

## **Ad-Hoc Expert Group Meeting**

7-9 October 2003, UNCC, Addis Ababa

# **Fiscal Policy and Growth in Africa: Fiscal Federalism, Decentralization and the Incidence of Taxation**

## **Theories of Federalism**

By

Ademola Ariyo



Economic Commission for Africa

---

## Introduction

It is generally acknowledged that Africa is, arguably the most richly endowed region of the world in terms of human and non-human resources. It is similarly acknowledged that it is the least developed region of the world. This contrast of poverty of development in the midst of abundant resource endowment has been attributed primarily to institutional deficiencies. Specifically, inappropriate intra and inter-governmental relations and processes have been flagged as the root cause of this institutional deficiency. It is in view thereof that there has been renewed interest in the discussion of issues relating to federalism.

This paper, therefore, aims to contribute to ongoing efforts aimed at enhancing the understanding of the theories and philosophy of federalism, with a view to ensuring ...

## Concept of federalism

The concept referred to as federalism is concerned about the need for an orderly arrangement of relationship among different tiers of government in a nation. Wheare (1963) notes that federalism refers to the method of dividing powers so that general and regional governments are each, within a sphere, coordinate and independent. It is a principle of organization and practice whose ultimate test is how the federal system operates". Elaborating on this, Friedrich (1963) explains that federalism seems the most suitable term by which to designate the process of federalizing a political community, that is to say, the process by which a number of separate political organizations, be they states or any kind of association, enter into agreements for working out solutions, adopting joint policies and making decisions on joint problems. Elaborating further, Tamuno (1998:13) states that federalism, as I understand it, is that form of government where the component units of a political organization participate in sharing powers and functions in a cooperative manner though the combined forces of ethnic pluralism and cultural diversity, among others, tend to pull their people apart.

## Categories of federalism

Federalism is usually categorised by function, whether political or economic. However, regardless of this, it is important to note that federalism is about nature and form of association among interacting states. It is therefore more appropriate to categorise the concept by the means through which it can evolve, regardless of the functions performed if and when the association comes into being.

Given this, we can identify two major modes of federalism. The first mode is the **dual** federalism. In this case, the constitution allows for the creation of two

---

separate and independent tiers of government with their own clearly defined areas of responsibility. The second is the **cooperative** federalism, on the other hand, simply refers to making federalism work through cooperation between various levels of government. It emphasizes the partnership between the different levels of government providing effective public service for the nation. According to Commings and Wise (1974) the various levels of governments, under this arrangement, are seen as related parts of a single government system, and characterized more by cooperation and shared functions than by conflict and competition.

## **Context of and rationale for federalism**

It is interesting to note that federalism exists even in a unitary system of government. This suggests an inherent desirability of the arrangement in any polity. This should be appreciated when we understand better the essence of government. Specifically, it is important to note that the basic responsibility of any government is to identify and evolve an appropriate framework for the satisfaction of the needs and preferences of the citizenry. This can be achieved through effective mobilization, optional allocation and efficient utilization of national resources. Appropriate division of labour among all stakeholders is a necessary condition for the realization of these goals.

At the national level, institutional arrangement is usually put in place for sharing of responsibilities among the different tiers of government towards meeting these goals. This arrangement anchored on the concept of division of labour in economic theory is synonymous with the concept of federalism in practical terms. This paper, therefore, attempts to clarify the concept, and the types of themes underlying the concept of federalism.

This philosophy of division of labour in the management of national resources and processes gave rise to the concept of federalism. That is, a single tier of government may not be in a position to cope with the management and coordination conducive to the achievement of the development aspirations of the nation. The concept suggests that each identifiable tier of government has unique attributes, challenges and resources, and therefore has comparative advantage in the performance of some functions and in carrying out specific responsibilities for meeting adequately the needs of the citizenry.

Federalism is justifiable on political, economic and socio-cultural grounds. Politically, federalism could arise from constitutional development in the process of nation building, among states or communities (Ramphal 1979). Federalism can therefore be viewed as a process of unifying power within a cluster of states and decentralizing power within the unified state.

