Decentralisation of health care and its impact on health outcomes

Dolores Jiménez[1], Peter C. Smith[1,2]

January 2005

Abstract

This paper explores the impact of health care decentralisation on a characteristic of human development: the overall level of a population's health. While much of the literature on decentralisation in health care has stressed the advantages of sub national provision of health services, in the absence of a quantitative measure of the magnitude of the effect of decentralisation, there is little that can be said in terms of its benefits and costs for the health sector. The purpose of this study is therefore to contribute to the limited empirical literature on this issue by investigating the hypothesis that shifts towards more decentralisation would be accompanied by improvements in population health. The analysis draws on a theoretical model of local government's public finance applied to health. We use the ten provinces of Canada as a case study. Apart from being one of the most decentralised countries in the world, Canadian data required to estimate our model was found to be one of the best. The results of the empirical analysis suggest that decentralisation in Canada has had a positive and substantial influence on the effectiveness of public policy in improving population's health.

JEL codes: I12 H77
Key words: Fiscal Decentralisation, health outcomes, Canada

We wish to thank the members of the York Seminars in Health Econometrics (YSHE), and in special to Prof. Hugh Gravelle and Diane Dawson, for helpful comments that led to substantial improvement of the paper. Jiménez Rubio is also grateful for funding from Fundación Ramón Areces (Madrid, Spain).

[1]Department of Economics and Related Studies, University of York, Heslington, York, YO10 5DD, United Kingdom. Tel:+44 (0) 1904 433784; Fax: +44 (0) 1904 433759. Email: pcs1@york.ac.uk
[2]Centre for Health Economics, University of York, United Kingdom.

1 Introduction

Fiscal decentralisation[1] reforms are producing an ongoing restructuring of the public sector all over the world. On the one hand, major decentralisation efforts can be appreciated in a variety of industrialised countries. In the United States, for instance, the primary responsibility for a number of social programs has been shifted back to the states[2]; in the United Kingdom decentralisation movements have brought about the foundation of Scotland and Wales' own parliaments; and in Italy, Spain, and other countries, there has been increasing fiscal powers for regional and local authorities. The traditional theory of fiscal federalism, the economic literature on decentralisation, identifies advantages that have encouraged these reforms (Oates, 1999). On the other hand, a great deal of interest in the fiscal decentralisation issue has also emerged in the developing world. In this case, decentralisation is mainly regarded as a political alternative to the central planning failure to achieve continuous economic growth (Akin et al, 2001).

As maintained by the fiscal federalism theory, decentralisation of public goods and services with localised effects is likely to produce efficiency gains (Oates, 1972). However, this prescription is a very general one, since what is considered as "local" is expected to vary across settings. In the health care sector, in particular, there is little guidance concerning the most efficient level of provision of health goods and services. But a trend towards health care decentralisation is becoming evident in many nations. Thus, in the United States, Medicaid is one of the programs for which important allowances of federal au-

[1] In this study decentralisation is synonymous with devolution because it is merely concerned with the impact of political decentralisation. Devolution is a political reform designed to promote autonomy at the local level. See Hunter et al. (1998, p.311-3) for a detailed classification of the different types of decentralisation, namely deconcentration, delegation, and devolution.

[2] In 1996 a set of measures were passed in the United States that replaced highly regulated matching grants to the states for various welfare programmes by a system of block grants with few strings attached.

thority have been devolved to the states[3]; in the United Kingdom health services are one of the basic responsibilities of the new Scottish and Welsh parliaments; in Spain and Italy, legislative powers have been combined with an augmented fiscal autonomy in the health care area. In developing countries, on the other hand, the increasing decentralisation of health care services has been mostly a response to the impetus in the promotion of primary health care by international donor organisations, such as the World Health Organisation (WHO) or UNICEF[4](Akin et al, 2001).

At the other extreme are countries such as Portugal, that has traditionally been a very centralised country, or Norway where legislation introduced in 2002 moved powers over hospitals away from local governments (WHO, 2005). And even in federal countries like Australia or Canada, the central government often preserves substantial oversight and regulation autonomy (Lazar et al., 2002). Yet the most appropriate level for the decentralisation of health policy is an important unresolved issue in the research literature. In spite of this, it is surprising the little attention that has been paid to the evaluation of decentralisation in the health care sector, as opposed to the relatively broader literature about the effect of decentralisation on government size, economic growth or government quality. Because decentralisation is often politically motivated, much of the literature has stressed the advantages of sub national provision of health services and its possible limitations.

This paper attempts to address the lack of formal analysis of the effects of decentralisation in health care. In the next section we discuss decentralisation from the perspective of the fiscal federalism literature. Section three explores

[3]Medicaid is the public health care insurance agency for low income individuals in the United States.

[4]Policy documents from these institutions such as the Primary Health Care Declaration of Alma Ata (1978) or the Health for All by the Year 2000 (1981) emphasised the importance of primary health care and the role of community participation in planning and providing health services. Authors such as Collins and Green (1994), among others, stress the incompatibility between promotion of primary health care and centralised health systems.

the main implications for the health care sector that follow from this literature, while section four present some evidence on the issue of decentralisation and health care. In section five we present the theoretical model of public finance that we use to test the main hypothesis of the fiscal federalism theory applied to health care. The remainder of the paper concentrates on our empirical analysis.

2 The theory of fiscal federalism

The economic literature on decentralisation is usually referred to as the fiscal federalism theory. This theory basically analyses the vertical structure of the public sector, that is, the optimal assignment of functions to different levels of government, and the most appropriate fiscal instruments for carrying out these functions.

While Hayek (1945) and Tiebout (1956) led the earlier discussion on some of the key benefits of decentralisation[5], the foundation for most of the conventional literature of fiscal federalism is the study of the public sector carried out by Richard Musgrave (1959) within a welfare economics' framework. According to Musgrave's analysis, the public sector should intervene in the economy to address the market inability to: attain the most equitable distribution of income (*distribution* function); maintain a high level of employment and stable prices (*stabilisation* function); and establish an efficient pattern of resource use (*allocation* function). The main conclusion from Musgrave's study is that an economic case for a federal structure of the public sector exists. Thus, while the stabilisation and the redistribution functions are traditionally assumed to be best placed at the central government's level[6], decentralised tiers of government are left with the primary responsibility of providing *local* public goods and services[7]. This proposition was later formulated by Oates (1972, p.54) into the

[5] Hayek (1945) emphasised the ability of sub national governments to make decisions concerning local circumstances. Tiebout (1956) stressed the role of competition among local governments in allowing citizens to match their preferences with a particular menu of local public goods (*people vote with their feet*).

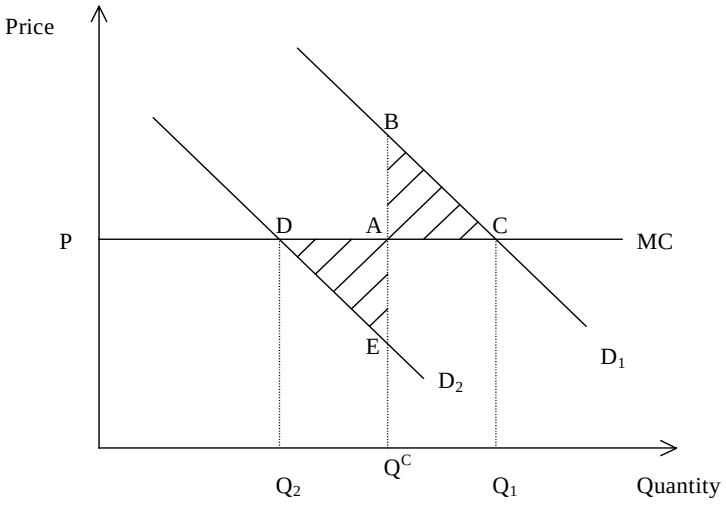
[6] With respect to the stabilisation function, local governments are believed to have limited means to impose a macroeconomic control of their economies (due to their poorer knowledge of the relevant economic variables, their lower capacity to use automatic stabilisation instruments, such as progressive income taxes, etc). Decentralised programs to redistribute income may result in sub optimal levels of support for the low-income individuals in the presence of mobility of economic units.

[7] Local goods and services are those for which the sum of the marginal benefits to all the residents in a jurisdiction equals their marginal costs. Public goods and services incorporating substantial spillover effects are assumed to be inadequately provided by decentralised levels of government, as they entail a potential for free riding behaviour which might result in an

Decentralisation theorem.

The *Decentralisation theorem* basically postulates, on grounds of economic efficiency, a presumption in favour of sub national provision of local public goods and services: given that local preferences and costs of a local public good or service are likely to vary across jurisdictions, decentralisation could increase economic welfare in the society as a whole. The key point is that sub national governments have access to better information about local circumstances than central authorities, and therefore can use this information to tailor services and spending patterns to citizen's needs. In contrast, centralised government structures face significant informational and political constraints that are likely to prevent them from providing an efficient level of a local public good or service.

Figure 1: The welfare losses from centralisation



inefficient level of provision.

