Redistribution and Local Public Finance*

Masayoshi Hayashi

Associate Professor, Hitotsubashi University

Abstract

This essay reviews theoretical arguments concerning the issue of decentralized redistribution. We make a representative examination of what are regarded as the merits and demerits of fiscal decentralization from a redistributive perspective, emphasizing the concepts of social solidarity. We proceed to take a closer look at the decentralization theorem, and argue that the celebrated theorem is not an appropriate framework for the issue of redistribution. We then review other merits as well, and argue that those merits largely concern the cost-efficiency of public service delivery and could be in principle exploited without compromising the principle of horizontal equity, through two forms of administrative decentralization, delegation and de-concentration. However, the local incentive problem still remains. The design of transfers and personnel control mechanism will be a crucial factor to attain the cost-efficiency of local administration.

I. Introduction

What governments do, directly or indirectly, involves redistribution of one kind or another. ‘Redistributive policy’ here refers to public policy whose primal purpose is to directly redress inequalities, both ex-post and ex-ante.1 Differences in earning capabilities and endowments of households result in unequal outcomes of allocations in the market economy. Progressive taxation, welfare payments, and various targeted transfers are intended to alleviate such ex-post inequalities. Even worse, opportunities are also unequal. Such ex-ante inequalities create disadvantages an individual cannot control, including being born with innate disabilities or poor health as well as being in a disadvantaged family, time or place. Redistribution guards against such ex-ante disadvantages with policies that put individuals on more equal terms, among which education is arguably the

* This article is based on a study first reported by the author as ‘Saibunpai Seisaku to Chihozaisei’, Financial Review, Vol.82, pp.138-160, 2006 (in Japanese). The comments by Robin Boadway and Motohiro Sato, and the assistance by Yukako Iwata and Yosuke Sunahara are gratefully appreciated. Of course, the usual caveat applies.

1 This typology, including social insurance, draws from Boadway and Keen (2000). Following Gramrich (1993), we may also categorize the ex-post type as direct transfers geared to short-run welfare, and the ex-ante type as human capital type programs geared to long-run welfare.
most important. Health care and other types of social services for the handicapped can also contribute to this cause.

Indeed, bad luck occurs even among the equal (in the ex-ante sense), taking many forms like becoming unemployed, injured, sick or disabled. In some cases such bad luck can be adequately handled in the market, but in other cases this is not so. Such cases are taken care of by social insurance. The effects of social insurance, typically public health insurance, disability pensions, public pensions, and unemployment insurance, are both ex-post and ex-ante. The risk-averse benefit from social insurance as it reduces their future risks, which yields ex-ante benefits. Since social insurance compensates the unlucky for a loss caused by bad luck with funds transferred from the lucky, it in effect constitutes an ex-post redistributive policy.

We may even argue that the whole system of redistribution (i.e., progressive taxation, welfare payments, various targeted transfers, public education, health and social care) as a package constitutes social insurance for individuals in “the veil of ignorance.” Assume that we are in the veil of ignorance and do not know which place, time, family or individual we are born into. Some are lucky enough to be born in good circumstances, but some are unlucky enough to be born in poor circumstances. The system of redistribution then gives ex-ante benefits to all risk-averse individuals in the veil of ignorance and ex-post benefits to those who have turned to be unlucky after their birth.

Defined as such, modern governments are largely institutions for redistribution (Boadway and Keen 2000). In the last couple of decades, many countries have experienced the trends towards decentralization, to which Japan is not an exception. Fiscal decentralization necessarily involves elements of redistribution, since it re-assigns taxing and spending responsibilities to different levels of government. Indeed, the large share of redistributive expenditures does not allow us to discuss decentralization without due consideration of the redistributive elements of public policy.

The issue is even more important in the Japanese case. First, Japan is a unitary country governed by a Constitution which defines the rights and the obligations applicable to all the Japanese wherever they may reside within the country. Such constitutional rights and obligations are in fact what makes modern government largely redistributive. In Japan, it is mainly local governments that actually implement such rights and obligations through their expenditures. It then seems natural that the central government ensures that all locally implemented programs meet national standards so that all citizens can enjoy their constitutional entitlements, as most Japanese local expenditures are mandated in one way or another and are funded partially or fully by the central government.

Second, OECD data show that Japanese local governments do much more than most of the other OECD countries do. While the share of local revenues (26 percent) is higher than the relevant OECD average (21.9 percent), the share of local expenditures (40.7 percent) is even higher than the relevant OECD average (32.2

---

2 Social insurance programs are called as such in recognition of the fact that they are intended to compensate for misfortune (hence ‘insurance’) and to insure for misfortunes that would otherwise not be privately insured (hence ‘social’).

3 A study on selected OECD countries (Mitchell 1991) shows that the Gini coefficients decrease for income after taxes, although the changes in the coefficients vary, with France being the largest (39.8%) and the Netherlands the smallest (8.4%). In addition, a majority of public expenditures are redistributive. If we consider expenditures for health, education and social protection ‘redistributive’, the shares of such spending for selected OECD countries range from 60.16% (Iceland) to 86.15% (Luxembourg).
percent). In addition, the local expenditure share hikes up to 86 percent if the total expenditures are restricted to tax-financed redistributive spending. The local share in Japan is one of the highest among the unitary countries, being only second to Denmark. It is also noteworthy that they are even comparable to those of the federal states where provincial or state governments have constitutionally assigned redistributive functions.

The purpose of this paper is to provide a set of theoretical reviews on the literature on decentralization and redistribution and shed some light on policy issues in the Japanese reform towards decentralization. Economics literature (e.g., Boadway and Keen 2000) typically frames redistribution in terms of social justice (social welfare), efficiency (Pareto improvement) and exploitation (political economy). Social justice here relates to evaluating public policies with some consistent set of explicit rules of distributive ethics. Efficiency refers to attaining Pareto improvement where there is a chance for redistributive policy to make everyone better off. There are no tradeoffs between efficiency and redistribution here. Exploitation may refer to public choice that favors specific sets of social groups and makes other groups worse off. We basically follow this scheme, but as our intent is to provide a normative framework of analysis, we skip the topic of the issue of exploitation, or political economy, concerning redistributive policy.

We also discuss the issue of the “race to the bottom” caused by a decentralized system of redistribution. As the literature on fiscal competition shows, economists tend to agree that free mobility of production factors (i.e., labor and capital) makes tax rates and redistributive expenditures too low (recall the threats of globalization to the welfare state!). We will provide some selective review on the related literature and summarize related theoretical and empirical results.

