Cross Border Health Care Provision: Who Gains, Who Loses

Rosella Levaggi and Francesco Menoncin

Abstract The diffusion of the welfare state has produced a widespread involvement of the public sector in financing the production of private goods for paternalistic reasons. In this chapter we model the production of health care as a merit impure local public good whose consumption is subsidized and whose access is free, but not unlimited. The impure local public good aspect means that the production of health care spreads its benefits beyond the geographical boundaries of the Region where it is produced. Finally, we include the (optional) provision of an equalization grant that allows reduction of fiscal imbalance among Regions. In this framework we study the possible effects of cross border provision of health care. We assume that information is complete and symmetric and that there is no comparative advantage in local provision. In this context devolution is always suboptimal for the whole community: the lack of coordination means that the impure public good is under-provided. However, more efficient Regions may be better off because of the impure public good nature of health care.

Keywords Health care provision · Decentralization · Impure public goods · Cross border supply

JEL Classification H77 · H4

R. Levaggi (✉) · F. Menoncin
Department of Economics and Management, University of Brescia, Via S. Faustino 74/b
25122 Brescia, Italy
e-mail: levaggi@eco.unibs.it

F. Menoncin
e-mail: menoncin@eco.unibs.it

R. Levaggi and M. Montefiori (eds.), *Health Care Provision and Patient Mobility*,
Developments in Health Economics and Public Policy 12,
DOI: 10.1007/978-88-470-5480-6_10,

1 Introduction

The diffusion of the welfare state has produced a widespread involvement of the public sector in financing the production of private goods for paternalistic reasons (Schnellenbach 2012). Because of the parallel process of devolution, the provision of such goods has often been delegated to local Governments and provided mainly through (partial or total) public subsidies and expenditure-based equalization grants.

Health care is one of the most relevant examples in this class of goods; its expenditure has been steadily growing since the inception of the Welfare State and it is expected to grow in the future. Whatever the cause for such a growth, the recession that started in 2008 calls for all possible efforts to reduce its cost. Telemedicine and the widespread diffusion of internet technology may allow substantial reduction of diagnostic costs. In the US, where the cost of health care has reached about 18 % of GDP, a growing number of insurance companies and Health Maintenance Organisations (HMOs) hire physicians in developing countries and use them to make diagnoses.

Smith et al. (2009) present the four models of service delivery involving cross country trade in health care which are regulated by GATS (General Agreement on Trade and Service).

The first model is the supply of cross border health services, an emerging trade that has been made possible by the advance in information technology. It consists in remotely providing a service from a provider in one country to an overseas recipient. It covers a widening basket of services ranging from diagnostics (teleradiology and laboratory testing) to treatment (remote surgery and teleconsultation). The second model is about consumption of services abroad and it comprises what is normally meant by "patient mobility" in all the forms that have been presented in this book. The third model can be defined as "foreign direct investment" and is somehow related to the second model. In order to deliver services with a standard level of quality, new hospitals devoted to treating non-residents are built in developing countries by joint ventures between domestic and foreign partners. Finally, the fourth model involves health professionals that move from one country to another. The first model is becoming popular in the US where the cost of health care is particularly high. In order to reduce the insurance bills to their clients, some companies have hired specialists in emerging countries. The tests are carried out in the US by nurses who send the images abroad. In the emerging country doctors examine them and send the referral back to the US. In this way the cost of the medical staff is sensibly reduced without lowering the quality. In Europe the second and the fourth model of trade described above are the most common. In this chapter we concentrate on the first two models and we build a framework that allows us to study cross border health services in a context where equity issues are also taken into account. In this chapter we model the production of health care as a merit impure local public good whose consumption is subsidised and whose access is free, but not unlimited. The impure local public good

aspect means that the production of health care spreads its benefits beyond the geographical boundaries of the Region where it is produced. Finally, we include the (optional) provision of an equalization grant that allows reduction of fiscal imbalance among Regions. In this framework we study the possible effects of cross border provision of health care.

Our benchmark model is *decentralization*, where provision at local level is made by each Region separately and where cross border mobility is not allowed; we will then study the welfare implications of allowing mobility and some form of coordination in the expenditure decisions among the Regions. We assume that information is complete and symmetric and that there is no comparative advantage in local provision.[1] In this context devolution is always sub-optimal for the whole community: the lack of coordination means that the impure public good is under-provided. Opening to international trade allows the less efficient Region to improve its welfare, but the most efficient one may experience a welfare loss due to the fact that the price for cross border shopping may be lower than the production cost. This effect is due to the impure public good characteristic of health care. On the other hand, international trade and more coordination among Regions (either through bilateral agreements or the intervention of an upper Government level) will ensure that spillovers are taken into account. In general, health care expenditure will increase and so will the equalization grant. The richer Region may suffer a welfare loss in this case because more resources will be required to finance the grant to the poorer Region. This opens an interesting policy debate about cross border supply and federalism. Unless this policy is imposed by an upper level, it is not likely to be the outcome of coordination between local authorities. This may partially explain the present debate at EU level about patient mobility across countries (Legido-Quigley et al. 2007, 2012; Glinos et al. 2010; Brekke et al. 2011).

This chapter is organised as follows: in Sect. 2 we present the general framework. In Sect. 3 the first best centralized solution is computed, while Sect. 4 presents devolution. A discussion and a numerical simulation are presented in Sect. 5. Section 6 concludes.

2 The Model

We model health care as an impure local public good with spillovers. Impure public goods belong to a fairly heterogeneous category, varying from impure public goods in their most traditional definition (Musgrave and Musgrave 1989) to

[1] The traditional literature assumes that there is a comparative advantage in producing at local level. Asymmetry of information and spillovers may however reduce this advantage (see Koethenbuerger 2008; Tommasi and Weinschelbaum 2007).

paternalistic goods with spillovers.[2] In our model, this characteristics of the service is captured by the form of the subsidy. For an impure public good, the usual form is a user charge, i.e. the consumer is asked to pay a fraction of the price of the service produced. When the impure public good is also a paternalistic good, it is usually supplied free of charge, but not necessarily to the entire population; this is the approach we will be using in this work. Let us consider a community, whose population is normalized to one and equally divided into two Regions $i \in \{1, 2\}$ (which can be two Local Authorities in the same country or two Member States in a Union). Each individual has an exogenous money income, M^k in the range $\left[\underline{M}_i, \overline{M}_i\right]$, with density function $\varphi_i(M^k)$. Then, total income in local authority i is:

$$Y_i = \frac{1}{2}\int_{\underline{M}_i}^{\overline{M}_i} M^k \varphi_i(M^k) dM^k.$$

Income it is used to buy private commodities and one or zero unit of a paternalistic local public good (health care). Each inhabitant has a preference towards such a good, which is defined by the parameter $\alpha \in [0, \beta]$ which also measures the utility of each unit consumed. We assume α is uniformly distributed between agents (i.e. its density function is β^{-1}).

