Incentive Schemes for Local Government: Theory and Evidence from Comprehensive Performance Assessment in England

By Ben Lockwood and Francesco Porcelli*

This paper studies Comprehensive Performance Assessment, an explicit incentive scheme for local government in England. Motivated by a theoretical political agency model, we predict that CPA should increase service quality and local taxation, but have an ambiguous effect on the efficiency of service provision. We test these predictions using Welsh local governments as a control group: CPA increased the property tax, and our index of service quality, but had no significant effect on efficiency overall. There is evidence of a heterogenous effect of CPA: CPA impacted more on councils where electoral competition was initially weak, in line with our theory.
JEL: H10,H70,H77,C21
Keywords: local government, incentives, efficiency, difference in difference, DEA

In recent years, explicit incentive schemes for public organizations, based on quantitative measurement of outputs, have become increasingly commonly used in the UK[1]. For example, school league tables, hospital star ratings, and various schemes for local government, such as Comprehensive Performance Assessment (CPA), have been introduced in the last twenty years or so. The focus of this paper is on CPA, the most important such scheme for local government. This scheme, introduced in 2001, rated local governments in England on the quality of service in six major areas: education, housing; social care; environment; libraries and leisure; use of resources. Hundreds of performance indicators and a variety of audit and inspection reports were collected, summarized, weighted, and categorized so as to arrive at final star ratings between 0 and 4[2].

As well as an evaluation scheme, CPA was also an incentive scheme. The stated objective of the CPA was to target support at those councils that need it

* Lockwood: University of Warwick, Coventry CV4 7AL, England; Email: B.Lockwood@warwick.ac.uk. Porcelli: University of Warwick, Coventry CV4 7AL, England; Email: f.porcelli@warwick.ac.uk. We would like to thank Fabian Waldinger, Nicolas Serrano-Verlade, and participants at the conference "New issues on fiscal federalism and tax competition", Lyon, the 2010 CESifo Political Economy conference in Dresden, and seminars at IFS, Bristol, Uppsala and Lancaster for helpful comments.

[1]Schemes of this type have been little used outside the UK. There are exceptions: in the US, for example, the No Child Left Behind legislation punishes schools financially for poor test results, which are made public to parents.

[2]In fact, from 2002-5, the rankings were designated: "excellent", "good", "fair", "weak", and "poor", changing to zero to four stars during the latter part of CPA - see Tables A1 and A3 below. But, for simplicity, we refer to star ratings throughout.

most, and to offer a number of benefits for better-performing councils, including elimination of "ring-fencing" grants, and a three-year exemption from subsequent audit inspections[3]. Moreover, because the results of the CPA were widely disseminated in the media, it was also an exercise in providing voters with more information about the performance of their local council, both absolutely, and relative to other councils. In turn, this, in principle, provides *indirect* incentives for good performance. Indeed, there is evidence that councils which performed poorly on CPA were punished by the voters at subsequent elections[4].

CPA is of particular interest because it is, to our knowledge, the only explicit evaluation scheme to date, worldwide, that numerically scores and rewards *elected representatives*, as opposed to public service managers. The purpose of this paper is to assess the impact of CPA on local government in three dimensions: quality of service delivery, taxation policy, and the efficiency with which services were provided.

Figure 1 below shows the average CPA score achieved by English local authorities from the beginning to the end of the CPA experience together with average current local expenditure per capita. There is clearly a steady upward trend in average CPA star ratings. Indeed, the Audit Commission declared officially in 2009 (Audit Commission(2009)) that the CPA had done its job stimulating a continuous improvement in local government performance. However, Figure 1 also shows that at the same time, expenditure by local governments went up, more or less in line with CPA scores.

So, the key problem is that *we do not observe the counterfactual*; given the large increases in local government spending over this period, it may be that service delivery would have improved anyway, even in the absence of the CPA. To address this, we treat the CPA as a natural experiment by exploiting the fact that it was only introduced in England, whereas in Wales, where the structure of local government is the same, a much weaker performance management scheme was introduced (Haubrich and McLean (2006b) Martin, Downe and Grace (2010)). In particular, in Wales, there were no quantitative rankings, much less information published, and authorities also had a say with regard to the type of inspections they would like to see for specific services. So, we use local authorities in Wales as a control group when assessing the impact of CPA on the treatment group, the English councils.

What would we expect the effects of a scheme such as CPA to be on service quality, tax levels, and efficiency? We develop a simple two-period political agency model in the paper to look specifically on the effect on taxation and efficiency of

[3] "High scoring" councils were councils that were performing well under CPA would enjoy reduced audit and inspection regimes, and their associated fees, and be granted greater flexibilities and borrowing freedoms by central government. At the other end of the performance spectrum, a combination of audit, inspection and other improvement work was to be commissioned as an outcome of the CPA process, with the aim of transforming failing or poorly performing authorities." (Audit Commission(2009)).

[4] Revelli (2008) finds that an increase in one star rating increases the probability that the incumbent party retains control of the council by seven percentage points, and Boyne et al. (2009) find "a low CPA score (0 or 1 star) increases the likelihood of a change in political control".

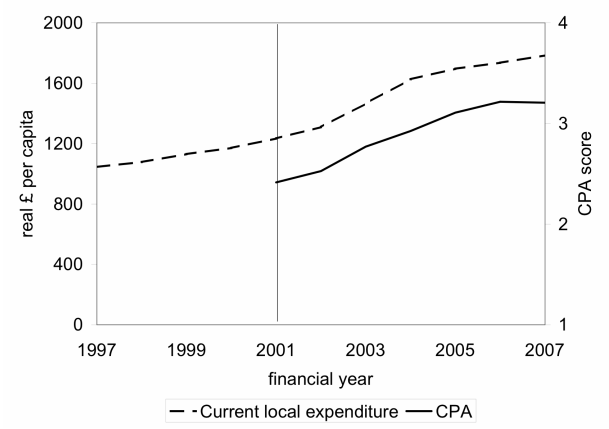

FIGURE 1. CPA SCORES AND EXPENDITURES.

an incentive scheme that both rewards service quality and provides information about this quality to voters. In any period, the quality of a public good or service is determined by the politician's ability, effort, and tax revenue. In this environment, efficiency measures the level of service quality that can be produced at a given level of tax revenue. Voters value service quality and dislike taxes, and thus they care about both service quality and efficiency. The incumbent faces an election against a randomly selected challenger at the end of the first period.

Our key predictions are as follows. The larger the direct reward, or the better the information, in the incentive scheme, the more the incumbent taxes, and the higher the effort he makes. While higher effort is not surprising, the prediction of higher taxation, which voters dislike, is a distinctive feature of our theoretical analysis. As both effort and taxes rise, service quality is unambiguously increased by an incentive scheme. But, the effect of either a larger direct reward, or better information, on *efficiency* is ambiguous, because inputs, purchased by taxes, are also higher.

We then test these predictions, using Wales as a control group. Our results broadly confirm the predictions of the theory. First, looking across a number of different measures of revenue, the introduction of CPA appears to have raised council tax revenues in England relative to Wales. For example, we see that the introduction of CPA raised the effective band D council tax rate by about £40, or about 4%, in percent in England. To test the effects of CPA on quality of output and efficiency of local councils, we used specially constructed indices of both, described in more detail below (see also Porcelli(2010) on the efficiency index). We find, consistently with the theory, that the CPA raised our quality of output index by roughly 4% above what it would have been, had English local councils also been subject to the same regime as in Wales.

But, again consistently with the theory, we find that CPA had no significant effect on efficiency. So, our finding is consistent with the story that local authori-

ties reacted to CPA by performing better, but also spending more. Therefore we conclude that CPA did not boost efficiency overall. This is in stark contrast to the view of the Audit Commission (2009) that CPA has "done its job" effectively.

We then look more closely at the impact of CPA on English councils. In particular, our theory suggests that CPA should be a *substitute* for electoral competition, with weaker competition (as measured by the incumbent's margin of victory) unambiguously positively associated with; (i) a larger impact of an incentive scheme on quantity or quality of output; (ii) a more positive impact of an incentive scheme on efficiency. Empirically, we find some evidence of this. In councils where electoral competition was initially weak, it appears that CPA significantly increased both output and efficiency, leaving the level of local taxes unchanged.

The rest of the paper is arranged as follows. Section 1 surveys related literature. Section 2 gives a brief overview of CPA. Section 3 develops the theoretical framework, and Section 4 describes our empirical strategy. Sections 5 and 6 give the baseline results and results on the impact of electoral competition respectively, and Section 7 concludes.

I. Related Literature

There are several related literatures. First, there is an academic literature on the CPA itself. Boyne(2009) and Revelli(2008) have already been mentioned in the introduction. Revelli (2010) is perhaps the most closely related. In this last paper Revelli studies the link between council spending and CPA scores. In particular, he finds that spending in excess of the standard set by central government (standard spending assessment), can have a negative effect on the CPA score. His theoretical explanation for this is that some councils are more efficient than others in transforming expenditure into CPA scores. Moreover, he assumes that all councils have the same relative preferences for CPA scores and spending. In this environment, other things equal, a more efficient council will both spend less and achieve a higher CPA score.

In contrast, our paper constructs an explicit index of efficiency, *independent* of CPA, and asks how the introduction of CPA affects the efficiency of English councils, relative to Welsh ones. So, the two papers are quite different; we are more interested in the incentive effects of CPA, whereas Revelli(2010) is focussing on CPA as a measure of performance or outcome. Basically, as explained in the Section 2 below, we do not believe that the CPA is a good measure of either output or efficiency: rather, we are studying how it performed as an incentive scheme.

A second related literature is the wider one on incentives in the public sector. Most relevant to our study is the very recent and independent work by Burgess, Wilson and Worth (2010). They use the abolition of school league tables in Wales (but not in England) in 2002 as a natural experiment to estimate the effect of league tables on secondary school performance. This is related to our study because one of the output indicators we use is the proportion of secondary school

pupils achieving GCSE grades A to C in the local authority. Clearly, as CPA was introduced in England in the same year as school league tables were abolished in Wales, we cannot separately identify the effect of both reforms on school "output". To deal with this problem, we also test whether CPA increased our output index excluding education. We find that the effect of CPA is still significantly positive but smaller in magnitude. This is consistent with a story where both CPA and school league tables have positive effects on output.

Also related are Propper et al. (2008), (2008a), and Besley, Bevan and Burchardi(2009), which are papers investigating the effect of the hospital star rating regime in England over 2001-5 on waiting times for hospital treatment, using either Scotland and Wales as control groups. The hospital star rating regime is similar in form to CPA, with good performance closely tied to reducing waiting lists. All three of these papers find strong evidence that the scheme had the desired effect on the targeted "output" i.e. waiting times were reduced in England relative to Scotland and Wales, although waiting times also fell everywhere due to higher spending. Note also that all the papers just discussed only focus on single dimensions of local government "output"; unlike us, they do not address efficiency issues, or look at taxation.