When considering the local assignment of redistributive functions, it is important how burdens are shared and benefits are delivered. From this perspective, we examine the merits of fiscal decentralization and emphasize the concepts of cost efficiency of public service provision. Like education and personal social services, redistributive programs involve face-to-face service delivery or transfers in kind. Even cash payments frequently entail face-to-face interaction such as means testing and counseling services for the needy. Such services require local implementation conducted either by local branches of the central government or by local governments themselves. We review local merits in delivering such services, and argue that those merits largely concern cost-efficiency and could be exploited without compromising horizontal equity. However, the local incentive problem still remains. The design of transfers and personnel control mechanisms will be a crucial factor to attain the cost-efficiency of locally administrated redistributive programs.

II. Redistribution as Social Justice

II.1. The Social Welfare Function

Despite its criticism, the dominant approach for economists to social justice is the welfarist one (Boadway and Keen 2000). The welfarist approach evaluates public policy based on its impact on individual utilities (or

---

4 The figures are based on Joumard and Kongsrud (2003).
welfare), and aggregate the individual effects through the conceptual device of the social welfare function (SWF). Since this necessitates the interpersonal comparison of individual utility, it is not immune from criticism. But with the advent of “new” new welfare economics (Stiglitz 1987), the interpersonal comparison of utility has now been recognized as a necessary assumption for the economic analysis of redistribution (e.g., Feldstein 1972; Boadway 1974, 1976; Bergson 1980), and the use of the SWF is now the standard practice in applied research (e.g., Ahmad and Stern 1984, Sandmo 1998, Dahlby 1998, Frankhauser et al. 1997, Azar 1999, Johansson-Stenman 2000).

The desirability of alternative social states is assessed in terms of the following SWF of Bergson-Samuelson type:

\[ W = W(v) \]  \hspace{1cm} (2.1)

where \( v = [v_1, \ldots, v_i, \ldots, v_N] \) is a vector of utilities of \( N \) members of a society. The social utility function is typically assumed to satisfy the Pareto principle, the anonymity principle, and the inequality aversion. The former principle simply assumes that a Pareto improvement always leads to an increase in social welfare, i.e., \( \frac{\partial W}{\partial v_i} > 0 \) \( \forall i \). In other words, a Pareto improvement is sufficient for an increase in social welfare. The anonymity principle implies that only the combination, but not the permutation, of individual utilities affects social welfare. This means that reshuffling of the elements of \( v \) (or changing the arguments of \( W \)) does not affect the level of \( W \). The SWF is also assumed to be inequality averse, which requires \( W(v) \) to be concave in \( v \).

If lump sum transfers are possible, the implications of the SWF defined as above (2.1) are straightforward. Assume for simplicity that all individuals have the same utility functions \( v_i = v(y_i) \) where \( y_i \) is fixed income and \( v'(\cdot) > 0, \ v''(\cdot) < 0 \). With the SWF now given as \( W = W(v(y_1), \ldots, v(y_N)) \), the level of social welfare is maximized if

\[ \frac{\partial W}{\partial v_i} \frac{\partial v_i}{\partial y_i} = \frac{\partial W}{\partial v_j} \frac{\partial v_j}{\partial y_j} \quad \forall i \neq j. \] \hspace{1cm} (2.2)

When the standard assumptions are satisfied, we then know that the first-best optimum is attained with equal distribution of income \( (y_i = y^* = N^{-1} \sum y_i) \forall i \). Note that this result holds whether the SWF is Benthamite \( (W = \sum v) \), Laxwian \( (W = \min \{v_i\}) \), or somewhere between the two, as in Nash \( (W = \prod v) \). Of course, the choice made in the real economy is not the first-best. Almost all tax bases are elastic. Redistributive taxes and expenditures require detailed information on consumers, but there are always asymmetries of information between the public sector and consumers. These distortions work as constraints on the economy and we can only hope for second-best solutions, where the equal distribution of income is not expected.

In the context of local public finance, the central issue should be the identification of the “society,” to which the identical SFW is applied. The “society” here relates to the concept of a sharing community. In a redistributive program, there are recipients (beneficiaries) and contributors (benefactors). The aggregate amount of contributions and that for receipts in a redistributive program are to match at some level of grouping, which is

---

5 For more on welfarism, see Boadway and Bruce (1985).
6 For more assumptions on the social welfare function, see Chapter 5 of Boadway and Bruce (1984).
7 For the sharing community or community sharing, see Banting and Boadway (2004).
called a sharing community. In a sharing community, the net contributions by benefactors will just equal the net receipts by the beneficiaries in budgetary terms. If redistributive programs are applied nationally, the sharing community is the nation as a whole. If the program is financed and applied within a region, the sharing community is that region. Since benefits and burdens do not match in a sub-region of a sharing community, it is meaningless to discuss the match of benefits and burden at that level.

The most fundamental issue for redistribution policies is thus how to draw lines among communities for a specific redistributive program. We may argue that the largest possible sharing community is the nation as a whole. However, how we draw lines within a nation depends on how the nation is defined and works, which may be embodied in a Constitution. If a country is unitary like Japan is, we could safely argue that the sharing community is the country as a whole. But, if a country is federal, the answer may not be that straightforward.

II.2. Horizontal equity

The concept of community sharing is related to the concept of horizontal equity. The principle of horizontal equity dictates that those who are identical in “relevant” characteristics should be treated identically. Note that we do not always agree on what should be included and excluded in the “relevant characteristics” as they are not prescribed à priori by the principle of horizontal equity itself. Therefore, the statement is conditional. If the place of residence is considered “relevant,” people with comparable characteristics do not have to be treated identically no matter where they reside. However, since every individual has a uniform set of rights and obligations in a unitary state, the place of residence should not be included in the relevant characteristics. And for the same reason, the sharing community should be the nation as a whole.