The above diagram shows that under centralisation the quantity of the public good provided to individuals is given by Q^C (Oates, 1989). This quantity is lower than the amount demanded by representative individual 1 but more than would be demanded by representative individual 2. As a consequence, each of these individuals will experience welfare losses as shown by triangles ABC and ADE. Triangle ABC indicates the loss caused by individual 1 not consuming as much as she should if area one could decide on the amount of the good to be provided. Individual 1 is willing to pay $Q^C BCQ^1$ to get $Q^C Q^1$, even though these would only cost $Q^C ACQ^1$ to be available.Triangle ADE shows the loss experienced by individual 2 because he is consuming more than he would otherwise choose. This individual is paying $Q^2 DAQ^C$ for consuming $Q^2 Q^C$ units of the good while valuing them at only $Q^2 DEQ^C$. Welfare deadweight losses from centralisation are greater the greater the heterogeneity and the more inelastic the demand curves are.

A corollary to the Decentralisation theorem states that the gains in allocative efficiency are further enhanced by the increase in competition among local governments that decentralisation might bring about (Oates, 1972). At the same time, competition is expected to increase productive efficiency as a result of the greater experimentation and innovation in the production of public goods and services than if those goods or services were provided by the central government[8]. As a consequence, production costs (and therefore, prices) could be lowered and the quality increased.

Potential gains to be realised from decentralisation are however conditional on the existence of decentralisation of political decision-making authority. In

[8] One of the reasons for this is that decentralised governments have more freedom in implementing new production methods, whereas the central government would not embark in a new production technique unless it has gained acceptance in all the local areas. The increased competitiveness among local governments when one of them implements a new production method that turns out to be superior might also serve as a stimulus to innovation (King D, 1984).

particular, effective channels for the individuals to express their preferences, and incentives for the policymakers to respond to those preferences are implicit assumptions in most of the proposed benefits of decentralisation. For this reason many authors are sceptical about the successful implementation of decentralisation in less developing countries, given their weaker administrative capacity and their lower initial levels of democracy as compared to developed countries (Khaleghian, 2003).

In a well-known paper, Tiebout (1956) argued that citizens "vote with their feet" and choose to reside in the jurisdiction that provided the service mix best suited to their tastes. Whether the mobility of citizens characterises the European setting as well as the US one is a matter for debate (Oates, 1999). However, although the gains from decentralisation are enhanced by such mobility, they are not fully dependent on them. Even in the absence of mobility the efficient provision of a local public good will be determined by the condition that the sum of the marginal rates of substitution equals the marginal cost, a condition that will usually differ among jurisdictions (Oates, 2005).

The literature of fiscal federalism has identified several arguments in favour of centralisation of the provision of local public goods and services. Some of these arguments contradict those above discussed. For instance, it has been argued that in decentralised settings information can be distorted and oversight weakened. Heavy dependence on transfers may discourage fiscal discipline at lower levels of government, as central governments are more likely to be held responsible for any services' failures (Rodden, 2003). Moreover, local governments may claim high spending needs in order to secure a higher share of central funding. On the other hand, too much financial autonomy given to local authorities may result in inefficient levels of provision under decentralisation if competition is exercised on tax rates rather than on services (Oates, 1999). The existence

of economies of scale and/or externalities in the provision of a public good has also been often put forward as an economic argument for a certain central control. On the equity side, it has often been claimed that likely differences in tax bases among jurisdictions would inevitably result in inequalities across them unless the central government carries out a strong redistribution of resources from richer to poorer areas.

To provide local public goods and services, the central government transfers some taxes to local governments and grants them a certain taxing power. Usually, taxes with a mobile base (e.g. corporate taxes, value added taxes) remain under central government's control, whereas taxes with a relatively immobile base (e.g. property tax) are levied at the sub national level and constitute local government's own source of revenue. The theory of fiscal federalism identifies three arguments whereby central grants are necessary to guarantee an adequate provision of a local public good: vertical fiscal imbalances, horizontal fiscal imbalances and externalities[9]. At the same time, this literature recognizes the risks of an excessive dependency on transfers from the central government. First of all because, since transfers often come with strings attached, a disproportionate reliance upon them might result in an unnecessary interference from the central government. And in second place, because systems heavily dependent on grants place little pressure on local governments to manage spending efficiently (Oates, 1993). In particular, a common finding by the literature in this issue is that local spending is much more responsive to increases in intergovernmental trans-

[9] The horizontal fiscal imbalances argument emphasizes that because local jurisdictions have different tax bases at their disposal, the need to ensure that citizens have access to a roughly equal level of public services will require some degree of redistribution among jurisdictions by means of central grants. The vertical fiscal imbalance argument is referred to the fact that in most nations, central governments have excess of revenues, while local governments are not self sufficient to cover the costs of the services that they provide. To solve this problem the central government provides the local government with transfers to help them to cover their costs. Finally, the externality argument stresses that some local public goods may have spill over effects that are not taken account in their provision. Therefore, central matching grants may be required to ensure an efficient level of supply of the local public good.

fers than to equal increases in private income, a phenomenon which has become known as the *flypaper effect*[10]. Thus, at the margin, an additional pound spent in public goods and services does not seem to be equivalent to the benefit of an equal reduction in taxation.

Although most of the literature on federalism relies upon the principles of welfare economics there is a more recent school of thought that examines decentralisation with a strong emphasis on political economy. Unlike the conventional school of thought, these so-called *second generation theories* assume that governments are not necessarily interested in maximising social welfare[11]. However, as Oates (2005) points out these new theories do not contradict the old ones, but provide new perspectives on how to think about the centralisation versus decentralisation issue in the public sector.

[10] Since transfers are politically inexpensive sources of revenue for local governments, local politicians will not be encouraged to reduce taxes as a response to increases in grants, but will have more incentives to do so if local taxation revenue rises as a consequence of an increase in local income. Moreover, by breaking the links between costs and benefits, transfers make it difficult for voters to identify and penalize the causes of local inefficiencies in the use of the resources (Rodden, 2003).

[11] Some of the issues addressed by this new school of thought are the extent of political participation in a decentralised government (Inman and Rubinfield, 1997), or the role of decentralisation in containing the size of the public sector (Brennan and Buchanan, 1980). For a review of both the new and the old literature on fiscal federalism see Oates (1999).

3 Decentralisation and the health services

Economic theory provides only with a limited guidance in deciding how to allocate expenditure responsibilities among different levels of government. In principle, the central government should be responsible for national public goods, redistributive and stabilisation policies, whereas lower levels of government should provide local public goods. However, in reality most of the goods provided by the public sector do not correspond exactly to any of these categories, and the territorial limits are difficult to specify. Health care constitutes an example of goods with a mixed nature. In addition to individual benefits, health care provision generates important social externalities (Ahmad and Craig, 1997). Examples of policies in the health field with consequences for citizens in all areas include disease control and environmental pollution regulations. Also, public health interventions amongst younger people benefit other areas where these individuals reallocate later in their lives (Levaggi and Smith, 2005). Many other health policies, such as food hygiene or water fluoridation regulation, mainly affect local areas.

The existence of externalities in health care do not necessarily imply centralised provision as a superior alternative, since there might still be welfare gains from decentralised provision relative to a centrally determined level of health care services (Oates, 2005). Moreover, as discussed in the previous section providing local governments with subsidies may encourage efficient levels of health services to the point where the marginal social benefits for society as a whole from the provision of health care equals marginal costs. Following the fiscal federalism position the main argument for decentralising decision making in health is that local decision makers have greater knowledge of the health needs of their populations and of local conditions that affect the production of health care than national policy makers. According to the Decentralisation theorem

the potential efficiency gains from decentralising the health services might lead to an improvement in the health of the population if decentralisation of health services enables an increase in the quality of health inputs, and if these health inputs adjust to the particular preferences/needs of the local citizens. However, although decentralisation can result in greater total health gains, it may also lead to increased inequalities in health care.

The extent to which the Tiebout characterisation enhances welfare gains in health care is likely to be marginal. Mobility of citizens to areas that provide their preferred health care system is generally limited to patients with chronic conditions (HIV, diabetes) or very old people with high level of health care needs. For these patients closeness to health services of high quality (or low levels of user charges) might be an important consideration in choosing their jurisdiction of residence (Levaggi and Smith, 2005).

On the other hand, there are also some economic arguments put forward for centralisation that are relevant for health care. Regarding economies of scale, central intervention is considered as necessary to prevent inefficient location of facilities such as hospitals by local decision makers accountable to local electors (Gravelle, 2003). Another argument frequently adduced for some central intervention in health care is the more efficient pricing of inputs by a single purchaser of health care. National bargaining is believed to secure more favourable contracts with service providers as compared with a situation in which local purchasers may have to accept the prices set by monopoly suppliers. A further concern related to the provision of health services is that local authorities, under pressures to raise their own revenues, may have to rely on user fees to finance their services. Central intervention is again required in this case to guarantee that local authorities are able to provide a similar level of health care services to the citizens in their constituencies.