Finally, our theoretical model extends a literature on principal-agent problems where the agent has several tasks to perform, initiated by the classic paper of Holmstrom-Milgrom(1991). Holmstrom and Milgrom, however, restrict attention to a static framework, where monetary incentives can be used in an unrestricted way, and where the agent's payoff is exponential in money. Dewatripont, Jewitt, and Tirole(1999) extend that analysis to a career concerns framework, i.e. where the agent is rewarded not explicitly, but in proportion to their ability as inferred by the principal. There have been a few extensions[5] of the multi-task career concerns framework to political principal-agent problems, notably Gersbach(2008) and Alesina and Tabellini(2008). However, unlike us, neither of these papers allow for a specific reward being offered for one task[6].

II. The CPA - A Brief Overview

Local governments in England and Wales are of two types, unitary and two-tier. Unitary councils are responsible for primary and secondary education, social care, housing and housing benefit payments, waste disposal, transport, and environment, planning, and culture. Two-tier governments (counties) have the same

[5]Less closely related contributions include Besley(2004), Caselli and Morelli(2004), Messner and Polborn(2004), Mattozzi and Merlo(2008). These papers mostly focus on the effect of pay (fixed, not performance-related) on the incentive for different types of politicians to run for office. Besley(2004) also looks at the effect of varying pay on incentive and selection effects of elections.

[6]Alesina and Tabellini study a sequence of models where the incumbent politician assigns effort to two tasks, and the level of performance on each task is fully observable, and depends additively on effort and ability, as in our setting. But, the main focus is on redistributive policies; each of two voter groups only benefits from the performance on one task, and the politician can make a transfer between these two groups. Finally, Gersbach(2008) considers a political agency model with moral hazard only, i.e. where politicians do not differ in ability, and where voters are able to pre-commit to a re-election rule.

responsibilities, except for housing and housing benefit, and environment, where responsibilities are shared with district councils.

In this institutional setting, the precursor to CPA, introduced in the Local Government Act 1999, was the "Best Value" framework, which "provides a framework for the planning, delivery and continuous improvement of local authority services. The overriding purpose is to establish a culture of good management in local government for the delivery of efficient, effective and economic services that meet the users' needs." (http://www.idea.gov.uk/idk). A key part of this framework were the Best Value Performance Indicators (BVPIs), which were numerical scores measuring the quality of the above services provided by individual councils on various dimensions. Importantly for our purposes, BVPIs were calculated for both English and Welsh councils.

CPA, which started in the 2001/02 financial year, was a move to a stricter assessment regime within the general Best Value framework. In the first three rounds, the method for assessing the current performance of a council was the following. Performance of councils was assessed in seven categories[7] (social care; environment; libraries and leisure; use of resources; education; housing; housing benefit payments). Where available, performance was assessed through already existing judgements from inspectorates and auditors, such as those by Office for Standards in Education (Ofsted) and Department for Education and Skills (DfES) for education. These were augmented with BVPIs. All this information was aggregated to obtain a score between 1 and 4 for each of the service blocks (with 1 being the lowest and 4 the highest). The performance scores were then aggregated across service blocks[8] to produce a performance rating of between 1 and 4 for each authority as shown in Table A3 of the Appendix. This score was then combined with an estimate of the councils' ability to improve (1 to 4) as explained in the Table A3 of the Appendix to produce the final CPA score.

In 2005, a new methodology, the "harder test" was introduced. The current performance of the LA was now assessed in the same categories with the exclusion of education, which was dropped. The main innovation, however, involved the aggregation procedure, where the ability to improve was replaced by the corporate assessment, a three year period assessment of the council's ability "to lead its local community having clearly identified its needs and set clear ambitions and priorities" (Audit Commission, 2009).

So, what are CPA scores really measuring? Along with some commentators e.g. McLean, Haubrich and Gutierrez-Romero (2007), we take the view that CPA is a hybrid measure, partly measuring levels of service quality (thorough the BVPIs), partly measuring operational efficiency (use of resources) and partly

[7]The CPA did not evaluate transport and planning.

[8]The scores were weighted so that the scores for education and social services count four times, housing and environmental services twice, with the remaining blocks counting only once. These were then added up to produce a performance score of between 15 and 60 points, or 12 and 48 points for shire county councils (because they do not provide, and are therefore not assessed on, housing or benefits services).

broader aspects of corporate health or effectiveness (ability to improve). As McLean, Haubrich and Gutierrez-Romero (2007) point out, there may also be "categorization errors" in the aggregation procedure in Table A3, where fine numerical scores are compressed into just four categories. So, we take the view that CPA scores are measuring both service levels (output) and efficiency, and are doing so with some error. In this paper, we are not interested in as CPA as a *measurement system*, but as an *incentive scheme*. That is why we construct our own, independent, measures of output and efficiency for local councils, with the aim of studying the effect of the CPA regime on those measures, along with taxation.

III. A Theoretical Framework

A. The Environment

In each of two periods $t = 1, 2$ an incumbent politician produces a local public good at quality level Q_t. This quality level depends on resources, in the form of tax revenue T_t, chosen by the incumbent[9], plus an exogenous grant from central government G_t, the effort input of the incumbent, a_t, and also his ability parameter η_t :

$$Q_t = f(a_t, T_t) + \eta_t \tag{1}$$

where G_t is taken as exogenous by the incumbent and so can be suppressed[10].

We assume $f(a_t, T_t)$ strictly increasing in both arguments, and strictly concave. Strict concavity is useful for obtaining clean results, and can be justified conceptually as follows. The CPA scheme only operated over six years, and there are factors of production in public services (e.g. school buildings) which can reasonably be taken as fixed or not easily adjustable over this time frame. Moreover, empirically, estimates of production functions for public services, especially school education and refuse collection, often find decreasing or constant returns to scale[11]. This is relevant for us as these are both important activities of local councils in the UK.

[9]This captures the stylized fact that the Council Tax (a residential property tax) is an autonomous tax instrument - indeed, the only one - for local government in England and Wales. Over the sample period, due to the Labour government's decision to abandon "rate-capping", local authorities have had in practice considerable autonomy to set their council taxes. The government reserves the right to direct an authority to set a lower budget requirement if it considers that the Council Tax has been increased excessively. However capping took place only in 2004/05 and 2005/06 for 6 and 8 local authorities respectively.

[10]In principle, if G_t is time-varying, $f(a, T)$ should also be time-varying i.e. $f_t(a, T)$ is the production function at time $t = 1, 2$. This generalization has no impact on the results, however, and is notationally burdensome, so we work with a non-time-varying production function in what follows.

[11]For example, in a survey of studies of education production functions at the school level (which focus on student performance and school size, conditioning on other variables), Andrews, Duncombe, and Yinger(2002) find that almost all studies find constant or decreasing returns to scale. Callan and Thomas(2001) find evidence of constant returns to scale in refuse disposal services in Massachusetts.

Also, following Rogoff and Sibert(1988), Alesina and Tabellini(2008), we assume that η_t follows a moving average process i.e. $\eta_t = \theta_t + \theta_{t-1}$ where θ_t is a random draw from a symmetric distribution with mean zero, and support $[-\overline{\theta}, \overline{\theta}]$, and $\theta_0 = 0$. Symmetry, zero mean, and $\theta_0 = 0$ are assumed for convenience only. At the beginning of $t = 1$, both the incumbent and voters know θ_0, but neither know θ_1 or θ_2. Thus, we are in the career concerns framework of Holmstrom; when policy is chosen, the incumbent and voters have the same information about ability[12].

There is a continuum of measure 1 of voters. Voter $i \in [0, 1]$ has linear payoffs over Q_t and tax T_t of the form

$$u_{it} = Q_t - \mu_i T_t, \ t = 1, 2 \tag{2}$$

where μ_i is $i's$ tax price of public spending, and may differ across voters. In England and Wales, the only local tax is the property tax, so the natural interpretation of μ_i would be as that voter's property value relative to the average.

The incumbent politician, while in office, gains some office-related benefits, R, and also incurs a cost of effort c. We also assume he puts some weight $\omega \geq 0$ on a weighted average of voter payoffs, either because he himself is a tax-payer and consumer of the local public good, as would be natural in a citizen-candidate setting (Besley and Coate(1997)), or because he is lobbied by special interest groups, or because he cares about his legacy (Maskin and Tirole(2004))[13]. This of course nests the purely office-seeking politician as a special case where $\omega = 0$. So, the politician in office in period t has payoff

$$\omega(Q_t - \mu T_t) + R - ca_t, \ t = 1, 2 \tag{3}$$

where μ is the weighted average of the μ_i. As the unweighted average of the μ_i is one, $\mu < 1$ if the politicians put more weight on poorer groups, for example. Following Maskin and Tirole(2004), we assume that when out of office, the politician has zero payoff. Finally, for simplicity, we assume that the incumbent does not discount future payoffs.

There is an election at the end of period 1, described in more detail below. Also, the incentive scheme is only used in period 1 and is described in more detail below. This simplifies the exposition, and in the two-period model, is without much loss of generality[14].

B. Political Equilibrium

[12]The alternative would be to suppose that θ_t is known by the incumbent at the beginning of period t. In this case, the signaling behavior of the incumbent would become more complicated, and there would be many equilibria, depending on out-of-equilibrium beliefs of the voters. As our intention is to develop testable predictions, we abstract from such complications.

[13]It is beyond the scope of this paper to provide micro-foundations for these processes.

[14]An infinite-horizon version of this model is available on request where it is possible to distinguish between temporary and permanent incentive schemes; the qualitative effects of the two are similar.

Equilibrium in Period 2. — Substituting (1) into (3), we see that in period 2, the politician's expected payoff is

$$\omega(f(a_2, T_2) + \theta_1 - \mu T_2) + R - ca_2 \tag{4}$$

So, the incumbent politician chooses a_2, T_2 to maximize (4), giving rise to a continuation payoff in period 2, conditional on θ_1, of

$$V(\theta_1) = \max_{a_2, T_2} \{\omega(f(a_2, T_2) + \theta_1 - \mu T_2) + R - ca_2\} \equiv V + \omega\theta_1 \tag{5}$$

The Incentive Scheme and Equilibrium in Period 1. — We begin by describing the incentive scheme. The politician gets a bonus B per unit of output, i.e. BQ_1. This can be interpreted as monetary or psychological. Obviously, the second interpretation is appropriate in the case of CPA, as local officials - elected or not - do not get any direct personal payment as a result of a good CPA score. Moreover, in view of the important role in practice that CPA and other incentive schemes play in giving voters better information, we assume that in period 1, *voters only observe output Q_1 before the election with probability q*. We suppose that this q can be increased by the incentive scheme; we refer to this as the *information effect* of the scheme. As Q_1 appears in the utility function, voters must observe it after the election i.e. at the end of period 1, if they do not observe it earlier[15]. Finally, it is assumed that voters always observe T_1 before the election, reflecting the fact that local property taxes are highly "visible".

From now on, without ambiguity, we can drop period 1 subscripts, so $\theta_1 = \theta$, $a_1 = a$, etc. The order of events in period 1 is then as follows. First, politicians choose a, T, knowing θ_0. Then, voters vote for incumbent or challenger, having observed T and, with probability q, Q. The challenger's productivity is randomly drawn from the same distribution as the incumbent's.