The principle thus dictates that all the people who are identical relevant characteristics, except for the place of residence, should be treated identically no matter where they reside within the sharing community.8 A decentralized system of redistribution then undermines horizontal equity to the extent that local governments behave differently. If decentralization implies independent exercise of fiscal responsibility by local governments, horizontal equity cannot be maintained as long as different governments want to behave differently. Forcing them to treat all individuals of a given kind identically across all jurisdictions essentially involves abrogating their independence and contradicts the notion of decentralization. Thus, to the extent that decentralization is maintained, horizontal equity inevitably must be compromised. In other words, decentralized redistribution within a unitary state may thus be an untenable proposition. Decentralizing redistribution fragments the

---

8 There are several interpretations for horizontal equity. The first is that the principle may be the result of some social welfare maximizing rule (Atkinson and Stigliz 1980). However, it is straightforward to set up a counter example where maximizing social welfare does not lead to what is regarded as a horizontally equitable condition (Atkinson and Stigliz 1980, Broadway and Bruce 1985). The second view regards horizontal equity as an independent principle of justice, which has to be set into balance alongside the welfarist approach (Atkinson and Stigliz 1980). Welfarism only cares about the post-intervention distribution of individual welfare, which may sometimes be unacceptable unless some additional principle is utilized to evaluate the policy intervention. The third view sees the principle of horizontal equity as a constraint on the public sector. We are taking this third interpretation in the text.
nation-wide sharing community,\textsuperscript{9} and inevitably compromises horizontal equity.

Since redistribution coercively takes away private resources and disburses them to the needy, a sense of social solidarity is a necessary ingredient of successful redistributive programs. Social solidarity welds a mass of individuals into a united whole, and involves a sense of shared common destiny, a sense of belonging, and sympathy for the needs of strangers’ within the nation (Ignatieff 1984). It may be then argued that social solidarity is a prerequisite for successful implementation of redistributive programs.\textsuperscript{10} In addition, a consistent application of a set of redistributive standards is another necessary element. As Breton (2002) argues, citizens may simply want a nationally standardized provision of goods and services. “Horizontal equity is the one clear principle in a context where the alternative is chaos (Usher 1995, pp. 97-99).” Indeed, fairness and predictability are virtues themselves. And if we live in the same community where we share the burden, we simply would like the same rules to apply to every community member. In this sense, the principle of horizontal equity has its virtue especially in relation to redistributive policy.

As Musgrave (1959) characterizes, horizontal equity thus works as “a safeguard against capricious discrimination.” The principle puts constraints on the central government in a unitary country when it designs the system of local public finance, requiring national policy to satisfy uniform standards. Note that “uniformity” here only implies consistent applications of standards where comparable levels and types of services and transfers are provided to individuals with comparable characteristics. We emphasize that it does not refer to a single level of service or transfer provision to the whole nation, as a typical model of decentralization conceptualizes in Section 2 below. In particular, it is quite absurd to discuss a uniform level when we discuss redistributive policy since redistribution necessarily involves different levels of treatment for different types of individuals. For example, a redistributive program distinguishes among individuals according to their characteristics that are considered ‘relevant’ to the program. If a program can take account of individual characteristics, it can also take account of local needs. When we say that different regions have different needs for redistribution, we do not really mean that a specific place or an aggregate of people has such a need. Only individuals have needs. What is implied is a conspicuous concentration of those who share certain needs in a given place, and that such common characteristics are differentiated among regions. The “uniformity” is thus more flexible than it may look at first glance.

III. Efficient Redistribution

Redistribution does not always trade off with efficiency. We could expect Pareto improvements through redistribution in an economy, employing reasoning based on altruism, risk-sharing, dynamic efficiency and

\textsuperscript{9} Decentralization can have many meanings. Fiscal decentralization refers to the case where the units hold full discretion on taxing and spending and can set their own equity standards.

\textsuperscript{10} It may also be argued, as histories of industrialized countries show, that centralized redistribution is a cause, rather than an effect, of social solidarity. Redistribution was in fact employed as an instrument for forging solidarity in the periods of national crisis like war and depression. It may even be claimed that this may in turn have contributed to nation-building by forging nationalism. This perspective may especially be relevant for some of the developing countries where societal divisions and conflicts are still prevalent.
non-convexities (Boadway and Keen 2002). More relevant among these in regard to local public finance are altruism and risk-sharing.

III.1. Altruism

The altruistic individual \( i \) obtains more utility when they see the poor become happier. His utility function may be expressed as

\[ U_i = U(x_i, U^p) \]

where \( x \) is his consumption of a composite good, \( U^p \) is the utility level of the poor (\( P \)), and \( \partial U_i / \partial U^p > 0 \) since he is altruistic. The utility of the poor increases only in their private consumption \( x^p \):

\[ U^p = U(x^p) \]

which shows that altruist’s utility is dependent on \( P \)’s consumption

\[ U_i = v(x_i, x^p) = U(x_i, U(x^p)) \]  

If altruist \( i \) transfers \( g_i \) of his income to the poor so that the latter’s consumption is given as \( x^p = w^p + \sum g_j \), the former’s utility is rewritten as

\[ U_i = v(w_i - g_i, \sum g_j) = U(x_i, U(w^p + \sum g_j)) \]  

(3.1)

where \( w^p \) is \( i \)’s income and \( G = \sum g_j \) is the total amount of transfers given by each of the altruists. Note that \( G = \sum g_j \), and therefore \( g_i \), works as a public good among altruists.

A voluntary transfer from altruist \( i \) is then characterized by the first order condition

\[ 1 = \frac{\partial v(x_i, G) / \partial G}{\partial v(x_i, G) / \partial x_i} \]  

(3.2)

whereas the optimal level of total transfer \( G = \sum g_j \) is characterized by

\[ 1 = \sum_i \frac{\partial v(x_i, G) / \partial G}{\partial v(x_i, G) / \partial x_i} \]  

(3.3)

By comparing (3.2) with (3.3), we see that the voluntary transfers are suboptimal, leaving room for Pareto improvements.

This suboptimal level of transfers (i.e., redistribution) may be mitigated by public interventions through taxes and subsidies. What matters here is the level of government that conducts such redistributive policy. The classic argument regards redistribution as the responsibility of the center (Musgrave 1959). Against this classic argument, however, Pauly (1973) argues that local government could implement redistribution more efficiently.

The upshot of the argument is that, since programs administered by the center tend to be uniform across the country, lower level governments would be better able to design their own programs to suit the preferences of local residents. This is a restatement of the celebrated decentralization theorem (Oates 1972) applied to redistribution, which may be described as follows. Redistribution should take place when people benefit from redistribution. Altruists feel happy when they see that the poor are better off. The risk-averse gain ex-ante benefits if a system of social insurance is in place. Or people have some ethical preference for a specific distribution of income and individual utilities, which transcends the determinants of individual behavior. Regardless of these
differences in motives, people benefit from redistribution.

