

# DISCUSSION PAPER SERIES

No. 5856

## ASSESSING THE EFFECTS OF LOCAL TAXATION USING MICROGEOGRAPHIC DATA

Gilles Duranton, Laurent Gobillon  
and Henry G. Overman

*INTERNATIONAL TRADE*



**C**entre for **E**conomic **P**olicy **R**esearch

[www.cepr.org](http://www.cepr.org)

Available online at:

[www.cepr.org/pubs/dps/DP5856.asp](http://www.cepr.org/pubs/dps/DP5856.asp)

# ASSESSING THE EFFECTS OF LOCAL TAXATION USING MICROGEOGRAPHIC DATA

**Gilles Duranton**, University of Toronto and CEPR  
**Laurent Gobillon**, Institut National d'Etudes Demographiques (INED)  
**Henry G. Overman**, London School of Economics (LSE) and CEPR

Discussion Paper No. 5856  
September 2006

Centre for Economic Policy Research  
90–98 Goswell Rd, London EC1V 7RR, UK  
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999  
Email: [cepr@cepr.org](mailto:cepr@cepr.org), Website: [www.cepr.org](http://www.cepr.org)

This Discussion Paper is issued under the auspices of the Centre's research programme in **INTERNATIONAL TRADE**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as a private educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions. Institutional (core) finance for the Centre has been provided through major grants from the Economic and Social Research Council, under which an ESRC Resource Centre operates within CEPR; the Esmée Fairbairn Charitable Trust; and the Bank of England. These organizations do not give prior review to the Centre's publications, nor do they necessarily endorse the views expressed therein.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Gilles Duranton, Laurent Gobillon and Henry G. Overman

## ABSTRACT

### Assessing the Effects of Local Taxation Using Microgeographic Data\*

We study the impact of local taxation on the location and growth of firms. Our empirical methodology pairs establishments across jurisdictional boundaries to estimate the impact of taxation. Our approach improves on existing work as it corrects for unobserved establishment heterogeneity, for unobserved time-varying site specific effects, and for the endogeneity of local taxation. Applied to data for English manufacturing establishments we find that local taxation has a negative impact on employment growth, but no effect on entry.

JEL Classification: H22, H71 and R38

Keywords: borders, local taxation, regression discontinuity and spatial differencing

Gilles Duranton  
Department of Economics  
University of Toronto  
150 Saint George Street  
Toronto ON  
CANADA M5S 3G7  
Tel: (1 416) 978 2678  
Fax: (1 416) 978 6713  
Email: gilles.duranton@utoronto.ca

Laurent Gobillon  
National Institute of Demographic  
(INED)  
133 Davout Boulevard  
75980 Paris Cedex 20  
FRANCE  
Tel: (33 1) 56 06 20 16  
Fax: (33 1) 56 06 21 99  
Email: Laurent.gobillon@ined.fr

For further Discussion Papers by this author see:  
[www.cepr.org/pubs/new-dps/dplist.asp?authorid=119196](http://www.cepr.org/pubs/new-dps/dplist.asp?authorid=119196)

For further Discussion Papers by this author see:  
[www.cepr.org/pubs/new-dps/dplist.asp?authorid=157695](http://www.cepr.org/pubs/new-dps/dplist.asp?authorid=157695)

Henry G. Overman  
Department of Geography and  
Environment  
London School of Economics  
Houghton Street  
London  
WC2A 2AE  
Tel: (44 20) 7955 6581  
Fax: (44 20) 7955 7412  
Email: h.g.overman@lse.ac.uk

For further Discussion Papers by this author see:  
[www.cepr.org/pubs/new-dps/dplist.asp?authorid=143443](http://www.cepr.org/pubs/new-dps/dplist.asp?authorid=143443)

\* Thanks to Tom Davidoff, Tom Holmes, Thomas Klier, Thierry Mayer, John Quigley, Helen Simpson and to seminar and conferences participants in Berkeley, Boston, Ghent, Glasgow, Hamburg, Kansas City, Kiel, Kyoto, London, Paris, Seattle, Vancouver, and Venice. We also owe special gratitude to Tim Besley. Financial support from the Economic and Social Research Council (ESRC Grant R000239878) and the Leverhulme Trust is gratefully acknowledged. This paper uses data that are Crown Copyright and reproduced with the permission of the HMSO Controller and the Queen's Printer for Scotland. However, the use of these data does not imply ONS endorsement of either the interpretation or the analysis in this paper.

Submitted 16 August 2006

## 1. Introduction

This paper develops new empirical methodologies to provide evidence on the effects of local taxation on the location and growth of firms. This issue has been the focus of an extensive theoretical literature and our paper is clearly not the first paper to consider these issues empirically.<sup>1</sup> Bartik (1991) summarises results from the earlier literature. Evidence from the 1960s and 1970s suggested there was no effect of taxes on firm location decisions. Bartik's own work focusing on the 1980s suggested that there was a negative relationship and a number of subsequent papers have confirmed that general finding. Much of this work used fairly large spatial units (mostly US states). More recent work (e.g., Guimaraes, Figueiredo, and Woodward, 2004) has started to move towards smaller spatial units such as US counties with similar results.

The existing literature, however, has failed to satisfactorily resolve three main problems when attempting to assess this impact. First, firms are faced with a choice of a large number of heterogenous locations when deciding where and how to produce. Many of these site characteristics are unobserved and are likely to be correlated with other explanatory variables such as plant characteristics and local taxation, thus biasing the results. Second, firms themselves are also heterogenous. Again much of this heterogeneity is unobservable so that the sorting of firms according to these characteristics provides another possible source of bias. Third, some aspects of the tax system may be endogenous to firm production and location decisions, which may lead to a reverse causality bias.

This paper attempts to deal with all three of these problems. The crux of our approach is to use spatial differencing between neighbouring firms across borders to control for the first source of bias. Incorporating standard panel and instrumental variable techniques

---

<sup>1</sup>One can identify at least three strands of theoretical literature. The issue of tax competition has received considerable attention with the race to the bottom argument prominent in many discussions. See Wilson (1999) for a review. A similarly lengthy debate has centred around the issue of local public good provision and the Tiebout (1956) hypothesis that inter-jurisdictional competition helps achieve the efficient provision of local public goods. See Epple and Nechyba (2004) for a recent review. Finally, capitalisation of local taxes and the possibility of the efficient taxation of land are key concerns for urban economists. This has been a subject of discussion since George (1884). See Fujita (1989) for a thorough presentation of the arguments and Arnott (2004) for a recent discussion of its applicability.

then allows us to deal with the other two possible biases. When applying the resulting methodology to data for English manufacturing establishments we find that local taxation has a negative impact on employment growth, but no effect on entry. We also show that standard methodologies that do not address these three key problems give substantively different results.

The rest of the paper is structured as follows. In section 2 we outline our methodology and relate it to the existing literature. Section 3 outlines our data. Section 4 presents our findings on the impact of local taxation on employment while section 5 presents results on entries. Section 6 concludes.

## **2. Methodology**

### *A. Heterogenous locations*

When locating, establishments are faced with a large number of heterogenous sites.<sup>2</sup> These sites come at very different rental prices. For instance, Thompson and Tsolacos (2001) document a sixfold difference in the rental price per square metre between industrial sites close to Heathrow airport and those in the suburbs of Leeds. Once located, many variables may then impact on production decisions. Traditional empirical approaches have worried about the fact that the costs of labour and other factors of production differ across geographical regions. We will address this issue, but are also concerned with heterogeneity at a much finer geographical scale. A wide variety of factors will affect both the attractiveness of particular sites and the success of establishments once they have chosen their site. For example, the attractiveness of a site may depend on access to the road network while changes to that network may affect the performance of establishments at that site. Hence, we expect considerable site heterogeneity, possibly at a very fine spatial scale, and this heterogeneity may well vary over time.

---

<sup>2</sup>We use the term establishments to refer to our basic units of observation. In the data that we use, these are sometimes part of larger multi-establishment firms. We ignore the complications this introduces.