First, consider the voter choice between the incumbent and challenger. Because distributional concerns, measured by μ, are fixed and the same across all politicians, a fully rational voter only cares about the ability of the incumbent and challenger; such a voter wishes to re-elect the incumbent if the incumbent's expected ability, θ^e, exceeds zero, the expected ability of the challenger.

However, we are also interested in the impact of the unpredictability of voter behavior on the effectiveness of the incentive scheme. So, we assume that voters are not fully rational: following the probabilistic voting framework of Persson and Tabellini(2000), we assume that a voter votes for the incumbent if

$$\theta^e + \varepsilon + v \geq 0 \tag{6}$$

[15]A more general assumption would be that Q_1 is observed with some error ε, and the variance of this error is decreased by the incentive scheme. Under the additional assumption that θ_1, ε are Normally distributed, it can be shown that a reduction in the variance of ε has a qualitatively identical impact on a_1, T_1 to an increase in q; the details are more complex however, and are available on request.

where ε is a voter-specific shock, distributed uniformly on $[-1, 1]$ and independent across voters, v is an aggregate popularity shock, common to all voters, which has zero mean and is symmetric. The distributions of ε and v are known to the incumbent. Also, v, θ, ε are mutually independent. The individual shocks ensure that the margin of victory for either incumbent or challenger is non-trivial, i.e. not always 100%. This in turn allows us to make empirical predictions about the interaction of an incentive scheme with electoral competition. The role of the variance of the aggregate shock v is that it can be interpreted as a measure of the *salience* of the incumbent's ability for the voters, relative to other factors that they may care about (e.g. candidate charisma or other, unmodelled, factors).

We now turn to the characterization of θ^e. If Q *is not* observed, voters cannot make any inference about the incumbent's ability, and cannot do better than assume that $\theta^e = \theta_0 = 0$. Then, from (6), a voter with shock ε votes for the incumbent if $\varepsilon + v \geq 0$; then, it is easy to see that the incumbent is re-elected with probability 0.5 and moreover, his expected vote share is 0.5.

If Q *is* observed, however, the voters infer that $\theta^e = Q - f(a^e, T)$ where a^e is the voters' conjecture about the effort choice by the incumbent. Combining this with the production relationship $Q = f(a, T) + \theta$ and eliminating Q, θ^e can be written

$$\theta^e = f(a, T) - f(a^e, T) + \theta \tag{7}$$

So, combining (6) and (7), we see that a voter with shock ε votes for the incumbent if

$$\theta + \varepsilon + v \geq f(a^e, T) - f(a, T) \tag{8}$$

Note that from (8), given a fixed belief a^e by the voters, a higher effort a by the incumbent will increase his re-election probability.

Then, the maximization problem of the incumbent is to choose a, T subject to the re-election constraint (8) and taking a^e as given. Now, for tractability, and following Alesina and Tabellini(2008), we assume that v, θ are Normal with variances $\sigma_\theta^2, \sigma_v^2$. It can then be shown that (see Appendix A1) the resulting first-order conditions to this problem, evaluated in equilibrium, where $a^e = a$, are;

$$(\omega + B + qVh(0))f_a = c, \tag{9}$$

$$(\omega + B)f_T = \omega\mu \tag{10}$$

Here, V is defined in (5), and $h(0)$ is the density of the distribution of $\theta + v$ at zero. Recalling that θ, v are both normally distributed with mean zero and

variances $\sigma_\theta^2, \sigma_v^2$, this is:

$$h(0) = \frac{1}{\sqrt{(\sigma_\theta^2 + \sigma_v^2)(2\pi)}} \tag{11}$$

So, given B and other parameters, the endogenous variables a, T are simultaneously determined from the two equations (9), (10) and V is residually defined by (5). This constitutes a political equilibrium.

Note that (9) says that there are three motivations for the incumbent to supply effort; some preference congruence with the electorate ($\omega > 0$), career concerns, measured by $qVh(0)$, and finally the incentive scheme, B. Note also the asymmetry; career concerns affect the choice of effort, but not tax, ultimately because the voters can *directly* observe tax. So, changing the tax rate does not affect the voter's inference about ability θ and hence does not affect the re-election probability.

C. An Alternative Interpretation

Although this model has been presented as one of an elected representative being motivated by voters via an election, in the British context, there is an alternative, and possibly more plausible, interpretation[16]. Councils in England and Wales have the following management structure; strategic decision-making is undertaken by an executive comprised of elected officials, typically in the form of a cabinet with the leader elected by council members, with day-to-day operations headed by a full-time CEO. One could argue that CPA is also a management tool for the executive to monitor the CEO. One can therefore re-interpret our model as follows.

Voters can be plausibly re-interpreted as councillors, who live in the council district and who therefore have similar preferences to voters. The "politician" can be re-interpreted as the council CEO, who can be fired or otherwise sanctioned for poor performance. Thus, the election can be reinterpreted as any action that the executive can take to discipline the CEO. CPA is of value to councillors either because it gives them more information about CPA performance (higher q), or because there are direct benefits to the CEO of a higher CPA score i.e. earned autonomy. This re-interpretation is of course, applicable to other local government contexts where there is also a clear division between legislative and executive functions, such council–manager local government in the US.

D. Effects of an Incentive Scheme

Here, we establish our main results of the effects of an incentive scheme. We consider the effects of small changes in both B and q on taxation, T, output, the

[16]This was suggested to us by Tim Besley.

expected value of Q, and also on "efficiency", defined more precisely below. Note that up to a constant, the expected value of Q is simply $f(a,T) \equiv f$. Our first result, proved in the Appendix A1, is:

Proposition 1. *If T, a are weak complements i.e. $f_{aT} \geq 0$ then: (a) $\frac{da}{dB} > 0, \frac{dT}{dB} > 0$, and so $\frac{df}{dB} > 0$; (b) $\frac{da}{dq} > 0, \frac{dT}{dq} \geq 0$, and so $\frac{df}{dq} > 0$.*

So, we see that a stronger incentive scheme, interpreted as an increase in B and/or q, will unambiguously increase both taxes and expected output. The intuition for this result is simple. First, from (9) and (10), an increase in B raises the perceived marginal benefit of both higher effort, a, and higher tax $, T$, on the part of the incumbent. Moreover, from (9), an increase in q increases perceived marginal benefit of higher effort, a, via its effect on the career concerns term $qVh(0)$. Then, if T, a are independent, this has no effect on T, but if T, a are strict complements i.e. $f_{aT} > 0$, T will also rise.

Note also that this result does not depend on the relative size of the direct effect and the information effect of the incentive scheme. This is important, because in the empirical work, we cannot estimate the effects of B and q separately.

We now turn to look at efficiency. In our setting, the natural measure of efficiency, and the one that will be used in our empirical work, is the expected output f, minus the cost of inputs, T;

$$e \equiv f(a,T) - T \tag{12}$$

From (12), the effect of B or q on efficiency is :

$$\frac{de}{dB} = f_a \frac{da}{dB} + (f_T - 1)\frac{dT}{dB}, \ \frac{de}{dq} = f_a \frac{da}{dq} + (f_T - 1)\frac{dT}{dq} \tag{13}$$

So, we see immediately that an increase in B or q has a an *ambiguous* effect on efficiency; there is a positive effect via a, but an effect that can be negative via T. Specifically, this effect will be negative if the incumbent is already collecting too much tax revenue at the margin $f_T < 1$. In turn, from (10), we can see that $f_T = \frac{\omega\mu}{(\omega+B)}$. So, $f_T < 1$ if the bonus B is already large, there are weak concerns for voter welfare (low ω), or μ, the politician's disutility of tax, is small enough. To pin down these effects more precisely, assume that f is Cobb-Douglas. Then we can prove:

Proposition 2. *Assume $f = a^\alpha T^\beta$, $\alpha, \beta > 0$, $\alpha + \beta < 1$. An increase in q increases efficiency iff $\frac{\mu\omega}{\omega+B} > \beta$. An increase in B increases efficiency iff*

$$\frac{\mu\omega}{\omega + B} > \frac{\beta(\omega + B) + \beta(1-\alpha)qVh(0)}{(\alpha+\beta)(\omega+B) + \beta qVh(0)} \tag{14}$$

Note also that by straightforward computation[17], the right-hand side of (14) is increasing in $qVh(0)$. So, as $q, V, h(0)$ only appear on right-hand side of (14), we see that introducing a small incentive scheme B, or increasing the incentive scheme by a small amount, can *decrease* efficiency if: (i) career concerns V are strong e.g. from (5), the ego-rent R from office is high; (ii) if μ is low, reflecting a e.g. low tax price facing the median voter; (iii) the incentive scheme is more informative to voters i.e. large q, or (iv) if $h(0)$ is large. In turn, from(11), $h(0)$ is large when the variance of the ability parameter, θ is low, or when voter behavior is predictable in the aggregate i.e. the variance of the aggregate shock v is small.

Proposition 2 also implies that increasing the publicity effect q of the incentive also can decrease efficiency if: (i) μ is low, reflecting a low tax price facing the median voter; (ii) the lower is ω, the weight on welfare in politician payoffs (suggesting that the scheme is less likely to be effective where it is most needed).

Some of these results may appear counter-intuitive, but they all arise from the fact that the higher are "career concerns", measured by $qVh(0)$, the greater is the tendency to set a high tax in the first period, in order to boost output and get re-elected. In turn, from (13), if the tax is high enough, it can lead to lower efficiency.

We now focus specifically on the effect of uncertainty about incumbent ability, σ_θ, or of the salience of θ for the voters - measured inversely by σ_v - on the output and efficiency impacts of the incentive scheme. Let $\sigma = \sigma_\theta + \sigma_v$. We can then state:

Proposition 3. *(i)* $\frac{d^2 f}{dBd\sigma} > 0$. *That is, the positive effect of an incentive scheme on output is bigger when either* σ_v *or* σ_θ *is higher i.e. when there is larger uncertainty about incumbent ability, or when incumbent ability is less salient for voters. (ii) If* $\sigma' > \sigma$, $\frac{de}{dB}\big|_\sigma > 0 \implies \frac{de}{dB}\big|_{\sigma'} > 0$. *That is, an incentive scheme is more likely to increase efficiency when either* σ_v *or* σ_θ *is higher i.e. when there is larger uncertainty about incumbent ability, or when incumbent ability is less salient for voters.*

Note that the result that an incentive scheme is more likely to increase efficiency when either σ_v or σ_θ is higher follows directly from the above discussion, which has established that increased efficiency is less likely when $h(0)$ is large, plus the fact that $h(0)$ is decreasing in σ. The result that $\frac{d^2 f}{dBd\sigma} > 0$ is proved in the Appendix A1.

