Decentralization with national standards. The case of the Netherlands.

Paper prepared for the 2011 Copenhagen Workshop "Normative Frameworks of Decentralisation and Intergovernmental Fiscal Relations: Efficient Instruments for Desired Balance between Decentralization and Merit", jointly organized by the Danish Ministry of the Interior and Health and Korea Institute of Public Finance

Maarten A. Allers
Professor of Economics of Sub-National Government
Director, Centre for Research on Local Government Economics (COELO)
University of Groningen

1. Introduction

One of the classic reasons for governments to provide goods and services to their citizens is that policymakers consider some goods or services essential or beneficial. Examples of such merit goods are education, health and social protection. When left to themselves, citizens would not spend enough on these goods, according to policymakers. Therefore, the government steps in.

This approach may conflict with decentralization of government, however. If subnational governments are free to choose local service levels, these might be too low, in the view of national policymakers, in the case of merit goods. Thus, in some countries, national standards apply for a number of public services. But if national standards apply to locally run spending programs, local jurisdictions must be enabled to finance them. How can this be organized?

This paper first describes the institutional and normative environment in the Netherlands, highlighting the strong emphasis on equity considerations and the supply of merit goods in the Netherlands. The second part of the paper zooms in on a particular program, which aims to ensure equal access to welfare benefits throughout the country, while at the same time making local governments responsible for welfare administration. Local governments receive intergovernmental grant to finance welfare. These grants are not related to actual welfare payments. Rather, they reflect local economic and demographic conditions which influence the probability that inhabitants will become welfare dependent. This program has been rather successful in reaping the benefits of decentralization while at the same time ensuring a high degree of similarity in service standards across the country.

2. Institutional background

Subnational government in the Netherlands

There are three territorial levels of government in the Netherlands: central government, provinces (12) and municipalities (418 in 2011). Each level covers the whole country. All provinces and all municipalities more or less face the same responsibilities and have the same tax autonomy. Dutch municipalities, with 40,000 inhabitants on average, are relatively large compared to those in other countries. Municipalities, which spend about eleven percent of GDP, provide many of the services that are of daily importance to citizens: from sewers to refuse collection, from local roads to poverty relief.

Provinces spend relatively little (1.4 percent of GDP). However, provinces have a number of tasks that do not require spending a lot of money, like spatial planning and coordinating regional public transport. There is a third local government level in the Netherlands: the water boards (25 in 2011). These originate from the 12th century, and are responsible for keeping the country, an important part of which is below sea level, dry and safe from floods.

Water boards, municipalities, and provinces are each ruled by subnational governments controlled by parliaments which are elected every four years.

Municipalities

Dutch municipalities finance their spending through specific (18% in 2010) and general grants (34%) from the central government, municipal levies (14%) and income from property and market activities (33%).

Specific grants are received from different central government departments and are earmarked to finance local government tasks imposed by the central government, whereas general grants are used to finance the autonomous tasks of local governments. In practice, the distinction between specific and general grants has been blurred somewhat in recent years. Specific grants have become broader, giving local governments more spending discretion, while general grants are increasingly tied to specific spending programs, be it mostly implicitly. The number of specific grants, and the amounts involved, have decreased significantly in the first decade of this century. This reflects the growing insight that local governments should not depend too much on earmarked revenues, because an appropriate local trade-off between the costs and benefits of public services can only be made if municipalities have enough spending autonomy.

General grants are formula-based and independent of local spending and taxation. These are equalizing grants: the formula takes into account spending needs and tax capacity of municipalities. The rather ambitious aim is to enable all municipalities to finance equivalent service levels at equivalent tax rates. Municipalities are free to choose higher or lower service levels and corresponding tax rates. The Dutch grant allocation system, which consists of 60 criteria, is one of the most complicated systems in the world. It reflects the great importance the Dutch attach to equity. Citizens in similar circumstances are considered to have equal rights to government services, independent of where they reside. In, e.g., the United Kingdom, a similar aspiration applies. Nearby countries like Belgium and Germany take a more relaxed attitude, however.