4 Evidence on decentralisation and health outcomes

There is little evidence that countries with a more decentralised health system have better health outcomes. So far only a limited number of studies have attempted to measure the magnitude of the effect of public sector decentralisation on health outcome indicators. On the whole these studies find a beneficial effect of decentralisation on indicators of health outcomes.

Mahal et al. (2000) use data from rural villages in India for 1994 to test the hypothesis that decentralisation is positively associated with child mortality once the effect of socio economic factors, civil society organisations, and so on, are controlled for. They have used dummy variables for states that have significantly moved towards decentralisation during the period 1970-94, and the frequency of local body elections during the same period as proxies for decentralisation. While the estimated coefficients for decentralised states have the expected positive signs, the election frequency variable is statistically insignificant. The study by Asfaw et al. (2004) corroborates the previous results for rural India using an index of fiscal decentralisation obtained by factor analysis[12] on the basis of three variables[13] for the period 1990-1997. Their results also show that the effectiveness of fiscal decentralisation increases with the level of political decentralisation[14].

In the study by Robalino et al. (2001), a panel data of low and high income countries is used to test how a measure of fiscal decentralisation -the propor-

[12] Factor analysis is a statistical technique that can be used to summarise a set of correlated variables into a single measure.

[13] These are: the share of local (rural) expenditure on the total state (intermediate government tier) expenditure, the total local expenditure per rural population, and the share of local own revenue from the total local expenditure.

[14] Political decentralisation is measured by an index constructed on the basis of factorial analysis from total voter's turnout, women's participation in polls and the number of polling stations per elector in each state.

tion of sub national government spending over central government spending-, affects infant mortality rates over the period 1970-1995. After controlling by a set of structural variables (GDP per capita, corruption, ethno-linguistic fractionalisation, etc.), one of the main results of the fixed effects estimation is that decentralisation is associated with lower infant mortality rates. Interestingly, the marginal benefit from decentralisation is found to be greater at low-income levels.

Using a panel data of 29 Chinese provinces for the period 1980-1993, Yee (2001) examines the relationship between several indicators of health care performance -the number of doctors per 10.000 people, mortality rates, hospital beds per 10.000 people, and local health care expenditures-, and various measures of decentralisation. These include two indicators of fiscal decentralisation –the ratio of local government expenditure to central government expenditure, and the ratio of local government expenditure to total government expenditure-, and two other indicators of political decentralisation[15]. The results of the regressions, based on either fixed effects or random effects estimations, show that fiscal decentralisation has been beneficial to the health sector in terms of reducing mortality rates and increasing local expenditure on health care.

Ebel and Yilmaz (2001) employ an intervention analysis[16] to evaluate the outcomes of decentralisation in terms of immunisation rates for DPT[17] and measles of children under 12 months in six developing countries (Argentina, Brazil, Colombia, Philippines, South Africa and Venezuela) during the period 1970-1999. The results of the estimated fixed effects model suggest that inter-

[15] The political decentralisation measures are: bureaucratic distance –an index of top provincial officials proximity to the province-, and state industry decentralisation –the proportion of industrial output from state owned enterprises controlled by local government on the total industrial output from all state owned enterprises in a province-. These measures were found to be statistically insignificant in explaining variations in health care performance indicators over time.

[16] An intervention analysis involves a test of the change in the mean of a variable as a result of a policy reform.

[17] A series of three vaccines against diphtheria, pertussis (whooping cough) and tetanus.

vention by sub national governments has been associated with an increase in the coverage of children immunised for measles.

The study by Habibi el al. (2001) shows that the percent of revenue raised locally and the proportion of controlled revenue over the total have a negative and significant association with infant mortality rates for a panel of Argentinean provinces over the period 1970-1994. In addition, the authors find that during the period of decentralisation reforms studied, regional inequalities were considerably reduced.

Khaleghian (2003) examines the association between decentralisation and immunisation coverage rates for DPT3[18] and measles of children at one year of age in 140 low and middle income countries during the period 1980-1997. The main indicator of fiscal decentralisation used in this study is a binary variable defined as the presence of taxing, spending, or regulatory authority on the part of sub national authorities. Two other decentralisation indicators were used to double check the results: the share of sub national expenditures on total government expenditures, and the share of health spending on total sub national expenditures. The model also included several control variables (GDP per capita, illiteracy rate, democracy score, ethnic tension, etc). The findings suggest that decentralisation improves coverage rates only in low-income countries[19].

[18] The third and last vaccine of the DPT vaccine series.

[19] From a theoretical perspective, decentralisation of immunisation services –services with important externalities-, is expected to encourage local jurisdictions to "free-ride" on the immunisation status of their neighbours. The result could be a sub optimal disease protection level provided in the country as a whole. Contrary to these predictions, in both Ebel and Yilmaz´s and Khalegian´s studies, a positive association between decentralisation and immunisation coverage rates is found on their sample of developing countries. Khaleghian (2003, p. 27) has pointed out that this finding "may reflect a salutary balance between the proximity of local authorities to the community, and the preservation of central influence and bureaucratic autonomy, both of which are essential to the effective functioning of an immunisation program".

5 A theoretical model of decentralisation and health outcomes

To conceptualise the relationship between decentralisation and health outcomes we have developed in this section a simple model of decentralisation in health care taken from the public finance literature (e.g. Albi et al., 1992). The theoretical foundation of this model is that a local policy maker wishes to maximise the level of utility in his community, where utility depends on the consumption of a public and a private good. For the purposes of this study, we consider individual utility to be determined by consumption of a private good and health status, where health status is a function of expenditure in health care, social capital and decentralisation. Although some of the premises of this model may be unrealistic, it defines a first best scenario that can be used as a reference for an empirical analysis about the effects of decentralisation.

The primary assumptions of the model are:

- Each locality consists of N individuals.
- Individual utility depends on health outcomes and expenditure on a private good:

$$U = U(H, x) \tag{1}$$

$$\frac{\partial U}{\partial H} > 0; \frac{\partial U}{\partial x} > 0; \frac{\partial^2 U}{\partial H^2} < 0; \frac{\partial^2 U}{\partial x^2} < 0$$

where:

H: health outcomes

x: expenditure on a private good

$U(.)$ can be interpreted either as the preferences of the representative consumer if all the individuals are considered as identical, or as those of the median voter.

- Health outcomes depend in turn on health expenditure –local government health expenditure (Y_l) and other non-local government health expenditure (Y_{nl})-, social capital (S), and the level of decentralisation (D):

$$H = H\left(Y_l, Y_{nl}, S, D\right) \tag{2}$$

$$\frac{\partial H}{\partial Y_l} > 0; \frac{\partial H}{\partial Y_{nl}} > 0; \frac{\partial H}{\partial S} > 0; \frac{\partial H}{\partial D} \gtreqless 0; \frac{\partial^2 H}{\partial Y_l^2} < 0; \frac{\partial^2 H}{\partial Y_{nl}^2} < 0; \frac{\partial^2 H}{\partial S^2} < 0; \frac{\partial^2 H}{\partial D^2} \lesseqgtr 0$$

- Local governments can identify individual's preferences over health and the private good and use this information to maximise overall welfare.
- Local health care services are financed by means of local lump sum taxes and by transfers granted by the central government.

The initial resource constraint in the locality is given by:

$$Y_l + X = I \tag{3}$$

where:

I: total income in the locality

X: income available for spending in the private good (I – local taxes)

While local income (I) is fixed, the amount spent in the private good (X) depends on the consumer's preferences for health care. In the absence of central government's transfers, the amount of local taxes must be equal to the local spending in health care, i.e., $Y_l = I - X$.

After central government's transfers (M) the budget constraint becomes:

$$Y_l + X = I + M \tag{4}$$

The maximisation problem faced by the local government can then be defined as follows:

$$\begin{aligned} &\underset{Y_l}{Max.} U(H(Y_l, Y_{nl}, S, D), X) \\ &subject\ to: Y_l + X = I + M \end{aligned} \tag{5}$$

Maximising $U(.)$ with respect to $H(.)$ and subject to 4 gives:

$$Y_l^* = f(I, M, Y_{nl}, S, D) \tag{6}$$

Now,

$$H^* = g(Y_l^*, Y_{nl}, S, D) \tag{7}$$

And hence,

$$H^* = g(I, M, Y_{nl}, S, D) \tag{8}$$

Either solution 7 or 8 of the maximisation problem provide a useful starting point to model the relationship between health care decentralisation and health outcomes.

