Società italiana di economia pubblica
XX CONFERENZA

ECONOMIA DELLA TASSAZIONE
Sistemi tributari, pressione fiscale, crescita

Pavia, Università, 25 - 26 settembre 2008

SPEND MORE, GET MORE?
AN INQUIRY INTO ENGLISH LOCAL GOVERNMENT PERFORMANCE

FEDERICO REVELLI

pubblicazione internet realizzata con contributo della
COMPAGNIA di San Paolo

Via Po, 53 – 10124 Torino (Italy)
Tel. (+39) 011 6704917 - Fax (+39) 011 6703895
URL: http//www.de.unito.it

WORKING PAPER SERIES

Spend more, get more?
An inquiry into English local government performance

Federico Revelli

Dipartimento di Economia “S. Cognetti de Martiis”

Working paper No. 4/2008

Università di Torino

Spend more, get more? An inquiry into English local government performance*

Federico Revelli†

April, 2008

Abstract

Based on a unique measure of performance of English local governments in the provision of public services (Comprehensive Performance Assessment, CPA), this paper develops a simple analytical framework that fully encompasses the institutional features of the British system of local government finance in order to model the process of performance determination, and uses panel data (2002-2007) to identify the determinants of local government performance. Due to the nature of CPA ratings - measured on a five category (poor to excellent) scale - the empirical work relies on an ordered response approach allowing for cross-sectional heterogeneity. Maximum likelihood estimation of a random effects ordered probit model provides no evidence in support of the "spend more, get more" hypothesis, but rather suggests that spending in excess of centrally set standards has a detrimental effect on local public service performance.

JEL classification: C23; C25; H72.

Key words: local public expenditures; performance rating; random effects ordered probit.

*Preliminary, comments welcome

†Address: Department of Economics, University of Torino, Via Po 53, 10124 Torino (Italy); e-mail: federico.revelli@unito.it.

1 Introduction

The rising role of decentralized governments in the provision of public services in both developed and developing countries has spurred a growing research into the determinants of variously defined measures of subnational government performance.[1] Of particular interest from a public finance standpoint is the evaluation of the impact of public resources on local government performance.

A well established strand of the literature exploits the decentralized provision of education that is observed in many countries and estimates production functions for education where performance (be it at the state, school district or school level) is typically proxied by gross educational output. While the review in Hanushek [18] provides little evidence in favour of the hypothesis that public resources (in terms of class size, teacher qualification and experience, and expenditure per pupil) have a positive and significant impact on pupils' achievements (in terms of standardized test scores, pass rates and drop-out rates), the most recent studies (Papke [26], Barankay and Lockwood [6], Leuven et al. [24]) report some more mixed results.[2] On the other hand, some studies in the economics of education literature proxy school performance by the marginal effect of schools on educational outcomes, and extract the value-added of schools from the residuals of a school gross output equation

[1] According to Joumard and Kongsrud [22], the average share of subnational government spending increased in the majority of OECD countries in the past two decades and reached $\frac{1}{3}$ in total public spending by the early 2000s. They also provide evidence, though, that national governments countered that tendency by imposing stricter norms and minimum quality standards on the locally provided public goods.

[2] Papke [26] considers Michigan schools and exploits the dramatic changes in funding schemes brought about by the centralizing school finance reform in 1994, and finds that spending has a significant positive effect on student achievement. Barankay and Lockwood [6] consider Switzerland and the heterogeneous degree of decentralization of the education sector, and find a positive impact of the degree of decentralization and of the level of spending per pupil on educational outcomes. On the other hand, Leuven et al. [24] find negative effects of targeted subsidies at schools with large proportions of disadvantaged students on nationwide exam achievements in the Netherlands.

that controls for demographic, ethnic and socio-economic composition of the student body (Grosskopf et al. [15], [16]).

As far as health care services are concerned, in spite of the fact that measurement of health care performance occurs in almost all OECD countries, the detailed review in Propper and Wilson [27] points to the technical difficulties in constructing meaningful measures of performance and value-added outcomes and to the lack of rigorous evaluation of the impact of performance measures on the response by decentralized health-care providers, and urges a rethinking of the existing performance measurement schemes.[3]

Finally, a strand of the literature aims at measuring the overall performance of multi-purpose decentralized governments. In those studies, local government performance is typically proxied by the degree of technical efficiency and is estimated via a number of stochastic or non-stochastic techniques.[4] Recent works in this area include Hayes et al. [20], Geys [12], Revelli and Tovmo [30] and Afonso and Fernandes [1], where the issue of overall local government efficiency is confronted with data on municipal governments in Illinois, Belgium, Norway and Portugal respectively. Grossman et al. [17] consider a sample of US central cities and, based on the argument that local government efficiency is capitalized into property values, take the latter as a measure of the output of local government activity, and a recent paper by Hauner [19] estimates the efficiency of public expenditures on health, education and social protection by the regions of the Russian Federation.

This paper aims at contributing to the existing literature on decentralized government performance in the following ways. First, based on the properties of a unique measure of local government performance that has been available for English authorities since 2002, this paper provides a simple theoretical framework that fully encompasses the institutional features of the British

[3]For a cross-country comparison of national government performance with respect to health, education and public infrastructure outcomes, see Afonso et al. [2].

[4]An early analysis in this spirit, though restricted to the evaluation of the efficiency of US police departments in producing a single output (crime rate), is Davis and Hayes [9].

system of local government finance in order to highlight the effect of public expenditure on performance. In particular, the theoretical set-up takes as the crucial input variable in the performance determination process the deviations of actual spending decisions from centrally set spending standards, and exploits the mandatory nature of local public service provision requirements to model the effect of public resources on performance. Second, the paper uses panel data on institutional, financial and socio-economic characteristics of the 148 main local authorities in England in the 2002-2007 time span to identify the determinants of performance ratings.