The problem is that the preferred level of redistribution (or ethical preferences) diverges among individuals. The literature argues that central governments tend to provide uniform programs. Uniformity in the presence of divergent preferences results in the classic problem of local public finance: to the extent that individual preferences differ, the uniformity yields a welfare loss. If the amount of public service provision is to be unique, it has to be decided somehow through the political process. Once it is determined, every individual in the society has no choice but to accept that amount. Consider an example where the amount is decided by majority voting and in which the median voter wins. The supply level then will be the bliss point of the median voter, but is not so for those with different preferences. Then, non-median voters lose their welfare to the extent that the bliss point of the median voter is different from those of theirs. In other words, the more preferences are divergent, the larger welfare losses will be. If decentralization implies a smaller unit of decision with smaller divergence of preferences, we may expect welfare losses to be reduced in a decentralized system. The decentralization theorem thus argues that decentralization yields welfare gain.

Let us consider a simple example where an economy is made up of three types of residents indexed by \( i = 1, 2, 3 \) of equal number which can be normalized to unity. Each type has different preferences over the degree of redistribution. Figure 1 measures the degree of redistribution on the horizontal axis and marginal benefits \( MB_i \) for \( i = 1, 2, 3 \) on the vertical axis. The costs for redistribution are characterized as identical constant marginal costs \( MC \). All three types reside in the same community, and the uniform level of redistribution is decided by majority vote. Since the median voter is type 2, redistribution will be given at A where \( MB_2 \) and \( MC \) intersect. While the degree is optimal for type 2, types 1 and 3 suffer from welfare losses ADE and ABC. We then assume the community is divided into three parts, each being inhabited by individuals of a single type. This will yield welfare gain, since each community will be redistributing at a point which is optimal for its residents, namely, point D for type 1 and point C for type 3, without entailing any welfare losses.

However, this line of argument should be considered with reservations. First, in the ideal world of the decentralization theorem, no redistribution is necessary. The limiting implication of the theorem leads to a system of local jurisdictions that consist of homogenous individuals. By definition, no redistribution is carried out since everyone is identical in such a community.

Second, poverty persists in communities of identical poor. After the breakup of the original community, residents are sorted to form separate communities of rich and poor. Residents in rich communities will gain, since there is no need for them to redistribute their income. People in the poor communities will lose, since they do not receive transfers after the breakup since there are no rich to count on in their community any more. It then follows that fragmenting a community does not yield Pareto superior outcome, unless there are intergovernmental transfers from richer communities to poorer communities.

For example, Hamlin (1991) formulates the theory that central governments are “constrained to adopt policies that apply uniformly across the entire population.” Crémer et al. (1995) also assume that “the central government is … institutionally compelled to adopt a uniformity constraint in service delivery.” Furthermore, Seabright (1996), while rejecting the original assumption as “empirically false,” claims for a weaker version that “central political systems do tend to implement policies that are regionally more uniform than decentralized ones.”

This of course assumes that some sort of redistribution was in place before the decentralization.
Third, the rich may also suffer if they are altruistic, since the poor remain poor in the fragmented communities without assistance in the decentralized world. Even if there are voluntary donations among communities, we see a case analogous to the voluntary provision of benefits with spillovers. In such a setting, the literature suggests that the total amount of voluntary giving would be undersupplied. Recall that one of the key assumptions for the decentralization theorem is that there are no benefit spillovers among communities, which implies that people do not care what happens to the poor in other communities. In the world of the decentralization theorem, “the spatial horizon of altruism” (Gramlich 1993) does not extend beyond community borders, and there is no sense of community across the nation. In other words, the theorem fails to the extent that redistribution is a national concern. We emphasize that the spatial horizon of altruism is one of the important factors in considering the proper assignment of redistributive functions. This point should be related to the concept of social solidarity we discussed in II.2.

### III.2. Social insurance

**Risk pooling**

Risk-averse individuals attain Pareto improvements through insurance. Such individuals prefer a certain level of income to an uncertain level of income even if the former is smaller than the latter. They are even willing to pay as premiums to obtain that certainty through insurance. Insurance thus attains Pareto improvements, giving “safety” in exchange of premiums. Of course, markets can provide insurance. But as Rothchild and Stiglitz (1976) show, if insurers cannot observe the characteristics of the insured, markets would fail and insurance may not be provided. In such a case, the public sector can intervene and provide social insurance, possibly by exerting

---

13 It might be similar to the current system of nation-states. Many people in developed economies do not seem to care about people in some LDCs even if they are starving to death or forced to move out of their settlements.

14 Ladd and Doolittle (1982) argue the case of poverty as a national concern.
its taxing powers, to attain Pareto improvements from an allocation with no insurance.

The question is again which level of government should provide social insurance. Insurance requires risk-pooling. Risk can be pooled to the extent that the income changes of participants in a program are idiosyncratic and the number of the participants is large enough. For example, consider an economy where individual $i$ obtains income $w_i$ which is a random variable with mean $\mu$ and variance $\sigma^2$. For an economy which consists of $N$ individuals, its guaranteed income is the average value of individual incomes, i.e. $\overline{w} = \sum w_i/N$. This guaranteed income will fluctuate with its variance

$$\text{var}(\overline{w} - \mu) = \frac{\sigma^2}{N} + \frac{\sigma^2}{N^2} \sum_{i,j} \text{cov}(w_i, w_j)$$

where $\text{cov}(w_i, w_j)$ is covariance between the incomes of individual $i$ and $j$. We then see that the larger size ($N$) contributes to a decrease in uncertainty (i.e., variance) of the guaranteed level of income $\overline{w}$. In other words, risks are pooled and the guaranteed income fluctuates less in a larger economy. The expression also suggests that less correlation among individuals’ incomes induces less fluctuation in the guaranteed income. In the limits, as $N \to \infty$ and $\text{cov}(w_i, w_j) \to 0$, $\text{var}(\overline{w} - \mu) \to 0$. The larger size of the economies then contributes to such idiosyncratic variations and, by definition, to the large size of shared individuals, and therefore to risk-pooling.

This line of argument clearly points to the advantage of national pooling of risks. First, the size of $N$ is by definition larger for central government than that for local government. Second, the smaller the space a jurisdiction covers, the more correlated their inhabitants’ incomes are. For example, labor markets are usually formed around a sub-sector of a country, and aggregate shocks are likely to happen at the regional level. Therefore, it should be difficult for local governments to pool the regional shocks. On the other hand, the center can better pool such regional shocks to some extent, albeit not perfectly. Of course, the center cannot handle national aggregate shocks, but neither the local governments.