Graphically, the generic model can be represented as follows:

Figure 2: Individual's optimal consumption

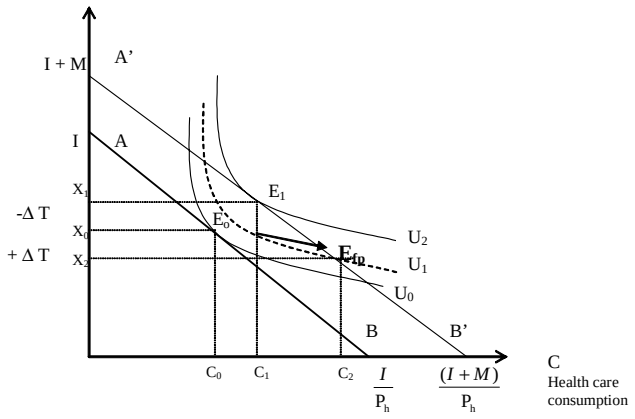


Figure 2 shows the equilibrium solution when the local governments are provided with a grant from the central government. Line AB represents the initial budget constraint of the jurisdiction, as described in equation 3. Without central government's grants, the equilibrium is achieved at point E_0, where the indifference curve is tangent to the budget line AB. At this point, the optimal local health care spending (Yl^*) is $P_h * C_0$, and the optimal private consumption X_0.

After central government transfers, the budget constraint shifts outwards in a quantity equal to the amount of the transfers (M). Line $A'B'$ represents the post-grant budget constraint of the locality, as shown in equation 4. If health care and the private good can be considered as normal goods, the post-grant equilibrium should be achieved at a point such as E_1, where the consumption of both goods increase (from C_0 to C_1, and from X_0 to X_1, respectively). Grants release private funds by the amount $\triangle T$, inducing individuals to consume more of the private good in order to maximise their utility.

However, researchers often find that the post-grant equilibrium is achieved in a point like E_{fp} rather than in a point like E_1. Equilibrium point E_{fp} is sub-optimal compared to E_1, since $U_1 < U_2$. This phenomenon, known as the flypaper effect, refers to the fact that, unlike income, grants induce an excessive increase in local government consumption[20]. As shown in figure 2, the new equilibrium point E_{fp} is characterised by an increase in taxation by the local government. At this equilibrium solution, health care consumption increases (from C_0 to C_2), while private consumption decreases[21] (from X_0 to X_2). Therefore, grants would not substitute for tax revenues collected by local authorities.

In addition to study the relationship between decentralisation and health outcomes, the above analytical framework allows us to examine whether decentralisation stimulates the growth of the health sector through the flypaper effect. This can be made by comparing the coefficients of income and transfers after estimating equation 6 above. A higher elasticity of transfers than of income with respect to provincial expenditure would indicate the existence of the flypaper effect.

[20] For a review of the empirical literature on the flypaper effect see Hines and Thaler (1995).
[21] Note that a flypaper effect could also exist even if private consumption remains constant following an increase in grants.

6 Measuring the extent of decentralisation of the health care system

All existing empirical studies on the relationship between decentralisation and health outcomes have used overall indicators of public sector decentralisation. This study departs from previous ones in examining the isolated effect of health sector decentralisation on health measures. We therefore focus on the effectiveness of health policy in improving a population's health. In doing so this work explores whether the recent trend of devolving health care responsibilities to the local level has an economic justification, independently of the decentralisation status of the remaining public policies.

A precise measure of health care decentralisation is however difficult to develop. Health care decentralisation is a complex phenomenon embracing a number of political, fiscal and administrative dimensions. Many of these aspects are, as yet, not easy to measure empirically, e.g., who determines the range of the services to be covered, who sets the regulatory framework, or who decides the financing mechanism of the system as a whole. The core question is to what extent health care policy is decided centrally or locally (Banting and Corbett 2002, p.6). However, up to now the only available quantitative measure of health care decentralisation is a fiscal one: the ratio of sub national health spending to the total health spending for all the levels of government. In the absence of more appropriate measures of decentralisation, similar fiscal decentralisation indicators (aggregated for all the public sector activities) have been widely used by the researchers in this field following Oates' pioneer work in 1972. According to Oates (1972, p.197), the extent of fiscal activities at each level of government is a major component in determining its influence on the allocation of resources. Moreover, in contrast to dichotomous indicators of decentralisation

or federalism[22], fiscal data reflect the *continuum* dimension of the decentralisation process.

The main source of the fiscal data is the Government Finance Statistics (GFS) of the International Monetary Fund (IMF). Although these statistics compile information for over 100 countries, however, a cross- country comparison on health care decentralisation based on IMF data is limited. The reason is that there is double counting of intergovernmental transfers in the estimates of IMF health spending. Thus, health transfers from higher levels of government are included both as a spending of this level of government, and as a spending of the recipient governments. When these transfers are equalisation grants provided to autonomous local governments, including these as expenditure by the donor underestimates the real level of local autonomy in health care decisions.

On the other hand, overall fiscal decentralisation indicators can be computed net of intergovernmental transfers. This is because in the IMF statistics there is an entry that identifies global transfers from higher levels of government to local governments. But even so, as several authors have pointed out (Ebel and Yilmaz, 2004; Rodden, 2003), overall fiscal decentralisation indicators based on IMF data are not without their problems. The main one is that spending and revenue measures of decentralisation may overestimate the level of real local autonomy. Local spending statistics include not only expenditures in functions controlled exclusively by the local jurisdictions, but also expenditures in functions controlled by higher levels of government (through directives, conditional grants, etc.). On the revenue side, the GFS classifies shared taxes, piggybacked taxes, and taxes in which the tax rate and/or base are decided by the central government as sub national own-source revenue. Recently, the OECD (1999) have made major efforts to improve the revenue information available for a set

[22] An example of this type of decentralisation measures is the binary variable used by Khaleghian (2003) that represents the presence of taxing, spending or regulatory authority on the part of sub national governments.

of countries by classifying taxes according to the tax autonomy entitled to local governments[23]. A major caveat of this data is that it only covers a small sample of countries and that it does not examine changes over time. In any case, neither the OECD nor the IMF revenue data are useful for the purposes of this study given that they do not disaggregate expenditures/revenues by function.

Given the limitations to make cross national comparisons on the basis of IMF health expenditure data, the focus of this study is on the outcomes of decentralisation of health care services in one of the most decentralised countries in the world: Canada. The main reason for choosing Canada as the case study is that for this country, a reasonable measure of health care decentralisation can be readily constructed. First of all, because the government health spending data, derived from the Canadian Institute for Health Information (CIHI), include federal transfers to provinces only as a spending of the provincial governments[24]. This means that unlike IMF based health decentralisation measures, our measure of health care decentralisation for Canada does not underestimate local autonomy. In addition, both the CIHI and the Canadian Statistics produce a great amount of data on government health spending –including health transfers- and health outcomes at the provincial level of analysis. By controlling for health transfers in our econometric estimations, we further reduce the potential problem of *overestimation* of local autonomy of expenditure-based indicators of decentralisation. In the Canadian context, however, this is not likely to be a serious problem given that since 1977 federal health transfers consist of

[23] The OECD (1999) provides a classification of sub national tax revenues ranging from a.) where the central government can set both the rate of taxation and the tax base, to e.) where sub national governments set both the tax base and the tax rate. Tax sharing agreements are further arranged into four categories from d.1.) where the sub national governments can determine the revenue split to d.4.) where the national government can unilaterally determine the revenue split.

[24] Instead of being included as an expenditure of the municipal government (on the basis of responsibility for payment), provincial transfers to municipalities for health purposes are included as a spending of provincial governments. This is however a minor problem in this context, given that provincial transfers to municipalities represent a very small proportion in total health spending.

a block funding with few conditions attached. Therefore, including them as a spending of provincial governments, and consequently as a positive determinant of health care decentralisation is not likely to provide a misleading picture of the autonomous decision making of Canadian provincial and municipal governments.

Table 1: Advantages and disadvantages of main fiscal decentralisation indicators

Variable	Source	Advantages	Disadvantages
Overall decentralisation			
Expenditures			
Proportion of sub national spending over the total	IMF	Long time series and cross sectional information	*Overestimation* of real level of local autonomy if local expenditures are tightly controlled by the centre
Revenues			
Proportion of sub national own revenues over the total	IMF	Long time series and cross sectional information	*Overestimation* of real level of local autonomy if tax revenues controlled by the centre are classified as sub national own-source revenue
Proportion of sub national own revenues over the total	OECD	Tax revenues can be classified according to the tax autonomy attributed to local governments	Small coverage: cross sectional information (1995) for 19 OECD countries
Health care decentralisation			
Expenditures			
Proportion of sub national health spending over the total	IMF	Long time series and cross sectional information	Double counting of health transfers as expenditure of both donor and recipient governments can lead to an *underestimation* of local autonomy if transfers constitute equalisation grants

7 An overview of Canadian health care system

Canada is a confederation of ten provinces[25] and three territories[26]. Health care services are mostly publicly financed and they offer comprehensive and universal insurance to Canadian citizens. Provision is left under private control.

Since Canada became a nation, following the Constitution Act of 1867, provinces have borne the primary responsibility for health care. Thus, among other functions, provinces regulate hospitals and other health institutions, they decide the financing schedules with health professionals, and they set global budgets for hospitals. Provincial governments are also responsible for the final health care costs of their jurisdiction (Banting and Corbett, 2002).