---

The economic rationale for the existence of multilevel government arises from the existence of public resources, goods and services with differing geographical spread of benefits. Hence, government can be perceived as concerned with the provision of three categories of goods: national, regional (state) and local (community) public goods and services. Some social goods are national in that their benefits spread across the entire nation (e.g., national defence, medical research findings, macroeconomic stability, national pride, etc.), while others are geographically limited (e.g. local fire service or street light). For the latter, clearly, the benefits are limited to residents within the radius of the benefit circle in which the facility is located. Some other commodities have spillover effects such that a larger unit of authority is required to coordinate their supply such as interstate and interlocal government roads and bridges. Thus, the rationale for multilevel government is the existence of benefit regions of diverse geographical sizes. This directly raises the issue of the optimal community size. Once this is settled the assignment problem and the allocation function become much easier. However, this is not part of the focus of this paper.

Musgrave and Musgrave (1989) have, however, observed that the concrete problems of fiscal federalism are embedded in their historical setting. Yet, it makes sense to investigate the possibility of other meaningful conceptions and principles. In other words, are there any other perspectives different from the conventional view guided by some general principles on the formation of multilevel/multiunit political systems with efficiency and equity in view? If there is, then, even if such perspectives may involve discretionary devising, rather than just being a consequence of the miscellany of historical exigencies, which rarely meet the two fundamental criteria of efficiency and equity in resource allocation, they are worthy of serious consideration.

Regardless of the basis, federalism revolves around three major functions of the state for promoting the welfare of the citizenry. These are allocation, distribution and stabilization. The execution of these functions is to be shared among the various development partners within each polity, especially the government and the private sector. However, the scope of the private sector depends on the role assigned to the market in the allocation of national resources.

There are three variants of this market. At one extreme is a centrally planned (or “command”) economy, in which no role assigned to the private sector. In a mixed economy arrangement, there is room for sharing of roles between the government and private sector, although the limit of the role of each partner is not clear and usually ill-defined. The other extreme case is that of a market economy in which the market largely influences the allocation of national resources. Regardless of the nature and importance of the market, the government has the residual responsibility of correcting for market failure and fill the gap in the provision of welfare-promoting goods and services that will not attract private investment. For our discussion in this paper, we will assume a

---

market-driven economy in the content of which the role of government will be amplified.

## **Types of public goods**

The nature of market failure can also be properly conceptualized by an understanding of the various goods that need to be provided towards satisfying the welfare needs of the citizenry. These are private goods, semi-public goods and full public goods. Ability and willingness determine the enjoyment of purely private goods, which by implication, are produced by private sector operations. The enjoyment of such goods is similarly specific to and of the discretion of the individual buyer. Semi-public goods are those that could be enjoyed both individually and collectively upon payment thereof. Among these are public highways on which tolls are collected, or official services such as education and health whose provision to the people are subsidized by the government. The purely public goods are those that are collectively consumed and not contingent upon any direct payment, such as fresh air or national defence.

The problem of market failure is manifest mainly in respect of pure public goods and the desire to mitigate its consequences is one of the reasons that have prompted increased interest in the issue of federalism.

## **Federalism and provision of public goods**

There are two types of pure public goods in terms of the spatial incidence of the benefits enjoyed, and this distinction is important for analytical purposes as would become clear later on. First, we have national public goods whose spatial incidence covers the entire nation, e.g. defence, international affairs, space exploration and the benefits of research and development. In the second group are local public goods whose spatial incidence is limited to particular geographic areas, e.g., regional transportation systems, street lighting and local fire engines. The issue in focus is the determination of the tier of government that should undertake the provision of either group of goods.

Theoretically, the central government has been adjudged as most suited for performing the distribution and stabilization functions and provides national public goods (Oates 1972: 3-38; Musgrave and Musgrave 1989: 445-455; Cremer et al 1994). The contention that lower levels of government could not perform these functions as efficiently as the central government were preached on the arguments. These are: difficulty in appropriating the full social benefits of the programmes undertaken at that level, and the tendency towards the free rider, problem. With regard to the former (Taiwo, 19..) notes that the local government often considers only its own marginal costs and benefits when deciding on its level of provision, and ignores the benefits conferred on other local

---

governments. Under this situation, local government may operate at a nationally desired level.