**Scale and scope economies in premium collection**

In addition, scale economies in social insurance may also make a case against decentralized redistribution. For example, tax administration may exhibit scale as well as scope economies. It may be argued that premiums (payroll taxes) can be levied along with personal taxes more efficiently by a unitary government with standardized procedures than by independent local authorities with different tax codes. However, for certain reasons, premiums have to be collected without depending on existing tax schemes as in Japan. In that case, the premium collections are likely to involve some face-to-face interaction with subscribers, and as such, levies by local governments may contribute to the cost-efficiency of the administration. We will turn to the issue of cost-efficiency in V.

**IV. Migration and Redistribution**

Musgrave (1959) lays out a framework for assignment of fiscal functions to different levels of government and contends that redistribution should be a national responsibility. The primary reason for this contention is the mobility of tax bases. An aggressive local program for the support of low-income households would encourage
the exodus of those with higher incomes who must bear the tax burden, in addition to an influx of the poor. The mobility of tax bases thus constrains attempts to redistribute economic resources. In fact, the theoretical support for decentralized redistribution by Pauly (1973) assumes that individuals, both rich and poor, are immobile. However, as Pauly himself demonstrates, if the poor move beyond community borders, redistribution with no coordination among local governments will result in an underprovided level of transfers.

**IV.1. Theoretical developments**

The reduced redistribution due to tax base mobility is sometimes called a *race to the bottom*, which can be intuited as follows. Assume there are two regions where local governments redistribute money from the rich to the poor in their own jurisdictions. Assume furthermore that the poor are mobile between the regions. Policy makers in both regions have social welfare functions which include the utilities of the rich and poor in their region. Less poor will lead to a reduction in the taxes on the rich, thereby enhancing regional welfare. If one region increases its welfare benefits, the other region benefits since the poor move to the former region. In other words, an increase in welfare benefits in one region generates an external economy in the other region. We then know from the theory of externalities that what generates positive externalities tends to be underprovided. This line of argument started to be seriously studied by Brown and Oates (1987) and Johnson (1988), followed by a number of theoretical studies during the nineties, which more or less implies that decentralized redistribution results in the suboptimal level of redistribution (Crémer et al. 1996, Crémer and Pesteau 2004).

**The benchmark**

In what follows, we first follow Crémer and Pesteau (2004) to provide a bench mark model to illuminate the issues in the theoretical literature. Assume a country which consists of multiple regions. Each region is endowed with a fixed amount of capital $K$, high-skilled labor $M$, and low skilled labor $L$. Regional production is a function of these three factors $Y = F(M, L, K)$ with constant returns to scale. Assume that residents provide one unit of labor so that $M$ and $L$ respectively represent the number of high-skilled and low-skilled residents. The low skilled are endowed only with labor. The high skilled are endowed both with labor and capital. The utility of the residents is expressed as $u(c_i)$ where $c_i$ stands for consumption with $i = s$ for the low skilled and $i = u$ for high-skilled. An equilibrium in the consumption good market implies that $Y = M_s c_s + L_u c_u$. Local government maximizes a social welfare function $W(c_s, c_u) = M_s u(c_s) + L_u u(c_u)$ by redistributing consumption between the high- and low-skilled within its jurisdiction. The social welfare function is Benthamite, weighted with initial factor endowments $M = M^*$ and $L = L^*$ as distributional weights. The instruments for the redistribution are taxes on the production factors. With unit tax rate $\bar{t}$, the budget constraint for the government is given as $\bar{t}M + \bar{t}K + \bar{t}L = 0$.

Given this setup, equal consumption among different classes of labor ($c_s = c_u$) hold if production factors are immobile, a variant of the result shown in II.1. This implies that positive taxes on the high-skilled $\bar{t}M^* + \bar{t}K^* > 0$ and negative taxes on (i.e., positive benefits for) the low-skilled $\bar{t}L^* < 0$. However, when the low-skilled are free to move among regions, this result does not hold if each local governments set its taxes in the belief that its choice does not affect the tax setting of the other governments. It is shown that in a symmetric equilibrium where every
region is identical, the mobility of the low-skilled reduces the degree of redistribution. As the number of regions increases and approaches to infinity, the level of redistribution decreases to zero in the limits. An open economy and a large number of jurisdictions thus constitute a case for the race to the bottom. A number of modifications can be made to this benchmark, some of which are shown as below.

Asymmetric regions

The benchmark model assumes that all regions are identical and so are the equilibrium results. Introducing heterogeneity into the regions somewhat changes the results. In an asymmetric equilibrium with different regional endowments of factors, it would be straightforward to show that production efficiency does not hold due to different marginal products across the country. In addition, the reduction in redistribution may not hold in every region (e.g., Leite-Monteiro 1997, Wildasin 1991, 1994, 1996, Epple and Romer 1991). For example, assume that local governments have different objective functions in a two-region economy. One government maximizes the aggregate income of initial residents in region A, whereas the other government has an objective function of risk-averse type in region B. Furthermore, assume that only the high-skilled are mobile. Then, it can be shown that more redistribution is possible through the inflow of the high-skilled, whereas less redistribution is observed due to the outflow of the high-skilled.

Unemployment

In addition, the benchmark model assumes no unemployment. All production factors, including low-skilled labor, are fully utilized. However, given the facts that many actual instruments for redistribution are targeted at the unemployed and that local governments compete for capital to reduce unemployment in their regions, incorporating unemployment into the benchmark model may be required. Lejour and Verbon (1994) consider the case of opening capital markets while assuming the total supply of labor is fixed. Unemployment benefits are financed through payroll taxes. While the results depend on specifics of the model, they show that opening capital markets leads to a reduction in payroll taxes if capital is mobile. In a similar vein, Lozachmeur (2001) extends the celebrated Harris-Todaro model to the case of fiscal competition and shows that population mobility reduces the level of unemployment benefits.

Capital mobility

As shown above, studies that deal with capital mobility proceed their analysis by holding the level of another factor, i.e., labor, constant. However, Crémer and Pesteau (2004) consider the case where both factors move across jurisdictions. They show in a symmetric equilibrium that taxes on capital will be nil when the number of regions is large enough, showing another case of the race to the bottom even when all factors are mobile. Also, Lee (2002) examines the effect of making labor mobile in the presence of capital mobility with an asymmetric two-region model. His result contrasts to the race to the bottom in the sense that it shows that additionally allowing labor to be mobile in the face of capital mobility may lead to an increase in redistribution. Local taxation on capital causes its outflow, which reduces wages if its decreases the marginal product of labor. Therefore, even if the taxes on capital are used for wage subsidies, there may not be a new inflow of labor, since
the reduction in wages does not necessarily counteract an increase in the subsidies. If that is the case, we instead observe the outflow of labor, which then leads to a reduction of benefits for the recipients. Since the cost of redistribution will be lessened, we may see an increase of per capita benefits transferred to labor.