The performance measure that is analyzed here - the CPA (Comprehensive Performance Assessment) rating - has a number of attractive features. First, it is built by an independent Commission - the Audit Commission - which annually reports on its findings (mainly based on audit and inspection activity) and categorises English local authorities in a consistent, comparable and transparent way. Second, the CPA performance measure has the unique feature of combining indicators of service quality with measures of per unit costs of services, thereby approximating an ideal performance measure that promises to be superior both to indices built on gross output only and to crude measures of technical efficiency. Third, CPA is a single judgement of the overall performance of local governments on a wide range of important local public services (including education, personal social services, transportation and environmental management and protection) on a five category scale (poor to excellent). In spite of the computational complications arising from the categorical ordered nature of the rating (requiring an ordered response latent variable econometric model), the nature of the CPA rating has the advantage of summarizing the overall activity of each local government in a simply understood index.

Moreover, an attractive feature of the British system of local government finance is that central government sets standard spending levels for each local authority based on assessed spending needs, thereby allowing us to estimate

the impact of local expenditures relative to centrally assessed standards on performance. In other words, it allows us to test whether local authorities spending in excess of the centrally set standards are able to provide more and better services to their citizens.

Maximum likelihood estimation of a random effects ordered probit model that accounts for the categorical nature of CPA ratings provides no evidence that higher public spending translates into better service performance. Rather, the estimate of the effect of excess spending on performance is negative. Moreover, controlling for correlation between jurisdiction-specific effects and regressors and for fixed characteristics of a locality - including institutional structure and socio-economic complexion - provides further evidence in support of the hypothesis that public expenditures in excess of centrally set standards have a detrimental effect on performance.

The remainder of the paper is organized as follows. Section **2** develops a simple theoretical framework for the analysis of the performance determination process, and models the link between public spending and performance by thoroughly exploiting the unique features of the British system of local government finance. Section **3** turns to the empirical analysis based on a panel data set of the 148 main local authorities in England over the 2002-2007 period, and section **4** concludes.

2 Some simple performance geometrics

Let the performance of local government i (π_i) be a function of the level of public services provided to the final user (s_i) and of the unitary cost of those services (c_i):[5]

$$\pi_i = \pi(s_i, c_i) \tag{1}$$

[5]This and the subsequent hypotheses are meant to capture the fundamental spirit of the performance evaluation system described in more detail in the Appendix, as well as the key institutional features of the British system of local government finance.

with: $\partial\pi_i/\partial s_i > 0$, $\partial\pi_i/\partial c_i < 0$, $\partial^2\pi_i/\partial s_i^2 < 0$, $\partial^2\pi_i/\partial c_i^2 < 0$, and where s_i and c_i are defined as:

$$s_i \equiv \frac{S_i}{n_i^u} \tag{2}$$

$$c_i \equiv \frac{E_i}{S_i} \tag{3}$$

where S_i is total provision of public services that we assume here to be rival in consumption, E_i is total public expenditures, and n_i^u is the number of users. n_i^u is assumed to be made of a deterministic ($\overline{n}_i^u$) and of a random (ε_i) component:

$$n_i^u = \overline{n}_i^u + \varepsilon_i \tag{4}$$

with: $E(\varepsilon_i) = 0$ and $Var(\varepsilon_i) = \sigma_\varepsilon^2$. We make the crucial assumption that policy-makers decide on public spending before observing the realization of ε_i, and make their policy decisions based on its expected value. Consequently, ε is orthogonal to E.

Local governments are heterogeneous with respect to an exogenous degree of "inefficiency" α_i, i.e., the fraction of total spending that is wasteful, in the sense that it does not contribute to the production of public services. As a result, the cost function for the production of local public services can be expressed as:

$$c_i = \frac{1}{1-\alpha_i} c(s_i) \tag{5}$$

with: $c(0) = 0$, $c' > 0$, $c'' > 0$. By representing the locus of (s_i, c_i) pairs that can be obtained by varying the level of public expenditures, equation (5) can be interpreted as a performance production frontier (PPF) for local government i, conditional on its exogenous degree of inefficiency.

Moreover, let central government set a "standard spending level" E_i^*, i.e., the level of expenditure at standard levels of inefficiency α^* that allows each government, given its expected spending needs $E(n_i^u) = \overline{n}_i^u$, to provide a standard level of public services s^* at the unitary cost c^*:

$$c^* \equiv \frac{E_i^*}{s^* \overline{n}_i^u} = \frac{1}{1-\alpha^*} c(s^*) \tag{6}$$

The standard level of public services is mandatory, in the sense that local authorities have to provide at least s^*:

$$s_i \geq s^* \tag{7}$$

Based on the standard spending assessment E_i^*, central government sets up a grant distribution scheme that allows each government to achieve the standard level of expenditure by exerting a standard tax effort ($\tau_i = \tau^*$) on an exogenously fixed tax base (B_i). Consequently, the grant distribution system effectively equalizes all non-stochastic differences in tax base (B_i) and spending needs ($\overline{n}_i^u$). Finally, the government in jurisdiction i sets the level of spending E_i in order to maximize π_i subject to (5) and (7), and conditional on its exogenous degree of inefficiency α_i.

Figure 1 offers a geometric representation of the constrained optimization problem of local government i. The performance production frontiers (PPF) in figure 1 show the pairs of s_i and c_i that can be obtained at any given level of inefficiency by varying the level of public expenditures E_i. Lower (higher) inefficiency is represented by upward (downward) shifts of the PPF. Let $PPF(\alpha^*)$ be the performance production frontier at the standard level of inefficiency and (s^*, c^*) the public services and cost pair when expenditures equal the centrally set standard E_i^*.

Figure 1 also shows iso-performance curves π_i, i.e., the locus of (s_i, c_i) pairs generating the same level of performance. Assume that point A^* in figure 1 - corresponding to $E_i = E_i^*$ - attains the highest performance π^* conditional on the inefficiency level α^*. This implies that, at point A^*:

$$\frac{1-\alpha^*}{c'(s^*)} = -\frac{\dfrac{\partial \pi_i}{\partial c_i}}{\dfrac{\partial \pi_i}{\partial s_i}} \tag{8}$$

with the slope of the performance production frontier PPF^* equaling the slope of the iso-performance curve.