As for the Canadian territories, health services have been directly managed and delivered by the federal government until beginning of the 1980s (CFHC, 2002). The territories also have constitutional arrangements, determined by the Parliament of Canada, that differ from those of the provinces. Other distinctive characteristics of the Canadian territories include: large proportion of aboriginal population[27], sparse populations and a persistent shortage of resources that have required territorial governments to concentrate in the provision of primary services. Secondary services are often delivered through contractual arrangements with the provinces.

The federal government's role in the system is limited to the direct provision of health services to specific sectors of the population[28], and to the management of the activities of health protection, disease prevention, and health promotion (WHO, 2005). Federal influence has been mainly exercised through financial as-

[25] Newfoundland and Labrador, Prince Edward Island, Nova Scotia, New Brunswick, Québec, Ontario, Manitoba, Saskatchewan, Alberta, British Columbia.

[26] Yukon, Northwest territories, Nunavut.

[27] 20% of the population in the Yukon, 50% in the Northwest Territories, and 85% of Nunavut's population.

[28] These include veterans, native Canadians living on reserves, military personnel, inmates of federal penitentiaries and the Royal Canadian Mounted Police.

sistance to the provinces (Banting and Corbett, 2002). Adherence to some basic principles –*Canada Health Act*- in return from federal support has enabled the creation of a national plan for the health care system. These principles are: *universal coverage*; *public administration*[29]; coverage of all "medically necessary" services (*comprehensiveness*); *portability* of coverage outside the province; and prohibition of financial barriers to access health services, such as user fees or extra billing by physicians (*accessibility*). Within this broad framework, provinces have scope for determining the health policy of their insurance plans. This contrasts with the situation in other similar federations, where transfers from the central government usually come with specific conditions attached. In addition, the shift from conditional matching grants to a block funding grant for health and postsecondary services in 1977 –*Established Program Financing* (EPF)- gave provinces more autonomy in their health related spending decisions. On the other hand, since the introduction of the block funding the federal government has unilaterally reduced the amount of the transfer payments to the provinces (Armstrong and Armstrong, 1999). The most severe cutback to federal transfers took place in 1996 with the combination of funding for health care, postsecondary services and social services in a single block: the *Canadian Health and Social Transfer*[30] (CHST).

Provinces have faced the federal cutbacks by restricting coverage for new and

[29] This principle implies the prohibition of private insurance of services already covered by provincial insurance. For-profit coverage is limited to supplementary services such as pharmaceuticals, vision care, dental care, and chiropractors and podiatrists' services (WHO, 1996).

[30] During the EPF era the federal government allocated separate cheques for health care and post secondary education services. Although provincial governments did not always follow federal's spending patterns, these allocations are used as rough approximations of federal contributions to each of the sectors. With the introduction of the CHST the federal government stopped the allocation of different cheques among programs. As a consequence, it is no longer possible to know the exact proportion of federal transfers in provincial health budgets. Some estimates show that the proportion of transfers in provincial health expenditure has fallen from 26.9% at the beginning of the block funding period (1976/77) to 10.2% in 1998/99, three years after the introduction of the CHST (Ministry of Health and Long Term Care, 2000).

existing treatments and services[31] and by discharging responsibility for some services to municipalities, so far mostly in charge for public health. Provincial governments, with the exception of Ontario, have also responded to restrictions in federal funds by devolving control of some aspects of the system to the recently created regional boards. The specific nature of devolved authority diverges considerably among provinces. However, in all the cases the level of autonomy given to the boards is still highly restricted for two reasons. Firstly, regions receive budgets determined by provinces on the basis of historical spending patterns and have no revenue raising powers. Secondly, regional decision-making is constrained by provincial guidelines and by provincial determination of key health services –physician services and drugs- (Lomas et al., 1997).

[31] Most of reduction in public benefits has however taken place on voluntarily provided supplementary health services (WHO, 1996)

8 Data and methodology

To examine the model we developed in section 5, we use a panel of the ten provinces of Canada for the period 1979-1995[32]. Given their special status, and in order to obtain unbiased estimates, Canadian territories have been excluded from our analysis. We have used infant mortality rates from the Canadian Statistics as the measure of health status. Infant mortality has been considered as the single most exhaustive indicator of health in a society. It reflects child's health and pregnant women's health, in addition to the state of health development within the society. Moreover, infant mortality is superior to life expectancy, our alternative measure of health status, for two main reasons**.** Firstly, because infant mortality is more reliably measured than life expectancy. Infant mortality figures are based on actual data, whereas life expectancy figures are based on extrapolations from child mortality data and assumed life tables. Secondly, because infant mortality is more sensitive to policy reforms such as decentralisation than life expectancy. In fact, disparities in the risk of infant death are higher than disparities in life expectancy in Canada over the period studied[33].

Health decentralisation is defined as the proportion of sub national health spending in Canada –municipal, provincial, and Worker's Compensation Boards[34] (WCB)- over the total –municipal, provincial, federal, and WCB-. The source of all the health spending data, as well as GDP per capita and population, is the *Canadian Institute for Health Information* (CIHI). The remaining control variables included in the model were taken from the Canadian Statistics. These

[32] This time frame is determined by the availability of data on health transfers (see section 7).

[33] The standard deviation of infant mortality is almost as twice as the standard deviation of life expectancy (1.9 and 1.1 respectively).

[34] The Worker Compensation Boards are province-based institutions in Canada that provide employed people with financial and health care assistance following work-related injuries or diseases (WHO, 1996).

include an indicator of social capital –education-, and a measure of needs -low birth weight-. Low birth weight was included as a control variable because it is considered to be an important determinant of infant survival. Table 4 in the Appendix provides a comprehensive definition of all the variables used in the empirical study.

On the basis of the theoretical analysis in section 5, two alternative ways of examining our model are possible. Given that in both methods the results should be roughly equivalent, we explore both of them and compare the results. The first method (*Model 1*) is based on equation 8:

$$INFMORT_{it} = \alpha_0 + \alpha_1 INC_{it} + \alpha_2 TRANSF_{it} + \alpha_3 DEC_{it} + \alpha_4 PRIV_{it}$$
$$+\alpha_5 FEDEXP_{it} + \alpha_6 MUNEXP_{it} + \alpha_7 EDUC_{it} + \alpha_8 LOWBIRTH_{it} + \lambda_t + \nu_i + E_{it}$$

The second approach (*Model 2*) is based on equations 6 and 7:

$$PROVEXP_{it} = \beta_0 + \beta_1 INC_{it} + \beta_2 TRANSF_{it} + \beta_3 DEC_{it} + \beta_4 PRIV_{it} +$$
$$\beta_5 FEDEXP_{it} + \beta_6 MUNEXPit + \beta_7 EDUC_{it} + \beta_8 LOWBIRTH_{it} + \lambda_t + \nu_i + E_{it}$$

$$INFMORT_{it} = \beta_0 + \beta_1 PROVEXPHAT_{it} + \beta_2 DEC_{it} + \beta_3 PRIV_{it} +$$
$$\beta_4 FEDEXP_{it} + \beta_5 MUNEXPit + \beta_6 EDUC_{it} + \beta_7 LOWBIRTH_{it} + \lambda_t + \nu_i + E_{it}$$

where:

i: 1,...10 (provinces)

t: 1,...,17 (1979-1995)

$INFMORT$: infant mortality rate per 1.000 live births

INC: provincial income per capita

$TRANSF$: per capita health care block grants from the federal government

DEC: health care decentralisation status

$PRIV$: private expenditure in health care per capita

$FEDEXP$: direct (non-grant) per capita federal expenditure in health care

$PROVEXP$: provincial expenditure in health care per capita

$MUNEXP$: municipal expenditure in health care per capita

$EDUC$: educational level

$LOWBIRTH$: proportion of low birth weights in all live births

$PROVEXPHAT$: predicted provincial expenditure in health care

λ_t: annual dummies

ν_i: provincial specific effects

E_{it}: disturbance term

The main advantage of using panel data estimation techniques is the attenuation of the problem of omitted variables. Panel data models control for individual heterogeneity, that is, inherent characteristics of the population of interest that are either unobservable or non-measurable (e.g. preferences, managerial skills)[35]. Fixed effects and random effects are the two most usual panel data methods. In our case, since the data exhausts the population –provinces- and the inferences are made with respect to the sample, the fixed effects version of the panel data estimator is the most convenient (Wooldridge, 2000)[36].

The use of a fixed effects panel data model, together with a wide range of control variables, intend to overcome the problem of oversimplification in modelling the complexities of decentralisation (Khaleghian, 2003). In addition, a series of year dummies has been included to account for the impact of period specific shocks (e.g. variations in local tax shares).

Reverse causation of some of the variables is a concern in each of the regression equations. In the *health outcomes* model INC, DEC, $PRIV$,

[35] For a more detailed discussion of the advantages and disadvantages of panel data see Baltagi (1995, p.3-7).

[36] The main difference between the fixed and the random effects models is that the former considers differences across units as constant, while the later considers these differences as randomly distributed across cross sectional units. The random effects approach would be more convenient if the cross sectional units were known to be drawn from a large population. For the sample of Canadian provinces, however, it is more reasonable to assume that the model is constant over the period studied.