Municipal levies consist of user charges (7% of total revenue; mainly for sewerage and garbage collection) and local taxes (7%). Local tax revenue is dominated by the property tax. The only other local tax of any importance is the parking tax, which cannot be raised easily for fear of discouraging potential visitors of local businesses. Since user charges are not allowed to exceed (budgeted) costs and municipal budgets must be balanced, higher service levels can only be funded by raising taxes, i.e., by setting higher property tax rates.

Although municipalities are free, in theory, to choose local service levels and tax rates, deviation from service levels or tax rates in other municipalities is only tolerated to a limited degree. Deviation from the norm may easily result in public resistance. Members of the (national) parliament are quick to ask the national government to intervene. Pressure groups use municipalities which spend relatively much on their hobby horses to set a norm with which other municipalities can be compared. It is no surprise, then, that municipalities have been shown to mimic expenditure and tax levels of neighboring jurisdictions (Allers and Elhorst 2011).

Decentralization

Dutch law prescribes that the national government must stimulate decentralization. In the last decade, several important tasks have been decentralized to municipalities. The present government, which came to office in 2010, is set to continue this policy. In the recent past, large scale decentralizations occurred in welfare and social services. In the next few years, youth care will be decentralized from (mainly) the provinces to the municipalities, along with some other social services. Moreover, locally administered welfare is to be merged with entitlement programs for disabled workers, which are currently the responsibility of the central government.

Decentralization is promoted because it is thought to give citizens more control over public services which they need in daily life. Municipalities can tailor services to local preferences and needs. That should result in higher welfare than uniform service levels across the country. Another reason to decentralize is that municipalities are supposed to be more efficient. The decentralization of welfare has indeed been rather successful in that respect. However, the central government systematically allocates smaller budgets to municipalities than those used for the services involved before they were decentralized. This puts pressure on the local governments' budgets.

Another problem is that small municipalities find it hard to cope with the continuing stream of new responsibilities. Often, they do not have enough staff to allow specialization. Many observers question the ability of small local governments to absorb new tasks. As a result, small municipalities merge and amalgamate at an alarming rate. In 1980, there were 811 municipalities. In 2011, there are 418 left, and the end of this process in not in sight. Both Denmark and Belgium applied a big bang approach to amalgamation, in 2007 and 1976, respectively. In the Netherlands, however, amalgamation is an ad hoc process. There is no grand design which guides the amalgamation process. Hardly a year goes by without some municipalities being merged. No end goal has been defined, e.g., a desired number of municipalities or a desired minimum size.

There seems to be a paradox here. Decentralization is partly driven by the desire to move decision making to the lowest possible level of government, in order to allow optimal citizens' control. But because decentralization leads to ever growing municipalities, decision making is moved steadily to larger units at greater distance from the citizen.

3. Welfare: decentralization with uniform standards

The drive at decentralization, combined with a strong attachment to equal rights to government services, is exemplified by the way welfare is administered and financed in the Netherlands.

Welfare financing

In many countries, welfare is administered locally, but financed by the central government. Decentralization allows public services to be tailored to local preferences. Decentralization may also be more efficient, as knowledge of local circumstances is needed to successfully run a welfare program. However, decentralized finance of redistributive programs is likely to break down as a result of the migration patterns it brings about. Jurisdictions with high welfare dependencies need to levy high taxes if they are to finance welfare themselves. This drives away richer inhabitants, as a result of which tax rates need to be raised even higher. Therefore, income redistribution is generally considered the responsibility of the central government. As a result, welfare is usually administered locally, but financed centrally. This raises the question as to how the center can induce local administrators to administer welfare efficiently. If the money comes from elsewhere, why bother restricting welfare benefits to the truly needy?

Matching grant versus block grant

Until 2001, each Dutch municipality financed 10 percent of welfare benefits it paid out from its own coffers, while 90 percent was reimbursed by the central government through an open-ended matching grant. Clearly, this did not provide a strong incentive to limit welfare payments by helping recipients find work or by clamping down on fraud. Matching grants are not efficient because they reduce the costs to the local government of an extra welfare beneficiary. Therefore, they reduce the incentive of the local administration to keep welfare dependency at a minimum. On the other hand, matching grants are equitable because they guarantee that the central government shoulders an equal share of every local government's welfare burden. As a result, jurisdictions with high welfare spending needs due to exogenous circumstances receive a larger grant.