IV.2. Empirical evidence

The theoretical studies above indicate that redistribution would be too low at least in some regions of a country. Even if that was not the case, the literature unambiguously implies that resource allocation under decentralized redistribution is distorted. The key elements here include the mobility of labor (population) and the strategic behavior of local government which responds to the factor mobility. Since population mobility is a prerequisite for the race to the bottom, it is important to empirically examine if population responds to differences in local expenditures and taxes (or especially welfare benefits in this case). Even if consumers do not respond to the fiscal differences, however, local government may perceive such potential moves and act strategically. As such, it should also be important to empirically investigate the existence of such regional policy interdependence.

Migration responses

Starting as studies that examines the hypothesis of people voting with their feet (Tiebout, 1956), a series of empirical studies have been conducted to examine the effects of regional differences in taxes, user fees, welfare benefits and other publicly provided services (Brueckner 2000). The more responsive migration is to regional differences, the greater its effects on redistribution would be. While there are few studies that examine such effects in Japan, analogous studies are abundant in the United States where regional governments (states) have the power to redistribute.

Table 1 provides a selective list of studies that empirically examine the effect of the Aid to Family with Dependent Children (AFDC). The AFDC consists of programs jointly administrated by the federal and state governments where each state provides different transfer payments. The empirical studies took advantage of such differences in welfare benefits to see if they affect the migration behavior of the recipients. The AFDC was replaced by the Temporal Assistance to Needy Families (TANF) in 90s, and it was this occasion that spurred empirical analysis on the AFDC. As shown in the table, although there are a few studies that indicate no effects of welfare differences on migration (Shroder 1995, Levine and Zimmerman 1999), most of the studies show that the difference in welfare payments would affect the choice of residence location (Southwick 1981, Gramlich and Laren 1984, Blank 1988, Shroder 1995, Enchautegui 1997, Borjas 1999).

Indeed, there is additional evidence in favor of the race to a bottom. Brueckner (2000) surveys empirical studies (including those in the table) on the welfare program in the United States and concludes that the evidence appears consistent with the claimed propositions. Also, Feldstein and Wrobel (1998) present evidence that the attempts to redistribute income at the state level are largely unsuccessful. Progressive state income taxes in the U.S. have had little impact on the net-of-tax relative wage rates of skilled versus non-skilled workers. Their claim is that the mobility of workers across state borders undoes the effects of redistribution. The result is no redistribution, and only welfare losses from inefficient location decisions.
Table 1. Empirical studies on the AFDC

<table>
<thead>
<tr>
<th>Data sets</th>
<th>Programs</th>
<th>Who moves?</th>
<th>Does migration respond to welfare benefits difference?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Southwick (1981)</td>
<td>Nine regions within the US, 1967</td>
<td>AFDC</td>
<td>Female recipients</td>
</tr>
<tr>
<td>Levie-Zimmerman (1999)</td>
<td>Household data, National Longitudinal Survey of Youth, 1979-92</td>
<td>AFDC, food stamps</td>
<td>Non-working female with more than 2 years of welfare receipts</td>
</tr>
</tbody>
</table>

Notes. AFDC: Aid to Families with Dependent Children; SSI: Supplemental Security Income

**Strategic interactions**

There is also a growing literature since the late nineties that examines the strategic interdependence among decentralized redistributive policies of local governments. If there is such an interdependence among governments, the choice, on say, welfare benefits \( g_i \) made by government \( i \) is influenced by the analogous choice of other governments \( g_{-i} \), yielding a reaction function \( g_i = f(g_{-i}, M) \) where \( M \) denotes a set of variables for \( i \)'s characteristics that affect its choice of benefits. Then empirical studies estimate the following regression function for the reaction function:

\[
g_i = f(g_{-i}, M) + \epsilon
\]

where \( \epsilon \) is an error term. The econometric issues for estimating this reaction function includes the simultaneity of \( g_i \) and \( g_{-i} \), unobserved heterogeneity, and the spatial and serial interdependence of the error term. If we can properly estimate the parameters that characterize \( \partial f / \partial g_i \) as positive and statistically significant by allowing for the issues just described, we see that strategic interdependence does exist.

Let us continue with the AFDC studies. Saavedra (2000) estimates the slope of the reaction function \( \partial f / \partial g_i \) with statistical significance but with positive signs, which he claims contradicts a popular expectation that welfare expenditures in one region would be crowded out by an increase in those in other regions. Figlio et al. (1999) then specify a model that differentiates local responses depending whether they are responding to an increase or a
decrease in welfare benefits of other regions. They also estimate positive signs when state governments are responding to a decrease in other states’ welfare benefit. But negative signs are obtained when they are responding to a welfare increase in other regions. Furthermore, Smith (1997) shows that allowing lags in response results in insignificant coefficients for the reaction function. However, he also sets up an asymmetric response model for estimation to show that the responses are estimated negative if welfare benefits in other states are higher than those in the state under consideration. Otherwise, the slope is estimated to be zero: no reaction is expected for a state if its welfare benefits are higher than its neighbors. These results support the existence of some sort of strategic interaction among US states concerning the AFDC benefits.

V. Cost Efficiency and Local Discretion

The arguments so far have supported Musgrave’s classic proposition and made cases against decentralized redistribution. At the least, we could argue that raising revenues for redistribution within a nation as well as setting up national standards are the role of the central government. Even ardent supporters of decentralized redistribution seem to accept the view that redistributive programs should be financed out of national taxes (Rao 2002). Then, the question remains whether or not the expenditure side of redistribution should be decentralized. If decentralized redistributive expenditures have advantages, the combination of centralized taxes and decentralized expenditures supplemented by intergovernmental transfers seems to be the venue to pursue. However, before jumping to conclusions, let us closely examine the several aspects of decentralized expenditures.

V.1. The merits of local discretion in expenditure

The literature indeed provides cases for the decentralized provision of redistributive services. We follow Banting and Boadway (2004) to summarize the advantages of decentralized redistribution in terms of (1) information acquisition, (2) agent control, (3) innovation acquisition and (4) intergovernmental competition.