Of course, σ_v or σ_θ are unobservable. However, it turns out that they are related to the margin of victory, m of the incumbent, which *is* observable. Specifically, define m to be the expected fraction of voters who vote for the incumbent, minus 0.5. Using (A1) in the Appendix A1, substituting $\theta^e = \theta$ in equilibrium, and recalling that Q is only observed with probability q, the expected fraction of

[17]Let the RHS of (14) be $f(x) = \beta\frac{1+(1-\alpha)x}{\alpha+\beta+\beta x}$, $x = \frac{qVh(0)}{(\omega+B)}$. Then $f'(x)$ has the sign of $1-\alpha-\beta$, which is positive.

voters voting for the incumbent can be calculated as

$$\frac{1}{2}(1-q)+\frac{1}{2}\left(1+E_{v,\theta}[\theta+v\,|\theta+v\geq 0]\right)q$$

So, m is this fraction minus one-half, or:

$$\begin{aligned} m &= \frac{q}{2}E_{v,\theta}[\theta+v\,|\theta+v\geq 0] \\ &= \frac{q}{2}\sqrt{(\sigma_\theta^2+\sigma_v^2)}\frac{\phi\left(0\right)}{1-\Phi\left(0\right)} \end{aligned} \tag{15}$$

where in the second line, we have used a standard formula for the mean of the truncated Normal distribution, and $\phi\left(0\right), 1-\Phi\left(0\right)$ are the cdf and density of the standard Normal. So, we see from (15) that m is increasing in σ_v or σ_θ. This is intuitive; when either σ_v, σ_θ increase, there is a mean-preserving spread in the fraction of the vote share for the incumbent, with the mean constant at zero. So, conditional on this fraction being greater than 0.5, it must go up.

So, we have shown that electoral competition and an incentive scheme are substitutes: that is, the incentive scheme *will have a smaller impact on output*, and *is less likely to increase efficiency when there is strong electoral competition i.e. the incumbent's margin of victory is low.* These are testable predictions, and are investigated in Section 6 below. This positive relationship can be generated either by variation in σ_θ the variance of incumbent ability, or σ_v, the randomness in voter behavior.

IV. Empirical Strategy

A. *Empirical Specification*

Our empirical approach is to estimate the impact of CPA on efficiency in a quasi-experimental setting through difference-in-difference estimation, using Wales, where CPA was not used, as a control group. Welsh local government performance was assessed by an evaluation program called the Welsh Program for Improvement (WPI) since 2001.[18] We believe that Welsh councils can be used to address the counterfactual question of what would have been the path of English councils after 2001 if CPA league tables would not have been produced, for the following reasons. First, Welsh and English local authorities have the same structure and functions.[19] Second, the mean values of our control variables

[18]Information and data about the Welsh Program for Improvement can be collected from the web site of the Wales Audit Office www.wao.gov.uk.

[19]All Welsh local authorities are unitary, but they have the same responsibilities as English local authorities, and until 2007, the same funding structure.

and the input and output variables used to construct our service quality and efficiency indices are very similar in the two countries (see Tables A5 and A6 in the Appendix). Third, as documented by Haubrich and McLean(2006b), WPI was, compared to CPA, a much less prescriptive and elaborate assessment regime since only confidential assessments were produced, the evaluation criteria were based only on local self-assessment without quantitative rankings, and no formal rewards or punishments were specified.[20]

As a further check, we examine whether our identification assumption holds by testing whether our dependent variables i.e. the Y_{it} in (16) follow a common time path in the years before the introduction of CPA in 2001. First, as a "visual" test, Figures A1, A2, A3, and A4 in Appendix show the common trend followed by the main dependent variables in England and Wales before the introduction of CPA. Moreover, using a formal test, also reported in Appendix A3, we find that with a few exceptions, we cannot reject the hypothesis that they did follow a common time path.

Finally, we have to address the question of whether the lack of "treatment" of Welsh local authorities was a truly exogenous event, or whether it was specifically related to the performance (in the setting of taxes or provision of public services) of Welsh LAs. First, the ability of Wales to determine a separate regulatory regime was ultimately determined by the creation of self-government in Wales, in particular the creation of the Welsh National Assembly in 1998. Ultimately, support for devolution was determined by cultural factors, and can very reasonably be taken as exogenous. Second, as Haubrich and McLean(2006a) make clear, the main reason why the Welsh government did not adopt CPA was due to the smaller size of the country, which again is exogenous; "the relationship between auditor, local government department, and authority can be more intimate than in England".

So, we proceed by estimating the following;

$$Y_{it} = \alpha(CPA_t \times D_i) + \beta' \mathbf{X}_{it} + \eta_t + u_i + \varepsilon_{it} \tag{16}$$

where CPA_t is a dummy that takes value 1 after 2001 and D_i is a dummy that takes value one for English councils. Also, $X_{\mathbf{it}}$ is a vector of controls, and η_t is a set of year dummies. Finally, $Y_{it} = T_{it}, Q_{it}, e_{it}$, where T_{it} is a measure of revenue collected from the council tax, Q_{it} a measure of service quality, and e_{it} a measure of efficiency. The main parameter of interest here is α which captures the treatment effect of the CPA. The theory suggests that if $Y_{it} = T_{it},\ Q_{it}$, then $\alpha > 0$ but if $Y_{it} = e_{it}, \alpha$ has an ambiguous sign theoretically.

We treat u_i as a council fixed effect. Finally, some of the variables (all the

[20]It is important to note that the greater regional autonomy obtained in Wales at the end of the 1990s does not affect our analysis since the The National Assembly for Wales was created by the Government of Wales Act in 1998 and gained limited primary legislative powers only in 2007.

outputs, and one of the tax variables) are between zero and 1, so also, as a robustness check, we estimate a non-linear model where the dependent variable is transformed to lie between zero and 1 as follows:

$$T_{it} = \Phi[\alpha(CPA_t \times D_i) + \beta' \mathbf{X}_{it} + \eta_t + u_i] + v_{it}$$

using a pooled Bernoulli quasi-MLE,[21]. Standard errors in (16) are clustered at the council level, allowing for serial correlation in the v_{it}.

B. Data

First, we discuss our choice of measures of T_{it}, Q_{it}, e_{it} for English and Welsh councils over the period 1997-2007. In the theory, T_{it} is property tax revenue. The closest empirical proxy for this is the tax requirement in the official statistics (CIPFA(2008a)) which is total (real) current spending in the financial year minus revenue from the revenue support grant and other grants, and revenue from the business tax rate[22]. We use the tax requirement, both as a raw figure, and normalized in several ways. Specifically, we divide the tax requirement by the number of equivalent band D dwellings to get an effective council tax rate. Finally, we also measure T_{it} as a percentage of the tax requirement to the budget requirement, where the latter is actual current expenditure that has to be financed by formula grants (which includes the police grant) and property tax revenue.

Next, we turn to the measurement of service quality Q_{it}. We need to construct an index of service quality consistently across both English and Welsh local governments. To that end, the BVPIs published by the Audit Commission for England and the Audit Office for Wales are the best source of information: first they are broadly accepted by the local governments as measures of output quality; second we are very confident about the comparability of these measures across local authorities since BVPIs were also chosen as one of the building blocks of the CPA procedure.

The first problem to solve was the absence of BVPIs for the housing and benefit sector in case of the counties, where this function is managed by districts. As DEA requires a balanced production function with the same inputs and the same outputs for all units in all years, the only possible solution was to drop this sector from the efficiency analysis. A further problem is the short life of many BVPIs. Despite the fact that there are more than 250 BVPIs published on the website of the Audit Commission, almost all of them have been subject to some changes after three or four years, and in many cases replaced with new indicators. There is also the problem that after 2001-2, BVPIs were defined and measured separately

[21] We are using the methodology proposed by Papke and Wooldridge (2008) to tackle the possibility of non linearity in case of fractional dependent variable. In the non-linear model we also include council-type fixed effects.

[22] Note that in England and Wales, local authorities can borrow only to finance capital spending, not current spending, and thus the difference between current spending and formula grants must be own revenues, principally the council tax.

in both England and Wales, and there was very little overlap. In the end, only five indicators could be used to measure the quality of output consistently for England and Wales; these measure aspects of education, social care of the elderly and children, waste disposal, and central services.[23] These variables are defined in Table A4 of the Appendix, and summary statistics are given in Table A5. But, it is important to note that expenditure on these categories accounts for fully 57% of the total local government expenditure, on average[24].

As is clear from that table, four of the five BVPIs are already expressed as percentages; we converted BVPI54 to a percentage also, and averaged it with BVPI49, thus giving an aggregate index for social services. We then calculated Q_{it} as the weighted average of these four indices, where the weights used were the relative expenditures on the four services in real £ per capita; all monetary amounts were deflated using the 2005 CPI. Summary statistics on these expenditures are given in the bottom panel of Table A5. The source for the expenditure data is from the Finance and General Statistics (FGS) and Local Government Comparative Statistics (LGCS), available on the website of the Chartered Institute of Public Finance and Accountancy (CIPFA) from the 1997/98 to the 2007/08 financial years (CIPFA (2008a) and CIPFA (2008b)).

Our efficiency index e_{it} is constructed as follows. We estimate a Debreu-Farrell[25] efficiency index e_{it} for each council and year in the sample using data envelopment analysis (DEA hereafter).[26] As output measures, we use the same five BVPIs used to construct the output index. As inputs, we use the expenditures already mentioned, corresponding to those outputs. Further details are given in our working paper, Lockwood and Porcelli(2011). DEA generates two indices. The first, the input index, $e_{it}^{IN} \in (0,1]$, has the following intuitive interpretation. If council i was using the technology efficiently at time t, its inputs could all be scaled down by a fraction $1-e_{it}^{IN}$ and it would still be able to produce the same vector of outputs. The second, the output index, $e_{it}^{OUT} \in (0,1]$ has a similar interpretation: if council i was using the technology efficiently at time t, its outputs could all be scaled up by an amount $\frac{1}{e_{jt}^{OUT}} - 1$, whilst using the same vector of inputs. Some descriptive statistics for $e_{it}^{IN}, e_{it}^{OUT}$ are provided in Table A6. It is important to stress that input-based and output-based approaches to the evaluation of efficiency do not need to produce the same results. In particular, the input and output approach indices of efficiency are equivalent only in the restrictive case of constant returns to scale. Hence, in our analysis, the use of two indices can be considered a sort of robustness check. The fact that our

[23]The presence of some missing and inaccurate values in some BVPI indicators is producing a reduction in the sample size.

[24]Remarkably, if one takes a less demanding view, and only requires identical BVPIs measured in England and Wales in only *one* year before, and *one* year after, the introduction of CPA, which is a minimal requirement for difference in difference analysis, there was just one additional BVPI available, the percentage of recycled household waste that was used to generate energy.

[25]Debreu (1951) and Farrell (1957).

[26]DEA was first developed by Charnes, Cooper, and Rhodes (1978); a survey can be found in Ali and Seiford (1993).

general conclusions are qualitatively the same with the output and the input approach indicates that our results are robust, even if the estimates, as reported in Table 3, are not exactly the same. Finally, it is also explained in Appendix A2 that $e_{it}^{IN}, e_{it}^{OUT}$ will generally be upward biased. So, as also explained there, we used bootstrap methods to correct for that bias, yielding bias-corrected versions of both $e_{it}^{IN}, e_{it}^{OUT}$. All four efficiency measures are used in our regressions.