In order to examine the incentives and constraints generated by the performance rating scheme, consider the case of two governments of inefficiency α_1 and α_2 respectively, with $\alpha_1 < \alpha^* < \alpha_2$. Since the grant system equalizes all non-stochastic differences in the number of users of public services, the only source of cross-jurisdictional heterogeneity in this context is the rate of inefficiency α_i, that we assume here to be an exogenous parameter that - at least in the short run - is not under control of politicians and may be determined by the structure and organization of the local bureaucracy. The α_1-government will end up in point A_1, attaining performance π_1. The relatively efficient government is subject to two effects when setting public expenditures. First, being relatively more efficient than the average ($\alpha_1 < \alpha^*$), it faces a lower cost of production, thereby having an incentive to supply more local public services (a substitution effect). Second, the lower degree of inefficiency also generates an income effect pushing in the direction of higher s_i, due to the fact that, relatively to an α^*-government, the mandatory level of services can be delivered at a lower cost. While lower inefficiency definitely implies higher supply of public services, the effect on total public spending is ambiguous.

On the other hand, an unconstrained α_2-government would end up in point A_2 with performance π_2. However, due to the constraint (7), a relatively inefficient government facing the performance production frontier $PPF(\alpha_2)$ has to raise spending up to point A_3, attaining the mandatory level of services s^* and the performance $\pi_3 < \pi_2$.

To sum up, governments that are more efficient than the average will provide higher than standard public services, and their level of spending might be either higher or lower than standard spending. Governments that are less efficient than the standard need to spend more than the standard, achieve the mandatory level s^* and attain low performances. In a way, inefficient governments are subject to two contradictory sets of incentives. In the presence of high inefficiency, the performance-maximizing strategy requires limiting

public expenditures (point A_2 in figure 1): since the government is unable to provide public services efficiently, the best policy consists in doing as little as possible (public service level s_2). On the other hand, the mandatory requirement on the minimum level of services to be provided (s^*) forces inefficient governments to spend up to point A_3, raise the unitary cost of public services to c_3, and reveal themselves as bad performers.

3 Spend more, get less?

In 2002, a system of rating of local government performance - the CPA (Comprehensive Performance Assessment) - was introduced in the United Kingdom in order to measure how well Councils deliver services for local communities.[6] An independent body (the Audit Commision) assesses the performance of the 150 English authorities that are responsible for the bulk of local public expenditures (including education, social care, roads and transport, and environmental services).[7] The Audit Commission annually delivers CPA ratings based both on its own audit and inspection activity and on the assessments provided by other independent Commissions (Audit Commission [4]).

CPA aims at looking at performance from a range of perspectives and, based both on existing service performance information through national indicators and on relevant inspections, it provides a simply understood rating on a five category (poor to excellent) scale of the performance of the English local authorities in exercising their functions.

The CPA framework has two main components: an assessment on the use of resources ("value for money") and an assessment on the level and quality of public services. The distinct ratings earned by local authorities on those two dimensions are then combined with an evaluation of the organization and

[6]For a critical assessment of the ability of CPA to properly capture public service quality, see Andrews et al. [3].

[7]Those authorities comprise 34 Counties and 47 Unitary Authorities in non-metropolitan areas, and 69 Authorities (Boroughs) in metropolitan areas.

direction of travel of the Council to generate one of five rating categories.[8]

Since its introduction, the CPA system has produced six waves of ratings of local governments (2002 to 2007), as shown in table 1.[9] As far as overall performance is concerned, the rating system appears to have had a significant positive effect. Table 1 shows that Council performances increased significantly, with more than half of the authorities exhibiting a performance improvement since the start of the system. Moreover, Councils achieving excellent performance rose from 21 in 2002 (14%) to 55 in 2007 (37%), and no Council was judged to perform poorly after 2005 (Audit Commission [5]).

Due to the nature of the CPA system, we estimate an ordered response empirical model based on the idea that the size of a continuous latent variable (π^*_{it}) determines the observed categorical outcome (π_{it}). In particular, let the unobserved variable originating the observed performance rating in jurisdiction i be expressed as a linear function of a vector of observed time-varying local characteristics including fixed year effects (z_{it}) and of a vector of time-invariant structural local characteristics (q_i), plus an error term η_{it}:

$$\pi^*_{it} = z'_{it}\beta + q'_i\gamma + \eta_{it} \tag{9}$$

Depending on the realized value of π^*_{it}, the observed performance ends up into one of five ratings: poor, weak, fair, good, excellent.

Due to the fact that the poor performance category includes a small and fading number of authorities along the six years (23 observations, with zero counts in 2006 and 2007), it seems preferable to pool the poor and weak categories and implement the model according to the following thresholds:

$$\pi_{it} = \begin{cases} \text{poor/weak} & \\ \text{fair} & \text{if} \\ \text{good} & \\ \text{excellent} & \end{cases} \begin{array}{c} \pi^*_{it} \leq \Pi_1 \\ \Pi_1 < \pi^*_{it} \leq \Pi_2 \\ \Pi_2 < \pi^*_{it} \leq \Pi_3 \\ \pi^*_{it} > \Pi_3 \end{array} \tag{10}$$

[8] See the Appendix for a more detailed description of the CPA system.

[9] Two of the 150 main English authorities (the City of London and the Isles of Scilly) are excluded from the empirical analysis because of their peculiar characteristics.

where Π_j $(j = 1, ..., 3)$ are estimable parameters (thresholds) that define the observed performance rating π_{it}.[10]

As far as the stochastic component of the model is concerned, we employ a conventional random effects ordered probit specification (Greene [14]):

$$\eta_{it} = g_i + \omega_{it} \tag{11}$$

where g_i is a random jurisdiction-specific effect, with $E(g_i|z_{it}, q_i) = E(\omega_{it}|z_{it}, q_i) = 0$, and g_i and ω_{it} are normally distributed and orthogonal to each other.