First, since local governments are ‘closer to people,’ they are better informed of the preferences and needs of their residents (information acquisition). This will be especially important for redistributive programs, which require information about the needs of individuals. Second, local governments can be more efficient at administering programs. Most redistributive services are delivered by institutions like welfare agencies and hospitals. Their management should be more effective at the local rather than the national level (agent control). Third, to the extent that there are different local programs, there will be more opportunities for local governments to devise effective programs (innovation acquisition). Experimentation at the local level will allow for trial and error that would lead to both successes and disappointments. It is better to have disappointments on a small scale than at the national level. A small success can easily be mimicked by other localities to yield a nation-wide impact. Lastly, local governments compete with one another to enhance efficiency (intergovernmental competition). Moreover, citizens are better able to hold their governments to account if they can observe what is going on in neighboring jurisdictions, a phenomenon described as ‘yardstick competition.’

We argue that these cases for decentralization should be reframed in terms of cost efficiency. Cost efficiency
is defined for a given level of public service. Assume a case where public service \( z \) is produced with input vector \( x \), with production technology \( z = f(x) \). The optimal choice of the input mix for a given level of \( z \) is given as the input vector \( x \) that minimizes expenditure \( w \cdot x \) subject to \( z = f(x) \) where \( w \) is a factor price vector. The minimized cost is thus given as \( c(w,z) = \min_x \{ w \cdot x \mid f(x) = z \} \), which is influenced by two components. One is the choice of a mix of inputs \( x \) utilized to produce a given level of public service \( z \). Cost efficiency is then measured with the distance between \( c(w,z) \) and the actual expenditure required to produce a given level of \( z \). The other component concerns the production technology \( f(x) \). Let us redefine the production function \( f(x, \theta) \) such that \( \frac{\partial f}{\partial \theta} > 0 \), where \( \theta \) is a productivity index which can be a parameter to be given or a control variable decision makers can choose.

With the above formulation at our disposal, the four cases for decentralizing redistribution may be interpreted as follows. First, the phrase ‘better informed’ in the case of information acquisition probably means that relevant information is obtained with ease. Cost minimization requires observing relevant technology \( f(x, \theta) \). However, index \( \theta \) may be local information that the central government can observe only with substantive costs. If so, allowing ‘better informed’ local governments to choose \( x \) freely results in an efficiency gain. Second, in the cases of effective control of agents, the efficiency index \( \theta \) is now a control variable rather than a parameter, which pertains to the behavior of the agents. The case presumes that local governments are able to control \( \theta \) better than the center. Third, the case for innovation acquisition assumes that \( \theta \) is an innovation that spills in from other jurisdiction. The case presumably argues that innovations are returns to a portfolio that consists of local projects. If we assume that the projects are diversified to the extent that local discretions are allowed, we could spread the risk of failing projects and free ride on the fruits of successful projects. Lastly, as for the case of intergovernmental competition, the literature has conceptualized intergovernmental competition as a device to tighten slacks in local expenditures. The case then assumes that \( \theta \) is a control variable, like effort, which tends to be suboptimal unless there exists competition among localities.

Since these advantages are related to the choice of technology, we argue that the determination of policy output \( z \) is irrelevant here. Recall that cost efficiency is defined for a given level of output, which is independent of who decides the output, nationally or sub-nationally. It then follows that those advantages should hold even when the national government dictates local governments to attain a given set of outputs, if the former allows the latter to freely choose inputs and technologies to produce the designated outputs. More importantly, if the national government consistently mandates required sets of public services, horizontal equity can be attained. In other words, the four merits of decentralization could be exploited without compromising the principle of horizontal equity.

15 For more formal definitions of the basic terminologies, see Kumbhakar and Lovell (2000).
16 Empirical analysis examines the effects of intergovernmental competition on efficiency using the frontier analysis. The slack is conceptualized as the distance from the frontier of the cost or production function (i.e., inefficiency term), and the effects are investigated as the impact of intergovernmental competition on the inefficiency term. See, for example, Davis and Hayes (1993), Duncombe et al. (1997), and Grossman (1999).
17 Of course, the first case (i.e., catering to local needs) does directly pertain to the level of output \( z \), as we discussed in terms of the Decentralization Theorem.
V.2 The demerits of local discretion

Note that the mere existence of local advantages does not guarantee the listed advantages. First, even if local governments are ‘closer to the people’ and better informed of the needs of local residents, they may not have incentives to utilize their advantages. For example, localization may bring officials in close contact with residents, open up opportunities for local interests to creep in and hereby breed corruption and favoritism. This tendency may even be stronger for redistributive programs. The same may apply to the advantages gained from the closer relation between the managers (local governments) and agents (welfare agents and hospitals). The agents themselves might turn out to be vested interests, forging a basis for corruption in the local administration.

Second, experiments and innovations may be underprovided or may not be provided at all. If information on innovation spills over, its acquisition is less costly than its invention. Since experiments involve a series of trials and errors which are costly for local governments, local governments may want to free ride on the efforts of other local governments. As the standard theory of voluntary contributions to public goods implies, the result would be underprovision or even non-provision of innovative activities of local governments.

Third, competition among local governments may not necessarily result in efficiency gains. As we have discussed, there is a possibility of the race to the bottom: independent redistributive spending at the local level may cause externalities through movements of potential recipients and contributors of the redistributive services, resulting in a suboptimal level of redistribution. The tax competition literature also shows that if local governments are benevolent (accountable for their citizens’ welfare), local government sets tax rates too low so that local public goods are underprovided when tax bases are mobile among jurisdictions. And the theory of fiscal externalities implies that the free mobility of residents results in inefficient allocation of the regional population (Flatters et al. 1974). In the cases where the output is fixed by the central government, analogous adverse effects may accrue to the input choice of local governments.

Fourth, the common pool problem may rise. When decentralized expenditures are combined with centralized taxes in an unevenly developed economy, localities will depend on national taxes as commons for redistribution. Since they in effect depend on revenue from outside through transfers, their perceived marginal cost of taxes would be smaller than actual marginal costs, resulting in excessive redistributive expenditures. However, this problem of excessive expenditures may well be counteracted by central controls and mandates, if our concerns are only for cost efficiency.