Finally, our control variables X_{it} are described in Table A6 of the Appendix, and can be subdivided as follows. First, there are demographic variables, such as the percentage of the total population below the age of 16 and above the age of 75, the percentage of population that declare themselves religious, the percentage of white people, the population density, and finally the percentage of households who own their house, the number of band D equivalent dwellings per capita that correspond to the tax base of the council tax and has been included as a proxy of the demand for local public services. The second category includes a set of dummy variables to capture the impact of the ruling party and the features of the electoral system ("all out" election every four years, or "by thirds" system which involves more frequent elections). The third group of variables is related to the structure of the local economy and includes: the amount of real per-capita revenue support grant received every year by each council,[27] average household disposable income, the percentage of the workforce claiming unemployment-related benefits, the percentage of people below 65 claiming disability living allowance, the percentage of VAT tax payers in the financial and real estate sector, the percentage of high qualified workforce, and the percentage of workforce self employed[28].

V. Empirical Results

A. Taxes

The first empirical prediction of the theoretical model is that CPA should increase council tax revenues (Proposition 1). So, we first estimate (16) with $Y_{it} = T_{it}$. As a first pass, Figure A1 of Appendix shows that the council effective tax rate (the tax requirement per equivalent band D dwelling) exhibits a clear increase in England relative to Wales after 2002. So, we would expect α to be significantly positive. For each of the three tax measures described above we estimate the empirical model in (16) as described in Section 4A above.

Table 1 shows that with all three tax measures, α is positive and significant at the 1% level. First, the introduction of CPA raised the tax requirement by about £23 per capita, or 7% in England relative to Wales. It also raised the

[27]It is important to stress that both the English and the Welsh grant system were based on the same rules during the period of our analysis. Only after 2007 some differences appeared in the English system. In particular, in both countries the system is formula based; as a result grants can be considered exogenous in relation to the behavior of local governments, since they are mainly determined by local demographic and income characteristics.

[28]Due to the absence of some data on control variables and output measures in some years, the panel is unbalanced.

effective council tax rate by about £46, corresponding roughly to a 4 percent increase in England relative to Wales[29]. Finally, it raised the tax requirement as a percentage of the budget requirement by about 6 percent in England relative to Wales. For the non-linear model, the average partial effect is reported, which is the percentage change in the dependent variable caused by the treatment. So, in this model, the effect of CPA is somewhat smaller.

Table 1—Point estimates of the treatment effect of CPA on council tax revenues.

Model	Tax requirement (real £ per capita) (A)	Effective council tax rate (real £ per band D equivalent dwelling) (B)	Tax requirement (% of budget requirement) (C)
FE (linear)	23.05*** (5.30)	45.99*** (10.81)	6.30** (1.24)
GLM (non linear) (1)	n.a.	n.a.	4.96*** (0.79)
Observations	1850	1850	1810
Number of councils	170	170	170
Control variables	yes	yes	yes
Year dummies	yes	yes	yes

Note: Clustered standard errors in brackets.
*** significant at 1%, ** significant at 5%,* significant at 10%.
(1) Point estimates are expressed as average partial effects.

B. Outputs

Second, we estimate α in (16) when the dependent variable is our service quality index. As a first step, Figure A2 and A3 in Appendix show clearly that the output index rose faster in England than in Wales after the introduction of CPA, so we would expect $\alpha > 0$. Point estimates of α are reported in the first column of Table 2 using the same econometric specifications as in Table 1. Also in this case, irrespective of the estimation method, it is possible to observe a positive and statistically significant effect of CPA on the level of outputs: on average, after the introduction of CPA, the aggregate output increased by less than 4% in English councils compared to Welsh local authorities.

But, as remarked in Section 1, a concern for us is that secondary school performance, as measured by the percentage of pupils achieving between A and C in GCSE exams, is a major component of our output index (with a weight of 63%).

[29]Note that the estimates of £24 and £45 are broadly consistent, using the fact that there are on average, according to the latest statistics (CIPFA (2008a)), about 2.3 persons per dwelling in England.

Burgess, Wilson and Worth (2010) show that this measure of performance was impacted by school league tables, which were abolished in Wales in the same year in which CPA was introduced in England. To deal with this problem, we also test whether CPA increased our output index excluding education. The results are in column 2 of Table 2: we find that in our baseline fixed effects specification, the effect of CPA is still significantly positive but smaller in magnitude. We believe that this is evidence that both CPA and other "service-specific" performance indicators such as school league tables can have positive effects on output.

Table 2—Point estimates of the treatment effect of CPA on service quality.

Model	Output measures					
	Aggregated output	Aggregated output (no education)	Education	Social services	Central services	Environment
FE	3.67***	0.60***	5.69***	2.77**	1.57	-6.29***
(linear)	(0.56)	(0.22)	(0.66)	(0.73)	(1.63)	(0.88)
GLM	4.16***	0.64***	6.22***	2.72***	2.05	-6.28***
(non linear) (1)	(0.79)	(0.23)	(0.79)	(0.54)	(1.39)	(0.74)
Observations	1397	1428	1669	1463	1808	1747
No. of councils	141	141	158	141	170	166
Control variables	yes	yes	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes	yes	yes

Note: Clustered standard errors in brackets.
*** significant at 1%, ** significant at 5%,* significant at 10%.
Point estimates are interpreted as percentage change in output index due to CPA.
(1) Point estimates are expressed as average partial effects.

The other columns of Table 2 display the results of the same exercise conducted considering the quality measures of each sector. So, for English local authorities, in general, there is empirical evidence of a positive effect of CPA on all quality measures, with the exception of the percentage of household waste recycled. Thus, our results are again broadly consistent with the theory.

C. Efficiency

We turn to look at the effect of CPA on our efficiency indices. Figure A4 of Appendix shows the path of the efficiency index in England and Wales (average between input and output approach) between 1997 and 2007. In both countries the initial decreasing trend in efficiency reversed its course after the introduction of CPA, and although the initial gap between Welsh and English councils is completely closed in the last year, there is no clear evidence of a positive impact of

CPA on the efficiency of English local authorities. This suggests an insignificant α, which is in fact what we find. Our econometric specifications are the same as in the previous two tables, except for the non-linear specification. In this case, to account for the possibility of non-linearity, we exploit the fact that the DEA indices of efficiency have an ordinal meaning; therefore we use as a dependent variable a binary indicator that will take value one if the council is ranked above the 50th percentile in the distribution of the DEA efficiency scores, and zero otherwise. This gives a random effect probit model estimated using the unconditional MLE estimator.[30]

Looking at Table 3, there is no empirical evidence in favor of a an impact of CPA on the efficiency of English councils. The coefficient of the treatment effect is statistically significant only in case of RE probit model in relation to the input approach, however the magnitude of the estimate tell us that after the introduction of CPA the probability of observing a council ranked in the upper 50th percentile of the efficiency index distribution increased by 2%, a very small number that leads us to the conclusion that the introduction of CPA did not stimulate any change in the efficiency of English local authorities in delivering public services.

Table 3—Point estimates of the treatment effect of CPA on efficiency.

Model	Input approach		Output approach	
	no bootstrap	bootstrap	no bootstrap	bootstrap
FE	-0.49	-0.26	0.53	-0.48
(linear) (1)	(2.00)	(2.18)	(0.76)	(1.13)
RE probit	1.91***	0.87	0.30	-0.06
(non linear) (2)	(0.72)	(1.34)	(0.39)	(0.61)
Observations	1245	790	1245	748
Number of councils	141	140	141	139
Control variables	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes

Note: Clustered standard errors in brackets.
*** significant at 1%, ** significant at 5%,* significant at 10%.
(1) Coefficient point estimates are interpreted as percentage change in efficiency index due to CPA.
(2) Dependent variable is 1 in year t iff council is ranked in the upper 50th percentile of the efficiency distribution in year t. Coefficient point estimates are interpreted as percentage change in the probability of being ranked in the 50th percentile of the efficiency index distribution.

[30]It is important to note that in this case the Mundlak (1978) approach, of including time-average of time-varying regressors as additional regressors, will be followed in order to tackle the possibility that the unobserved heterogeneity and the regressors may not be orthogonal.

D. Robustness Checks

Here, we report a number of robustness checks. First, as pointed out by Bertand, Duflo, and Mullainathan (2004) even with clustered standard errors, there can be downward bias in the standard error in α, leading to false rejection of the null hypothesis of no treatment effect. To deal with this, we follow the procedure recommended in that paper of collapsing the time dimension to before and after the treatment, and re-estimate all of our specifications. As can be seen by comparing Tables A7-A9 in Appendix to Tables 1-3 and 4, the results are robust to this alternative estimation method, providing evidence that serial correlation is not a problem.

A second robustness check is to allow for council-specific time trends (see e.g. Friedberg(1998)). To avoid collinearity problems, we just add linear time trends for each *type* of council (London borough, Metropolitan district, County, Unitary authority, Welsh Unitary authority). The addition of these effects does not significantly change our regression results, with the exception of the results of column C of Table 1[31]. However, this third empirical measure of tax revenue is less closely related to the theory, as it deflates tax revenue by expenditure. These additional results are available in the online appendix.

Finally, we run some placebo tests on the timing of the treatment. Specifically, we have re-estimated the effect of CPA on output, tax and efficiency, *assuming that the CPA program started in some other year than the year in which it actually occurred* i.e. the fiscal year 2001/2. The results of these tests are also available in the online appendix, but we summarize them here. In the placebo treatments where CPA was introduced "before" 2001/2, either the treatment effect is insignificant or it has the opposite sign to that predicted by the theory i.e. negative effects on taxes and output. In the placebo treatments where CPA was introduced "after" 2001/2, the treatment effect is mostly insignificant. However, we do observe significant positive treatment effects on taxes in the case where the placebo is one year after the true date of introduction. This could simply reflect the fact that councils reacted slowly to the introduction of the new regime.

VI. Electoral Competition and CPA

The effects of electoral competition on policy-makers' behavior are widely studied in the political science literature, and increasingly also by economists[32]. In Proposition 3 above and the subsequent discussion, we showed that an incentive

[31] These coefficients become significantly negative.

[32] See e.g. Holbrook and Dunk(1993) Griffin(2006) for recent political science studies. The study most closely related to ours in this respect is Besley and Preston(2007), who construct a measure of electoral districting bias for English local authorities. They find some evidence that a larger bias for the incumbent party (which protects the incumbent from electoral competition) allows the party more chance to pursue its policy preferences, which is lower expenditure and lower local government employment in the case of Conservatives, and the reverse in the case of Labour.

TABLE 4—HETEROGENEOUS TREATMENT EFFECT OF CPA ON TAXATION, AGGREGATE OUTPUT AND EFFICIENCY FOR ENGLISH LAS WITH INITIAL LOW LEVELS OF ELECTORAL COMPETITION.