A random effects ordered probit model is preferable in this context to a fixed effects specification for a number of reasons. First, while the full ordered probit model with fixed effects can in principle be estimated by unconditional maximum likelihood, this is not generally feasible in samples of this size. In fact, no theoretical results on the small T bias ($T = 6$ in our case) have been derived (Greene [14]). Second, the alternative estimation route represented by a fixed effects ordered logit model (Ferrer-i-Carbonell and Frijters [11]) has the important drawback of using only a fraction of the total information available in the data, with all units showing no change in the score variable being dropped. Moreover, similarly to the alternative fixed effects estimator based on the Chamberlain binary approach developed by Das and van Soest [8], fixed effects can only be identified by modelling unit-specific thresholds - an hypothesis that is untenable in our context of performance evaluation by strictly uniform criteria. Finally, the fixed effects ordered logit estimator entails the cost of losing the information needed to compute predicted probabilities and partial effects.

Consequently, we employ the widely used random effects ordered probit specification and try to control for fixed effects in two ways. First, we include a number of time-invariant jurisdictions' institutional characteristics (such as the one-tier *versus* two-tier structure of local government and the heteroge-

[10]The estimation results turn out to be similar when using all five rating categories and estimating four threshold parameters. However, computation time is longer and convergence of the likelihood function is obtained with more difficulty.

neous local electoral rules) as well as a number of Census 2001 variables as regressors - q_i in equation (9) - in order to capture fixed traits of a locality that might affect the performance of the government and be correlated with the time-varying regressors included in (9). Second, we implement the approach developed by Mundlak [25] and discussed in Greene [14], and specify an explicit (linear) relationship between the jurisdiction-specific effect (g_i) and the included time-varying regressors (z_{it}):

$$g_i = \overline{z}_i'\rho + r_i \tag{12}$$

with $r_i|z_{it} \sim N(0, \sigma_r^2)$ and $\overline{z}_i = \frac{1}{6}\sum_{t=1}^{6} z_{it}$, yielding:

$$\pi_{it}^* = z_{it}'\beta + q_i'\gamma + \overline{z}_i'\rho + r_i + \omega_{it} \tag{13}$$

Vector z_{it} includes the size of resident population, population density, property tax base per capita, a dummy variable that equals one for coalition governments, a dummy variable that equals one for Conservative governments, and real public spending per capita in excess of centrally set standard. The population and population density variables are included to allow for the possibility of economies of scale and congestion in the production and consumption of local public services respectively, while the property tax base variable should capture income effects on the demand for public services. The coalition government variable is included to account for the fact that "fragmented" governments tend to be weaker than one-party governments, and could therefore be less able to extract an efficient production of public services from the bureaucracy (Kalseth and Rattso [23]). The Conservative party dummy is included to allow for the possibility that, after controlling for observable performance determinants, party ideology might directly influence performance in the provision of public services.

Vector q_i includes authority class dummies (London Borough, Metropolitan Borough, Non-metropolitan Unitary Authority, Non-metropolitan County) and a dummy for the features of the local electoral system ("all out" elections

every fourth year *versus* yearly "by thirds" elections) to control for differences in the institutional framework across English localities.[11] Moreover, in an attempt to capture the underlying demand for public service performance and proxy the degree of control of the local polity on politicians' behaviour, vector q_i comprises a number of Census 2001 variables reflecting the composition of the local population and workforce: the age structure of the population (percentage of residents aged 0-16 and aged over 75); the qualification level (in terms of the percentage of highly qualified workers) and sectoral composition (percentage of employment in financial and real estate services) of the workforce; the percentage of self-employed, unemployed and disabled workers; indicators of ethnic composition (percentage of the population that is white) and religious affiliation (percentage of the population that is religious). Descriptive statistics for all the variables used in the analysis are reported in table 2.

The estimation results are reported in tables 3 to 6. Table 3, column (a) reports the results from estimating a parsimonious ordered probit model specification on pooled data. Column (b) adds a number of time-invariant institutional characteristics. Columns (c) and (d) show the estimation results of a random effects ordered probit model.

In all instances, spending in excess of the standard is estimated to have a negative and significant effect on performance. As for the other variables, population size, property tax base and Conservative control are estimated to affect performance positively, while population density and fragmentation of the Council have no significant effect in the random effects specifications of columns (c) and (d). As far as the institutional structure is concerned, Uni-

[11] About two thirds of the English local authorities (including all Counties and London Boroughs, plus a fraction of non-metropolitan Unitary Authorities) have *en bloc* elections every four years. In the other localities (including all Metropolitan Boroughs), elections take place "by thirds," in the sense that one third of the councillors are elected every year. In both all out and by thirds systems, councillors are elected on a "first past the post" basis and sit for a four-year period. Whether and in which direction might the electoral system affect the performance of a local government is unclear a priori.

tary, Metropolitan and London authorities appear to perform less well than the (reference) County authorities, while the effect of the all out electoral system is only weakly significant.

Table 4 reports the partial effects corresponding to the random effects ordered probit specification in column (d) of table 3. As far as the continuous variables are concerned, the partial effects are computed as marginal probability effects (MPE) of, say, regressor z_k on the probability of outcome $j = 1, ..., 4$ (poor/weak, fair, good, excellent):

$$\begin{aligned} MPE_{jk} &= \frac{\partial P(\pi_{it} = j | z_{it}, q_i, \overline{z}_i)}{\partial z_{ikt}} \\ &= \frac{\partial \left[\Phi(\Pi_j - z'_{it}\beta - q'_i\gamma - \overline{z}'_i\rho) - \Phi(\Pi_{j-1} - z'_{it}\beta - q'_i\gamma - \overline{z}'_i\rho)\right]}{\partial z_{ikt}} \\ &= \left[\phi(\Pi_{j-1} - z'_{it}\beta - q'_i\gamma - \overline{z}'_i\rho) - \phi(\Pi_j - z'_{it}\beta - q'_i\gamma - \overline{z}'_i\rho)\right]\beta_k \end{aligned} \tag{14}$$

and are evaluated at the sample means.