Health care is a paternalistic good whose access is restricted to individuals with $\alpha > p_i$, where p_i represents the marginal utility of the paternalistic good the decision-maker is willing to finance. It is financed using a linear income tax at rate τ_i. The utility function for a representative individual living in local authority i can be written as:

$$V_i^k = M^k(1-\tau_i) + \begin{cases} \alpha^k & \text{if } \alpha^k \geq p_i, \\ 0 & \text{if } \alpha^k < p_i. \end{cases} + \phi_i(S_i, S_j). \tag{1}$$

Finally, utility from the public good nature of health care depends on the quantity produced in each Region (S_i), which accrues welfare because it creates a local public good in terms of option to use health care if needed.[3] We assume that preferences for the impure public good have the following form:

$$\phi_i(S_i, S_j) = w_i\left(z_i S_i - S_i^2\right) + (1 - w_i)\left(z_i S_j - S_j^2\right),$$

where S_i and S_j are the quantity of good H produced in the two jurisdictions.

The function is increasing and concave in its arguments (decreasing marginal utility at community level), hence the utility of an additional unit depends on

[2] The basic difference between a merit good and an impure public good is that the former is actually a private good that is used to improve income redistribution or to pivot consumers' preferences towards the use of goods which the planner thinks they should use. We define as spurious merit good a class of services that have this *dual* characteristics, for example health, education and cultural activities.

[3] The literature had defined this an "option demand good", see Lindsay (1969).

where it is produced.[4] The level of publicness of the good depends on the value of parameter w_i. In particular:

1. for $w_i = \frac{1}{2}$, the good is a public good;
2. for $w_i = 1$, the good is a local public good;
3. for $0 < w_i < 1_i$, the good is a local public good with spillovers.

z determines the marginal utility of health care produced in each Region. The marginal utility is in fact equal to $z - 2S$ and z has to be sufficiently high to insure that $z - 2S > 0$.

The quantity of health care demanded for in each Region is

$$Q_i = \frac{1}{2}\int_{p_i}^{\beta} \frac{1}{\beta} d\alpha = \frac{\beta - p_i}{2\beta},$$

while the utility given by consuming Q_i is

$$\frac{1}{2}\int_{p_i}^{\beta} \frac{1}{\beta} \alpha d\alpha = \frac{\beta^2 - p_i^2}{4\beta}.$$

Total welfare of Region i is given by the sum of the net income, the utility from consumption of the locally produced good, and the utility from the public characteristics of the locally produced good:

$$W_i = Y_i(1 - \tau_i) + \frac{\beta^2 - p_i^2}{4\beta} + w_i\left(z_i S_i - S_i^2\right) + (1 - w_i)\left(z_i S_j - S_j^2\right). \qquad (2)$$

We assume:

1. $Y_1 > Y_2$, i.e. Region 1 is richer than 1;
2. the marginal cost to locally produce the good is constant and there is no fixed cost, but also in this case Region 1 is more efficient than Region 2.

Given the double nature of private and public good, the beneficiaries of the two characteristics may not coincide. The quantity demanded by residents in each local authority (Q_i) does not necessarily need to coincide with the quantity produced in the same area (S_i). In other words, we allow for cross border provision.[5] When cross border mobility is allowed, supply and demand are matched by Regions through negotiation of a transfer price which may not be equal to marginal cost, given the externality produced.

[4] For a distinction between global public goods and local public goods with spillovers, see Levaggi (2010).

[5] The analysis presented here neglects transportation costs in order to concentrate on the coordination problem among regions.

2.1 Equality Issues

In most countries income is unevenly distributed across geographical areas and the super-national authority may be called to introduce some form of interregional distribution among Regions. This function is certainly more important within a unitary State, but in Europe also the EU plays an important role in this process.[6] In our model we introduce horizontal equity which implies that a given tax effort should be rewarded with the provision of a uniform amount of public services. This objective can be pursued using an equalization grant G_i. Several forms of equalization exist; in this model we will use the lump-sum form as suggested by Dahlby and Wilson (1994) and Smart (1998). The idea behind this equalization grant is to virtually increase the tax base of the poorer Region so that the fiscal effort in terms of tax rate allows the same tax revenue to be obtained. The lump sum form is chosen in order to reduce the deadweight loss deriving from the intergovernmental grant, but this implies that the amount of the equalization grant depends on the total level of health care expenditure. The equalization grant G_i can be written as :

$$G_i = \frac{1}{2}\tau^m\left(\overline{Y} - Y_i\right),$$

where

$$\tau^m = \frac{\tau_1 Y_1 + \tau_2 Y_2}{Y_1 + Y_2},$$
$$\overline{Y} = \frac{Y_1 + Y_2}{2},$$

are the national average tax rate and the standardized tax base respectively. Both are invariant to each regional fiscal decision, i.e. local authorities do not perceive the effects that their tax rate has on the equalization grant. It is interesting to note that this form of equalization grant implies that $G_i = -G_j$. It is a form of horizontal equalization grant since it does not imply raising a tax at the super national level. This form is particularly suitable to study welfare in the EU where taxes are levied at the lower level.

Given the local tax τ_i and the grant G_i, if we assume that the total cost of producing the quantity Q_i is $v_i Q_i$, then the local authority constraint can be written as:

$$\tau_i = \frac{v_i Q_i - G_i}{Y_i}.$$

[6] About 30 % of the total EU budget is in fact devoted to Regional support (De la Fuente et al. 2010).