If this is the case, assessing the impact of local taxation will require us to control for both fixed and time-varying site characteristics since they are likely to be correlated with local taxes or other establishment characteristics. Some of these site characteristics may be observable and thus can be controlled for directly. However many are likely to be unobservable. Because we expect unobserved site characteristics to vary smoothly across space, neighbouring sites are likely to have similar values of these unobserved characteristics. Within jurisdictions, neighbouring sites also have similar tax rates. This raises the possibility that unobserved site characteristics may be correlated with the very variable in which we are interested. Unobserved site characteristics are also potentially correlated with establishment characteristics if, for instance, better establishments tend to locate on better sites. Thus, for example, if more productive owners help to ensure establishment survival and also choose better sites then age may be correlated with unobserved site characteristics. More complex interactions are also possible if establishments sort according to both site characteristics and taxation. For these reasons, dealing with these unobservable site characteristics is the first key step in developing a methodology to assess the impact of local taxation.

A typical solution to this problem would be to assume that the unobservable effects are common across sites *within* jurisdictions and to include dummy variables to condition out these effects. Using longitudinal data, these jurisdictional dummies should pick up any unobserved, time-invariant fixed effects common to all establishments in the jurisdiction.<sup>3</sup> For unobserved site characteristics that vary strongly across jurisdictions (as a result, for instance, of very different institutional contexts), the assumption that unobserved effects are constant within jurisdictions may be a natural one. Our approach will control for such differences. However, to assess the impact of *local* taxation within a given institutional context, this is unlikely to be sufficient since many, if not most, unobserved site characteristics are likely to vary within jurisdictions and across time.

---

<sup>3</sup>In the location literature this approach has proved difficult to implement in situations with a large number of jurisdictions, because non-linear models (e.g., conditional logit) do not allow one to difference out the large number of dummy variables before estimation. See section 5 for further discussion.

The novelty in our approach is to exploit the fact that, although unobserved site characteristics vary across space, they are likely to be highly spatially correlated. Using ideas from Holmes (1998) we identify establishments that are close to one another, but on different sides of jurisdictional boundaries.<sup>4</sup> Because of the spatial correlation in site characteristics, these establishments will have very similar unobserved site characteristics, but will face different tax rates. We look at the extent to which differences in employment between these pairs of establishments can be explained by different local tax rates. This spatial differencing allows us to control for unobserved site characteristics, be they fixed or time-varying. Hence, compared to standard approaches, ours has the added advantage that it can control for variations in site characteristics that occur within jurisdictions as well as the fact that these unobserved effects may vary over time.

### ***B. Heterogenous establishments***

We also expect production establishments to be very heterogenous and much of this heterogeneity to be unobservable. This is a second likely source of bias if establishments with different unobserved characteristics sort across jurisdictions with different tax rates.

Our approach to dealing with the existence of heterogenous establishments is more standard. Because we have a panel of establishment level data, we are able to control for observed time-varying characteristics of establishments, as well as condition out unobserved time-invariant characteristics through the inclusion of establishment fixed effects. Note that the inclusion of establishment fixed effects should also control for unobserved time-invariant site-specific effects (if establishments do not move) leaving our spatial differencing to control for unobserved time-varying site-specific effects. Although there is nothing particularly innovative in the method we use to deal with establishment

---

<sup>4</sup>See also Black (1999). The methodology proposed by Holmes (1998) and Black (1999) has been repeated elsewhere, but most applications only use cross-sectional data and do not address endogeneity issues. The two exceptions are Gibbons and Machin (2003) who consider endogeneity and Kahn (2004) who uses some longitudinal information. As will become clear below, our analysis improves on these existing methodologies by incorporating instrumented panel data techniques into the spatial discontinuity approach.

heterogeneity, as far as we are aware panel data techniques have not been used in the existing literature on the impact of local taxation.<sup>5</sup>

### *C. The endogeneity of taxation*

Another issue that we need to confront is the endogeneity of tax rates to employment and location decisions. To do this, we will argue that the nature of the UK business rates system (the local tax that we consider in this paper) allows us to instrument the levels of local taxation by local political variables. We provide further details below.

### *D. A model of the employment impact of local taxation*

To make this discussion concrete, we consider the impact of local taxation on establishment employment decisions (the main focus of this paper). Because the specific tax that we look at in this paper, the UK business rates, is essentially a property tax that applies to all commercial properties, we model it as changing the rental cost of buildings.

Take location as given.<sup>6</sup> The establishment occupies one of many sites indexed by  $z$ . The jurisdiction that sets the tax for this establishment depends on the site occupied. Jurisdictions are indexed by  $a$ .

Establishments use buildings,  $K$ , and labour,  $L$ , to produce a final good. We take the final good as numeraire and assume that the price is common across all jurisdictions. The wage rate,  $w_z$ , is site-specific. The rental cost of building comprises two elements: the rent,  $\tau_z$ , which is site-specific, and the tax,  $r_a$ , which may vary across jurisdictions. This specification mirrors the working of the UK rates system as described below. Note that this specification is very flexible and allows for establishments to face site-specific factor costs. Ignoring time subscripts for simplicity, the establishment's profit maximisation problem

---

<sup>5</sup>Berman and Bui (2001) use panel data methods to estimate the effects of local environmental regulations (but not taxes) on the labour demand of US oil refineries.

<sup>6</sup>Allowing establishments to choose location involves a simple extension. The problem can be solved backward through a two-stage maximisation. First, calculate maximum profits at each location. Second, maximise with respect to location. See section 5 for further details.

can be written as:

$$\text{Max}_{L_i, K_i} \Pi_i = A_i B_a F(L_i, K_i) - w_z L_i - (1 + r_a) \tau_z K_i \quad (1)$$

where  $F$  is the production function,  $A_i$  is an establishment-specific productivity effect, and  $B_a$  is a jurisdiction productivity effect. Under standard regularity conditions, a sufficient condition for profit maximisation is that the following first-order conditions hold:

$$\frac{\partial \Pi_i}{\partial L_i} = A_i B_a \frac{\partial F(L_i, K_i)}{\partial L_i} - w_z = 0 \quad (2)$$

$$\frac{\partial \Pi_i}{\partial K_i} = A_i B_a \frac{\partial F(L_i, K_i)}{\partial K_i} - (1 + r_a) \tau_z = 0 \quad (3)$$

Again, using standard regularity conditions and focusing only on labour demand we have

$$L_i = g(A_i^+ B_a^+, w_z^-, (1 + r_a)^? \tau_z^?) \quad (4)$$

where  $g$  is some (possibly non-linear) function and the expected signs of the cross-partial derivatives are indicated above each variable.

Several important issues arise. First, an increase in the cost of buildings, either through the rent,  $\tau$ , or through the tax,  $r$ , will have two effects on the demand for labour: a positive substitution effect and a negative output effect. The sign of the sum of these two effects is theoretically ambiguous. However, we expect building and labour to complement each other in most manufacturing industries. Situations where, on average, manufacturing establishments reduce their use of space and increase their employment following a higher cost of building strike us as implausible. Employment growth means that ultimately new employees need to be accommodated and more space is needed. Hence, we expect an increase in the cost of buildings to have a negative effect on employment.

Second, we could extend our framework to incorporate other factors of production. Taxes on buildings could clearly cause establishments to change the amount of these other inputs that they use, which in turn could impact on employment through changes to the marginal product of labour. Because of these omitted inputs, we need to be cautious about the interpretation of our results. Specifically, we are only able to estimate the overall

effect of taxation on employment even though more cross-factor effects may be at work. Note that we could also have the price of the additional inputs vary across locations or have productivity shifters that are site-specific. Provided that these factors are highly spatially correlated our empirical strategy deals with these issues as it controls for site-specific effects.

Third, it may also be the case that the effects of taxation vary across places. For instance, in Baldwin, Forslid, Martin, Ottaviano, and Robert-Nicoud (2004), the effects of local taxation are non-linear and depend on the local market potential. We leave these issues for future work and in this paper only estimate an average effect.

Finally, we also need to keep in mind that establishments are potentially mobile. Hence an increase in local taxation,  $r_a$  may imply a decrease in rents,  $\tau_z$  (for  $z \in a$ ). Put differently, an increase in local taxation may be capitalised into the rent. Full capitalisation occurs only under restrictive circumstances (Fujita, 1989). For the tax we consider, the first major obstacle to full capitalisation is that this is a tax on building (i.e., a property tax) and not only on land. Improvements (e.g., building expansions) were taxed whereas full capitalisation requires that they should not be. Further possible obstacles to full capitalisation are discussed in section 3. In relation to this, note that we can only assess the overall effect of a change in taxation on labour demand without being able to break it down in to a direct effect (an increase in tax keeping rents constant) and an indirect capitalisation effect.