For dummy variables, partial effects are computed as the change in the probability of outcome j when a dummy variable, say z^d_k, shifts from 0 to 1:

$$\Delta P_{jk} = P(\pi_{it} = j | z_{it}, q_i, \overline{z}_i, z^d_{ikt} = 0) - P(\pi_{it} = j | z_{it}, q_i, \overline{z}_i, z^d_{ikt} = 1) \tag{15}$$

and are evaluated at the sample means.

As far as the effect of public spending is concerned, table 4 shows that higher expenditures make good and excellent performances less likely, and weak and fair performances more likely. At mean values, an increase in local public spending per capita by, say, 1% (amounting to around 13 pounds) lowers the chances of achieving good and excellent performances by about 5 and 1 percentage points respectively, and raises the chances of getting fair or weak performances by 6 percentage points.

Table 5, column (e) shows the results of estimation of the random effects ordered probit model augmented with the Mundlak correction. When allowing for and explicitly modelling the correlation between jurisdiction-specific effects and the regressors as in equation (12), the effects of the included time-varying characteristics turn out to be only weakly significant (per capita tax

base and Conservative dummy) or virtually vanish. However, the effect of public spending on performance remains large and statistically significant.

A number of Census variables are added as controls in column (f) of table 5, and table 6 reports the corresponding partial effects. Most of the Census variables are estimated to have a significant impact on performance, with the proportion of highly qualified, white and religious people being associated with better government performance, and with the rate of unemployment and the fraction of employment in the financial and real estate services sector being associated with worse performances. It is interesting to notice that, once controlling for those Census variables, metropolitan authorities appear to be significantly more likely to achieve excellent performances than non-metropolitan ones. In fact, while table 1 shows that Counties achieve overall better performances than the other authorities (over 80% of good and excellent ratings against an average 60% of the rest of the authorities), table 2 reveals that Counties operate in a more favourable environment according to the socio-economic traits extracted from the Census and their estimated impact on performance (in terms, for instance, of lower unemployment and less diverse communities).

Finally, the inclusion of the within-groups averages ($\overline{z}_i$) among the right hand side variables of the performance determination equation in the Mundlak specification (13) makes the Census variables orthogonal to the time-varying regressors z_{it}. As a result, even when controlling for fixed socio-economic community characteristics as in column (f) of table 5, excess spending maintains a large detrimental effect on performance. On average, a 1% increase in spending lowers the chances of attaining good and excellent performances by about 3 and 1 percentage points respectively, and raises the chances of getting fair or weak performances by around 4 percentage points.

4 Concluding remarks

Based on the properties of a unique measure of local government performance that was introduced in Britain in 2002 in order to measure how well Councils provide services to their citizens, this paper has explored the determinants of local governments' performances. After setting up a simple analytical framework to model the performance determination process, we have used panel data on the 148 main local authorities in England in the 2002-2007 time span to estimate an ordered response model that accounts for the categorical nature of the performance ratings and for cross-jurisdictional heterogeneity.

The empirical analysis provides no evidence that higher public spending translates into better performance. Rather, the estimate of the effect of excess spending on performance is negative in a random effects ordered probit specification. Moreover, controlling for correlation between jurisdiction-specific effects and regressors and for fixed characteristics of a locality - including institutional structure and socio-economic complexion - provides further evidence in support of the hypothesis that public expenditures in excess of centrally set standards have a detrimental effect on performance.

Overall, the impressive upward trend in performance of English local governments since the introduction of the CPA system seems to suggest that the incentives generated by the CPA system itself - along with the mandatory requirement on the level of public services to be provided - were extremely powerful in inducing local governments to improve the production process of local public services and adopt measures to fight inefficiency. Due to the short time-series dimension of the available panel data set, though, an explicit analysis of the dynamic behaviour of local government performance and of the process of endogenous determination of efficiency - i.e., an investigation of the channels through which efficiency improvements are obtained and an exploration of the role of the bureaucracy - remains an issue for future research.

Moreover, this paper has not tackled the potentially important issue of

strategic budgeting and endogenous spending determination. In fact, the empirical analysis of the determinants of performance is based upon the hypothesis of orthogonality of shocks to performance with respect to actual budgeting decisions: the fact that shocks to performance cannot be predicted prevents local governments from strategically manoeuvring public spending in order to counter the effects of those shocks. While the relaxation of the above hypothesis within an explicit "political economy" approach to local government behavior could certainly contribute to our understanding of the performance determination process, it reasonably seems to go beyond the scope of this paper and is consequently left for future research.

References

[1] Afonso, A., Fernandes, S., Measuring local government spending efficiency: Evidence for the Lisbon region, Regional Studies 40 (2006) 39-53.

[2] Afonso, A., Schuknecht, L., Tanzi, V., Public sector efficiency: An international comparison, Public Choice 123 (2005) 321-347.

[3] Andrews, R., Boyne, G., Law, J., Walker, R., External constraints on local service standards: The case of Comprehensive Performance Assessment in English local government, Public Administration 83 (2005) 639-656.

[4] Audit Commission, CPA - The harder test framework for 2007, Audit Commission, London, 2007.

[5] Audit Commission, Comprehensive Performance Assessment: Scores and analysis of performance, Audit Commission, London, 2003-2008.

[6] Barankay, I., Lockwood, B., Decentralization and the productive efficiency of government: Evidence from Swiss cantons, Journal of Public Economics 91 (2007) 1197-1218.

[7] CIPFA, Finance and General Statistics, Chartered Institute of Public Finance and Accountancy, London, 2002-2007.

[8] Das, M. van Soest, A., A panel data model for subjective information on household income growth, Journal of Economic Behavior and Organization 40 (1999) 409-426.

[9] Davis, M.L., Hayes, K., The demand for good government, Review of Economics and Statistics 75 (1993) 148-152.