3 Decentralized Solution

Each Region sets its own tax rate and service production according to its preferences and resources and takes G_i as given. The upper government level sets the equalization grant; this actor is the last one to move, i.e. it sets the grant after the Regions have set their own level of expenditure and taxation.

This solution is not optimal from a welfare point of view because the Regions do not fully take into account the consequences of their actions on welfare (Petretto 2000). This is a well known result in the literature on fiscal federalism: the presence of spillovers means that the quantity produced in each local authority is always lower than the optimal one. Below we examine the two cases with and without cross border supply and we will then consider the implications for welfare of opening health care to trade.

3.1 Without Cross Border Supply

Let us now examine the optimal conditions in an environment where each Region maximizes its own welfare and cross border supply is not allowed. The problem can be written as:

$$\begin{aligned} &\max_{p_i} W_i(p_i, Q_i, Q_j) \\ &\text{s.t.} \\ &\tau_i = \tfrac{v_i Q_i - G_i}{Y_i}, \\ &Q_i = \tfrac{\beta - p_i}{2\beta}. \end{aligned} \tag{3}$$

The optimal provision for the problem are derived in Appendix and can be written as

$$\begin{aligned} \bar{p}_i &= \beta \frac{v_i + w_i(1 - z_i)}{\beta + w_i} = v_i - w_i \left(\frac{v_i + \beta(1 - z_i)}{\beta + w_i} \right), \\ \bar{Q}_i &= \frac{1}{2} - \frac{v_i - w_i(1 - z_i)}{2(\beta + w_i)}. \end{aligned} \tag{4}$$

The number of people that receive health care depends on the marginal cost of production (v_i) and the public good aspect of health care w_i. This implies that the cost effectiveness of the treatment has to include both the utility accruing to the patient and the increased benefit that society receives from the production of that specific unit of health care.

The quantity produced is not optimal because each Region does not take into account the positive spillover its production is creating on the neighbour Region and it does not exploit the higher productivity level of its neighbour.

Total welfare in this case can be written as:

$$\begin{aligned}\overline{W}_i &= Y_i(1-\bar{\tau}_i)+\frac{\beta^2-p_i^2}{4\beta}+w_i\left(z_i\bar{Q}_i-\bar{Q}_i^2\right)+(1-w_i)\left(z_i\bar{Q}_j-\bar{Q}_j^2\right),\\ \bar{W} &= \sum_{i=1,2}\overline{W}_i,\\ \bar{\tau}_i &= \frac{\bar{Q}_i v_i}{Y_i}-\frac{\left(\overline{Y}-Y_i\right)}{(Y_1+Y_2)Y_i}\sum_{i=1,2}\bar{Q}_i v_i.\end{aligned} \tag{5}$$

This framework can be considered our benchmark model. In the following sections we will study the effects of cross border supply of health care on this solution in terms of quantity of health care supplied, its regional distribution and its implications for welfare. Cross border mobility may result either from bilateral agreements between the Regions or from specific upper Government tier initiatives. In our model we do not distinguish between these two sources; in fact we concentrate on the implications of the decision process rather than on how it was originated.

In a neoclassical model, trade is always beneficial. Because of specialization, both trade partners will be better off (see, for instance, Frankel and Romer 1999). In this model this might not be the case owing to the combined effect of several factors: the impure public good characteristics of health care and the presence of the equalization grant.

3.2 With Cross Border Supply

Transnational health care is a new and emerging phenomenon in the European context, but it may assume several forms that needs to be studied separately. Levaggi and Levaggi (2014) show the different welfare implications of regulated patient mobility versus patient choice. In this paper we focus on a more specific form of patient choice by considering forms of collaboration between nations very close to one another or the flow of mobility across neighbouring regions within the same country. In both contexts we assume that the traveling costs are negligible and can be approximated to zero. In this way, we can concentrate on the welfare effects of opening the borders to competition. Usually, increased competition should improve welfare, but this may not be the case in a context where the good supplied is subsidized. In our model two forms of redistribution exist:

- interpersonal redistribution: the good is supplied for free to those that need health care and it is financed out of general taxation;
- inter-regional redistribution: expenditure is used to equalize resources across Regions.

In this context, total welfare will certainly increase (due to a more efficient use of resources), but the benefits between the two Regions may be distributed quite

unevenly and opening to international trade may not be a Pareto optimal solution. In this section we study such a case.

When the service is supplied across borders, the Region producing it is reimbursed at rate q (the so-called "transfer price"). Accordingly, the local authority i receives: (i) the amount of tax $Y_i\tau_i$, (ii) the grant G_i and, on the other side, must pay: (i) the production cost v_iS_i, and (ii) the reimbursement q proportional to the product supplied across the border $Q_i - S_i$. The constraint can be written as

$$Y_i\tau_i + G_i = v_iS_i + q(Q_i - S_i),$$

from which we have

$$\tau_i = \frac{v_iS_i + q(Q_i - S_i) - G_i}{Y_i}.$$

The problem faced by each local authority can thus be written as:

$$\begin{aligned} &\max_{p_i,S_i} W_i(p_i, S_i, S_j) \\ &\text{s.t.} \\ &\tau_i = \frac{v_iS_i + q(Q_i - S_i) - G_i}{Y_i}. \end{aligned} \tag{6}$$

The FOCs for the problem are derived in Appendix and can be written as

$$\begin{aligned} p_i &= q, \\ S_i &= \frac{q - v_i + z_iw_i}{2w_i}. \end{aligned} \tag{7}$$