Overall we expect a negative effect of taxation on employment, which may be quite small since taxes could be largely capitalised into rents.

### ***E. Econometric specification***

Denote by  $e_{it}$  the log employment of establishment  $i$  at time  $t$ , the empirical counterpart of  $L_i$  in equations (1)-(4). As can be seen from equation (4), employment will depend on the establishment's productivity effect, the wage rate and the rental cost of the building. To

estimate this labour demand equation, we assume a log-linear relationship and specify:<sup>7</sup>

$$e_{it} = \alpha r_{at} + X_{it}\beta + \mu_i + \gamma_a + \theta_{zt} + \epsilon_{it} \quad (5)$$

The main parameter of interest,  $\alpha$ , captures the (net) effect on employment of the (log) local tax,  $r_a$ . Then,  $\beta$  is a vector of parameters that captures the effect of time-varying establishment-specific observable variables,  $X_{it}$ . The establishment also experiences an establishment-specific productivity effect that is captured by a time-invariant establishment fixed effect  $\mu_i$ . This corresponds to the time-invariant component of the establishment effect  $A_i$  in equation (4).

Equation (5) also contains two local effects,  $\gamma_a$  and  $\theta_{zt}$ . The first,  $\gamma_a$ , is the effect of being located in jurisdiction  $a$ .<sup>8</sup> It is thus the direct empirical counterpart to the productivity term  $B_a$  in equation (4). This term will, for instance, capture the fact that establishments in (or around ) Cambridge may enjoy a productivity boost caused by the presence of a major research university nearby. The second local effect,  $\theta_{zt}$ , is both time-varying and specific to the site  $z$  that the establishment occupies. This empirical term is extremely important since it captures the effects of all the site-specific terms in the labour demand function (4): the rent,  $\tau_z$ , the site-specific wage,  $w_z$ , and any site-specific productivity effect omitted from (4).

Note that estimating (5) and ignoring the unobservable effects will give inconsistent estimators of the effect of both  $\alpha$  and  $\beta$ . To get consistent estimators we need first to eliminate the unobservable time-invariant establishment and jurisdictional effects and then to control for the local time-varying effects. To control for establishment and jurisdictional fixed effects we can use the panel dimension of our data to calculate the *within* estimator (alternatively, we could calculate the first-difference estimator). The *within* transformation is obtained, as usual, by centring all observations around their mean. For any variable  $y$ ,

---

<sup>7</sup>Equivalently, take the log-linearisation of the non-linear relationship around the equilibrium point.

<sup>8</sup>Empirically, we cannot separately identify the time-invariant establishment-specific and jurisdiction-specific effects because we condition both out using establishment fixed effects. Separately identifying these effects would be possible if we assumed that establishment-specific effects were independent of location and we observed sufficiently many establishments moving jurisdictions over time. As this will not affect our results on taxation, we do not consider this identification issue further.

for observation  $i$ , let  $\bar{y}_i$  denote the time average and define  $\tilde{y}_{it} \equiv y_{it} - \bar{y}_i$ . We can then rewrite equation (5) as:

$$\tilde{e}_{it} = \alpha \tilde{r}_{at} + \tilde{X}_{it}\beta + \tilde{\theta}_{zt} + \tilde{\epsilon}_{it} \quad (6)$$

Note that we still cannot get consistent estimators of  $\alpha$  and  $\beta$  from this equation if changes in the unobserved site specific variables  $\tilde{\theta}_{zt}$  are correlated with any of the other explanatory variables (e.g., the local tax rate). To solve this problem, we proceed as follows. Define  $\Delta_d$  as the spatial difference operator which takes the difference between each establishment and any other establishment located at distance  $d$  from that establishment.<sup>9</sup> Applying this spatial difference operator to (6) gives:

$$\Delta_d \tilde{e}_{it} = \alpha \Delta_d \tilde{r}_{at} + \Delta_d \tilde{X}_{it}\beta + \Delta_d \tilde{\theta}_{zt} + \Delta_d \tilde{\epsilon}_{it} \quad (7)$$

Now, we impose the crucial identifying assumption that site-specific effects change smoothly across space. That is, for  $d$  sufficiently small  $\Delta_d \tilde{\theta}_{zt} \approx 0$ . Noting also that taxes will be the same for establishments within the same jurisdiction, this gives us:

$$\Delta_d \tilde{e}_{it} = \Delta_d \tilde{X}_{it}\beta + \Delta_d \tilde{\epsilon}_{it} \quad (8)$$

for establishments in the same jurisdiction and:

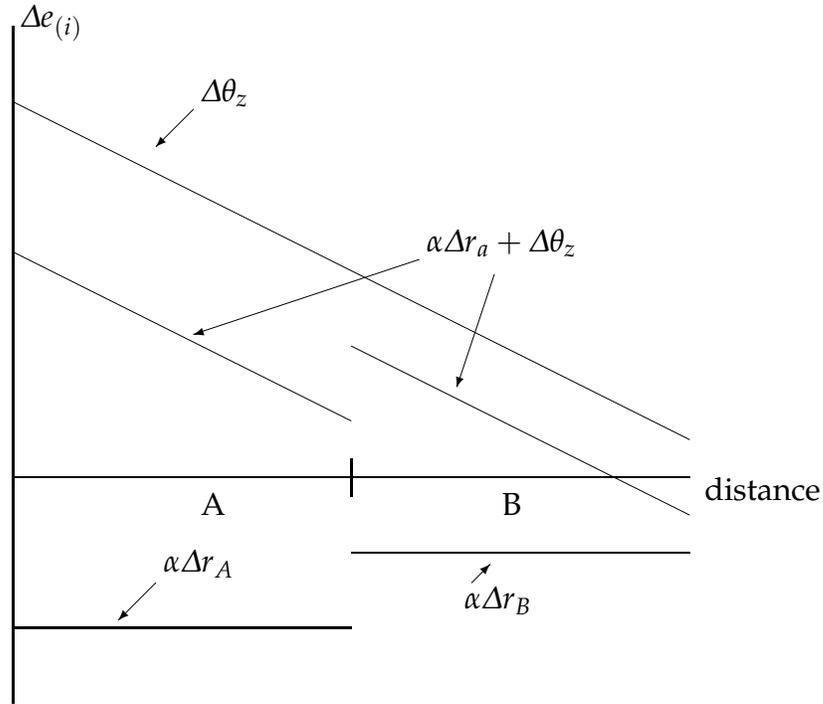
$$\Delta_d \tilde{e}_{it} = \alpha \Delta_d \tilde{r}_{at} + \beta \Delta_d \tilde{X}_{it} + \Delta_d \tilde{\epsilon}_{it} \quad (9)$$

for establishments across jurisdictional boundaries. This shows that we can use neighbouring establishments located across jurisdictional boundaries to identify the effects of local taxation. We can also use neighbouring establishments within the same jurisdiction to improve our estimates of the effect of establishment-specific variables.

The easiest way to understand how the methodology works is to consider two neighbouring jurisdictions (A, B) where all sites are located on a straight line running through the two jurisdictions. With distance from the left hand end of jurisdiction A represented on the horizontal axis, figure 1 shows the change in the site specific effect at each location. Assume that the increase in the site-specific effect is highest for sites at the left hand end

---

<sup>9</sup>Assume for the moment, that there is only one such establishment.



**Figure 1.** Boundary discontinuities

of jurisdiction A and lowest at the right hand end of B. The line  $\Delta\theta_z$  shows the effect on establishment employment at each location holding everything else equal. Everything else is not equal, however. Assume that jurisdictions base their tax increases on the *average* change in site characteristics in their jurisdiction. Jurisdiction A sees larger increases in taxes than jurisdiction B. The impact on employment (again *ceteris paribus*) is shown as  $\alpha\Delta r_a$ . The overall employment effect is then represented by  $\Delta\theta_z + \alpha\Delta r_a$ . Notice, that we have assumed that the impact of site-specific effects tends to outweigh the increase in taxation, so that *on average* employment growth in A is higher than in B. Thus, even after conditioning out establishment-specific effects, it appears that employment is positively related to taxes. Put differently, to estimate the effect of local taxation on employment one needs to condition out site-specific effects properly.