Table 2: Descriptive statistics for estimation sample

Variable (ln)		Mean	Std. Dev.	Min	Max
Infmort	overall	2.07	0.26	0.47	2.58
	between		0.11	1.91	2.23
	within		0.24	0.63	2.74
Inc	overall	10.03	0.23	9.57	10.56
	between		0.22	9.74	10.42
	within		0.09	9.79	10.19
Transf	overall	6.34	0.06	6.23	6.44
	between		0.01	6.33	6.35
	within		0.06	6.24	6.43
Dec	overall	-0.05	0.02	-0.12	-0.02
	between		0.02	-0.09	-0.03
	within		0.01	-0.10	-0.02
Priv	overall	6.37	0.18	5.88	6.87
	between		0.11	6.22	6.53
	within		0.14	5.98	6.72
Fedexp	overall	4.26	0.46	3.30	5.33
	between		0.38	3.74	4.96
	within		0.28	3.48	5.01
Provexphat	overall	7.36	0.12	7.03	7.59
	between		0.08	7.23	7.50
	within		0.08	7.15	7.47
Munexp	overall	2.03	1.37	0.00	4.92
	between		1.37	0.03	3.48
	within		0.43	-0.40	4.52
Educ	overall	-2.95	0.67	-4.00	-1.35
	between		0.66	-3.64	-1.62
	within		0.25	-3.50	-1.94
Lowbirth	overall	1.69	0.09	1.39	1.86
	between		0.07	1.58	1.81
	within		0.06	1.49	1.92

$TRANSF$, $FEDEXP$, and $MUNEXP$ are all regarded as suspected endogenous variables. It has been argued that while income leads to better health, good health may also contribute to improve living standards (e.g. Fogel, 1994). Although reverse causation between income and infant mortality is not likely to be crucial, INC has been considered as potentially endogenous in the first

instance. DEC could be endogenously determined because one of the main arguments for health care decentralisation –improvement in population's health- is also used in many contexts to claim for the implementation of decentralisation reforms. $PRIV$ could also be endogenous because voluntary spending decisions regarding health are based on the expectation of a high value for money (Filmer and Prichett, 1997). The likely endogeneity of $TRANSF$, $FEDEXP$ and $MUNEXP$ arises from the fact that larger amounts of money devoted to public health care might be driven, amongst other things, by the desire to reduce the aggregate levels of infant mortality in the population.

In the *health spending* model, there is a potential for reverse causation in variables DEC, $TRANSF$, $PRIV$, and $MUNEXP$. Simultaneity in the relationship between DEC and health expenditure makes sense since, by construction, an increase in provincial health expenditures leads to an increase in the DEC indicator. $PRIV$ and $MUNEXP$ are both likely to be endogenous since, as we have seen in section 3.3, in the face of financial hardships provincial governments have offloaded the responsibility for some health services to the municipal and the private sectors. A further reason that may explain causality of private health expenditures is the existence of a *quality effect* with respect to the aggregate spending in the public sector. This is explained by the fact that an insufficient level of public spending in health may cause a poor performance of the public sector, which in turn might lead to an increased reliance in the private sector, e.g. through more consumption of drugs as a substitute for direct care. Given that intergovernmental grants are usually related to the previous spending patterns of local governments, many empirical studies analysing the impact of transfers on recipients' governments spending have considered grants as endogenous variables[37]. However, in the Canadian setting the endogeneity

[37] The endogeneity problem of transfers is especially important under matching grant programs, since transfers are directly related to the local spending level through the matching rate (Gamkhar and Oates, 1996).

of transfers is not likely to be a very serious problem since from the start health transfers have been allocated to the provinces on a pure per capita basis (Ministry of Health and Long term Care, 2000). Moreover, the volume of the block transfer has been very unpredictable over the period.

If any of these variables is actually endogenous, standard *Ordinary Least Squares* (OLS) analysis will yield inconsistent estimates. In this situation *Instrumental Variables* (IV) techniques would be required in order to consistently estimate the parameters of this model (Wooldridge, 2000). On the first stage, the endogenous variable is regressed on a set of instruments. On the second stage, we apply OLS to the final equation where the endogenous variable is replaced by the corresponding "endogenity-purged" predicted value obtained in stage one.

Instruments must satisfy two requirements: high correlation with suspected endogenous variables and no correlation with the error term. In general, the consequence of excluded instruments with little explanatory power is an increased inconsistency in the IV estimates when instruments are not strictly exogenous (Shea, 1997). If the excluded instrument is exogenous but has low power, then conventional asymptotics fail. The relevance of the instruments in explaining the endogenous regressors has been checked by using the *Shea* R^2 statistic and the *Bound* R^2 statistic[38] (Baum et al., 2002). As for the second of the requisites, the validity of the instruments, we have computed the C statistic in addition to the *Hansen*-Sargan statistic[39].

[38] Baum et al. (2002) recommend the use of the *Shea* R^2 statistic when there are multiple endogenous regressors. This statistic can be interpreted as any other coefficient of determination. When there is a single endogenous regressor, the relevance of the instruments can be tested either by examining the joint significance of the excluded instruments in the first-stage regression or by computing the *Bound* R^2 statistic. The latter represents the squared partial correlation between the excluded instruments and the endogenous regressor.

[39] The *Hansen-Sargan* test evaluate the validity of the instruments, i.e. if the instruments are uncorrelated with the error term and if they are correctly excluded from the regression equation (null hypothesis). The *Hansen*'s J statistic is consistent in the presence of heteroskedasticity and (for heteroskedasticity and autocorrelation consistent estimation), autocorrelation, whereas the *Sargan* statistic is not. When the number of excluded instruments

In the context of time series data, there is a natural source of instruments: the lagged values of explanatory variables in the system (Wooldridge, 2002). The use of lags of the explanatory variables could be very useful for our empirical study, as it reduces the effort in finding such a large number of instruments for estimation. However, the use of internal instruments rules out autocorrelation in the error term. This is because in the presence of autocorrelation, lagged instruments are correlated with the current error term[40]. Therefore, our decision as to whether use internal instruments or not has followed the results of a test for autocorrelation.

On the other hand, employing IV when the suspected endogenous regressors are in reality exogenous is inefficient (Wooldridge, 2000). This is a consequence of the addition in the regression equation of additional sources of uncertainty: the instruments. In consequence, we have applied the C test to the potential endogenous regressors to test for the adequacy of OLS and the need to perform an IV estimation[41].

is very large, the *Hansen-Sargan* statistic may have weak power. In these cases, Baum et al. (2003) suggest the use of the C test. The C test is a general test for the adequacy of both included and excluded instruments. It has the null hypothesis that the specified instruments are valid.

[40] If we knew that autocorrelation was of order j, a solution would be to use lags $t+j+1$ of the explanatory variables as instruments. However, the explanatory power of such instruments is not likely to be high.

[41] When applied to suspected endogenous regressors, the C test can be interpreted as a *Hausman* type test with the null hypothesis that the OLS is a consistent estimator. In addition to being computationally easy, and by contrast to the Stata's Hausman test, the C test guarantees a non-negative statistic following a *robust* estimation (Baum et al., 2002).

9 Estimation results

In Table 3 we present the results from the estimation of the single equation model (*Model* 1), and the two-steps model (*Model* 2). $FEDEXP$ has been excluded from all regression equations given its high intercorrelation with the remaining regressors, especially with DEC[42] (Maddala, 2001). $FEDEXP$ has been instead used as an instrument for DEC.

The results for the health spending equation, the first equation of the two-stages model (*Model 2*), are reported in column 4 of Table 3. We have first run an IV regression technique. The *Durbin Watson* statistic (=0.78) prompted us to reject the null hypothesis of no autocorrelation of residuals. In consequence, we have not used lags of the suspected endogenous variables as instruments. Variables used as instruments for DEC, $TRANSF$, and $PRIV$ include[43]: $FEDEXP$, population and its squared value –POP and $SQPOP$-, population over 65 years –$POP65$-, the squared value of INC –$SQINC$-, and a measure of personal disposable income –PDI- (see Appendix). Only DEC turned out to be endogenous. Contrary to other studies on the impact of intergovernmental transfers in local expenditure (e.g. Gamkhar and Oates, 1996), we have not found endogeneity of $TRANSF$. However, as we discussed in the previous section, this finding is reasonable for the case of Canada.

We have then run another IV regression treating DEC as the only endogenously determined variable. The *Bound* R^2 (=0.82) reveals that the instruments used ($FEDEXP$, POP) have high power in explaining DEC. The *Hansen* statistic (p-value =0.15) suggests that instruments employed are adequate**.** Only

[42] The coefficient of determination of a regression of $FEDEXP$ against the other explanatory variables is 0.99. Given that the high level of multicollinearity made it difficult to estimate precisely the individual parameters we decided to drop $FEDEXP$, as we are mainly concerned about the coefficient of DEC.

[43] MUNEXP has been considered as exogenous because when it was included as an endogenous variable, the explanatory power of the instruments for the rest of endogenous variables was substantially reduced.