The local decision-maker does not take into account the spillover effect that its production creates on the neighbouring jurisdiction. Furthermore, in their maximization process, they take q as a given parameter, but in equilibrium only one value exists which clears the market. To reconcile devolution with market clearing conditions, it is necessary to find the q that satisfies the optimal conditions (7) and the market clearing constraint. The problem can be solved using a Nash game:

$$\begin{aligned} S_i &= \frac{q - v_i + z_iw_i}{2w_i}, \\ S_1 + S_2 &= Q_1 + Q_2 = 1 - \frac{q}{\beta}, \end{aligned} \tag{8}$$

and the level of transfer price that clears the market is equal to:

$$\begin{aligned} p_i &= q, \\ S_i &= \frac{q - v_i}{2w_i} + \frac{z_i}{2}, \\ q &= \frac{w_2(\beta + w_1)\bar{p}_1 + w_1(\beta + w_2)\bar{p}_2}{w_2(\beta + w_1) + w_1(\beta + w_2)}, \end{aligned} \tag{9}$$

where $\bar{p}_i = \beta \frac{v_i + w_i(1-z_i)}{\beta + w_i}$, $i = 1,2$ is the utility of the marginal patients receiving care in the decentralized solution. The first interesting result of this solution is that the level of production in the two Regions is now determined by the transfer price q. From Eq. (9) we can observe that the transfer price is a weighted average of $\bar{p}_i$. As in the previous model, the number of people receiving care depends on the marginal price of production and on the public good aspect of health care. This imply that the number of people that receive health care in the new equilibrium will increase in Region 2 and will decrease in Region 1. Overall the number of patients treated is however increasing because the average cost to treat them is decreasing. To determine the effects on demand of cross border mobility, the difference between Eqs. (4) and (9) should be evaluated. The latter depends on the combined effect of the difference in price and in the preferences. However, some conclusions can be drawn[7]:

- if the good has the same local public good characteristic ($w_i = w_j = w$), then $p_i > q$ if $z_j < z_i + \frac{v_j - v_i}{w}$;
- if the preferences for the public good are uniform ($z_i = z_j$), then $p_i > q$ if $v_i - v_j < 0$;
- in any case $q < \min(v_i, v_j)$ which means that health care is always subsidized.

The last point is quite interesting because it means that in a context of cross border mobility the Region that receives patients will have to finance its costs with local resources, an element that depends on the impure public good nature of health care, and that has an important role in the distribution of the benefits arising from cross border shopping.

After finding $\hat{q}$ that clears the market, it will be possible to obtain $\widehat{p}$, $\widehat{S_1}$ and $\widehat{S_2}$. Total welfare can be written as:

$$\begin{aligned} \hat{W} &= Y_i(1-\hat{\tau}_i) + \frac{\beta^2 - \hat{p}_i^2}{4\beta} + w_i\left(z_i\hat{S}_i - \hat{S}_i^2\right) + (1-w_i)\left(z_i\hat{S}_j - \hat{S}_j^2\right), \\ \hat{\tau}_i &= \frac{\hat{S}_i(v_i - \hat{q}) - \hat{Q}_i(\hat{p} - \hat{q})}{Y_i} \\ &\quad - \frac{\left(\overline{Y} - Y_i\right)}{(Y_1 + Y_2)Y_i} \sum\nolimits_{i=1,2} \left(\hat{S}_i(v_i - \hat{q}) - \hat{Q}_i(\hat{p} - \hat{q})\right). \end{aligned} \tag{10}$$

4 Welfare Improving Strategies

The solution presented in the previous section, although an improvement on the "no cross border shopping" case, does not represent a welfare First Best because of the positive externality caused by the production of health care. In our model a

[7] See the Appendix for a formal proof.

coordinated solution that allows a First Best optimal allocation may be obtained as a "bottom up" solution where the two Regions freely decide for a coordinated policy where quantities and mobility are set through a bargaining process or for a top down procedure where the supernational authority decides how much to produce and where. This solution can be attained only if specific conditions are met; in order to describe the process in the simplest way, we will start by presenting the optimal solution that in both cases derives from the maximization of the following welfare function:

$$\begin{gathered}\max_{p_1,p_2,S_1,S_2} W^* = W_1(p_1, S_1, S_2) + W_2(p_2, S_2, S_1) \\ \text{s.t.} \\ \tau_i = \frac{v_i S_i + q(Q_i - S_i) - G_i}{Y_i}, \\ S_1 + S_2 = Q_1 + Q_2 = Q.\end{gathered} \tag{11}$$

The solution is derived in the Appendix and can be written as:

$$\begin{aligned} p_i &= q, \\ S_i &= \frac{1}{2}\frac{q - v_i + z_i w_i + z_j(1 - w_j)}{w_i + 1 - w_j}, \quad j \neq i \in \{1, 2\}, \\ q &= \beta\frac{(w_1 - w_2)^2 + (z_2 w_2 - z_1 w_1)(w_1 - w_2) - 1}{(w_1 - w_2)^2 - (1 + \beta)} \\ &\quad + \frac{\beta}{2}\frac{(z_2 - v_1)(w_2 + (1 - w_1)) + (z_1 - v_2)(w_1 + (1 - w_2))}{(w_1 - w_2)^2 - (1 + \beta)} \end{aligned} \tag{12}$$

which can be interpreted as follows: the allocation of production between the two local authorities should follow an efficiency principle by balancing the need to reduce the cost of public provision with the utility both communities derive from the location of the production of that specific good.

In general $q^* < \hat{q}$ because in this case the quantity to be supplied takes into full account the effect of the spillover produced by the supply of health care. Production will be concentrated in the more productive Region, but in this case the full effect of the reduction in welfare in the sending Region is taken into account.

The quantity of the impure public good in the two local authorities is the same in equilibrium. p^* is chosen to equalize the marginal rate of substitution between income and the impure public good with the price ratio.

The optimal solution in terms of p^*, S_1^*, S_2^*, Q_1^* and Q_2^* can be substituted in the welfare function to obtain:

$$\begin{aligned} W^* &= Y_i\left(1 - \tau_i^*\right) + \frac{\beta^2 - p_i^{*2}}{4\beta} + w_i\left(z_i S_i^* - S_i^{*2}\right) + (1 - w_i)\left(z_i S_j^* - S_j^{*2}\right), \\ \hat{\tau}_i &= \frac{S_i^*(v_i - q^*) - Q_i^*(p^* - q^*)}{Y_i} - \frac{\left(\overline{Y} - Y_i\right)}{(Y_1 + Y_2)Y_i}\sum_{i=1,2}\left(S_i^*(v_i q^*) - Q_i^*(p^* - q^*)\right). \end{aligned} \tag{13}$$

Total welfare is clearly increasing in this equilibrium which represents First Best. However, this does not necessarily mean that both local authorities are better off. In fact the reimbursement price for the health care supplied across the border is lower than in the previous case and although the total quantity (hence the utility derived from the public characteristic of health care) increases, the level of taxation in the destination Region will increase. The second element that determines the difference in welfare is the equalization grant. The quantity of health care produced is increasing, the average cost is lower than in the decentralized case, but total expenditure is increasing. This implies that the equalization grant is increasing as well. In general this implies that more resources are flowing from the rich to the poor Region. If the rich Region is also the more productive one (i.e. the one receiving the flow of patients), both elements (cross border shopping and coordination) will have a depressing effect on local welfare. On the other hand, if the poor Region is the more efficient, the equalization grant will be able to partially compensate for the increased taxation due to cross border health care provision.