Model	Effective council tax rate (real £ per band D equivalent dwelling)	Aggregated output	Efficiency
	(A)	(B)	(C)
FE (linear)	5.59	4.37***	10.14***
	(23.32)	(0.75)	(2.31)
Observations	1329	1101	885
Number of councils	166	137	135
Control variables	yes	yes	yes
Year dummies	yes	yes	yes

Note: Clustered standard errors in brackets.
*** significant at 1%, ** significant at 5%,* significant at 10%.
Local authorities without a clear majority have been excluded from the sample.

scheme is less likely to increase efficiency, and will have a weaker effect on output, when there is strong electoral competition i.e. the incumbent's margin of victory is low. In other words, electoral competition and CPA are predicted to be *substitutes*. This is a testable prediction, which we investigate by estimating the following specification[33]:

$$Y_{it} = \alpha_1(CPA_t \times D_i) + \alpha_2(Z_{it} \times CPA_t) + \alpha_3(CPA_t \times D_i \times Z_{it}) + \beta' \mathbf{X}_{it} + \eta_t + u_i + \varepsilon_{it} \tag{17}$$

where $Y_{it} = Q_{it}, e_{it}$, and Z_{it} is a measure of electoral competition in LA i at time t. The parameter of interest is α_3. Note that this specification, via the inclusion of $Z_{it} \times CPA_t$, and $CPA_t \times D_i \times Z_{it}$ allows for characteristic Z_{it} to have separate effects on e_{it} in both England and Wales before and after CPA. Our measure is defined as $Z_{it} = 1$ if the margin of victory at the last election before the introduction of CPA was over 5%, and $Z_{it} = 0$ otherwise[34]. Councils where there was no overall control in that year are dropped from the regressions. So, Z_{it} is an inverse measure of electoral competition. As a result, from Proposition 3, we expect $\alpha_3 > 0$ for both $Y_{it} = Q_{it}, e_{it}$.

Table 4 shows the estimates for α_3 when the dependent variable is either the aggregate output index, or the efficiency index. The results show very clearly that low electoral competition has, consistently with the theory, a significantly

[33]Note that $\mathbf{X}_{it}$ includes also Z_{it}.
[34]As a robustness check, we also tried to use different thresholds and the results are not qualitatively different.

positive impact on both output and efficiency. For completeness, we also show the regression (17) where the dependent variable is the tax. In this case there is no significant effect.

So, the results are broadly in line with our hypothesis that CPA was a substitute for electoral competition; in councils where electoral competition was initially weak, it appears that CPA significantly increased both output and efficiency leaving the level of the property tax unchanged.

VII. Conclusions

This paper has studied Comprehensive Performance Assessment, an explicit incentive scheme for local government in England. A simple theoretical political agency model, which extends a literature on principal-agent problems where the agent has several tasks to perform, predicted that CPA should increase service quality and local taxation, but have an ambiguous effect on the efficiency of service provision. We tested these predictions using a difference in difference approach, using Welsh local authorities as a control group, exploiting the fact that local authorities in Wales were not subject to the same CPA regime. We also constructed indices of service quality and efficiency, using Best Value Performance Indicators as well as expenditures on different categories of services. We estimate that CPA increased the effective band D council tax rate in England relative to Wales by 4%, and increased our index of service quality output also by about 4%, but had no significant effect on our efficiency indices. Moreover, consistently with the theory, there is robust evidence that CPA can substitute for an initial lack of electoral competition in driving up output and efficiency. The main policy implication of these results is that an incentive scheme like CPA can fail in stimulating higher local government efficiency because is too output-oriented. An implication of our findings is that incentive schemes should be designed to place some substantial weight on efficiency, and not just reward output.

References

Alesina Alberto and Tabellini Guido, (2008). "Bureaucrats or politicians? Part II: Multiple policy tasks", Journal of Public Economics, 92:426-447.

Ali Agha Iqbal and Seiford Lawrence M., (1993). "The mathematical programming approach to efficiency analysis", in Fried, A. O. Lovell, A. K., Schmidt, S. S. (Eds), The measurement of Productive Efficiency, Oxford University Press, Ch. 3, 120-159.

Andrews Matthew, Duncombe William and Yinger John, (2002), "Revisiting Economies of Size in American education: are we any closer to a consensus?", Economics of Education Review, 21:245-262

Audit Commission, (2009)."Final score: The impact of CPA of local government, 2002-2008", www.audit-commission.gov.uk.

Bertrand Marianne, Duflo Esther and Mullainathan Sendhil, (2004). "How much should we trust differences-in-differences estimates?", The Quarterly Journal of Economics, 119(1):249-275.

Besley Timothy and Coate Stephen, (1997). "An Economic Model of Representative Democracy", The Quarterly Journal of Economics, 112(1):85-114.

Besley Timothy, (2004). "Paying politicians: theory and evidence", Journal of the European Economic Association, 2(2-3):193-215.

Besley Timothy J., Bevan Gwyn and Burchardi Konrad B., (2009), "Naming & Shaming: The impacts of different regimes on hospital waiting times in England and Wales", CEPR Discussion Papers 7306.

Besley Timothy and Preston Ian, (2007). "Electoral Bias and Policy Choice: Theory and Evidence", The Quarterly Journal of Economics, MIT Press 122(4):1473-1510.

Boyne George A., James Oliver, John Peter, and Petrovsky Nicolai, (2009). "Political and managerial succession and the performance of English local governments", The Journal of Politics, 71:1271-1284.

Burgess Simon, Wilson Deborah, and Worth Jack, (2010). "A natural experiment in school accountability: the impact of school performance information on pupil progress and sorting", CMPO Working Paper Series, No. 10/246.

Callan Scott J, and Thomas Janet M., (2001). "Economies of scale and scope: A cost analysis of municipal solid waste services", Land Economics, 77(4):548-60

Caselli Francesco and Morelli Massimo, (2004). "Bad politicians", Journal of Public Economics, 88(3-4):759-782

CIPFA, (2008a). Finance and General Statistics 2007/08, The Chartered Institute of Public Finance and Accountancy. London.

CIPFA, (2008b). Local Government Comparative Statistics 2008. The Chartered Institute of Public Finance and Accountancy. London.

Charnes A., Cooper, W. W. and Rhodes E., (1978). "Measuring the Efficiency of Decision-Making Units", European Journal of Operational Research, 2(6):429-444.

Debreu Gerard, (1951). "The coefficient of resource utilization", Econometrica, 19(3), 273-292.

DeGroot Morris H., (2004). Optimal Statistical Decisions, Wiley Classics

DewatripontMathias, Jewitt Ian and Tirole Jean, (1999). "The Economics of Career Concerns, Part I: Comparing Information Structures", Review of Economic Studies, 66(1):183-198.

Farrell, M. J., (1957). "The measurement of productive efficiency", Journal of the Royal Statistical Society, 120(3), 253-281.

Friedberg Leora, (1998). "Did Unilateral Divorce Raise Divorce Rates? Evidence from Panel Data", The American Economic Review, 88(3):608-627.

Gersbach Hans and Liessem Verena, (2008). "Incentive contracts and elections for politicians with multi-task problems", Journal of Economic Behavior & Organization, 68:401-411.

Griffin John D., (2006). "Electoral Competition and Democratic Responsiveness: A Defense of the Marginality Hypothesis", Journal of Politics, 68:911-921.

Haubrich Dirk and McLean Iain S., (2006a). "Assessing public service performance in local authorities through CPA - a research note in deprivation", NIER, 197:93-105.

Haubrich Dirk and McLean Iain S., (2006b). "Evaluating the performance of local government a comparison of the assessment regimes in England, Scotland and Wales", Policy Studies, 27(4):271-293.

Holbrook Thomas M. and Van Dunk Emily, (1993). "Electoral Competition in the American States", American Political Science Review 87:955-962.

Holmstrom Bengt and Milgrom Paul, (1991). "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design", Journal of Law, Economics and Organization, 7(0):24-52.

Kneip Alois, Park Byeong U. and Simar Leopold, (1998). "A note on the convergence of nonparametric DEA estimators for production efficiency scores", Econometric Theory, 14, 783-793.

Lockwood Ben and Porcelli Francesco, (2011). "Incentive Schemes for Local Government : Theory and Evidence from Comprehensive Performance Assessment in England," The Warwick Economics Research Paper Series 960, University of Warwick, Department of Economics.

Martin Steve, Downe James, Grace Clive and Nutley Sandra, (2010). "Validity, utilization and evidence-based policy: The development and impact of performance improvement regimes in local public services", Evaluation, 16(31):31-42.

Maskin Eric and Tirole Jean, (2004). "The Politician and the Judge: Accountability in Government", The American Economic Review, 94(4):1034-1054.

Mattozzi Andrea and Merlo Antonio, (2008). "Political careers or career politicians?", Journal of Public Economics, 92(3-4):597-608

McLean Iain, Haubrich Dirk and Gutierrez-Romero Roxana, (2007). "The Perils and Pitfalls of Performance Measurement: The CPA Regime for Local Authorities in England", Public Money and Management, April, 111-117

Messner Matthias and Polborn Mattias K., (2004). "Paying politicians", Journal of Public Economics, 88(12):2423-2445

Mundlak Yair, (1978). "On the Pooling of Time-Series and Cross-Section Data", Econometrica, 46:68-86.

Papke Leslie E. and Wooldridge Jeffrey M., (2008). "Panel data methods for fractional response variables with an application to test pass rates", Journal of Econometrics, 145:121-133.

Porcelli Francesco, (2010). "Can local governments buy a good performance evaluation? Theory and evidence from the comprehensive performance assessment of English local authorities", unpublished paper, University of Warwick

Propper Carol, Sutton Matt, Whitnall Carolyn and Windmeijer Frank, (2008). "Incentives and Targets in Hospital Care: Evidence from a Natural Experiment", CMPO Working Paper Series, No. 08/205.

Propper Carol, Sutton Matt, Whitnall Carolyn and Windmeijer Frank, (2008a). "Did targets and terror Reduce English Waiting Times for Elective Hospital Care?", CMPO Working Paper Series No. 07/179.

Revelli Federico, (2008). "Performance competition in local media markets", Journal of Public Economics, 92:1585-1594.

Revelli Federico, (2010). "Spend more, get more? An inquiry into English local government performance", Oxford Economic Papers, 62:185-207.

Rogoff Kenneth and Sibert Anne, (1988). "Elections and Macroeconomic Policy Cycles", Review of Economic Studies, 55(1):1-16.

Simar Leopold and Wilson Paul W., (1998). "Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models", Management Science, 44(1):49-61.

Simar Leopold and Wilson Paul W., (2000). "A general methodology for bootstrapping in nonparametric frontier", Journal of Applied Econometrics, 27(6):779-802.