In contrast, block grants are efficient, as they do not lower the cost of additional welfare recipients. But this comes at a price, as there is generally no guarantee that the welfare burden of every local government is shared by the central government to the same extent. Block grant financing may force local governments in economically backward regions to spend considerable sums of money from their own resources on welfare, while jurisdictions in affluent regions may not need to spend all the grant money on welfare. Welfare is an entitlement program; people who qualify cannot be denied welfare. The grant system should reflect this.

In order to improve the incentive for local government to administer welfare efficiently, the Netherlands reduced the match rate from 90 to 75 percent in 2001. As from 2004, no reimbursement takes place any more. Matching grants have been replaced by block grants. If a municipality spends more than its block grant, it bears the extra expenditures itself. If it spends less, it may use the balance as it sees fit. Block grants for welfare are used in other countries as well, e.g., the United States. However, the inequities inherent in the usual block grants are unacceptable in a country like the Netherlands. Therefore, a sophisticated allocation system was set up.

The Dutch block grant

In the Netherlands, block grants are allocated in such a way that municipalities that operate in a reasonably efficient way will not need to use own resources to finance welfare expenditures. The allocation formula is based on an econometric model. At the same time, total grants add up to no more than forecasted aggregated welfare expenditures.

The new financing arrangement was accompanied by greater local autonomy in the treatment of welfare recipients. However, it is important to stress that this new autonomy is limited to administration. Local governments have discretion over the programs they employ to assist recipients in moving from welfare to work, and the intensity of their fraud investigations. Eligibility rules and welfare benefit levels are still uniform and determined centrally, reflecting the strong Dutch preference for equity. This is an important difference with the 1996 welfare reform in the US. Municipal spending on welfare benefits can only be lowered by reducing caseloads. As assistance to the truly needy cannot be refused, caseloads can only be reduced by weeding out fraudulent beneficiaries and by helping *bona fide* recipients find work.

The nationwide budget available for welfare block grants, referred to as the *macro budget*, is calculated annually based on forecasts of the number of persons eligible for welfare. These forecasts are made by the independent Netherlands Bureau for Economic Policy Analysis, known by its Dutch acronym CPB. Forecasts are based on the number of welfare beneficiaries, the development of the number of unemployed in the previous years, and regulatory changes that may affect welfare volumes.

The macro budget is allocated over municipalities according to the following rules. For small municipalities (fewer than 25,000 inhabitants, where 9 percent of welfare recipients live), the share of the macro budget in year t depends on their share of welfare expenditures in year t-2. For large municipalities (40,000 inhabitants and more, 80 percent of welfare recipients), a formula applies, which includes both demographic characteristics and labor market characteristics. The allocation formula is updated annually. Because a formula that covers smaller municipalities reasonably well could not be derived, this method does not apply to them. For medium-sized municipalities, a hybrid system applies: their share is partly derived from their expenditure share in year t-2, and partly from the formula.

It has proven difficult to derive a stable allocation formula. Municipalities may see their calculated share of the macro budget rise or fall considerably from one year to the next. In order to insulate local governments from budgetary shocks too great to cope with, the difference between the block grant and actual welfare expenditures are limited both ex post and ex ante.

Presently, the welfare grant allocation formula contains 14 variables. Among these are the number of single parent households, the number of lowly educated people, employment growth in the region to which the municipality belongs, and the number of disability benefits. The weights of these variables are derived annually from a regression at the municipal level of welfare volumes on the determinants included in the formula.

Problems with the regression- based approach

This approach is not without problems. Municipalities operate at different levels of efficiency, but efficiency levels are not observed. Actual welfare volumes, used in the regression analysis, are a biased indicator of spending need, which is defined as the welfare spending a municipality would incur if it operated efficiently (as defined above). Since inefficiency is not observed, the regression has an omitted variable problem and thereby a bias.