[10] Dixit, A., Incentives and organizations in the public sector: An interpretative review, Journal of Human Resources 37 (2002) 696-727.

[11] Ferrer-i-Carbonell, A., Frijters, P., How important is methodology for the estimates of the determinants of happiness? Economic Journal 114 (2004) 641-659.

[12] Geys, B., Looking across borders: A test of spatial policy interdependence using local government efficiency ratings, Journal of Urban Economics 60 (2006) 443-462.

[13] Glewwe, P., Jacoby, H., An economic analysis of delayed primary school enrollment in a low income country: The role of early childhood nutrition, Review of Economics and Statistics 77 (1995) 156-169.

[14] Greene, W., Econometric Analysis. Prentice Hall, 2008.

[15] Grosskopf, S., Hayes, K., Taylor, L., Weber, W., Budget-constrained frontier measures of fiscal equality and efficiency in schooling, Review of Economics and Statistics 79 (1997) 116-124.

[16] Grosskopf, S., Hayes, K., Taylor, L., Weber, W., On the determinants of school district efficiency: Competition and monitoring, Journal of Urban Economics 49 (2001)453-478.

[17] Grossman, P.J., Mavros, P., Wassmer, R.W., Public sector technical inefficiency in large US cities, Journal of Urban Economics 46 (1999) 278-299.

[18] Hanushek, E., School resources, in: E. Hanushek and F. Welch (Eds.) Handbook of the Economics of Education, Elsevier, 2006, vol. 2, pp. 865-908.

[19] Hauner, D., Benchmarking the efficiency of public expenditure in the Russian Federation, IMF Working Paper 246 (2007).

[20] Hayes, K., Razzolini, L., Ross, L., Bureaucratic choice and nonoptimal provision of public goods: Theory and evidence, Public Choice 94 (1998) 1-20.

[21] Hoxby, C., The productivity of schools and other local public goods producers, Journal of Public Economics 74 (1999) 1-30.

[22] Joumard, I., Kongsrud, M., Fiscal relations across government levels, OECD Economics Department Working Paper 375 (2003).

[23] Kalseth, J., Rattso, J., Political control of administrative spending: The case of local governments in Norway, Economics and Politics 10 (1998) 63-83.

[24] Leuven, E., Lindahl, M., Oosterbeek, H., Webbink, D., The effect of extra funding for disadvantaged pupils on achievement, Review of Economics and Statistics 89 (2007) 721-736.

[25] Mundlak, Y., On the pooling of time series and cross section data, Econometrica 46 (1978) 69-85.

[26] Papke, L., The effects of spending on test pass rates: Evidence from Michigan, Journal of Public Economics 89 (2005) 821-839.

[27] Propper, C., Wilson, D., The use and usefulness of performance measures in the public sector, Oxford Review of Economic Policy 19 (2003) 250-267.

[28] Revelli, F., Performance rating and yardstick competition in social service provision, Journal of Public Economics 90 (2006) 458-474.

[29] Revelli, F., Performance competition in local media markets, Journal of Public Economics (2008) forthcoming.

[30] Revelli, F., Tovmo, P., Revealed yardstick competition: Local government efficiency patterns in Norway, Journal of Urban Economics 62 (2007) 121-134.

Table 1 Comprehensive Performance Assessment ratings

	poor	weak	fair	good	excellent	
2002	12	21	41	53	21	148
2003	9	19	39	56	25	148
2004	1	14	33	60	40	148
2005	1	8	35	65	39	148
2006	0	5	25	72	46	148
2007	0	2	23	68	55	148
PARTY CONTROL						
Conservative (%)						
2002-2007	1.7	3.9	19.0	41.1	34.2	100.0
Labour (%)						
2002-2007	2.7	10.5	20.8	37.7	28.3	100.0
Liberal Democrats (%)						
2002-2007	1.5	4.5	23.9	58.2	11.9	100.0
Fragmented (%)						
2002-2007	3.5	8.5	26.0	44.6	17.4	100.0
AUTHORITY TYPE						
Counties (%)						
2002-2007	1.0	3.9	13.7	41.7	39.7	100.0
Unitary authorities (%)						
2002-2007	3.3	5.8	27.5	45.7	17.7	100.0
Metropolitan Boroughs (%)						
2002-2007	2.4	10.6	24.5	37.5	25.0	100.0
London Boroughs (%)						
2002-2007	3.6	11.5	20.3	42.7	21.9	100.0

Table 2 Descriptive statistics

	obs.	mean	s.d.	min	max
Population (,000)	888	337.6	252.5	34.9	1371.3
Population density	888	24.5	27.3	0.6	153.6
Tax base per capita	888	34.2	5.2	22.5	63.7
Spending per capita (£)	888	1297.0	251.7	884.1	2564.6
Standard spending per capita (£)	888	1258.3	262.7	784.6	2586.9
Spending over standard per capita (£)	888	38.7	42.5	-297.0	135.5
Conservative	888	0.26	0.44	0	1
Labour	888	0.37	0.48	0	1
Liberal Democrats	888	0.08	0.26	0	1
Fragmented	888	0.29	0.45	0	1
All out elections	148	0.62	0.48	0	1
County	148	0.23	0.42	0	1
Unitary	148	0.31	0.46	0	1
Metropolitan	148	0.24	0.43	0	1
London	148	0.22	0.41	0	1
Census 2001 variables					
Age 0-16 (% population)	148	20.27	1.74	13.49	26.17
- County: Age 0-16	34	19.56	0.81	17.85	21.20
- Unitary: Age 0-16	46	20.40	1.81	16.65	25.23
- Metropolitan: Age 0-16	36	20.93	1.15	18.77	23.43
- London: Age 0-16	32	20.07	2.46	13.49	26.17
Age 75+ (% population)	148	7.27	1.57	3.97	12.06
- County: Age 75+	34	8.30	1.35	6.40	12.06
- Unitary: Age 75+	46	7.44	1.76	4.78	11.97
- Metropolitan: Age 75+	36	7.29	0.63	5.73	8.72
- London: Age 75+	32	5.90	1.23	3.97	8.24
Highly qualified (% workforce)	148	20.27	8.66	9.69	51.53
- County: Highly qualified	34	19.14	3.50	14.16	27.71
- Unitary: Highly qualified	46	17.44	5.51	9.87	30.40
- Metropolitan: Highly qualified	36	15.28	3.61	9.69	24.32
- London: Highly qualified	32	31.14	10.81	10.23	51.53