To do this, consider two sites on the border between A and B. By assumption, the establishments experience approximately the same site-specific effect. All else equal, their employment would grow by roughly the same amount. However, the establishment in

A sees a higher tax increase and this offsets the impact of the improved site. The overall effect on employment is given by  $\alpha\Delta r_a + \Delta\theta_z$  which is larger for the establishment in B than in A. The difference between the two is  $\alpha(\Delta r_a - \Delta r_b)$ . Since our procedure compares the growth in employment in the establishment on the boundary in A to the growth in employment in the establishment on the boundary in B, we now correctly identify the negative relationship between increased taxation and employment. Of course, this example is just one possibility, but it does serve to emphasise exactly how our methodology allows us to correctly identify the impact of local taxes on employment.

Finally, before turning to the implementation, note that space differencing and the *within* transformation have implications for the error structure. As usual, directly implementing (9) would yield consistent estimates for the coefficients but does not give the correct standard errors. Appendix A shows how to correct the standard errors when estimating (9).

To summarise, we use the panel dimension of our data to remove establishment and jurisdiction fixed effects. We then identify establishments with similar site-specific effects but that face different tax rates. These are establishments that are located close to one another but in different jurisdictions. We use these establishments to identify the impact of local taxation on employment. Later in the paper, we use similar ideas to consider the impact of local taxation on the entry of new establishments. As there are some differences in the implementation, we leave the details to section 5 and for now turn to outlining our data and the results for employment.

### 3. Data

To implement our methodology, our data needs to satisfy a number of requirements. First, we need to have a panel of individual establishment level data. Cross-sectional data do not allow us to use the *within* transformation to remove establishment and jurisdiction-specific effects. We then need to be able to precisely locate these establishments so that we can identify which pairs of establishments are neighbours. Finally, we need to identify

a local tax which is time-varying. We would prefer this local tax to be economically significant to increase the chances of detecting any impact on location and employment decisions. Data satisfying all of these requirements is available for England for the six year period from 1984 to 1989. We first describe the establishment level data set we use before turning to details of the particular local tax that we consider.

### *A. Establishment data*

Our empirical analysis uses establishment level data for years 1984 to 1989 from the Annual Respondent Database (ARD) which is the data underlying the Annual Census of Production in the UK. Collected by the Office for National Statistics (ONS), the ARD is an extremely rich data set which contains information about all UK establishments from 1973 onwards. We face two restrictions on time period. Changes to the tax system restrict our focus to years before 1990, while changes to the ARD restrict us to years after 1983. We restrict ourselves to English manufacturing establishments.<sup>10</sup> For every establishment, we know postcode, five-digit industrial classification, and number of employees. In our analysis of employment we use only the information coming from a sub-sample of ‘selected’ establishments. These are the establishments that are requested to make a detailed return in any year. These establishments are generally larger and their employment information is of better quality than for non selected establishments.<sup>11</sup> The precise sampling frame for selected establishments and further description of this data can be found in Griffith (1999). In our analysis of entry we use the exhaustive data (that includes ‘selected’ and ‘non-selected’ establishments).

The postcode reported in the ARD is very useful for locating establishments. In the UK, postcodes typically refer to one property or a very small group of dwellings. The CODE-POINT data set from the Ordnance Survey (OS) gives spatial coordinates (Eastings

---

<sup>10</sup>We ignore Scotland because it operated a different local tax system, Wales because it is not covered by the data set that provides our instruments, and Northern Ireland because special permission is required to access ARD data for establishments located there.

<sup>11</sup>The better quality of data is not the only reason to focus on selected establishments. Implementation of the algorithm to identify neighbours and calculate data for pairs of neighbouring establishments is infeasible with the large samples that include non-selected establishments.

and Northings) for all UK postcodes.<sup>12</sup> By merging this data with the ARD we can generate very detailed information about the location of all English manufacturing establishments. For all but a tiny percentage of matched establishments the OS acknowledges a potential location error below 100 metres. For the remaining observations, the maximum error is a few kilometres. Overall, we expect a very high level of precision for our location data (see Duranton and Overman, 2005, for further discussion).

As the data requirements for our spatial differencing methodology are already somewhat restrictive, we only consider establishment specific variables that can be calculated for all establishments. For the ARD, this means only controlling for establishment age. For establishments already in the panel in 1976 (the earliest year for which we have information), we are unable to assess their exact age. With our regressions running over 1984-1989, this truncation of the data implies that age is censored for establishments above 8 years in 1984. For consistency across years, we thus construct three age variables. One is a dummy for establishments that are 8 years or older. The other two are age and age squared defined only for establishments younger than 8 years old. The dummy captures the effect of being 'old' although it does not allow us to assess differences between old establishments and very old establishments. Age and its square for establishments less than 8 years old capture a non-linear effect of age for new establishments.

There are 25,579 establishments in the ARD that are located in England and that report employment at least once within our study period. We make several sample restrictions. We deleted 122 establishments where the Local Authority code was missing since we need this to append the tax rate and political variables. We then dropped 374 establishments where we could not identify coordinates based on postcode as we cannot identify

---

<sup>12</sup>Eastings and Northings give grid references taking as the origin a point located South West of the UK.

neighbours for these establishments.<sup>13</sup> A further 723 establishments change sector during the period and 2,547 move site (i.e., change coordinates). In both cases these changes frequently reflect coding errors rather than genuine changes in activity or location, so we drop these establishments. This leaves us with a sample of 21,813 establishments and a total of 61,785 observations.

## **B. Local taxation**

The local tax that we consider is a property tax on non-residential property known as the UK business rate, set and collected by jurisdictions called Local Authorities (hereafter LAs). Before 1990, tax rates varied over time and jurisdictions. We are fortunate because this tax consisted only of a flat rate that applied to all non-residential properties. This simplicity implies that the only source of variation is the tax rate itself.

The UK business rates were subject to a major reform in 1990 which essentially eliminated jurisdictional variation. This reform provides a large amount of exogenous variation in the tax rate. Unfortunately we cannot exploit it because all properties were also revalued in 1990. Since we do not know by how much each property was revalued, we cannot compute the change in taxation faced by each establishment. Hence, as mentioned above, while the production data restricts our study period to start in 1984, the reforms to the tax system mean it must end in 1989.

UK business rates are levied on the *occupiers* (not owners) of non-domestic properties according to the rental value of the property. These taxes represented a considerable tax on business. To give some idea of the economic significance of the tax, note that in 1990 the average annual rates bill per square foot was over £4, compared with an average rent bill per square foot of just over £13 (Bond, Denny, Hall, and McCluskey, 1996). Hence

---

<sup>13</sup>We directly identified Eastings and Northings for around 90% of establishments. The main problem for the remaining 10% results from the creation of new postcodes not reported in our version of CODE-POINT. To increase matching rates, we checked our data against a data set of postcode updates. For observations with postcodes still unmatched, we imputed co-ordinates as follows. In the UK, a postcode is either six or seven digits. We first drop the last digit and assign establishments the mean coordinates of all postcodes sharing the same truncated postcode. If this failed to produce a match, we dropped two, then three digits from the postcode and again matched on mean coordinates where possible. This left us with a small percentage (1.4%) of establishments that could not be given a grid reference.

rates bills make up roughly 25 per cent of the total occupancy costs of rented commercial property. In 1992 business paid £13bn in local rates. A figure that was almost equal to the £15bn that they paid in corporation tax for the same year.

It is important to note that, as a first approximation, none of this money was used to finance local services for businesses. The responsibilities of LAs include social and health services for their residents, community safety, some aspects of education, housing, arts, culture, and environment. Nothing is directly provided to businesses.<sup>14</sup> Even basic services such as refuse collection need to be separately organised by businesses.<sup>15</sup> The only area of responsibility of LAs that can affect businesses directly is planning (and some trading regulations that will affect mostly retail). Despite a mostly national set of planning regulations, LAs may differ in their speed of processing applications. It is hard to believe, however, that there were substantial within jurisdictional changes in the efficiency of local planning offices over the time period we consider. Instead, differences with respect to planning efficiency and restrictiveness, if they matter at all, are expected to be part of the fixed effect of these jurisdictions.