Appendix

A1. Proofs of Propositions and Other Results

The Incumbent's Problem. We drop period 1 subscripts, so $\theta_1 = \theta,\ a_1 = a$, etc. The incumbent knows that if Q *is not* observed, he will be elected with probability 0.5. If Q *is* observed, from (6), the expected fraction of voters who vote for the incumbent is

$$\Pr\left(\varepsilon \geq -\theta^e - v\right) = \frac{1}{2} + \frac{\theta^e + v}{2} \tag{A1}$$

Then, from (A1), the incumbent is re-elected if $\theta^e + v \geq 0$, or using (7), if

$$\theta + v \geq f(a^e, T) - f(a, T) \tag{A2}$$

Now, we can write the expected continuation payoff of the incumbent in period 1:

$$(1-q)0.5V + qE_{\theta,v}[V + \omega\theta \,|\theta + v \geq f(a^e, T) - f(a, T)] \tag{A3}$$

The first term in (A3) is the unconditional expectation of $V + \omega\theta$ when Q *is not* observed times the re-election probability 0.5 in that event, times $1 - q$, the probability that Q is not observed. The second term in (A3) is the expectation of $V + \omega\theta$ when Q *is* observed, conditional on re-election using $(A2)$, times q.

Next, defining $x = \theta + v$, $x \sim N(0, \sigma^2)$, $\sigma = \sigma_\theta + \sigma_v$, and denoting the density of x by $h(.)$, we have;

$$\begin{aligned} E[V + \omega\theta \,|\theta + v \geq f(a^e, T) - f(a, T)] &= \int_{f(a^e,T)-f(a,T)}^{\infty} E[V + \omega\theta \,|x]h(x)dx \\ &= \int_{f(a^e,T)-f(a,T)}^{\infty} \left(V + \omega E[\theta \,|\theta + v]\right) h(x)dx \\ &= \int_{f(a^e,T)-f(a,T)}^{\infty} \left(V + \omega\frac{\sigma_\theta}{\sigma_\theta + \sigma_v}x\right) h(x)dx \end{aligned}$$

where in the third line, we have used a standard formula (e.g. DeGroot (2004)) for $E[\theta \,|\theta + v]$ in the case of Normal variables. So, ignoring the constant term $(1-q)0.5V$ in (A3), the incumbent solves:

$$\max_{a,T}\left\{(\omega + B)f(a, T) - \omega\mu T - ca + q\int_{f(a^e,T)-f(a,T)}^{\infty} \left(V + \omega\frac{\sigma_\theta}{\sigma_\theta + \sigma_v}x\right) h(x)dx\right\}$$

Taking $a^e = a$, the first-order conditions to this problem, evaluated at equilibrium,

where $a^e = a$ are:

$$(\omega + B)f_a - c + q\left(V + \omega\frac{\sigma_\theta}{\sigma_\theta + \sigma_v}.0\right)h(0)f_a = 0$$
$$(\omega + B)f_T - \omega\mu = 0$$

which give (9), (10) as required. □

Proof of Proposition 1. From total differentiation of (9), (10), and application of Cramer's rule, we have:

$$\frac{da}{dB} = \frac{-f_a f_{TT}(\omega + B) + f_{aT}f_T(\omega + B + qVh(0))}{D} \tag{A4}$$
$$\frac{dT}{dB} = \frac{-f_T f_{aa}(\omega + B + qVh(0)) + f_{aT}f_a(\omega + B)}{D}$$

where $D = (f_{aa}f_{TT} - f_{aT}^2)(\omega + B)(\omega + B + qVh(0)) > 0$ by the assumption of strict concavity of $f(a, T)$. So, as $f_{aT} \geq 0$, we see from (A4) that $\frac{da}{dB}, \frac{dT}{dB} > 0$. Also, in the event of a change in q :

$$\frac{da}{dq} = \frac{-f_a f_{TT}(\omega + B)Vh(0)}{D}, \frac{dT}{dq} = \frac{f_{aT}f_a(\omega + B)Vh(0)}{D} \tag{A5}$$

From (A5) using $D, f_{aT} > 0$, we again see that $\frac{da}{dq}, \frac{dT}{dq} > 0$. □

Proof of Proposition 2. From (13), we can write:

$$\frac{de}{dq} \geq 0 \Leftrightarrow \frac{f_a\frac{da}{dq}}{f_T\frac{dT}{dq}} > \frac{1 - f_T}{f_T} \Leftrightarrow f_T > \frac{1}{1 + \frac{f_a\frac{da}{dq}}{f_T\frac{dT}{dq}}} \tag{A6}$$

But from (A5), and $f = a^\alpha T^\beta$,we have

$$\frac{f_a\frac{da}{dq}}{f_T\frac{dT}{dq}} = \frac{-f_a^2 f_{TT}}{f_T f_{aT} f_a} = \frac{1 - \beta}{\beta} \tag{A7}$$

Combining (A6), (A7) gives the condition $f_T > \beta$. Again from (13), we have:

$$\frac{de}{dB} \geq 0 \Leftrightarrow \frac{f_a\frac{da}{dB}}{f_T\frac{dT}{dB}} > \frac{1 - f_T}{f_T} \Leftrightarrow f_T > \frac{1}{1 + \frac{f_a\frac{da}{dB}}{f_T\frac{dT}{dB}}} \tag{A8}$$

But from (A4), and $f = a^{\alpha}T^{\beta}$,we have

$$\text{(A9)} \qquad \frac{f_a \frac{da}{dB}}{f_T \frac{dT}{dB}} = \frac{-f_a^2 f_{TT}(\omega + B) + f_a f_{aT} f_T(\omega + B + qVh(0))}{-f_T^2 f_{aa}(\omega + B + qVh(0)) + f_T f_{aT} f_a(\omega + B)} = \frac{\alpha(\omega + B) + \beta\alpha qVh(0)}{\beta(\omega + B) + \beta(1-\alpha)qVh(0)}$$

where in the second line, we have used the properties of the Cobb-Douglas production function. Combining (A8) and (A9), and using $f_T = \frac{\omega\mu}{\omega+B}$, gives (14), as required. □

Proof of Proposition 3. From total differentiation of $f(a, T)$, we have:

(A10)

$$\begin{aligned}\frac{df}{dB} &= f_a \frac{da}{dB} + f_T \frac{dT}{dB} \\ &= \frac{(f_T f_a f_{aT} - f_a^2 f_{TT})(\omega + B) + (f_T f_a f_{aT} - f_T^2 f_{aa})(\omega + B + C)}{(f_{aa} f_{TT} - f_{aT}^2)(\omega + B)(\omega + B + C)},\ C = qVh(0) \\ &= \frac{f}{1-\alpha-\beta}\left(\frac{\alpha}{\omega + B + C} + \frac{\beta}{\omega + B}\right)\end{aligned}$$

where in the second line, we have used (A4), and in the third, the properties of the Cobb-Douglas i.e. using $f = a^{\alpha}T^{\beta}$. But from (A10), it follows that $\frac{df}{dB}$ is decreasing in C and therefore $h(0)$. But as $h(0)$ is decreasing in σ, the result that $\frac{df}{dB}$ is increasing in σ is proved. □

A2. Bias Correction of the Efficiency Indices

It is well-known that DEA produces an upward-biased estimate of the true Debreu-Farrell measure of technical efficiency, due to the piece-wise shape of the DEA frontier that approximates the true unobserved frontier. As a result DEA underestimates the distance of all input/output combinations from the true frontier. Typically the bias, as well as the precision of the its estimation, become smaller as the number of observations increases and becomes larger as we increase the dimensions of the production function (see Kneip, Park and Simar 1998). In this study we have a large number of observations (more than 1200). Nevertheless, as a robustness check, we construct a "bias corrected" measure of efficiency, $\tilde{e}_{it}$, following the bootstrap methodology developed by Simar and Wilson (1998, 2000).

After estimating our bias corrected measure of efficiency, we found that the Spearman correlation between e_{it} and $\tilde{e}_{it}$ is 0.96 and 0.93 in cases of input and output approaches respectively. Therefore, given the large number of observations, the magnitude of the bias is not a big issue in this case, in fact e_{it} and $\tilde{e}_{it}$

provide very similar regression results.

The main concern is that $\tilde{e}_{it}$ may be imprecisely estimated. The precision of the estimate of $\tilde{e}_{it}$ is measured by the width of the 95% confidence interval CI_{it}. So, to check the robustness of our bias-corrected measure of efficiency, we drop observations where the efficiency index is too imprecisely measured. To do this, first calculate the quartiles of the distribution of the $\tilde{e}_{it}$. We then retain observation $\tilde{e}_{it}$ only if CI_{it} lies entirely in one quartile; otherwise, we drop it. As a result, we have constructed a sub-sample of statistically "significant" bias-corrected indices of efficiency. As shown in Table A1 is possible to keep 64% of the DEA bias-corrected efficiency indices in case of the input approach, and 60% in case of the output approach.

Table A1—Statistically significant efficiency scores.

Type of local authorities	Total observations	Input Approach		Output Approach	
	e_{it} NxT	Statist. Significant	%	Statist. Significant	%
English Counties	310	188	60%	193	62%
London Boroughs	249	186	75%	155	62%
English Metr. Districts	284	171	60%	169	60%
English Unit. Authorities	317	195	62%	196	62%
Welsh Unit. Authorities	85	50	59%	35	41%
Total	1245	790	63%	748	60%

A3. Testing the Common Trend Assumption

The fundamental identifying assumption underlying the validity of the quasi-experimental setting is that the variable of interest should follow the same time path in control and the treated group in the absence of the treatment. Figures A1, A2, A3 and A4 show that the effective council tax rate, the aggregate output, the output on education and the efficiency indices where following a similar path in England and Wales before the introduction of CPA. We test for this hypothesis more formally by running, for the pre-treatment period from 1997 to 2000, the regression

$$Y_{it} = \eta_t + \theta_t(\eta_t \times D_i) + \beta' \mathbf{X}_{it} + u_i + v_{it} \tag{A11}$$

In (A11) Y_{it} is the variable of interest, η_t is the set of year dummies, D_i is a dummy for English councils, and θ_t is the parameter of interest. So, given that CPA started in 2001, the null hypothesis that the variable of interest follows the same time path is simply $H_0 : \theta_{97}, \theta_{98}, \theta_{99}, \theta_{00} = 0$. As reported in the Table A2

below, the null hypothesis can not be rejected in most of our tests. As reported in the table, p-values were below the critical threshold of the 5% significance level only for the output variables related to central services and environment sector, for the tax requirement as a percentage of the budget requirement, and for the raw index of efficiency in case of output approach.

TABLE A2—P-VALUES RELATED TO THE NULL HYPOTHESIS $H_0 : \theta_{97}, \theta_{98}, \theta_{99}, \theta_{00} = 0$.

Variables	p-value*
Aggregate output	0.58
Aggregate output without education	0.22
Education (BVPI38)	0.12
Social services (BVPI54+BVPI49)	0.06
Environment (BVPI82a)	0.00
Central services (BVPI8)	0.01
Tax requirement (real £ per capita)	0.28
Tax requirement (% of budget requirement)	0.01
Effective council tax rate (real £ per dwelling)	0.15
e^{IN}	0.06
e^{OUT}	0.00
e^{IN}+ bootstrap procedure	0.68
e^{OUT}+ bootstrap procedure	0.73

Note: *Probability of rejecting the null hypothesis of similar time path between England and Wales in the pre-treatment period when the null is true.