Table 2 (continued)

Census 2001 variables					
Disabled (% workforce)	148	5.64	2.23	1.94	12.15
- County: Disabled	34	4.51	1.58	2.43	10.32
- Unitary: Disabled	46	5.49	2.21	1.94	10.76
- Metropolitan: Disabled	36	7.79	1.89	4.39	12.15
- London: Disabled	32	4.61	1.25	2.51	7.04
Unemployed (% workforce)	148	3.62	1.23	1.55	6.91
- County: Unemployed	34	2.59	0.52	1.73	3.80
- Unitary: Unemployed	46	3.47	1.18	1.55	6.23
- Metropolitan: Unemployed	36	4.13	0.92	2.47	6.25
- London: Unemployed	32	4.36	1.36	2.47	6.91
Self-employed (% workforce)	148	7.93	2.24	3.94	13.71
- County: Self-employed	34	9.66	1.54	5.38	12.60
- Unitary: Self-employed	46	7.37	2.21	3.94	13.71
- Metropolitan: Self-employed	36	6.06	1.20	4.22	8.39
- London: Self-employed	32	8.99	1.86	4.96	13.57
Financial and real estate (% employment)	148	18.39	7.16	9.34	45.34
- County: Financial and real estate	34	15.74	4.00	9.86	26.06
- Unitary: Financial and real estate	46	16.32	5.34	9.34	28.32
- Metropolitan: Financial and real estate	36	14.59	2.88	10.72	22.58
- London: Financial and real estate	32	28.47	6.21	20.06	45.34
Religious (% population)	148	77.42	4.62	63.25	87.49
- County: Religious	34	77.87	2.79	73.32	84.06
- Unitary: Religious	46	76.22	4.90	63.25	84.20
- Metropolitan: Religious	36	80.80	3.37	74.03	87.49
- London: Religious	32	74.85	4.66	65.76	82.59
White (% population)	148	89.19	12.64	39.41	99.27
- County: White	34	97.09	1.88	92.13	99.27
- Unitary: White	46	93.32	8.33	63.70	99.09
- Metropolitan: White	36	91.85	7.24	70.35	99.08
- London: White	32	71.85	13.34	39.41	95.17

Table 3 Performance determination equation: ordered probit estimates

	(a)	(b)	(c)	(d)
Population	0.0009	0.0005	0.0025	0.0013
	(5.30)	(2.18)	(6.37)	(3.25)
Density	-0.0068	-0.0032	-0.0012	0.0059
	(4.15)	(1.38)	(0.28)	(1.35)
Tax base	0.0449	0.0443	0.0319	0.0377
	(5.07)	(4.17)	(1.16)	(1.90)
Fragmented	-0.3801	-0.3738	-0.0440	-0.0727
	(4.24)	(4.14)	(0.29)	(0.47)
Conservative	-0.3847	-0.3870	0.1353	0.4646
	(3.36)	(3.35)	(0.52)	(2.43)
Excess spending	-0.0057	-0.0058	-0.0130	-0.0136
	(4.96)	(4.82)	(4.10)	(5.40)
All out elections		0.2736		-0.5986
		(2.03)		(2.02)
Unitary		-0.2180		-1.4531
		(1.25)		(4.19)
Metropolitan		0.0219		-1.5561
		(0.11)		(3.90)
London		-0.5071		-1.5267
		(2.47)		(4.26)
Π_1	0.257	0.188	-0.383	-2.034
Π_2	1.172	1.112	1.508	-0.126
Π_3	2.415	2.363	4.435	2.835
Random effects	-	-	yes	yes
Log likelihood	-1048.83	-1043.78	-742.57	-736.10
Observations	888	888	888	888

Notes: t statistics in parentheses; year effects included.

Table 4 Random effects ordered probit: partial effects

	poor/weak	fair	good	excellent
Continuous variables				
Population	-0.00001	-0.00045	0.00037	0.00009
Density	-0.00006	-0.00208	0.00170	0.00044
Tax base	-0.00036	-0.01326	0.01085	0.00278
Excess spending	0.00013	0.00478	-0.00391	-0.00100
Dichotomous variables				
Fragmented	0.00052	0.01834	-0.01518	-0.00368
Conservative	-0.00246	-0.10388	0.07549	0.03085
All out elections	0.00370	0.13882	-0.10449	-0.03753
Unitary	0.03165	0.40997	-0.38419	-0.05743
Metropolitan	0.04642	0.45304	-0.44954	-0.04992
London	0.04820	0.45029	-0.45337	-0.04511

Notes: partial effects are computed according to equation (14) for continuous variables and to equation (15) for dichotomous variables.