During our period of study, tax rates were set locally by the 366 LAs in England, which entirely covered its land area. Large cities comprise several such LAs (e.g., 33 for London and 10 for Manchester). The tax paid depended on the value of the buildings occupied. The tax rates were known as 'rate poundages' and the value of the buildings as the 'rateable value'. The tax rate was changed yearly. Rateable values of buildings were fixed in 1973 and did not change until 1990. The average rate for 1987 was 229.5 pence per pound of 1970 rateable value with a standard deviation of 32.2 pence. The ratio between the highest business rate (Oldham in Manchester) and the lowest (Kensington and Chelsea in London) was about three. These two LAs are of course far from each other and this

---

<sup>14</sup>Clearly, there may be some indirect effects through, for instance, crime or education. However we expect these effects to spill over across jurisdictions so that establishments on both sides of any border face similar conditions. This argument is reinforced by the fact that LAs had only very limited autonomy on crime, because police districts usually comprise several LAs.

<sup>15</sup>One of the key reasons why the Central government decided to take this tax setting power away from LAs in 1990 was that it acknowledged it was a taxation without representation nor services. Another reason for this reform is that large increases in business rates suggested that UK LAs were engaged in something very much like a 'race to the top' rather than the often feared 'race to the bottom'.

difference will not be taken into account in our estimation when we spatial difference. The largest ratio between two *neighbouring* LAs is nearly two. There was also significant time-series variation. There was for instance an average increase in the business rates of nearly 10 per cent between 1988 and 1989. Over the entire period (1984 – 89) the average increase was 45 per cent with a standard deviation of 0.15. Again looking at neighbouring LAs, we observe a 25 per cent decrease in Kensington and Chelsea against a 31 percent increase in neighbouring Hammersmith over the period. We think this is more than sufficient variation to perform our estimations.

#### 4. Results

We use log employment ( $e_{it}$ ) as the dependent variable. As explanatory variables we include log tax rate ( $r_{at}$ ), the three variables described above to capture the impact of age ( $X_{it}$ ) and a full set of industry-year dummies (two-digit industries) for which we do not report coefficients. To show the difference that our estimation strategy makes, we also estimate the effects of local taxation using standard techniques. Results for the non-spatial specifications (i.e., using standard techniques and ignoring the micro-spatial nature of the data) are reported in table 1. Table 2 reports results for comparable specifications that use spatial differencing. For the sake of comparison, we have restricted the sample to be the same for all specifications by only including those establishments that we use to estimate our preferred specification: the instrumented, fixed effect spatial-differencing specification discussed below (the final specification in table 2). This restriction requires establishments to be in a pair in which both establishments simultaneously report employment in at least two years.<sup>16</sup>

---

<sup>16</sup>Consider an establishment that reports employment for (say) only 1987 and 1988. It is in the sample if it has at least one neighbour and that neighbour also reports employment for these two years. It is not in the sample if it has no neighbours or if it has neighbours but they do not also report employment in 1987 and 1988. We impose this restriction because our implementation of spatial-differencing uses fixed effects for pairs, rather than estimating the fixed effect for establishments and then spatial differencing. Furthermore, an establishment can be part of more than one pair satisfying this condition. However, we only use each establishment once in estimating the non-spatial specifications.

	OLS	WITHIN	WITHIN IV
(log) tax rate	0.210 <sup>a</sup> (0.060)	0.071 <sup>b</sup> (0.030)	0.363 <sup>a</sup> (0.104)
age censored dummy	0.569 <sup>a</sup> (0.091)	0.214 <sup>a</sup> (0.033)	0.223 <sup>a</sup> (0.034)
age	0.023 (0.047)	0.036 <sup>a</sup> (0.011)	0.036 <sup>a</sup> (0.011)
age squared	0.005 (0.005)	-0.001 (0.001)	-0.001 (0.001)
Adjusted R-squared	0.12		
Number of observations	13490	13490	13490
Number of establishments	4414	4414	4414

Notes: Regression of (log) employment on (log) local tax rates and age variables. First column reports results from OLS, second column (WITHIN) allows for establishment specific fixed effects, third column (WITHIN IV) further instruments local tax using local political variables. Standard errors under coefficients. <sup>a</sup>, <sup>b</sup> and <sup>c</sup> denote significance at the 1%, 5% and 10% level respectively.

**Table 1.** Non-spatial regression results

Starting with 21,813 establishments, 7,938 of them are dropped because they only report employment once, independently of whether they are in a pair. Of the remaining 13,875 establishments, 8,792 have a neighbour within 1 kilometre (the distance threshold that we use for our reported results). However, only 4,414 of these establishments are in at least one pair where both establishments simultaneously report employment in at least two years. These establishments form the restricted sample for which we report results in the text. To show that this sample is representative, Appendix B reports results for the same specifications but without imposing this sample restriction.

The first regression in table 1 reports the results from estimating equation (5) using pooled OLS for all six years 1984-1989. The results on the age variables show that, as expected, older establishments have higher employment. Our main focus, however, is on the role of taxation and from our results we see that in the cross-section, higher tax rates are associated with higher employment.

As we noted above in our discussion of equation (5), one possible explanation of this positive correlation is that some establishments are larger than others for unobserved reasons and larger establishments happen to be located in higher tax jurisdictions. The

second specification in table 1 allows for this possibility by introducing an establishment-specific fixed effect and calculating the *within* estimator. From the table we can see that the coefficient on tax rate is divided by three suggesting that much of the positive correlation between employment and local tax rate is indeed due to unobserved characteristics of establishments. Note, however, that the effect remains positive and significant.

The remaining problem that we need to tackle is that the tax rate may be correlated with the error in equation (6). There are two possible sources for this correlation. First, there may be a feedback from local employment to local tax rate. This feedback will be positive when local politicians feel able to tax local business more when it is doing well. Alternatively, and working in the opposite direction, it could be that LAs can afford to keep taxes low during 'good times'. In the UK context, this alternative may arise because some redistribution takes place locally so that good times may imply a lesser need for social expenditure. This being said, we do expect the first effect to dominate and taxation to go up when local business is doing well. Second, there may be other time-varying characteristics that are positively correlated with both employment and tax rates and that we do not control for through the use of establishment-specific fixed effects. Recall that the establishment fixed effects only condition out time-invariant unobserved heterogeneity.

To solve these problems we take an instrumental variable approach. More specifically, we use local political variables to instrument for local tax rates.<sup>17</sup> The full set of instruments includes the share of local politicians affiliated with the Conservative, Labour and Liberal Democrat parties, a set of dummies indicating whether the LA is controlled by the Conservative, Labour or Liberals and a set of interactions giving the share of the Conservatives, Labour and Liberals if they control the LA.<sup>18</sup> The R-squared of the first stage regression is 69%. For the interested reader Appendix C provides further details on

---

<sup>17</sup>The data is from the British Local Election Database which is available through the UK Data Archive.

<sup>18</sup>These are the three main political parties. There are many smaller parties that play a role in local politics and we aggregate these in to 'other' and treat them as the omitted category.

the first stage regression.<sup>19</sup>

The third column in table 1 shows what happens when we use these instruments for the level of taxation in the fixed effects specification that we reported in column 2. Surprisingly, perhaps, the coefficient on local taxation actually increases after instrumenting. It would be tempting to conclude that, contrary to most priors, UK LAs taxed businesses according to jurisdictions (social) needs rather than local business' capacity to pay. That is, there is a negative correlation between local tax rates and the omitted variables so that instrumenting leads to a higher coefficient on local taxation. Nonetheless, a positive effect of local taxation on employment with an elasticity close to 0.4 strikes us as implausible. As we will see below it appears that this increase in the coefficient on local taxation following instrumentation occurs because the unobserved time-varying site-specific effect is correlated with both employment and local political variables.<sup>20</sup>

In table 2 we present two sets of results (with and without corrected standard errors) for three different spatial specifications that parallel those presented in table 1. In the first column, we spatial difference equation (5) and then estimate it using OLS. In the second, we spatial difference (6) and then estimate it using a fixed effect for each pair of establishments.<sup>21</sup> Finally in the third column, we instrument the tax rate in the spatially differenced *within* specification using spatially differenced political variables as in the non-spatial specification.

The results that we report use a distance threshold of 1 kilometre to identify neighbours. In our choice of threshold, we face a tradeoff between sample size and the extent to which the spatially varying site-specific effects are equal across neighbouring sites. We chose the minimum threshold which gives us sufficient observations to identify the effect of local

---

<sup>19</sup>One might be tempted to test for the validity of instruments given that we have more instruments than endogenous regressors. We would argue, however, that such a test is invalid because all our instruments are based on the same underlying assumption (that local politics is independent of changes to establishments' employment).