INC appears to have a significant role in determining provincial health spending. On average it is estimated that, other things equal, a 1 per cent increase in income leads to a 0.2 per cent increase in provincial health spending[44]. The coefficient of $TRANSF$ is not statistically different from zero. Therefore, it is likely to be zero. A priori, the coefficients of INC and $TRANSF$ are very similar in terms of their economic significance. Moreover, a formal test on the statistical equivalence between the two coefficients could not be rejected at any conventional significant level (p value = 0.33). The existence of a roughly equivalent spending response to INC and $TRANSF$ contradicts one of the main features of the *flypaper effect*: federal grants do not lead to an excessive spending propensity from provinces. In consequence, the shift from a matching grant to a block-funding grant seems to have been successful in containing health costs at both the federal and the provincial government levels.

With respect to the health outcomes equation the results for both models are reported in the first and second columns of Table 3. Given that the Durbin Watson statistic is over 2 in both OLS estimated equations, we cannot infer autocorrelation of residuals. Therefore we have used the lags of the suspected endogenous variables as instruments, together with $FEDEXP$, POP and $POP65$. Only $PROVEXPHAT$ in Model 2 was found to be causally related to $INFMORT$ (see Appendix). In consequence, we have relied on IV techniques to estimate this model. According to the Bound R^2 (=0.6) and the C test (=0.96) the variables used as instruments for $PROVEXPHAT$ ($PROVEXPHAT\_1$, $POP65$) perform well.

The estimates for DEC are statistically significant in both Model 1 and Model 2. It is estimated that, *ceteris paribus*, a 1 per cent increase in decentralisation is associated with approximately a 4 per cent reduction in infant

[44] Since we have used a log-log specification for every regression equation, the coefficients of each parameter can be interpreted directly as elasticities.

mortality. The estimates for $PRIV$ are weakly significant in Model 1, although the coefficient is small: on average, a 1 per cent increase in private health spending is expected to lead to a 0.3 per cent increase in infant mortality. The positive estimated association between $PRIV$ and $INFMORT$ may be explained by the fact that in provinces with poorer health care systems, people choose to spend more in private services, but because the core health services are provided publicly, a higher private spending in health may still be associated with a high infant mortality. In addition, as the estimates for DEC in the health spending equation show, health gains arising from decentralisation have not been at a cost of a higher provincial health spending.

$PROVEXPHAT$ is statistically significant in Model 2, and the magnitude of the effect is very similar to that obtained for DEC. In order to compare the spending impact on infant mortality at each level of government, we have introduced a third model (Table 3, third column) in which we have kept $PROVEXPHAT$ from *Model 2*, but DEC has been replaced by $FEDEXP$. As with the other health outcomes' IV regression estimations, the tests performed in this model show no autocorrelation of residuals (Durbin Watson test = 2.56). In consequence, we have used again lags of the suspected endogenous regressors as instruments. Again, as in Model 2 only $PROVEXPHAT$ was found to be endogenously determined (see Appendix).

Table 3: Estimation results[a]

Regressors (ln)	Regressand: infant mortality (ln)			Regressand: provincial health spending pc (ln)
	Model 1	Model 2	Model 3	Model 2
	OLS	IV[b]	IV[c]	IV[d]
INC	-0.46 [-1.6]	----	----	0.2** {2.5}
TRANSF	0.23 [0.2]	----	----	-0.22 {-0.5}
FEDEXP	----	----	0.16 [1.15]	-----
PROVEXPHAT	----	-3.2** [-2.0]	-3.8** [-2.4]	----
MUNEXP	-0.02 [-1.0]	0.03 [1.0]	0.03 [1.0]	0.01 {1.6}
PRIV	0.34* [1.8]	0.13 [1.1]	0.13 [1.0]	-0.02 {-0.3}
DEC	-4.8** [-2.5]	-3.26* [-1.8]	----	0.67 {1.0}
EDUC	-0.1 [-0.7]	-0.01 [-0.1]	-0.003 [-0.02]	0.06 {1.3}
LOWBIRTH	-0.33 [-0.7]	-0.36 [-0.8]	-0.35 [-0.8]	0.01 {0.2}
F-statistic	21.8	16.5	16.0	34.9
R^2	0.69	0.65	0.64	0.86
Durbin Watson test	2.54	2.59	2.56	0.41
Breusch Pagan test (p value)	0.00	0.00	0.00	0.00
Obs.	140	140	140	140

*** - significant at 1%; ** - significant at 5%; * - significant at 10%

{.}-t statistics computed with heteroskedatsicity and autocorrelation consistent (Newey-West) standard errors

[.]- t statistics computed with heteroskedasticity-robust (White) standard errors

a- All the estimations include time and provincial dummies

b,c-Excluded instruments are: PROVEXPHAT_1 and POP65

d- Excluded instruments are: FEDEXP and POP

According to the results of *Model 3*, it is only provincial expenditure in health that is statistically significant in explaining infant mortality in Canada between 1979 and 1995. It is estimated that, *ceteris paribus*, a 1 per cent increase in provincial expenditure in health care stimulates roughly a 3.8 per cent reduction in infant mortality. This result provides additional support for the negative association found between DEC and $INFMORT$, since only the absolute amounts of health spending at the primary level of decentralisation in health (provincial level) appear to be significant in reducing $INFMORT$. In addition, by finding a positive association between a component of public health spending and infant mortality, this paper contrasts with other studies on the determinants of health, where public resources devoted to health care are found to be statistically insignificant in explaining health outcomes[45].

[45] A brief summary of the literature about the public spending impact on health outcomes can be found in Or (2000).

10 Discussion

The theoretical literature on fiscal federalism predicts potential efficiency gains from placing responsibilities of local public goods at the local level. For the case of the public health services, these efficiency gains are manifested in an improvement of the population's health. However, in the empirical literature little attention has been paid to the evaluation of the outcomes of decentralisation in the health care public sector.

In this study we have explored the relationship between a measure of health care decentralisation –the proportion of local health spending on the total health spending for all the levels of government- and an indicator of health outcomes -infant mortality- in Canada. To formalise the linkages between health care decentralisation and health outcomes, we have used a simple theoretical framework that we have then estimated using panel data for the period 1979-95. In addition, this model has allowed us test for the existence of the flypaper effect in provincial's government spending behaviour.

The results of the econometric estimations for Canada suggest that decentralisation in Canada has had a positive and substantial influence on the effectiveness of public policy in improving population's health (in terms of infant mortality). Moreover, the efficiency gains from the particular decentralisation structure in Canada do not seem to be counteracted by the flypaper effect. However, some caution is required in interpreting these results. First of all, the indicator of health decentralisation used captures only one of the multiple dimensions of the health care decentralisation process: the fiscal one. Secondly, the measure of health outcomes employed does not fully reflect the underlying level of health in a society. In spite of these limitations, this research adds a new empirical perspective to the evaluation of the economic gains arising from decentralisation in health care.

References

[1] Ahmad E, Craig J (1997): "Intergovernmental transfers" in: Ter-Minassian T. (Ed.), *Fiscal Federalism in Theory and Practice*, Washington, D.C.

[2] Akin J, Hutchinson P, Strumpf K (2001): "Decentralization and Government Provision of Public Goods: The Public Health Sector in Uganda", Carolina Population Center University of North Carolina at Chapel Hill, Working Paper 01-35.

[3] Albi E, Contreras C, González-Paramo JM, Zubiri I (1992): "Federalismo Fiscal" in: *Teoria de la Hacienda Publica*, Ariel Economía, Barcelona.

[4] Armstrong P, Armstrong H (1999): *Decentralised health care in Canada*, British Medical Journal 318, 1201-1204.

[5] Asfaw A, Frohberg K, James KS, Juting J (2004): "Modelling the impact of fiscal decentralisation on health outcomes: empirical evidence from India", ZEF Discussion Paper 87, Bonn.

[6] Baltagi B (1995): *Econometric analysis of panel data*, Wiley, New York.

[7] Banting K, Corbett S (2002): "Multi-level Governance and Health Care: Health Policy in Five Federations", paper presented to the Meetings of the American Political Science Association, Ontario.

[8] Baum CF, Shaffer M, Stillman S (2002): "Instrumental Variables and GMM: Estimation and Testing", Department of Economics Working Paper No 545, Boston College.

[9] Commission on the Future of Health Care in Canada, CFHC (2002): "Building on values: the future of health care in Canada", Ottawa, available online at: http://www.hc-sc.gc.ca/english/pdf/romanow/pdfs/HCC_Final_Report.pdf

[10] Ebel RD, Yilmaz S (2001): "Fiscal Decentralisation: is it happening? How do we know?", paper prepared for the Conference on Public Finance in Developing Countries, Georgia State University.

[11] Ebel RD, Yilmaz S (2004): "On the measurement and impact of fiscal decentralisation", Tax Policy Center, Washington DC.

[12] Habibi N, Huang C, Miranda D, Murillo V, Ranis G, Sarkar M, Stewart F (2001): "Decentralization in Argentina", Center Discussion Paper No. 825, Economic Growth Centre, Yale University.