Table 5 Random effects ordered probit: Mundlak specification

	(e)		(f)	
	z_{it}	$\overline{z}_i$	z_{it}	$\overline{z}_i$
Population	-0.0181	0.0191	-0.0148	0.0190
	(1.51)	(1.59)	(1.21)	(1.55)
Density	-0.0199	-0.0083	-0.0252	0.0664
	(0.47)	(0.20)	(0.57)	(1.47)
Tax base	-0.1405	0.2672	-0.1395	0.2043
	(1.73)	(3.19)	(1.66)	(2.19)
Fragmented	0.2093	-2.0906	0.1829	-1.1937
	(1.17)	(7.06)	(1.01)	(4.27)
Conservative	0.4095	-2.3147	0.3950	-0.9622
	(1.67)	(6.20)	(1.60)	(2.63)
Excess spending	-0.0109	-0.0126	-0.0119	0.0015
	(2.23)	(2.27)	(2.41)	(0.27)
	q_i		q_i	
All out	-0.2178		1.3211	
	(0.93)		(5.32)	
Unitary	-0.9686		1.6260	
	(2.58)		(3.87)	
Metropolitan	-0.4948		1.7245	
	(1.20)		(4.43)	
London	0.0667		2.0134	
	(0.15)		(4.21)	
Age 0-16			0.1579	
			(1.38)	
Age 75+			-0.0661	
			(0.73)	
Highly qualified			0.1471	
			(5.22)	
Disabled			0.1869	
			(2.60)	
Unemployed			-0.6163	
			(4.92)	

Table 5 (continued)

	(e)	(f)
Self-employed		-0.1020
		(1.33)
Financial and RE		-0.1409
		(4.34)
Religious		0.1681
		(6.62)
White		0.0932
		(5.50)
Π_1	-0.590	26.358
Π_2	1.352	28.340
Π_3	4.391	31.538
Random effects	yes	yes
Log likelihood	-723.204	-704.530
Observations	888	888

Notes: t statistics in parentheses; year effects included.

Table 6 Partial effects

	poor/weak	fair	good	excellent
Continuous variables				
Population	0.00005	0.00388	-0.00273	-0.00120
Density	0.00008	0.00660	-0.00465	-0.00204
Tax base	0.00046	0.03660	-0.02577	-0.01129
Excess spending	0.0004	0.00311	-0.00219	-0.00096
Age 0-16	-0.00052	-0.04143	0.02917	0.01278
Age 75+	0.00022	0.01735	-0.01222	-0.00535
Highly qualified	-0.00048	-0.03858	0.02716	0.01190
Disabled	-0.00061	-0.04902	0.03452	0.01512
Unemployed	0.00202	0.16164	-0.11381	-0.04985
Self-employed	0.00033	0.02675	-0.01883	-0.00825
Financial and RE	0.00046	0.03696	-0.02602	-0.01140
Religious	-0.00055	-0.04409	0.03104	0.01360
White	-0.00031	-0.02445	0.01721	0.00754
Dichotomous variables				
Fragmented	0.00391	0.31309	-0.22045	-0.09655
Conservative	0.00315	0.25237	-0.17769	-0.07783
All out elections	-0.00836	-0.29183	0.23218	0.06802
Unitary	-0.00342	-0.22221	0.03471	0.19092
Metropolitan	-0.00263	-0.20120	-0.04046	0.24430
London	-0.00277	-0.20881	-0.12757	0.33915

Notes: partial effects are computed according to equation (14) for continuous variables and to equation (15) for dichotomous variables.

Appendix
The CPA system

Comprehensive Performance Assessment (CPA) was introduced in Britain in 2002 in order to measure how well local Councils deliver services for local communities (Audit Commission [4]). CPA aims at looking at performance from a range of perspectives and providing a simply understood rating on a poor to excellent scale. An independent body - the Audit Commision - assesses local government performance and annually delivers the CPA ratings based both on its own audit and inspection activity and on the assessments provided by other independent Commissions: the Commission by the Office for Standards in Education (OFSTED), the Commission for Social Care Inspection (CSCI) and the Benefit Fraud Inspectorate (BFI).

The CPA framework consists of two main dimensions across which Councils are evaluated yearly: the use of resources assessment and the service assessment.

The use of resources assessment is conducted annually and provides a judgement on how well a Council manages and uses its financial resources. The assessment focuses on the importance of having sound and strategic financial management to ensure that resources are available to support the Council's priorities. It covers the following themes: financial reporting; financial management; financial standing; internal control; value for money. For instance, the judgement criteria refer to whether the Council's accounts are prepared in accordance with regulatory requirements and accounting standards ("financial reporting"), whether the Council has put in place a sound medium-term financial strategy ("financial management"), whether the budget is balanced ("financial standing"), whether the Council conducts an annual review of the effectiveness of the system of internal control ("internal control") and whether costs are significantly higher than other Councils providing similar services ("value for money") (Audit Commission [4]). Judge-

ments are made by the Audit Commission for each theme on a 1-4 scale based on a number of criteria, and the scores are then combined to produce an overall use of resources score.

Service assessment scores are published annually and bring together existing service performance information through national performance indicators (PIs) and relevant service inspections. Service assessments are derived either by independent inspectorates or by the Commission itself. In particular, the Audit Commission delivers annual assessments for the service areas of environment, housing and culture. Examples of performance indicators include the percentage of pedestrian crossing with facilities for disabled people in the service area of "environment"; the average time spent by homeless people in temporary accommodation in the service area of "housing"; aggregate library opening hours per 1000 population in the service area of "culture" (Audit Commission [4]). The overall assessment score (1-4) for each service is determined by combining a score for the PI set with a score for any relevant service inspections.

In addition, each Council has a so-called "corporate assessment" every three years as well as a yearly "direction of travel" assessment. Corporate assessment aims at measuring how effectively the Council is working corporately, and with its partners, to deliver improved outcomes for local people. In particular, corporate assessment teams assess how effectively the Council is working, how well Councils understand their local communities, how this understanding translates into ambitions and priorities, and what, in practice, Councils are achieving. This is assessed under five themes: ambition for the community; prioritisation; capacity; performance management; achievement (sustainable communities and transport; safer and stronger communities; healthier communities; older people; children and young people). Each of the above corporate assessment themes is scored on a 1-4 scale and then aggregated in a single 1-4 score. The direction of travel assessment is a judgement which is meant to provide public assurance about whether a

Council is complying with its duty of making arrangements to secure continuous improvement. The assessment is scored through the use of four levels of judgement (improving strongly; improving well; improving adequately; not improving adequately or not improving).

The use of resources and service assessments scores are then combined with the corporate and direction of travel scores to generate one of five rating categories: poor, weak, fair, good and excellent performance.

Figure 1: Performance determination process