<sup>20</sup>One possible solution would be to use lagged values of the political variables as instruments. However, this would mean dropping years from a short panel and would only solve the problem if there is no correlation in the unobserved factors and the political variables over time. Instead, we use our spatial differencing strategy to control for these unobserved time-varying effects.

<sup>21</sup>Our pairing algorithm only matches establishments if they are in the same 2 digit sic, so we do not need to include industry-year dummies in our spatially differenced specifications.

taxes (remembering that the identification of the tax effect comes from cross jurisdiction border pairs). We use neighbouring establishments within the same jurisdiction to improve our estimates of the effect of establishment-specific variables.<sup>22</sup>

Uncorrected standard errors			
spatial difference of	OLS	WITHIN	WITHIN IV
(log) tax rate	0.846 <sup>a</sup> (0.225)	0.111 (0.119)	-1.024 <sup>a</sup> (0.314)
age censored dummy	0.738 <sup>a</sup> (0.076)	0.134 <sup>a</sup> (0.028)	0.132 <sup>a</sup> (0.028)
age	0.068 <sup>c</sup> (0.039)	0.042 <sup>a</sup> (0.009)	0.041 <sup>a</sup> (0.009)
age squared	-0.003 (0.004)	-0.003 <sup>a</sup> (0.001)	-0.003 <sup>a</sup> (0.001)
Corrected standard errors			
spatial difference of	OLS	WITHIN	WITHIN IV
(log) tax rate	0.846 <sup>b</sup> (0.379)	0.111 (0.167)	-1.024 <sup>b</sup> (0.420)
age censored dummy	0.738 <sup>a</sup> (0.138)	0.134 <sup>a</sup> (0.051)	0.132 <sup>b</sup> (0.049)
age	0.068 (0.071)	0.042 <sup>a</sup> (0.016)	0.041 <sup>a</sup> (0.015)
age squared	-0.003 (0.008)	-0.003 <sup>c</sup> (0.002)	-0.003 <sup>a</sup> (0.002)
Adjusted R-squared	0.04		
Number of observations	18370	18370	18370
Number of establishments	4414	4414	4414
Number of pairs	6087	6087	6087

Notes: Regression of spatial difference of (log) employment on spatial difference of (log) local tax rates and age variables. First column reports results from OLS, second column (WITHIN) allows for establishment specific fixed effects, third column (WITHIN IV) further instruments local tax using local political variables. Standard errors under coefficients. First block of results report uncorrected standard errors. Second block of results report standard errors corrected according to Appendix A. <sup>a</sup>, <sup>b</sup> and <sup>c</sup> denote significance at the 1%, 5% and 10% level respectively.

**Table 2.** Spatial differencing regression results

Note that, although the overall sample of establishments is restricted to be identical for both the spatial and non-spatial specifications, the number of observations is higher for the spatially differenced specifications (18,370 compared to 13,490). This is because each

<sup>22</sup>We get the same results if we restrict attention only to the 164 establishments that are part of cross-jurisdiction border pairs. The spatially differenced within IV specification gives a coefficient of  $-1.072$  as opposed to the  $-1.024$  reported in the text. Both coefficients are significant at the 1% level.

establishment can be involved in more than one pair. Specifically, we have 6,087 unique pairs as compared to 4,414 unique establishments suggesting that each establishment has on average three neighbours.<sup>23</sup>

Before turning to the individual coefficients, comparing the two blocks of results (with and without corrected standard errors), we see the corrections outlined in Appendix A generally increases estimates of standard errors by around 50%. In our particular context, this results in some minor changes in significance levels, but does not change our overall findings. In other contexts it clearly could, suggesting that the correction outlined in the appendix should usually be implemented. Turning to the coefficients, we see that, apart from changes in significance levels, the results on the age variables are essentially unchanged. As before, we focus on the impact of the local tax rate and note that after spatial differencing we get a higher correlation between (spatially differenced) local tax rate and (spatially differenced) employment than previously, with a coefficient of 0.846 compared to 0.210 with OLS. These coefficients have the same probability limit when there are no unobserved establishment or site-specific effects, or when these effects are uncorrelated with tax rates or other included explanatory variables. A possible reason for the higher correlation between employment and taxation after controlling out site-specific effects is that areas with poor sites had higher local tax rates thus biasing downward the coefficient on local taxes in the non-spatial OLS estimation which ignores these local site-specific effects. A correlation between having 'poor sites' (following bad local economic conditions during the period) and higher taxes is certainly believable given the political economy of UK LAs: de-industrialising LAs tended to vote for very left-wing councils that then greatly increased local taxation.<sup>24</sup>

However, this interpretation is based on the assumption that establishment fixed effects are either absent or uncorrelated with local tax rates. To account for the possible correl-

---

<sup>23</sup>Working against this, is the fact that both establishments in the pair must simultaneously report employment data in at least two years. For this distance threshold the first effect dominates. That need not be the case for other distance thresholds.

<sup>24</sup>For instance, Sheffield council under the leadership of the (then) leftwing firebrand David Blunkett ended up with the highest business rate in the country in 1990 and Liverpool council led by the equally notorious Derek Hatton ranked 12th.

ation between establishment fixed effects and local tax rates we use, as before, the panel dimension of our data. This controls for unobserved establishment heterogeneity. We do that in the column 2 of table 2 which reports results when we both spatially difference and allow for fixed effects for pairs of establishments. Note that these pair fixed effects not only control for time-invariant unobserved establishment heterogeneity but also for other time-invariant local effects such as the (possible though unlikely) propensity of some jurisdictions to provide better services and thus have consistently better performing establishments.

Comparing results across the non-spatial and spatially differenced specifications we see that after controlling for establishments fixed effects, the coefficient on local tax rate is again higher after we spatial difference. However because of higher standard errors, it is hard to provide a definitive interpretation for this comparison. More significantly, comparing across the spatially differenced specifications, we see that allowing for pair fixed effects reduces our estimate of the positive correlation between local taxation and employment relative to the spatially differenced OLS results. The coefficient even becomes insignificant. This confirms our finding from the non-spatial specifications that establishment fixed effects appear to be positively correlated with tax rates: LAs with 'good' establishments charge higher taxes.

As with the non-spatial fixed effects specifications, we still want to control for the fact that local tax rates may be endogenous. To do this, we instrument using the spatial difference of the same set of political variables that we use for the non-spatial specification. Results are shown in column 3 of table 2. We now see that local taxation has a negative effect on establishment employment. As argued above, the difference between the spatial and non-spatial results suggests that the unobserved time-varying site specific effects are correlated with both employment and local political variables. Spatial differencing removes these site-specific effects, ensuring that our instruments are valid and thus allowing us to identify the negative effect of local taxes on establishment employment.

Pulling the results together, we see that spatial differencing offers two improvements

over existing methodologies. First, comparing the non-instrumented spatial and non-spatial regressions allows us to identify the nature of the relationship between site-specific effects and local taxation. Second, and more importantly, because spatial differencing removes unobserved time-varying site-specific effects it makes it far easier to find a valid set of instruments that allows us to identify the negative relationship between local taxation and employment. Quantitatively, spatial differencing greatly affects the results. In what is arguably the best estimation using standard techniques (again one that does not control for time-varying site-specific effects), we find a positive elasticity around 0.4. Our preferred specification with spatial differencing reverses the sign of the coefficient and leads to an elasticity around  $-1$  (although it is not very precisely estimated). In a nutshell, instead of a positive effect of local taxation on employment, we find, thanks to our spatial differencing approach, a significant negative effect of local taxation on local employment.

The negative effect of local taxation on employment is possibly stronger than one would have thought since this tax is not far from an ideal land tax. There are, however, two differences with a non-distortionary land tax. The first is that building extensions were taxed. This could have had an especially important effect because expansions would typically lead to a complete building re-assessment with the risk of an increase in taxation on the already existing facilities as well as the new ones. The second distortion is more subtle. It finds its root in the reciprocal hold-up problem which occurs because most manufacturing facilities are fairly specific to their tenant and conversely. With a vast majority of manufacturing firms renting their facilities, this hold-up problem is mitigated in UK manufacturing by having long-term rental contracts (typically 20 years) that are renegotiated only every five years or so. These long lags between rent negotiations probably meant that tax increases would not capitalise very fast and thus fall primarily on the (mobile) occupier who had to pay them. We believe that both these distortions help explain our finding of a negative effect of local taxation on employment.