V.3 Central control and intergovernmental transfers

As we have seen, local discretion will not automatically solve the problem. The question here is how to design an administrative system that can exploit the merits of local discretion and constrain its demerits at the same time. Recall that, since cost efficiency is defined for a given level of redistribution, who determines the

---

18 Of course, this is an empirical issue (e.g., Rao 2002). But it being an empirical issue does not mean that we do not have to take caution against the possible danger of corruption due to decentralization.
levels of, or the standards for redistribution is irrelevant. Rather, as we argued when we discuss horizontal equity, the center should control the local outputs in some coherent manner if local redistributive expenditures are financed at the central level. Thus, the question relates to *administrative decentralization* rather than fiscal decentralization. While fiscal decentralization assigns sub-national governments the authority to legislate, design, and implement tax and/or expenditure policies, \(^{20}\) administrative decentralization concerns how *centrally financed* public programs are implemented throughout the nation. The concept is mainly defined on the expenditure side of fiscal activities, and the central financing could be partial or full. Following the literature (Rondinelli et al. 1983, Silverman 1992), administrative decentralization is categorized into several sub-concepts including *de-concentration* and *delegation*.\(^{21}\)

De-concentration refers to planning and/or implementation of national policies through local offices of the central ministries. In a de-concentrated administrative structure, central government employees are stationed at branch offices that are dispersed geographically, which is a common practice in the delivery of social services. For example, central governments set up local welfare offices and have them assess beneficiary eligibility and deliver social services like in the UK and Australia. De-concentration could be a mere spatial readjustment of workloads within the national government (i.e., a shifting of workload from centrally located officials to those located outside of the capital), or could be more than a readjustment.

While de-concentration only involves entities *within* national administrative structure, delegation takes advantage of entities *outside* of central administration.\(^{22}\) Such outside entities act as agents on behalf of the central government, which controls the financing and the overall design of a policy. The central government retains ultimate responsibility, provides funds if necessary, and specifies the functions and duties to be borne by the agents. Indeed, local governments can be agents of the central government. In fact, welfare programs and transfers to the needy are delegated to local governments in some unitary countries, a typical example of which is Japan. For example, the Japanese government established the Public Assistance Law, which sets the standards of

---

\(^{19}\) Different studies sometimes use the word ‘decentralization’ for different concepts, as shown by Yuliani (2004). For example, different contributors to the same single volume use the same terminologies in different ways (e.g., Ahmad and Tanzi 2002). Rondinelli et al. (1983) and Silverman (1992) classify the concept to include decentralization both within and outside of the public sector. Since we are considering decentralization within the structure of the public sector, our use of decentralization arguments exclude decentralization outside the public sector, namely, (i) decentralization to private sector entities without territorial jurisdictions (i.e., privatization), and (ii) decentralization to the people which includes the empowerment of local citizens and their representatives (i.e., democratization).

\(^{20}\) Decentralized redistribution then implies fiscal decentralization, since redistribution necessarily involves at least either taxation or spending, or both. Authority over one fiscal domain can be assigned to the center while another authority is assigned to local governments. For example, as mentioned above, spending may be decentralized while financing is kept centralized. Or some specific expenditure domain like education may be decentralized while others like infrastructure formation are kept centralized. Furthermore, authorities in some areas could be shared by both levels of governments.

\(^{21}\) In addition to these two types of administrative decentralization, there is another concept for administrative decentralization called ‘devolution.’ Rondinelli et al. (1983) and Silverman (1992) offer a definition where, under devolution, local units of government are autonomous and independent, and their legal status makes them separate or distinct from the central government, and that to the extent devolution advances, local governments do what they decide to do. However, again, the term devolution is sometimes differently used by different studies (Yuliani 2004).

\(^{22}\) Breton (2002) called this ‘field administration.’
public assistance. The national law requires local governments\textsuperscript{23} to set up welfare affairs offices to provide assistance to the needy. The welfare offices, which are within the administrative structure of local governments, follow the national standards to conduct means tests and provide assistance to the poor. The national government covers the expense through the categorical as well as general transfers.\textsuperscript{24}

These two concepts of administrative decentralization do not preclude local discretions. In fact, the central issue concerns the choice of types of organizations that act on behalf of the central government and the degree of discretion allowed for them. Local branches or local governments may be given some discretion to plan and implement programs. They may also be allowed to adjust central directives to their local conditions within national guidelines. To the extent that local discretions are allowed, policies could be implemented in a way that adjusts to local conditions. As a consequence, national policies, though formally uniform, could vary in their applications.

The difference between delegation and de-concentration also depends on the status of public officials, i.e., whether programs are implemented by public officials or national officials stationed locally. If delegation fails in attaining a given policy output, the national government may set up local branches to de-concentrate its redistributive policy. If “being closer to the people” is important, setting up local branches might simply suffice. If welfare agencies are effective when staffed with local personnel, it would suffice to employ local residents who have more local information. If policy experiments and competition are important, we may allow de-concentrated branches to compete.

Differences would be found in the incentives for public officials and the instruments for the central government, constrained by the administrative difference between the two. Gains from each of the two would depend on the existing structure and resources of the public sector. If local governments are harnessed with good administrative resources, delegation to local governments will be less expensive than de-concentration since the latter involves the setting up of regional offices of the central government. But, if programs are new and local governments lack resources to implement them, de-concentration may be desired instead.

One of the keys is the design of transfers. The literature offers some characterization of optimal transfers in the presence of principal-agent relations (e.g., Boadway et al. 1999) and fiscal externalities (e.g., Dahlby 1997, Boadway 2000, 2004). The case of de-concentration should have wider options for controlling locals than delegation since the former takes advantage of entities within the central administration, while the latter concerns those outside. When programs are administered by public officials employed by the national government, the center can in principle design incentives on a personnel basis, in addition to aggregate transfers to their local offices.

\textsuperscript{23} To be exact, towns and villages are not required to set up welfare offices. In case towns and villages do not set up such offices, prefectures assume the role to provide public assistance.

\textsuperscript{24} While the rest is to be covered by local taxes, general-purpose grants based on a gap-filling formula takes care of it if total local revenues fall short of the basic fiscal needs of that locality.
VI. Concluding Remarks

This paper has provided selected reviews on decentralized redistribution that may shed some light on policy issues towards decentralization in Japan. We framed causes for redistribution in terms of social justice (social welfare) and efficiency (Pareto improvement), and discussed several related issues. We have also reviewed local merits in delivering such services, and argue that those merits largely concern cost-efficiency and could be exploited without compromising horizontal equity. For this purpose, delegation or de-concentration may suffice. However, we have also pointed out that there still remain incentive problems, the key to which is the design of transfers and personnel control. The literature on fiscal federalism provides some clue in regard to this issue, but there should definitely be more to add to the design of optimally delegated and/or de-concentrated systems of redistributive policy.
References


Switzerland. (available at http://www.intercooperation.ch/offers/events/interlaken-2004/draft-def.pdf)