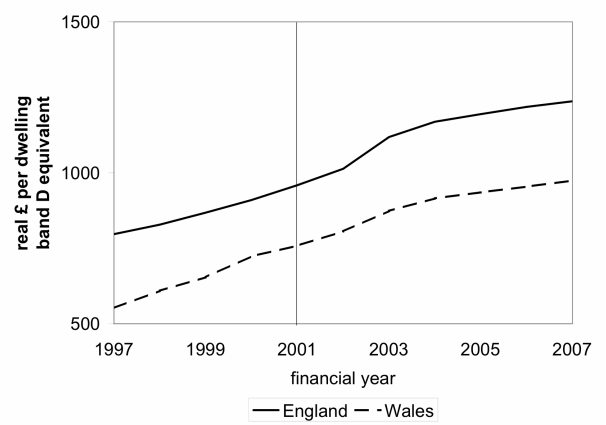

FIGURE A1. EFFECTIVE COUNCIL TAX RATE.

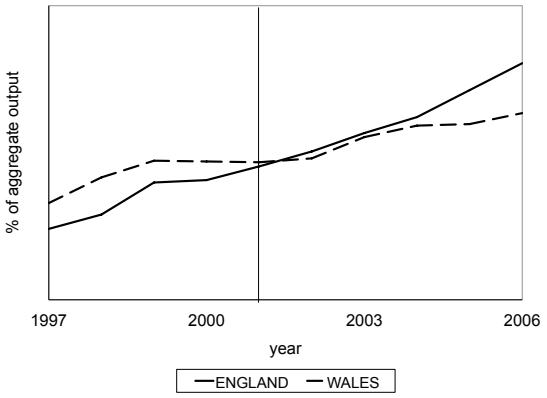

FIGURE A2. AGGREGATE OUTPUT.

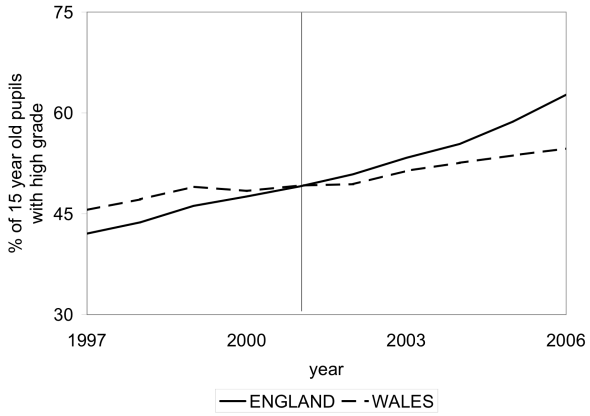

FIGURE A3. EDUCATION.

Note: Percentage of 15 year old pupils in schools maintained by the local education authority achieving five or more GCSEs at grades A*-C or equivalent.

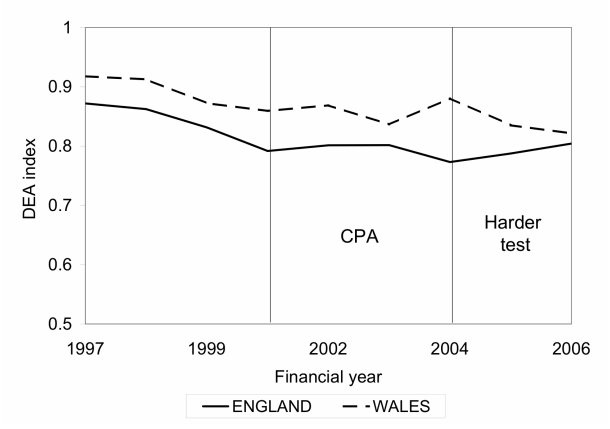

FIGURE A4. EFFICIENCY.

Note: Raw DEA indices, average between input and output approach.

Table A3—CPA aggregation rule, first three rounds.

Counties		London, MD, UA	
Performance score	Category score	Performance score	Category score
Less then 24 points	1	Less then 30 points	1
24 to 29 points	2	30 to 37 points	2
30 to 36 points	3	38 to 45 points	3
More than 36 points	4	More than 45 points	4

Councils' ability to improve	Councils' performance score on core services			
	1	2	3	4
1	poor	poor	weak	n.a.
2	poor	weak	fair	good
3	weak	fair	good	excellent
4	n.a.	good	excellent	excellent

Source: Audit Commission(2009)

TABLE A4—DESCRIPTION OF OUTPUT VARIABLES.

Service	BVPI code	Description	Period
Secondary Education	BVPI38	Percentage of 15 year old pupils in schools maintained by the local education authority achieving five or more GCSEs at grades A*-C or equivalent	Average over the current and the three following academic years
Social services (children)	BVPI49	The percentage of looked after children with no more than three placements during the last financial year (BVPI49)	Financial year
Social services (elderly)	BVPI54	Older people helped to to live at home per 1000 population aged 65 or over (BPVI54)	Financial year
Waste disposal	BVPI82a	Percentage of household waste arising which have been sent by the Authority for recycling (BPVI82a)	Financial year
Central services	BVPI8	Percentage of invoices paid by the Authority within 30 days of receipt or within the agreed payment terms (BVPI8)	Financial year

Source: Audit Commission(2009)

Table A5—Descriptive statistics, output and input variables.

	Mean		Std. Dev.		Observations	
	England	Wales	England	Wales	England	Wales
Output variables						
Secondary education	52.14	50.51	10.00	6.42	1505	184
Social service (children)	88.12	90.34	4.13	4.51	1352	122
Social service (adults)	86.33	100.27	27.63	39.01	1599	215
Waste disposal	12.44	16.03	6.16	10.11	1549	199
Central services	83.29	82.69	10.98	9.66	1596	213
Input variables						
Secondary education (real £ per capita)	508	551	106	81	1460	220
Social service, children and adults (real £ per capita)	250	267	78	61	1362	203
Waste disposal (real £ per capita)	21	20	7	9	1457	220
Central services (real £ per capita)	20	30	12	15	1606	242

Table A6—Control variables and dependent variables, descriptive statistics.

	Mean		Std. Dev.		Observations	
	England	Wales	England	Wales	England	Wales
Tax requirement (real £ per capita)	300	244	71	55	1608	242
Tax requirement (% of budget requirement)	32	19	12	2	1608	242
Effective council tax rate (real £ per dwelling)	1039	808	196	165	1589	221
e^{IN}	0.78	0.82	0.14	0.10	1160	85
e^{OUT}	0.94	0.95	0.04	0.03	1160	85
e^{IN}(after bootstrap)	0.69	0.74	0.12	0.09	740	50
e^{OUT}(after bootstrap)	0.92	0.94	0.04	0.03	713	35
% age 0 - 16	22.40	22.63	1.78	1.16	1628	242
% age over 65	15.36	17.86	2.92	2.07	1628	242
% religious	77.71	73.79	4.54	4.20	1628	242
% white	89.18	98.35	12.80	1.70	1628	242
% tenure (house ownership)	66.81	70.95	11.26	3.86	1628	242
Band D equivalent dwelling (% per capita)	33.91	35.46	5.25	4.43	1608	242
Population density (persons per hectare)	24.51	4.13	27.15	4.70	1608	242
Conservative dummy (majority of seats)	0.23	0.03	0.42	0.18	1628	242
Labour dummy (majority of seats)	0.40	0.52	0.49	0.50	1628	242
Lib. Dem. dummy (majority of seats)	0.07	0.00	0.26	0	1628	242
No overall control dummy	0.28	0.15	0.45	0.36	1628	242
Election by thirds dummy	0.37	0	0.48	0	1628	242
Revenue support grants (real £ per capita)	379	810	238	143	1608	242
Disposable income (real £ per capita)	12868	10918	3076	1124	1628	242
% firms in the financial sector	29.89	17.69	9.00	6.53	1628	242
% of unemployment related benefit	3.06	3.04	1.61	1.06	1628	242
% attendance allowance below age 65	4.67	7.94	1.93	2.51	1628	242
% high qualified workforce	5.12	3.52	2.33	1.19	1628	242
% self employed work force	7.92	8.04	2.24	3.46	1628	242

TABLE A7—TREATMENT EFFECT OF CPA ON COUNCIL TAX MEASURES, DATA COLLAPSED PRE- AND POST-REFORM PERIODS.

Model	Tax requirement (real £ per capita) (A)	Effective council tax rate (real £ per band D equivalent dwelling) (B)	Tax requirement (% of budget requirement) (C)
FE	72.61***	126.34**	11.31***
(linear)	(24.45)	(54.47)	(2.14)
GLM	n.a.	n.a.	1.69
(non linear) (1)			(1.56)
Observations	340	340	340
Number of councils	170	170	170
Control variables	yes	yes	yes

Note: Clustered standard errors in brackets.
*** significant at 1%; ** significant at 5%; * significant at 10%.
Coefficient point estimates are interpreted as follows: £ per capita in column (A), £ per dwelling in column (B), and % change in column (C).
(1) Point estimates are in terms of average partial effect.

TABLE A8—POINT ESTIMATES OF THE TREATMENT EFFECT OF CPA ON SERVICE QUALITY, DATA COLLAPSED PRE- AND POST-REFORM PERIODS.

Model	Output measures					
	Aggregated output	Aggregated output (no education)	Education	Social services	Central services	Environment
FE	4.24**	1.82***	5.07**	6.00***	2.86	-7.37**
(linear)	(1.89)	(0.54)	(2.55)	(1.57)	(6.58)	(2.97)
GLM	5.56	1.55	6.90	5.71***	5.78	-4.37
(non linear) (1)	(3.99)	(1.17)	(5.20)	(2.11)	(6.92)	(3.13)
Observations	280	281	316	281	340	332
No. of councils	141	141	158	141	170	166
Control variables	yes	yes	yes	yes	yes	yes

Note: Clustered standard errors in brackets.
*** significant at 1%; ** significant at 5%; * significant at 10%.
Coefficient point estimates are interpreted as percentage change in output index due to CPA.
(1) Point estimates are in terms of average partial effect.

TABLE A9—TREATMENT EFFECT OF CPA ON EFFICIENCY AND HETEROGENEOUS EFFECTS, DATA COLLAPSED PRE- AND POST-REFORM PERIODS (ONLY RAW DEA EFFICIENCY INDICES).

Model	Efficiency indices		Heterogeneous effect on English LAs with initial weak electoral competition		
	Input app.	Output app.	Taxation	Agg. output	Efficiency
FE (linear)	1.34	0.03	4.59	4.52***	6.32***
	(7.38)	(2.85)	(24.62)	(0.75)	(2.02)
Observations	280	280	314	256	254
No. of councils	141	141	166	137	137
Control variables	yes	yes	yes	yes	yes

Note: Clustered standard errors in brackets.
*** significant at 1%; ** significant at 5%; * significant at 10%.