## 5. Entries

### A. Methodology and data

We now turn to the effects of local taxation on entry decisions. This is an important issue for two reasons. First, the rate at which new establishments enter LAs is an important economic outcome that deserves specific attention. Second, spatial differencing provides solutions to the same three problems (establishment heterogeneity, site heterogeneity, and endogeneity) that we addressed when looking at employment growth. However, there are subtle differences which we believe are worth highlighting.

Consider establishment  $i$  that wishes to enter in year  $t$ . It can choose between all available sites, indexed by  $z$ . As before, the jurisdiction that sets the tax for the establishment depends on the site occupied, and is indexed by  $a$ . Profit maximisation can be performed in two stages. First, establishment  $i$  computes the highest profit it can achieve,  $\Pi_{iz}$ , at each site  $z$ . It then selects the site offering the highest profit.<sup>25</sup> We assume that the highest profit for establishment  $i$  entering in year  $t$  at site  $z$  can be written as:

$$\Pi_{izt} = \lambda r_{at} + Z_{it}\zeta + v_i + \kappa_a + \varphi_{zt} + \epsilon_{izt} \quad (10)$$

where  $v_i$  is an establishment fixed effect,  $\kappa_a$  is a jurisdiction fixed effect,  $Z_{it}$  are explanatory variables at the establishment level,  $\varphi_{zt}$  is a site-specific effect, and  $\epsilon_{izt}$  is an establishment site-specific shock.<sup>26</sup> Establishment  $i$  will choose the site  $z$  that gives the highest expected profit. When the shocks  $\epsilon_{iz}$  follow an appropriate *iid* extreme value distribution, the probability of choosing site  $z$ ,  $P_{izt}$ , is logistic and is given by

$$P_{izt} = \frac{\exp E(\Pi_{izt})}{\sum_{z=1}^Z \exp E(\Pi_{izt})} \quad (11)$$

---

<sup>25</sup>We ignore any possible interaction between the location decisions of entrants. Constraints on site choices aside, this seems reasonable in established manufacturing industries where existing establishments drive local wages, determine product market competition etc. As discussed in the text, the fact that the effect on profits of these factors will be highly correlated across neighbouring sites then justifies our approach.

<sup>26</sup>The establishment fixed effect,  $v_i$ , mirrors  $\mu_i$  in (5). Similarly the jurisdiction effect,  $\kappa_a$ , and the site-specific effect,  $\varphi_{zt}$ , are the counterparts of  $\gamma_a$  and  $\theta_{zt}$ . Finally both (10) and (5) contain coefficients for the effect of local taxation ( $\lambda$  and  $\alpha$ ) and establishment-level variables ( $\zeta$  and  $\beta$ ). Note that our approach for entries is also consistent with a more general setting where the specification includes variables that are individual-site specific,  $\phi_{izt}$ , in addition to individual effects,  $v_i$ , and site effects,  $\varphi_{zt}$ .

where  $E(\cdot)$  is the expectation operator and the summation is across all possible sites  $Z$ .

The standard approach to estimating the coefficients on tax rate and establishment specific variables ( $\lambda$  and  $\zeta$ ) is to ignore the site-specific effect,  $\varphi_{zt}$ , and estimate a conditional logit model. To do this, one creates a set of establishment-jurisdiction observations and defines  $c_{ia} = 1$  if establishment  $i$  locates in jurisdiction  $a$  and  $c_{ia} = 0$  otherwise. The coefficients can then be estimated by maximising the log likelihood of the conditional logit model:

$$\log L_{cl} = \sum_{i=1}^I \sum_{a=1}^A c_{ia} \log P_{ia} \quad (12)$$

where for simplicity we have dropped the time subscripts as establishments only enter once.

As is well recognised in the literature, application of the conditional logit model can be problematic when the set of possible jurisdictions is large. One possible solution is then to take a random sub-sample of jurisdictions, although this clearly has implications for the efficiency of the estimator and the small sample properties are unknown. Another possibility, recently proposed by Guimaraes, Figueiredo, and Woodward (2003) is to use the fact that, under certain conditions, the log likelihood of the Poisson model is identical to that of the conditional log likelihood. Estimating a Poisson regression is computationally much easier though the equivalence between the likelihoods only holds in the absence of establishment specific variables (i.e.,  $Z_{it}$ ).<sup>27</sup> In any case both solutions ignore the site-specific effects. As we now show, spatial differencing provides an alternative which controls for site-specific effects and which, in other contexts, would allow for the inclusion of establishment specific variables.

Our approach is as follows. Consider two neighbouring sites close to the border between two jurisdictions  $a_1$  and  $a_2$ . One site, denoted  $z_1$ , is located in jurisdiction  $a_1$ . The other,  $z_2$ , is located in jurisdiction  $a_2$ . Since the two locations are very close, we assume that  $\varphi_{z_1 t} \approx \varphi_{z_2 t}$ . This is the same identification assumption made in section 2 to derive our employment specification (9). To repeat, this assumption is justified by the fact that

---

<sup>27</sup>Guimaraes *et al.* (2003) show that the Poisson regression can also be used to simplify computation in specifications with variables, such as sector dummies, that are constant across groups of firms.

site-specific effects (labour market conditions, access to markets and major facilities, etc) vary smoothly across space. The probability of choosing  $z_1$  conditional on locating in one of these two neighbouring sites is:

$$P(i \in z_1 | i \in \{z_1, z_2\}) = \frac{P_{iz_1}}{P_{iz_1} + P_{iz_2}} \quad (13)$$

Note that unlike equation (11) above, this specification conditions out both establishment- and site-specific effects. This occurs because these effects are the same at both locations  $z_1$  and  $z_2$ .<sup>28</sup> In relation to this, recall that in standard conditional logit models observed site-specific factors are computationally hard to deal with. By contrast our approach directly conditions out both observed and unobserved site-specific factors in a way that is computationally easy to implement. In addition, we do not need to rely on the assumption of the Independence of Irrelevant Alternatives which underlies the conditional logit model. This is a distinct advantage as it is often argued that this assumption is unlikely to hold in spatial settings (see, for example, Head and Mayer, 2004).

When the shocks  $\epsilon_{izt}$  follow an appropriate *iid* extreme value distribution, the probability of choosing one of the sites is logistic and is given by:

$$P(i \in z_1 | i \in \{z_1, z_2\}) = \frac{1}{1 + \exp(\lambda(r_{a_2t} - r_{a_1t}) + \kappa_{a_2} - \kappa_{a_1})} \quad (14)$$

In our particular application, equation (14) only involves jurisdictional level variables so we can estimate  $\lambda$  directly from an aggregate logit model where the observation units are the (border-side, time) pairs. In the estimation, the observations are weighted by the number of entrants for consistency with equation (14).

In a nutshell, we select entrants located close to jurisdictional boundaries and examine their decision to choose to locate on one particular side of the border. The main conceptual difference with our employment regression (9) is that we consider the location decision of a new establishment choosing between neighbouring *jurisdictions* rather than comparing

---

<sup>28</sup>We assumed in (10) that the effect of establishment characteristics did not depend on location. However our spatial differentiation approach is more powerful than this since any interaction between establishment and site characteristics is conditioned out by spatial differencing provided the characteristics of the local environment vary smoothly over space.

the employment outcomes for (existing) neighbouring establishments. For entry, this is the appropriate way to control for both time invariant establishment specific effects and any unobserved site-specific effects common to both sides of a border.<sup>29</sup> As local tax rates do not vary smoothly across space at jurisdictional boundaries, we can use entrants on either side of these boundaries to identify the effect of taxes after conditioning out time-varying site-specific factors and time-invariant establishment-specific effects. Interestingly one step is enough to eliminate both establishment and site effects for entry, whereas two steps are necessary when studying employment. This reflects the fact that instead of comparing each establishment with a matched establishment on the other side of the border we compare each establishment with itself on the two sides of the border making the *within* transformation redundant.