A second problem with the regression method is that it gives rise to perverse incentives. Matching grants reduce the marginal cost of welfare spending and thereby increase the attractiveness of working inefficiently. Standard block grants do not have this property. However, the sophisticated block grants discussed here have the property that higher expenditures increase future grants. That is because higher expenditures influence the outcome of the regression analysis on which the future grant is based. This provides perverse incentives to local administrators. Greater efficiency in the past results in lower welfare expenditures, which translate into lower weights in the regression formula for the variables reflecting the characteristics of the municipality, and therefore into a lower grant. As a result, bad behavior in the past is rewarded. This results in perverse incentives which distort efficiency.

The best of both worlds?

Because the allocation formula is updated annually, policymakers assume that as the new grant design improves efficiency across the board, the regression bias will gradually disappear, and efficiency will improve everywhere. If that would be the case, the Dutch would have the best of both worlds: the efficiency of block grants, without the inequities usually associated with them.

Unfortunately, this does not seem to be the case. Toolsema and Allers (2011) show that the bias in the regression-based allocation system will only disappear in extreme circumstances. As long as some municipalities continue to operate below the maximum efficiency level (i.e., as long as welfare dependency is not reduced to the absolute minimum in every local jurisdiction), the bias continues to exist. However, they also show that the perverse incentive is rather limited, at least in the Netherlands. This is generally the case where the number of local jurisdictions is large compared to the number of parameters in the econometric model (Toolsema and Allers, 2011). As a result, the sophisticated block grant presently in use in the Netherlands is rather efficient.

Another problem associated with the regression-based approach is that a perfect econometric model to forecast welfare dependency cannot be found in practice. Welfare dependency seems to depend for an important part on unobserved variables. As a result, some municipalities seem to receive considerable less grant money than they should, while others get too much. This is another reason that, apart from being less than perfectly efficient, the Dutch system is also less than perfectly equitable.

Of course, nobody would expect a real-world program to be perfect. The question is whether it is good enough. Clearly, the Dutch think their welfare financing system works pretty well. The introduction of the new financing system in 2004 was followed by a significant fall in welfare dependency (Kok et al. 2007; Van Es and Van Vuuren 2010). Opposition from municipalities which reckon they receive too little grant money is vocal, but not strong enough to alter welfare financing policy. Presently, the government plans to set up a similar system for related social protection programs which are about to be decentralized.

4. Summary and conclusions

The desire to provide merit goods may clash with decentralization of government, as decentralization may lead to different service levels, while merit good considerations compel national governments to set nationwide service standards. This is especially apparent in the Netherlands, a country with a strong egalitarian tradition and an active and decentralizing government.

A case in point is the Dutch welfare program. Although it is administered by local governments, eligibility rules and benefit levels are uniform across the country. This raises a principal-agent problem: how can local governments be persuaded to administer welfare efficiently, i.e., to reduce welfare dependency?

Dutch local governments receive an intergovernmental grant to finance welfare, which is not related to actual welfare payments. Grant allocation is based on regression analysis and reflects local economic and demographic conditions which influence the probability that inhabitants will become welfare dependent. Thus, every local government in theory receives enough to finance welfare benefits to those who truly need them. If they pay out more benefits, they bear the associated costs themselves. This program has been rather successful in reaping the efficiency benefits of decentralization while at the same time ensuring a high degree of similarity in service standards across the country.

References

M.A. Allers, Elhorst, J.P. (2011), A Simultaneous Equations Model of Fiscal Policy Interactions, *Journal of Regional Science*, 51, 271-291.

Kok, L., Groot, I., Güler, D. (2007), *Kwantitatief effect WWB*, Amsterdam: SEO.

Toolsema, L.A., Allers, M.A. (2010), *Welfare financing: Grant allocation and efficiency*, unpublished paper.

Van Es, F., Van Vuuren, D.J. (2010), Minder uitkeringen door decentralisering bijstand, *Economisch Statistische Berichten*, 95 (4589), July 9.