5 Discussion and Numerical Example

The model presented in this chapter is a first attempt to study the welfare properties of opening international trade in the health care sector. We study the problem from the perspective of local or national governments where health care is supplied by the public sector and has the double characteristics of being a paternalistic good and a local public good with spillovers. The first characteristics means that the good will be supplied for free at the point of use, and its cost will be financed through a linear income tax. The Government rations the quantity of health care available by determining the marginal individual that is allowed to receive care. The latter depends on the utility derived from health care, on its production cost. However, utility has two components: the benefit received by the individual in terms of accrued health and a public good element that depends on the quantity of care produced in each Region. This implies that the benefit of the marginal patients will always be lower than the production cost of health care. When cross border supply is not allowed, the only effect we observe is a redistribution of income from non users of the service to users. In fact the former pay taxes that will then be used to finance health care. When cross border shopping is allowed this redistribution effect may spread to the regional level.[8]

International trade allows the less efficient Region to buy health care in the more efficient one; in this way the number of people allowed to receive health care increases, and its average production cost decreases. Although health care

[8] The optimal conditions for the provision of an impure public good imply that the production cost is equal to the marginal private benefit increased by the utility deriving from the public good nature of the good. See Musgrave and Musgrave (1989) for a formal proof.

production increases, the quantity of health care locally produced decreases; this implies that utility from the public good is likely to decrease. In general, these effects produce an increase in welfare. On the other hand, the more efficient Region may not necessarily gain in this process. Fewer resident patients may be allowed to receive health services owing to an increase in the threshold to be eligible (which optimally equates q, the price that clears the market). More health care is produced in the efficient Region, which implies an increase in welfare, but the price (q) for cross border shopping is going to be lower than the marginal cost, i.e. the tax rate will have to increase. The equalization grant has a countervailing effect in this case: the total cost to produce health care may decrease and the tax rate in the more efficient local authority certainly increases.

In general we can state that when two Regions cooperate, total welfare increases because the spillovers effects are correctly taken into account, but the gain may be unequally split between the two.

The analytic analysis of this case is quite cumbersome. Below we present some simulations that provide some numerical insights into the working of our model. The general form of the welfare function is:

$$W_i = Y_i(1-\tau) + \frac{\beta^2 - p_i^2}{4\beta} + w_i\left(z_i S_i - S_i^2\right) + (1-w_i)\left(z_i S_j - S_j^2\right).$$

We have evaluated the different solutions for the following initial parameters $Y_1 = 1.5, Y_2 = 1, w_1 = 0.65, w_2 = 0.65, v_1 = 5, v_2 = 6, \beta = 20, z_1 = 1.5, z_2 = 1.5$. The solution has been computed for two systems, with and without an expenditure-based equalization grant. The latter solution should correspond to a bilateral agreement between two national States while the first one is more likely to represent the case of an upper tier which is also interested in local income redistribution.

The results are presented in Table 1. Welfare reaches its maximum in a system where cross border shopping is allowed and the level of care produced is jointly determined, as one might expect. In this case production is concentrated in local authority 1, but the quantity demanded is the same in both local authorities. The tax rate and the grant depends on q, the price for mobility, as much as welfare of the two local authorities. The decentralized solution with cross border shopping is characterised by a higher price for the quantity traded, which creates an increase in the quantity of the local good produced in 1, but a reduction in total demand $Q_1 + Q_1$. This solution is characterised by the lowest average tax rate which implies that the equalization grant is minimum.

The welfare comparisons for this case are quite interesting. Let us start with the case where an equalization grant is present. For Region 1 opening to trade is not welfare improving. This is due to the combined effects of several elements: from the expenditure side we note that although total quantity and local production increase, health care available to residents decreases; on the financial side, local taxes increase because the reduction in the cost to produce health care is mainly borne by Region 1. The welfare loss persists also in the case where an equalization grant is not foreseen. This means that in this case the welfare loss is mainly due to

Table 1 Simulation results

	No mobility	Mobility	First best
q		0.86854	0.68182
S_1		0.64888	0.59091
S_2		0.26427	0.34091
Q_1	0.46831	0.45657	0.46591
Q_2	0.44484	0.45657	0.46591
Q	0.91315	0.91315	0.93182
p_1	0.63380	0.86854	0.68182
p_2	1.10330	0.86854	0.68182
Bilateral agreement			
τ_1	0.31221	0.32124	0.33712
τ_2	0.66725	0.56343	0.59659
G_1	0	0	0
G_2	0	0	0
W_1	4	3.9726	3.9702
W_2	3.2765	3.3233	3.3366
Upper tier decision			
τ_1	0.38791	0.39092	0.41061
τ_2	0.55370	0.45890	0.48636
G_1	−0.11356	−0.10453	−0.11023
G_2	0.11356	0.10453	0.11023
W_1	3.8864	3.8680	3.8599
W_2	3.3901	3.4278	3.4469
$W_1 + W_2$	7.2765	7.2958	7.3068

The values of the parameters are: Y_1 and Y_2 are independent and uniformly distributed. Their average is equal to 1.5 and 1 respectively. $w_1 = 0.65$, $w_2 = 0.65$, $v_1 = 1$, $v_2 = 1.5$, $\beta = 10$, $z_1 = 1.5$, $z_2 = 1.5$

the effect of q on local tax rates. More coordination further reduces the welfare of Region 1, especially in the presence of an equalization grant. This opens an interesting policy debate about cross border supply and federalism. Unless the good is perceived as a local public good, this policy is imposed by an upper level, it is not likely to be the outcome of coordination between local authorities. This may partially explain the present debate at EU level about patient mobility across countries (Glinos et al. 2010; Brekke et al. 2011; Legido-Quigley et al. 2012). Most of the agreements that have been put forth are related to planned admissions, a type of activity where the public good aspect of health care is probably less important as well as the spillover effect. In these cases agreements may be feasible and are beneficial to both Regions because the price for non residents is going to be quite close (if not equal) to the marginal cost of production.