This problem may not arise if the central government was the decision-making unit, since all the associated costs and benefits would be internalized and taken into consideration. According to Taiwo (19..) the same argument prevails if it is external costs that are imposed on other localities. However, in this case, the locality's level of operation will be more, rather than less, than what is nationally acceptable. Related to this is the free rider problem, which arises when an individual or a group has the incentive to benefit from the provision made by others (Layard and Waiters 1978: 192).

The second relates to the distribution function, which can be typified by a proposed income maintenance programme by local government. Such a programme would inevitably require transfer of income or other resources from the rich to the poor in the locality. In the absence of any restriction on mobility of persons, the programme would attract poor residents from other localities to free ride. At the same time, the policy induces the emigration of rich residents from the area in an attempt to minimize their own tax burden. In the absence of a compensation scheme from the other localities, the particular locality that initiated the income maintenance programme would be unable to reap the full benefits of its initiative and such a programme is bound to eventually collapse. At the national level, these problems are likely to be avoided or minimized, since all movement of persons within the country would be considered and all the costs and benefits of the programme internationalized.

The same principle applies to the stabilization function as a means to an end. We can define the end as the desire by a local government, if we assume that to boost its level of economic activity, in terms of increase in the production of goods and services. Let us assume this is facilitated through changes in fiscal policy, especially, say increase in public expenditure. If the local government has to print money and finance the increase in its public expenditure through this window, other equally eligible localities would follow suit and simply buy up the extra goods and services made possible by the programme initiated by the other locality.

On the other hand, if the activity was financed by raising taxes and people are mobile, residents from other localities (outside this tax net) would migrate to the locality that initiated the programme and share in the benefits provided. If the goods were also mobile, the economy would become open whereby serious economic leakages would occur. Buttressing this scenario, Oates (1972: 4-5) notes that small local economies become more open as their constituency purchase a large portion of the goods and services they consume from other localities. This implies that there would a leakage from a marginal unit of private spending are likely to be quite large.

---

The programme could also be financed through borrowing, from internal and external either internally from local residents and institutions, or externally from other localities, or both. The Ricardian equivalence principle however implies that the choice between tax finance and debt finance is equivalent to a choice about the timing of those taxes (Buchanan 1960: 52). Tax finance involves immediate tax payments, whereas debt finance involves future tax payments. Borrowing would then increase the debt burden on local residents in the future. If the debt were internal, there would be transfer payments from the locality's taxpayers to creditor residents. The loss to the locality owing to the programme would then be restricted to the portion of the benefits that other localities could free ride on. But if the source of debt was external, there could be additional cost to free ride on benefit sharing, depending on the conditionalities attached to the debt.

A similar line of thought applies to issues on the provision of national public goods such as defence. For our discussion, we assume that: (i) it does not matter where the production facility is located, whether within the central government or under the auspices of the local government; and (ii) the cost of providing each good is the same regardless of location of facility. The latter assumption presumes uniform technology with identical returns to scale. Thirdly, we further assume that only one national public good is being produced at only one location.

Being a national public good, other localities will, as a right, derive benefit from its provision. This scenario raises the issue of optimality of scale or level of output by the locality. This is attainable if, and only if there is an agreement between all the localities to share the cost of providing the service. Achieving this will be a tall order because the other localities cannot be excluded from enjoying the service provided, hence the tendency to free ride on the enjoyment of the benefits.

If the locality decides to continue to provide this national public good, the level of output will obviously be suboptimal. It will be confined to that level at which the marginal cost is equal to the value of the marginal product for each producer within the locality. However, the estimation of the marginal value will ignore the benefit. This desire for efficiency underscores the necessity for the central government to produce and/or provides national public goods.