[13] Khaleghian, P (2003): "Decentralization and Public Services: The Case of Immunization", World Bank Policy Research Working Paper No. 2989, Washington DC.

[14] Lazar H, Banting K, Boadway R, Cameron D, St Hilaire F (2002): "Federal-provincial relations and health care: reconstructing the partnership". Commission for the Future of Health Care in Canada, Ottawa.

[15] Levaggi R, Smith PC (2005): "Decentralization in health care: lessons from public economics", in: Smith PC, Ginnelly L, Sculpher M (Eds.), *Health Policy and Economics*, Open University Press, Maidenhead.

[16] Lomas J, Woods J, Veenstra G (1997): "Devolving authority for health care in Canada's provinces: An introduction to the issues", Canadian Medical Association Journal, 156 (3), p. 371-377.

[17] Filmer D, Pritchett L (1997): "Child mortality and public spending on health: How much does the money matter?", World Bank Policy Research Working Paper No. 1864, Washington DC.

[18] Fogel RW (1994): "Economic Growth, Population Theory and Psychology: The Bearing of Long-Term Process on the Making of Economic Policy", American Economic Review 84 (3), p. 369-395.

[19] Gamkhar S, Oates WE (1996): "Asymmetries in the response to increases and decreases in inter-governmental grants: some empirical findings", National Tax Journal, (49) 4, p. 501-512.

[20] Gravelle H (2003): "A comment on Weale's paper from an economic perspective", in: *Equity in health and health care*, proceedings from a meeting of the health equity network, The Nuffield Trust, London.

[21] Maddala GS (2001): *Introduction to Econometrics,* Wiley, Chichester.

[22] Mahal A, Srivastava V, Sanan D (2000): "Decentralization and its impact on public service provision on health and education sectors: the case of India" in: Dethier J (Ed.), *Governance, Decentralization and Reform in China, India and Russia*, Kluwer Academic Publishers and ZEF, London.

[23] Ministry of Health and Long term Care (2000): "Understanding Canada's Health Care Costs", Toronto.

[24] Musgrave RA (1959): *The Theory of Public Finance*, Mc Graw Hill, New York.

[25] Oates, WE (1972): *Fiscal Federalism*, Harcourt Brace Jovanovich, New York.

[26] Oates WE (1989): "An economist's perspective on fiscal federalism", in: *Study Guide and Readings for Stiglitz's Economics of the Public Sector*, Edward C. Kienzle, New York.

[27] Oates WE (1993): "Fiscal decentralization and economic development", National Tax Journal, 46 (2), p. 237-243.

[28] Oates, WE (1999): "An Essay on Fiscal Federalism", Journal of Economic Literature, 37, p. 1120-1149.

[29] Oates WE (2005): "Toward A Second-Generation Theory of Fiscal Federalism", International Tax and Public Finance, 12 (4), p. 349-373.

[30] OECD (1999): "Taxing Powers of State and Local Government", OECD Tax Policy Studies, Paris.

[31] Robalino DA, Picazo OF, Voetberg A (2001): "Does fiscal decentralization improve health outcomes? Evidence from a cross-country analysis". Policy Research Working Paper 2565, World Bank, Washington DC.

[32] Rodden J (2003): "Reviving Leviathan: Fiscal Federalism and the Growth of Government", International Organization 57, p. 695-729.

[33] Shea, J (1997): "Instrument Relevance in Multivariate Linear Models: A Simple Measure". Review of Economics and Statistics, 49 (2), p. 348-352.

[34] Tiebout C (1956): "A pure theory of local expenditure", Journal of Political Economy, 64 (5), p. 416-424.

[35] Wooldridge JM (2000): *Introductory Econometrics,* South Western College.

[36] Wooldridge JM (2002): *Econometric analysis of cross section and panel data*, The MIT Press, London.

[37] World Bank (2005): Decentralisation and Subnational Regional Economics website.

[38] World Health Organization (1996): "Health Care Systems in Transition: Canada", Copenhagen.

[39] World Health Organization (2005): European Observatory on Health Care Systems website.

[40] Yee E. (2001): The Effects of Fiscal Decentralisation on Health Care in China. University Avenue Undergraduate Journal of Economics, Princeton University.

Appendix

Table 4: Description of the variables used in the empirical estimations

Variable	Description	Coverage	Source
MUN	Municipal government health spending in per capita constant (1997) dollars. MUN includes health care spending by municipal governments for institutional services; public health; capital construction and equipment; and, dental services provided by municipalities in the provinces of Nova Scotia, Manitoba and British Columbia. Designated funds transferred by provincial governments for health purposes are not included in the municipal sector, but are included with provincial government expenditure.	1979-1995	CIHI
PROV	Provincial government health spending in per capita constant (1997) dollars. This includes provincial government funds for health care, federal health transfers to the provinces, and provincial government health transfers to municipal governments.	1979-1995	CIHI
FED	Direct federal health care spending in per capita constant (1997) dollars. FED includes health expenditure by the federal government in relation to health care services for special groups such as Aboriginals, the Armed Forces and veterans, as well as expenditures for health research, health promotion and health protection. FED does not include federal health transfers to the provinces.	1979-1995	CIHI

Table 4: Description of the variables used in the empirical estimations

SOCSEC	Medical aid spending by workers' compensation boards in per capita constant (1997) dollars.	1979-1995	CIHI
DEC	Measure of health decentralisation: sub national public health spending (municipal, provincial and social security funds) over total public health spending (municipal, provincial, federal and social security funds).	1979-1995	Own calc.
INC	GDP per capita in 1997 dollars	1981-1995	CIHI
POP	Population by province	1979-1995	CIHI
POP65	Population over 65 years as a proportion of all the population of a province	1979-1995	Statistics Canada
INFMORT	Number of infants who die within the first year of life, expressed as a rate per 1.000 births	1979-1995	Statistics Canada
LOWBIRTH	Proportion of low birth weights (less than 2500 grams) in all live births	1979-1995	Statistics Canada
EDUC	Full-time enrolments and graduates in postsecondary community college programs as a proportion of the population aged 18 to 24	1979-1995	Statistics Canada
PDI	Personal disposable income	1979-1995	Statistics Canada

Table 5: Analysis of endogeneity, Provincial health care expenditure: Model 2

Supected endogenous variable	**Exogeneity C test (p value)**	**Instruments performance**		
		Hansen test (p value): 0.35		
		Shea R^2	**Exogeneity C test (p value)**	
			fedexp	0.95
lndec	**0.05**	0.52	pdi	0.38
lntransf	0.53	0.13	sqinc	0.64
lnpriv	0.35	0.17	pop65	0.84
all	0.13		pop	0.34
			sqpop	0.41
Endogenous variable	**Exogeneity C test (p value)**	**Instruments performance**		
		Bound R^2	**Hansen test (p value)**	
lndec	**0.02**	0.82	0.15	

Table 6: Analysis of endogeneity, Infant mortality: Model 1

Supected endogenous variable	**Exogeneity C test (p value)**	**Instruments performance**		
		Hansen test (p value): 0.37		
		Shea R^2	**Exogeneity C test (p value)**	
lndec	0.18	0.49	lngdpt_1	0.26
lngdp	0.63	0.77	lntransft_1	0.16
lntransf	0.40	0.25	lnprivt_1	0.57
lnpriv	0.54	0.53	lnmunexpt_1	0.18
lnmunexp	0.45	0.15	lnfedexp	0.21
all	0.49		pop	0.16
			pop65	0.87

Table 7: Analysis of endogeneity, Infant mortality: Model 2

Supected endogenous variable	**Exogeneity C test (p value)**	**Instruments performance**		
		Hansen test (p value): 0.41		
		Shea R^2	**Exogeneity C test (p value)**	
lndec	0.24	0.59	lnprovexphat_1	0.49
lnpriv	0.27	0.54	lnprivt_1	1.00
lnmunexp	0.27	0.18	lnmunexpt_1	0.29
lnprovexphat	**0.05**	0.64	lnfedexp	0.18
all	0.22		pop	0.18
			pop65	0.73
Endogenous variable	**Exogeneity C test (p value)**	**Instruments performance**		
		Bound R^2	**Hansen test (p value)**	
lnprovexphat	**0.05**	0.6	0.96	

Table 8: Analysis of endogeneity, Infant mortality: Model 3

Supected endogenous variable	**Exogeneity C test (p value)**	**Instruments performance**		
		Hansen test (p value): 0.37		
		Shea R^2	**Exogeneity C test (p value)**	
lnfedexp	0.21	0.53	lnprovexphat_1	0.63
lnpriv	0.43	0.53	lnprivt_1	0.54
lnmunexp	0.33	0.17	lnmunexpt_1	0.32
lnprovexphat	**0.06**	0.67	dec	0.16
all	0.09		pop	0.16
			pop65	0.62
Endogenous variable	**Exogeneity C test (p value)**	**Instruments performance**		
		Bound R^2	**Hansen test (p value)**	
lnprovexphat	**0.03**	0.62	0.96	

Figure 3: Health decentralisation in Canadian provinces, 1979-95