In order to study the effects of the cross border price on welfare, in Table 2 we present a second simulation. The parameters are the same, but in this case Region 2 (the poorer) is the most efficient. Also in this case the most efficient Region is the loser in terms of welfare, but now the equalization grant has a mitigating effect on

Table 2 Simulation results

	No mobility	Mobility	First best
q		0.86854	0.68182
S_1		0.64888	0.59091
S_2		0.26427	0.34091
Q_1	0.44484	0.45657	0.46591
Q_2	0.46831	0.45657	0.46591
Q	0.91315	0.91315	0.93182
p_1	1.10330	0.86854	0.68182
p_2	0.63380	0.86854	0.68182
Bilateral bargain			
τ_1	0.44840	0.32124	0.39773
τ_2	0.46831	0.56343	0.50568
G_1	0	0	0
G_2	0	0	0
W_1	3.7765	3.9726	3.8366
W_2	3.5000	3.3233	3.4702
Upper tier decision			
τ_1	0.52054	0.4453	0.47121
τ_2	0.35475	0.37732	0.39545
G_1	−0.11356	−0.10453	−0.11023
G_2	0.11356	0.10453	0.11023
W_1	3.6630	3.7187	3.7264
W_2	3.6135	3.5771	3.5804
$W_1 + W_2$	7.2765	7.2958	7.3068

The values of the parameters are: Y_1 and Y_2 are independent and uniformly distributed with average 1.5 and 1 respectively. $w_1 = 0.65$, $w_2 = 0.65$, $v_1 = 1$, $v_2 = 1.5$, $\beta = 10$, $z_1 = 1.5$, $z_2 = 1.5$

the loss, as one might expect. It is also interesting to note that in this case more coordination is preferred to a simple trade agreement.

These examples show that it is not possible to determine the effects of international trade on welfare distribution. Total welfare will certainly increase, but international trade might not be a Pareto superior solution. The distribution of the welfare between the two Regions depends on several factors: the importance of spillovers (measured by w), the relative importance of the public good aspect of health care (measured by z) and, if an equalization grant is foreseen, the difference in income in the two Regions.

To show this in our model, we have determined the value of w (the spillover parameter) for which trade and cooperation is a Pareto efficient solution. For the first simulation (Region 1 richer and more efficient) trade would be welfare improving for both Regions for $w = 0.79$; if $w > 0.85$ even coordination improves welfare. These values are lower (0.72 and 0.81, respectively) without equalization grant.

For the second simulation (Region 1 richer and Region 2 more efficient) the value of w is equal to 0.81 for trade agreement and to 0.79 for a coordinated solution. These values are however higher without an equalization grant.

Finally, if health care is perceived mainly as a private good, international trade is more likely to have positive effects. In this case, in fact, the price q for cross border supply comes closer to the marginal of the more efficient Region. This Region will then experience a (relatively) lower increase in the local tax rate. This may explain why some countries have set agreements for cross border supply of non acute hospital care. These services are quite similar to private goods and, in this case, international trade may be a Pareto superior solution.

6 Conclusions

The development of trade and international relations has put increasing pressure on regulatory authorities to define new rules for health care cross border shopping. The process of globalization and the introduction of new technologies means that the market for health care is becoming more and more international. Telemedicine and telediagnosis allow patients to be treated and monitored at an increasing distance, often beyond the national boundaries. Fidler et al. (2009) argue that while for international trade rules are well defined by GATT and other treaties, for health care settlements they are not so clear and only about 40 % of international trade is carried out under these rules. International trade challenges the governance of health care systems in any country, but it is going to put extra pressure on public health care systems where the access to health care needs to be strictly regulated. In a global context these restrictions may be more difficult to enforce and may be challenged on legal grounds.[9]

In Europe economic integration has meant that also health care systems have to face the challenge of cross border shopping, both at demand and supply level. Although in Europe health care expenditure related to cross border shopping is relatively small, the number of people traveling abroad to receive care is increasing and shows a positive trend. The EU has started regulating these flows (Greer et al. 2013), but very little is known on the economic impact of such mobility.[10]

In this chapter we show that cross border shopping and coordinated efforts among Regions may not be compatible with local welfare maximization. This is because health care may be often used to redistribute income both among individuals and among Regions. If this is the case, cross border shopping allows reduction of the total cost of health care produced, but it may allocate the benefits of such reduction in a very uneven way. This is due to two different effects: (i) the presence of an equalisation grant and (ii) the spillover effect produced by health care.

[9] For example, the restrictions on the number of providers that can deliver health care is a controversial matter because this rule introduces restrictions to international trade.

[10] For a discussion of the open issues see Mackenbach et al. (2013).

This second element is particularly important because its role has been often overlooked. The public good aspect of health care implies that both the internal and the international price for health care are lower than the production cost. This means that if a country receives patients from outside, the application of the First Best rules may imply that the price (q) may be well below the production cost. All else being equal, the local authority receiving patients may experience an increase in its tax rate. Its utility will increase since utility derived from the option good aspect of health care is increasing, but this might not be enough to compensate for the increased tax rate.

In our model, mobility improves welfare for both Regions only if health care is a paternalistic good and there are no equalization grants. In all the other cases, the effect may be ambiguous and this means that a coordinated solution where cross border mobility is allowed may not be the outcome of this game. Our results allow us to explain why lower Government tiers (both within the same country or in the context of international trade agreements) may be reluctant to allow their patients to travel abroad to receive health care. The problem is particularly important in countries where income is unevenly distributed across jurisdictions so that the equalization grant plays a very important role in financing expenditure.

The EU directive on cross border mobility is at present foreseeing a reimbursement based on the cost to produce the service. This criterion represents a more favorable arrangement for the more efficient Region that will not have to subsidise the services produced for non residents, but it might not represent a First Best solution in a context where health care is an impure public good that produces spillovers effects on other Regions.