A similar argument can be made regarding the issue of equitable access to national public goods, since local governments producing same localities may initiate some prescriptive steps to deny residents of other location in a desperate attempt to minimize the problem of free-ride. We now focus on the analysis of production of local public goods and adapting the same set of assumptions. We also define local public goods as those with limited spatial incidence of the benefits enjoyed, such as the development of a gravity-fed pipe-borne water in Malawi (Hill and Ntawali 1989) for a small geographic group. We also assume the

---

existence of consumer groups with similar tastes and preferences in which that of one group differ from that of other groups.

Now, let us assume that the local public good is provided by the central government. It will normally decide to provide a uniform level and type of the good for all jurisdictions, in spite of the variations in tastes. This amounts to a suppression of individual or group preferences, thus imposing high implicit cost on the citizens, whether advertently or inadvertently. Also, this uniform provision forces some localities to overconsume, while others under-consume the good. The welfare, efficiency or deadweight loss associated with such an outcome is illustrated in Fig. (Taiwo, 19..)

The figure divides the country into two localities, A and B, whose respective demands are  $D^A$  and  $D^B$ , and each group of individuals is presumed to have identical tastes and income. The goods could also be supplied at both tiers of government at a constant price,  $P_o$ . At this price, localities A and B would wish to consume  $X_1$  and  $X_2$  of the good. If there is uniform provision of the good, at  $X_o$ , locality A would be over-consuming to the tune of  $(x_o - x_1)$ . The cost of this additional consumption is  $P_o(x_o - x_1)$ , or the area covered between  $e_1bx_o x_1$ .

In contrast, the benefit associated with the additional consumption is the area  $e_1cx_o x_1$  under the demand curve  $D^A$ . The locality loses to the tune of the area  $e_1bc$ . In the case of locality B, there is under-consumption to the tune of  $(x_2 - x_o)$ . The loss of benefit to the locality as a result of the uniform provision is the area  $ae_2x_2x_o$  under the demand curve  $D^B$ . In contrast, the gain is the amount that would have been paid for the shortfall in quantity, i.e., the area  $be_2x_2x_o$ . The locality's net loss is the area  $ae_2b$ . So both localities would lose although the losses may be unequal.

According to Taiwo (19..) flexibility is therefore required to bring both localities to their efficient levels. In this regard, a system must evolve to reacquire the differences in preferences and tastes and provide them accordingly. However, such information and production arrangement would be costly, and there may be reluctance to use it. The preferred option is to allow each locality to cater for its own tastes. By this, A would have the opportunity of consuming at  $X_1$  and locality B at  $X_2$ . The efficiency losses would simply disappear, as the supply of local goods would be matched to local tastes (Taiwo, 19..; Woller and Phillips (1998:139-140).

To capture practical realities, we can relax some of the assumption of uniform technology and thereby allow for differences in economies of scale. However, this does not obliterate the substance of the above discussion. Furthermore, despite these assumptions, credible arrangement exists in favour of decentralization of government (public) activities. Similar reasons articulated by Gboyega (1994); Tanzi (1995) and Amuwo (1998) include the following: The



---

first is that decentralization promotes competition both in terms of product differentiation (to cater for various tastes) and cost efficiency. Second, it encourages embedded technological progress among the localities in order to stay competitive. Third, promotes accountability and responsive governance, because the policy making unit is nearer to the grassroots. When local activities are funded by the central government, the tendency would be to overprovide and overconsume since a substantial part of the burden would be borne by other localities. It is the same free rider problem that would be at work here. Finally, decentralization provides the political glue for countries with regional ethnic, racial, linguistic or cultural diversity as in Ethiopia and Nigeria.

Based on the above, we can conclude as follows: The central government has comparative advantage in the provision of national public goods and in performing the distribution and stabilization functions. On the other hand, local governments have comparative advantage in the provision of local public goods, especially those whose preferences vary geographically and for which there are no substantial economies of scale. To reap these advantages, both tiers of government should coexist and be allocated functions in which each has comparative advantage.

