even capital owners living in the subsidized region might not gain in the terms of utility, but the workers living in the same region might win. The workers living in the larger, not subsidized region still face a loss in welfare.

### 4.2 Subsidizing profits with revenues from value added taxes

In this subsection we assume that the profit subsidy is financed by a value-added tax on the industrial goods. It can be shown that in case of the uniform value-added tax on both the traditional and manufacturing goods and in the both regions the effects are identical to these of an uniform income tax (with $1 - i^{inc} = \frac{1}{1 + t_V^{inc}}$), thus, the results coincide with those of Dupont & Martin (2006). In order to analyse the welfare effects of asymmetric taxation, we assume that the value-added tax can be different on
goods sold in different regions. We assume no arbitrage: it is not possible to bring the goods from the lower-tax-region and to sell these without paying the higher tax in the other region. Also, we assume that only the goods of the increasing returns sector are taxed.\footnote{The results would not change considerably if also the consumption of traditional goods were taxed at the same rates.} The relation between the taxes in the two regions is determined by the equation \( 1 + t^* = \rho(1 + t) \).

The VAT creates distortions in the goods market: the consumer and producer prices of the industrial goods are not equal any more. The industrial firms continue to set the prices according to the monopolistic pricing rule, but the consumers face prices that are higher. This affects the profits that the firms can earn as the consumers reduce their demand. Thus, the profit equations (4) have to be replaced by

\[
\begin{align*}
\pi &= b \frac{E^w}{K^w(1+t)} \left[ \frac{s_E}{s_n + \phi(1-s_n)} + \frac{\phi(1-s_E)}{\rho(\phi s_n + 1 - s_n)} \right], \\
\pi^* &= b \frac{E^w}{K^w(1+t)} \left[ \frac{\phi s_E}{s_n + \phi(1-s_n)} + \frac{1 - s_E}{\rho(\phi s_n + 1 - s_n)} \right].
\end{align*}
\]

The share of firms in region \( A \) that equalizes the after-subsidy profits (found from the condition that \( \pi = (1 + z^*)\pi^* \)) is now

\[
s_n = \frac{s_E \rho - (1 + z^*)(1 - s_E(1 - \rho)) \phi + (1 - s_E) \phi^2}{(1 - \phi)(1 + z^* - \phi - s_E(z^* + (1 - \rho)(1 - \phi) + z^* \rho \phi))}.
\]

Thus, given the share of expenditures, the share of firms in region \( A \) depends also on the VAT rate and the ratio of these taxes in the two regions.

The effect of the changes in the ratio of the VAT in the two regions (\( \rho \)) is shown at Figure 2. If the taxes are much more higher in the subsidized region, the subsidy might not be sufficient to attract firms: the local demand-diminishing effect of taxes can dominate the attractiveness of the subsidies for low trade freeness. Thus, local financing of the subsidies might not result only in the decreased welfare of the residents of the smaller (subsidized) region, but might also push some firms to move out of the region. Nevertheless, if the taxes are more similar in the region or the trade is freer, the subsidies manage to attract the modern firms to the smaller region.

The equations for the total national and regional expenditure stay as in the basic footloose capital model. Thus, the total national expenditure and the expenditure in region \( A \) are

\[
E^w = L^w + \pi K^w
\]

and

\[
E = s_L L^w + \pi s_K K^w,
\]

\footnote{We choose this specification instead of \( t^* = \rho t \) as in the above case the resulting expressions are simpler. Qualitatively the results do not differ.}
Figure 2: The tomahawk diagram in case of different value added tax in the two regions, based on equation (20). Assumptions: $s_E = 0.55$, $z^* = 0.1$, blue line: $\phi = 1$, green line: $\phi = 0.98$, red line: $\phi = 1.1$. Phi: trade freeness.

respectively.

The resource market equilibrium condition is now

$$L^w = E^w \left[ 1 - \mu + \frac{\mu (1 - s_E(1 - \rho))(\sigma - 1)}{(1 + t)\rho} \right]$$

and the government budget constraint

$$\frac{t}{1 + t} \mu s_E E^w + \frac{t^*}{1 + t^*} \mu (1 - s_E) E^w = z^*(1 - s_n)\pi^* K^w.$$  \hspace{1cm} (24)

The profits and taxes keeping the resource market in the equilibrium are given by equations (25) and (26), respectively:

$$\pi = \frac{L^w}{K^w} \frac{(1 + z)\mu}{[\sigma - \mu + z(1 - \mu + s_n(\sigma - 1))]}$$

$$t = \frac{\sigma(1 - s_E)(1 - \rho)(1 + s_n z^*) + (1 - s_n) z^*(1 - s_E(1 - \rho))}{\rho \sigma(1 + s_n z^*)}.$$  \hspace{1cm} (26)

The income inequality or the equilibrium share of expenditures in region $A$ is

$$s_E - \frac{1}{2} = \frac{1}{2} \frac{\mu (1 + z^*) (2s_K - 1)}{z^*(1 - s_n + s_n \sigma) + \sigma},$$

where it is assumed that labour is distributed equally between the two regions ($s_L = 1/2$).
The effect of the policy combining VAT on modern goods and profit subsidies is not very clear. Even though the profits are subsidized, the nominal incomes of capital owners might in reality decrease as the tax creates distortions in the market. The welfare effects of such a policy are even more ambiguous than in the case for the income taxation. Nevertheless, now the capital owners face a larger tax burden than the workers: both groups have to buy more expensive goods, but the profits of the modern firms decrease due to the VAT.

5 Welfare effects of the tax policies

[To be added]

6 Conclusions

In the paper it was shown that combining taxes and subsidies enables to influence the spatial pattern of economic activity. The footloose capital model known from the literature of new economic geography was used as the basis of the analysis. It was assumed that the government subsidizes the profits of firms locating in the smaller region, while the money necessary for paying the subsidies is collected as the income tax - which is possibly different for labour and capital income - or as the value added tax on the modern goods. In the latter case the tax might be different in the two regions.

From the policy measures the subsidies are more important in influencing the location decisions of the modern firms, though through the taxes the share of expenditures in the regions changes. Smaller subsidies are necessary for attracting firms to locate to the smaller region if the trade is freer as then it is not so important to locate close to the larger market - the home market effect decreases if trade costs decrease.

There are four groups of economic agents in the economy: workers residing in the smaller region, workers residing in the larger region, capital owners residing in the smaller region and capital owners residing in the larger region. Each of these groups is affected differently by the implemented policy, with the tax structure having a very important role.

References