Appendix

Solution to the Problem Without Cross Border Supply

When mobility is not allowed, the quantity produced in each Region is equal to demand; in this case it is sufficient to find the utility for the marginal consumer allowed to use health care. The problem can be written as:

$$\begin{gathered}\max_{p_i} W_i\left(p_i, Q_i, Q_j\right) \\ \text{s.t.} \\ \tau_i = \frac{v_i Q_i - G_i}{Y_i}, \\ Q_i = \frac{\beta - p_i}{2\beta}.\end{gathered}$$

The constraints can be substituted back into the maximization problem that can be solved as an unconstrained maximization:

$$\max_{p_i} \overline{W}_i = Y_i - v_i\left(\frac{\beta - p_i}{2\beta}\right) + G_i + \frac{\beta^2 - p_i^2}{4\beta} + w_i\left(z_i\frac{\beta - p_i}{2\beta} - \left(\frac{\beta - p_i}{2\beta}\right)^2\right) + (1 - w_i)\left(z_i\frac{\beta - p_j}{2\beta} - \left(\frac{\beta - p_j}{2\beta}\right)^2\right).$$

The FOC can be written as:

$$\frac{\partial \overline{W}_i}{\partial p_i} = 0 \Rightarrow \frac{v_i - p_i}{2\beta} + \frac{\beta w_i(1 - z_i) - p_i w_i}{4\beta^2} = 0,$$

which is solved for the value $\overline{p}_i$ in the text.

Solution to the Problem with Cross Border Supply

When mobility is allowed, the quantity produced in each Region may not be equal to demand; in this case each Region maximizes for p_i and S_i, but in a subsequent step they will have to define a price for mobility and the number of patients allowed to cross the border. The problem can be written as:

$$\max_{p_i, S_i} W_i(p_i, S_i, S_j)$$
$$\text{s.t.}$$
$$\tau_i = \frac{v_i S_i + q(Q_i - S_i) - G_i}{Y_i},$$
$$Q_i = \frac{\beta - p_i}{2\beta}.$$

The constraint can be substituted back into the maximization problem that can be solved as an unconstrained maximization:

$$\max_{p_i, S_i} \hat{W}_i = Y_1\left(1 - \frac{v_i S_i + q(\frac{\beta - p_i}{2\beta} - S_i) - G_i}{Y_i}\right) + \frac{\beta^2 - p_i^2}{4\beta} + w_i(z_i S_i - S_i^2) + (1 - w_i)\left(z_i S_j - S_j^2\right).$$

The FOCs can be written as:

$$\frac{\partial \hat{W}_i}{\partial p_i} = 0 \Rightarrow \frac{q - p_i}{2\beta} = 0,$$
$$\frac{\partial \hat{W}_i}{\partial S_i} = 0 \Rightarrow -v_i + q + w_i z_i - 2w_i S_i = 0,$$

which allow to obtain the solutions in the text in terms of q. The second step consists of reconciling demand with supply as explained in the text.

Solution to the Central Government Problem

The more general problem can be written as:

$$\max_{p_i,S_i,q} W_1(p_1,S_1,S_2)+W_2(p_2,S_2,S_1)$$

s.t.

$$\tau_i=\frac{v_iS_i+q(Q_i-S_i)-G_i}{Y_i},$$
$$Q_i=\frac{\beta-p_i}{2\beta},$$
$$Q=S_1+S_2=Q_1+Q_2.$$

The first and the second constraints can be substituted in the maximization problem while the third one will be used to define the Lagrangian for the problem

$$\max_{p_i,S_i,q}\mathcal{L}=\sum_{i\neq j=1,2}Y_i\left(1-\frac{v_1S_1+q(\frac{\beta-p_i}{2\beta}-S_i)-G_i}{Y_1}\right)+\frac{\beta^2-p_i^2}{4\beta}$$
$$+w_i\left(z_iS_i-S_i^2\right)+(1-w_i)\left(z_iS_j-S_j^2\right)-\lambda\left(\sum_{i=1,2}1-\frac{\beta-p_i}{2\beta}-\sum_{i=1,2}S_i\right).$$

The FOC for the problem can be written as:

$$\frac{\partial\mathcal{L}}{\partial p_1}=0\Rightarrow-\frac{1}{2}\frac{-q+p_i-\lambda}{\beta}=0,$$
$$\frac{\partial\mathcal{L}}{\partial S_i}=0\Rightarrow z_iw_i+z_j(1-w_j)-2S_i\left(1+w_i-w_j\right)-v_i+q+\lambda=0,$$
$$\frac{\partial\mathcal{L}}{\partial\lambda}=0\Rightarrow\sum_{i=1,2}1-\frac{\beta-p_i}{2\beta}-\sum_{i=1,2}S_i=0,$$
$$\frac{\partial\mathcal{L}}{\partial q}=0\Rightarrow\sum_{i=1,2}Q_i-\sum_{i=1,2}S_i=0.$$

In this form the solution is not determined because the derivative for q is not defined.

However, let us now observe the constraint: if it is satisfied ($\lambda=0$), the market is in equilibrium, i.e. the demand for health care is exactly equal to the supply of health care. In any other case rationing exists either on the demand or on the supply. Using the same procedure described in Sect. 3.2 we can find q that clears the market by adding the constraint $\lambda=0$ to the problem described in the FOCs above. The q that clears the market is the solution of the following problem:

$$1-\frac{q}{\beta}-\frac{1}{2}\frac{q-v_1+w_1z_1+z_2(1-w_2)}{w_1+1-w_2}-\frac{1}{2}\frac{q-v_2+w_2z_2+z_1(1-w_1)}{w_2+1-w_1}.$$

Comparisons

Quantity of Health Care

In order to determine the quantity of health care produced, we can start by evaluating the difference in the marginal private benefit of health care in the different model which is represented by p.