## **Fiscal federalism**

It is now appropriate to focus on the financing aspect of allocation being discussed and the aspect relating to allocation of production activities/facilities among the various tiers of government. Specifically, having justified the need for decentralization of the output function, it is now important to determine how we should raise and share financial resources for executing the decentralized “production” functions. That is, having settled issues relating to the “real sector”, we now focus on the financial sector of the centralization or decentralization debate. This is what we refer to as fiscal federalism.

The focus of our discussion is the evaluation of the merits and demerits of the issue of centralization or decentralization of fund raising functions or a mixture thereof. That is having dissected the “role sector” aspect of federalism, it is now time to extend a similar treatment to its financial sector. For a start, we need to characterize the various methods of raising public revenue. The first and most enduring source of public revenue is taxation. Others including special levies into a general pool for development targets or earmarked for the execution of specific public projects or programmes. In this paper, we focus only on the conventional tax-based revenue after tax-based revenues are raised through the allocation of taxing powers among eligible authorities and/or institutions. In the context of our discussion, the eligible authorities are the centralized versus the local levels of government.

---

There is also a variety of allocation options. For example, the taxing power could be fully centralized, fully decentralized, or shared between or among the various tiers of government. Such decisions are guided by some criteria. The first and arguably the most important, is the nature and extent of “production” responsibilities assigned to each tier of government. The second is the issue of efficiency of revenue generation otherwise referred to as effectiveness and efficiency of tax administration. The urge for economies of scale, curtailment of free-rider problem whereby citizens can escape taxation by migrating from one tax environment to another, and the national desirability for a uniform tax system usually work in favour of centralized taxing function. The need to spend public revenue on the articulated needs of the grassroots, and its tendency to promote and sustain accountability are some of the points in favour of decentralized taxing function. Others include the desire for fiscal autonomy and the need to promote competition in tax policies and programmes, whereby there exists an implicit reward (penalty) function for competitive (non-competitive) tax regimes or localities.

However, centralization and decentralization are polar extremes along a spectrum of options that could be considered. In fact, an appropriate mixture of both options often obtains in practice. We could therefore discuss the modalities for ensuring an optimal mix of both in this section. An understanding of the multifacet nature of taxation provides a prima facie argument for considering a mixture of centralization and decentralization. For example, there is the needs-based conception of taxation. That is, decision on taxation does not exist in a vacuum, rather, it is or should be demand-driven. Second, once the need had been articulated, there is need for legislative approval of the taxation proposal. Third, it is important to consider the various options for the administration of the tax function, as well as the optimal allocation of the proceeds therefrom among competing demands except for project-specific special tax levy.

A good tax system should also have certain qualities like efficiency, equity, revenue adequacy, low administration cost and be able to promote economic stability (World Bank 1988). The monitoring and control function in every aspect also need to be considered. It therefore stands to reason to conclude that neither a fully centralized nor fully decentralized system would be able to perform excellently with respect to each facet of the taxing chain described above, as an example.

It is also necessary to identify two features to which each or most of them would apply. The first feature relates to the processes of revenue generation, which is within the purview of tax base sharing. The other feature relates to issues of allocation of proceeds of tax efforts, otherwise known as revenue sharing. Again, there is need for division of labour especially with respect to the former, and need for consensus building in whichever form with respect to the latter. Revenue raising strategy is also guided by two main principles of taxation. These are the (i) ability to pay principle; and (ii) the benefit principle.

---

The ability to pay principle applies to all categories of national public goods, and in most cases focuses on the conventional issues associated with civic responsibility. These include taxes such as poll tax, personal income tax and company income tax, among others. The benefit principle on the other hand, applies to taxes raised for the production of goods at the instance of the beneficiary community or whose benefit accrue directly to the tax payer. Within this group are community-based projects such as local drainage, electricity and water, for which the residents made contribution (special levy) for their provision and sustenance. Others include items such as education tax to build schools for a locality, among others. That is, they belong to the category of taxes whose proceeds are earmarked for demand-driven goods and services. It is therefore not difficult to conclude that taxes based on the ability-to-pay principle apply generally to the provision of national public goods, while the benefit principle applies more to local public goods.