Difference Between First Best and Cross Border Supply

In this case we know from the theory (Koethenbuerger 2008; Oates 2008) that the quantity in First Best increases. To show this, let us consider the difference between $\hat{q}$ and q^*which can be written as:

$$\begin{aligned}&\frac{w_2(v_1+w_1(1-z_1))+w_1(v_2+w_2(1-z_2))}{w_2(\beta+w_1)+w_1(\beta+w_2)}\beta\\&-\frac{1}{2}\beta\frac{(w_i-w_j-1)v_i-(w_i-w_j+1)v_j+2(w_i-w_j)^2}{-1-\beta+(w_i-w_j)^2}\\&-\frac{1}{2}\beta\frac{\left(z_i+z_j-2w_i^2z_i-2w_j^2z_j+(w_i-w_j)(z_i-z_j)+2w_iw_j(z_i+z_j)-2\right)}{-1-\beta+(w_i-w_j)^2}\end{aligned}$$

and, for $w_i=w_j=1$

$$\frac{v_i+2-z_i+v_j-z_j}{2\beta+2}\beta-\frac{\beta}{2}\frac{-v_i-v_j-2+z_i+z_j}{-1-\beta}=0,$$

as expected.

Total quantity Q is clearly bigger in FB than in the cross border supply equilibrium.

Difference Between the "No Cross Border Supply" and the Case When Cross Border Supply Is Allowed

In this case it can be written as:

$$\begin{aligned}p_i-q=&\beta\frac{v_i+w_i(1-z_i)}{\beta+w_i}\\&-\beta\frac{w_j(v_i+w_i(1-z_i))+w_i(v_j+w_j(1-z_j))}{w_j(\beta+w_i)+w_i(\beta+w_j)}.\end{aligned}$$

Let us start by assuming that $w_i=w_j=w$. In this case the difference can be written as:

$$p_i - q = \beta \frac{v_i + w(1 - z_i)}{\beta + w} - \frac{1}{2} \frac{(v_i + w(1 - z_i)) + \left(v_j + w\left(1 - z_j\right)\right)}{(\beta + w)} \beta.$$

The two expressions are equal for

$$z_j = z_i + \frac{v_j - v_i}{w}.$$

In this case total quantity depends on the interactions of the parameters. The difference can be written as:

$$\bar{Q} - \widehat{Q} = 1 - \frac{1}{2}\left(\frac{v_i + w_i(1 - z_i)}{\beta + w_i} - \frac{v_j + w_j\left(1 - z_j\right)}{\beta + w_j}\right) - \left[1 - \frac{w_j(v_i + w_i(1 - z_i)) + w_i\left(v_j + w_j\left(1 - z_j\right)\right)}{w_j(\beta + w_i) + w_i\left(\beta + w_j\right)}\right].$$

For $w_i = w_j = w$ total quantity is the same, i.e. $\bar{Q} = \widehat{Q}$.

References

Brekke, K. R., Levaggi, R., Siciliani, L., & Straume, O. R. (2011). Patient mobility, health care quality and welfare. CEPR Discussion Papers 8559, C.E.P.R. Discussion Papers.

Dahlby, B., & Wilson, L. S. (1994). Fiscal capacity, tax effort, and optimal equalization grants. *Canadian Journal of Economics, 27*(3), 657–672.

De la Fuente, A., Doménech, R., & Rant, V. (2010). Addressing the net balances problem as a prerequisite for eu budget reform: A proposal. *CESifo Economic Studies, 56*(2), 221–250.

Fidler, D. P., Drager, N., & Lee, K. (2009). Managing the pursuit of health and wealth: The key challenges. *Lancet, 373*(9660), 325–331.

Frankel, J., & Romer, D. (1999). Does trade cause convergence. *American Economic Review, 89*, 379–399.

Glinos, I., Baeten, R., & Maarse, H. (2010). Purchasing health services abroad: Practices of cross-border contracting and patient mobility in six European countries. *Health Policy, 95*(2), 103–112.

Greer, S. L., Hervey, T. K., Mackenbach, J. P., & McKee, M. (2013). Health law and policy in the European Union. *The Lancet, 381*(9872), 1135–1144.

Koethenbuerger, M. (2008). Revisiting the "decentralization theorem"—On the role of externalities. *Journal of Urban Economics, 64*(1), 116–122.

Legido-Quigley, H., Glinos, I., Baeten, R., & McKee, M. (2007). Patient mobility in the European Union. *British Medical Journal, 334*(7586), 188–190.

Legido-Quigley, H., Glinos, I. A., Baeten, R., McKee, M., & Busse, R. (2012). Analysing arrangements for cross-border mobility of patients in the European Union: A proposal for a framework. *Health Policy, 108*(1), 27–36.

Levaggi, R. (2010). From local to global public goods: How should we write the utility function. *Economic Modelling, 27*(5), 1040–1042.

Levaggi, L., & Levaggi, R. (2014). Patients mobility across borders: A welfare analysis. In Levaggi, R., & Montefiori, M. (Eds.). *Health care provision and patient mobility: Health integration in the European Union*. Berlin: Springer.

Lindsay, C. M. (1969). Option demand and consumer's surplus. *The Quarterly Journal of Economics, 83*(2), 344–346.

Mackenbach, J. P., Karanikolos, M., & McKee, M. (2013). The unequal health of Europeans: Successes and failures of policies. *Lancet, 381*(9872), 1125–1134.

Musgrave, R. A., & Musgrave, P. B. (1989). *Public finance in theory and practice*. Finance series. Singapore: McGraw-Hill.

Oates, W. E. (2008). On the evolution of fiscal federalism: Theory and institutions. *National Tax Journal, 61*, 313–334.

Petretto, A. (2000). On the cost-benefit of the regionalisation of the national health service. *Economics of Governance, 1*, 213–232. doi:10.1007/PL00011014.

Schnellenbach, J. (2012). Nudges and norms: On the political economy of soft paternalism. *European Journal of Political Economy, 28*(2), 266–277.

Smart, M. (1998). Taxation and deadweight loss in a system of intergovernmental transfer. *The Canadian Journal of Economics, 31*(1), 189–206.

Smith, R. D., Chanda, R., & Tangcharoensathien, V. (2009). Trade in health-related services. *Lancet, 373*(9663), 593–601.

Tommasi, M., & Weinschelbaum, F. (2007). Centralization vs. decentralization: A principal-agent analysis. *Journal of Public Economic Theory, 9*(2), 369–389.