Decision on revenue sharing can also be guided by several considerations (Taiwo and Afolabi 1986). First, it can be used to overcome horizontal fiscal disparities between localities, and to mitigate vertical fiscal imbalance between tiers of government. Second, a higher-level government may, for good reasons, impose certain services or standards on lower-level governments. The services or standards thus imposed become merit goods, which the lower-level governments may not be capable of or willing to finance. Third, intergovernmental grants may be used to offset the effects of externalities, with subsidies given to those localities that confer net external benefits on others. Lastly, it is sometimes advantageous to collect some taxes centrally on behalf of all tiers or lower-level tiers of government. In this case, an arrangement should be made to transfer the proceeds to deserving localities. These factors should also govern the formula for revenue sharing.

Given the above feature, the issue of interest is how to promote an efficient and equitable fiscal structure or federalism. This could be addressed by the adaptation of the modified Samuelson's condition, that the sum of marginal benefits equal to average cost relative to the specified quality of the good or service. To achieve cost effectiveness and allocative efficiency, budgeting should truly reflect the real cost of raising funds (resources) for each assigned function. These principles inherently take care of both intercommunity and interpersonal distribution and therefore satisfy both the efficiency and equity requirements.

## **Principles of Fiscal Federalism**

Towards this end, some principles have been evolved over time to guide that revenue sharing exercise among different tiers of government. These principles are summarized below:

| Group | Principle                         |                                                                                                                                                                                                                                                              |
|-------|-----------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| a.    | Diversity                         | The Fiscal factor should take cognizance of and this diversity in resources and this needs to be adequately accommodated to ensure the supply of required national, regional (state) and local public goods.                                                 |
| b.    | Equivalence                       | The spatial incidence of the various public goods differs, and allocative efficiency requires the equalization of locational advantages arising from interjurisdictional differences through fiscal policy measures, including taxation and revenue sharing. |
| c.    | Centralised Stabilisation         | Propagates the use of fiscal instruments for achieving macroeconomic policy objectives.                                                                                                                                                                      |
| d.    | Correction of Spillover effects   | Use of fiscal policy to correct for jurisdictional externalities with respect to costs imposed on other localities as well as free-ride on enjoyment of benefits.                                                                                            |
| e.    | Minimum Provision of Basic Needs. | Fiscal federalism should guarantee provision of basic needs of the citizens.                                                                                                                                                                                 |
| f.    | Fiscal Equalisation.              | To minimize the effect of large regional differences in resource endowment and in capacity of each tier of government.                                                                                                                                       |
| g.    | Efficiency                        | This principle focuses on efficiency with respect to allocation of resources at the intra-country level and among different tiers of government.                                                                                                             |
| h.    | Derivation                        | This principle suggests the need for localities to benefit from extra considerations, given the availability of some resources within their domain, especially for addressing some externalities specifically related to the exploitation of such resources. |
| i.    | Centralised Redistribution        | This principle emphasizes the redistribution function of central fiscal policy through progressive taxation and targeted expenditure programmes.                                                                                                             |

**Source:** Adapted from ....

It is obvious that these principles do overlap and could be mutually inconsistent or even contradictory. Hence, each or a combination thereof would have to be adapted to take care of country-specific circumstances.

In light of the foregoing, fiscal federalism refers to the financial relationships between and among existing tiers of government. Specifically, it is the system of transfers or grants by which the federal government shares its revenue with state

---

and local governments. This system is what is generally referred to as revenue allocation.

According to ..... the fiscal relationships between and among the constituents of the federation can be explained in terms of three main theories, namely,

- the theory of fiscal location which concerns the functions expected to be performed by each level of government in the fiscal allocation;
- the theory of interjurisdictional cooperation which refers to areas of shared responsibility by the national, state and local governments; and
- the theory of multijurisdictional community, in which case, each jurisdiction (state, region or zone) will provide services whose benefits will accrue to people within its boundaries and so, should use only such sources of finance as will internalize the costs. These theories could serve as guide to politics in the choice of these principles.