As with the employment specifications, endogeneity must also be addressed since the local rate of taxation may be simultaneously determined with the rate of entry. To control for this, we estimate a two-stage IV *logit* model instrumenting the difference in tax rate,  $r_{a_2t} - r_{a_1t}$ , in the logit specification, with the predicted difference in tax rate from a first stage regression using spatially-differenced local political variables. The simplicity of this method comes at a price because, as is well known, correcting the errors to allow for possible correlation between the residual of the instrumentation equation and that of the entry equation is far from straightforward.<sup>30</sup>

## **B. Results**

To construct the entry data, we need to detect all entrants in the ARD. Because of a change in 1984 in the way the registry of establishments was constructed, there is a large amount of artificial entry in 1984 and 1985 (i.e., establishments enter the data set for the first time

---

<sup>29</sup>To see this, note that when choosing between two neighbouring sites, the entrant compares profits between them. Time-varying site-specific factors, which vary smoothly across space, will affect profits at both sites in the same way. Hence, these factors do not enter into the location decision. A similar argument applies to unobserved time-invariant establishment characteristics.

<sup>30</sup>As an alternative, we used an IV *probit* model that estimates the entry and instrumenting equation simultaneously. This approach provides us with the correct standard errors but at the cost that it is not fully consistent with the theoretical specification. The results are the same as with the two-stage *logit* model.

during those two years even though they already existed prior to 1984). See Griffith (1999) for further discussion. As a result, we ignore entries for these two years and focus instead on all newly reporting establishments between 1986 and 1989. There are 81,042 entrants between 1986 and 1989.

For consistency with the employment regressions we would like to identify all entrants within 1 kilometre either side of jurisdiction boundaries. The easiest way to do this would be to draw 1 kilometre buffers around boundaries. Lacking a set of digital boundaries for this time period, we instead proceed indirectly and identify the set of border entrants from the ARD itself. To do this, for each entrant we searched for the closest establishment located in each of the neighbouring LAs and retained only those entrants that had such a neighbour within one kilometre. Since this detection procedure is only meant to compute distances to the LA border for each entrant (rather than find a match for a pair), we considered all possible establishments in all sectors and all years as potential neighbours. We expect this procedure to catch nearly all entrants that are located within one kilometre of a border.

As for employment, we implement both the standard methodology (i.e., conditional logit) and our spatial differencing approach. Results for two non-spatial specifications, are given in columns 1 and 2 of table 3 while columns 3 and 4 report results for two comparable spatial specifications. Again, for the sake of comparison, we have restricted the sample to be the same for all specifications. This restriction requires establishments to locate within 1 kilometre of a boundary between two English LAs between 1986 and 1989. Imposing this requirement leaves us with a sample of 19,337 establishments. To show that this sample is representative, Appendix B once again reports results for the same specifications but without imposing this restriction.

Starting with the conditional logit we see, from the results reported in column 1, that there appears to be a positive effect of tax rates on entry. As discussed above this positive relationship may be driven by endogeneity. Column 2 shows what happens when we correct for this by replacing the actual tax rates with the predicted tax rates from a first stage

	CL	CL IV	LOGIT	LOGIT IV
(log) tax rate	0.397 <sup>a</sup>	0.521	0.108	0.809
	(0.079)	(0.883)	(0.177)	(0.921)
Number of entrants	19,337	19,337	19,337	19,337

Notes: Number of entrants as a function of local tax rates. First column (CL) reports results for conditional logit, second column (CL IV) instruments local taxes using political variables, third column (LOGIT) reports results from a logit model for spatially differenced variables; fourth column (LOGIT IV) instruments local taxes. Standard errors under coefficients. <sup>a</sup>, denotes significance at the 1% level. Estimates in the first two columns are from a Poisson regression using the equivalence result from Guimaraes *et al.* (2003).

**Table 3.** Results for entries

regression of tax rates on political variables. Once instrumenting, we find no significant relationship between tax rates and entry. Correcting the standard errors to allow for the fact that we are instrumenting would only reinforce this finding.<sup>31</sup> Columns 3 and 4 show that we reach exactly the same conclusions on the irrelevance of tax rates for entry using our spatially differenced approach.

This lack of effect of local taxation on entry is consistent with the interpretation we gave above for our employment results. Recall that the negative effect of local taxation on firm growth was attributed to imperfect capitalisation of the local tax rate into property rental prices. In turn, imperfect capitalisation was attributed to the imperfect mobility of existing firms and the lags between rent negotiations. Unlike existing firms, new entrants are mobile and negotiate their rents upon entry so that the capitalisation of local taxation is expected to be near perfect. In such case, the burden of local taxation should fall only on the landlords and leave the behaviour of firms undistorted.

## 6. Conclusion

We propose a new approach to assess the effects of local taxation. Our results point to the importance of controlling for both unobserved establishment-specific and unobserved site-specific characteristics and possible simultaneity. Simple OLS results sug-

---

<sup>31</sup>Hence our decision only to implement the theoretically consistent two stage conditional logit procedure rather than an instrumented probit specification.

gest a positive relationship between employment and taxes. Allowing for unobserved establishment-specific effects and instrumenting for local taxation, we still find a positive relationship between employment and taxes. Allowing for unobserved location-specific effects and instrumenting for local taxation, we find a negative significant relationship between employment and taxes. By contrast we find that local taxation has no effect on the entry of new establishments.

Beyond our methodological contribution, this analysis also suggests that the study of local taxation and, more broadly, that of decentralised public intervention faces serious endogeneity problems whereby local public decisions depend strongly on very local conditions, which are extremely difficult to control for. As shown here, properly controlling for such local conditions is a necessary condition to obtain reliable estimates.

The second broad lesson is that even taxes that are seemingly close to an 'ideal' tax that would be free of distortion can in practice generate significant distortions. Despite the fact that the UK business rates were close to George's 1884 'pure' land tax, frictions in the rental market implied that increases in local taxation were not fully capitalised into the rents and thus had an adverse effect on employment.

## References

- Arnott, Richard. 2004. Does the Henry George Theorem provide a practical guide to optimal city size? *American Journal of Economics and Sociology* 63(5):1057–1090.
- Baldwin, Richard E., Rikard Forslid, Philippe Martin, Gianmarco I. P. Ottaviano, and Frederic Robert-Nicoud. 2004. *Economic Geography and Public Policy*. New Jersey: Princeton University Press.
- Bartik, Timothy. 1991. *Who benefits from state and local economic development policies?* Kalamazoo, Michigan: W. E. Upjohn Institute for Employment Research.
- Berman, Eli and Linda Bui. 2001. Environmental regulation and labor demand: evidence from the south coast air basin. *Journal of Public Economics* 79(2):265–295.
- Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics* 119(1):249–275.
- Black, Sandra E. 1999. Do better schools matter? Parental valuation of elementary education. *Quarterly Journal of Economics* 114(2):587–599.
- Bond, Stephen, Kevin Denny, John Hall, and William McCluskey. 1996. Who pays business rates? *Fiscal Studies* 17(1):19–35.
- Duranton, Gilles and Henry G. Overman. 2005. Testing for localization using micro-geographic data. *Review of Economic Studies* 72(4):1077–1106.
- Epple, Dennis and Thomas Nechyba. 2004. Fiscal decentralization. In Vernon Henderson and Jacques-François Thisse (eds.) *Handbook of Regional and Urban Economics*, volume 4. Amsterdam: North-Holland.
- Fujita, Masahisa. 1989. *Urban Economic Theory: Land Use and City Size*. Cambridge: Cambridge University Press.
- George, Henry. 1884. *Progress and Poverty: An Inquiry into Causes of Industrial Depressions, and of Increase of Want with Increase of Wealth. The Remedy*. London: W. Reeves.
- Gibbons, Steve and Stephen Machin. 2003. Valuing english primary schools. *Journal of Urban Economics* 53(2):197–219.
- Griffith, Rachel. 1999. Using the ARD establishment level data: an application to estimating production functions. *Economic Journal* 109(456):F416–F442.
- Guimaraes, Paulo, Octavio Figueiredo, and Douglas Woodward. 2003. A tractable approach to the firm location decision problem. *Review of Economics and Statistics* 85(1):201–204.
- Guimaraes, Paulo, Octavio Figueiredo, and Douglas Woodward. 2004. Industrial location modeling: Extending the random utility framework. *Journal of Regional Science* 44(1):1–20.

- Head, Keith and Thierry Mayer. 2004. Market potential and the location of Japanese investment in Europe. *Review of Economics and Statistics* 86(4):959–972.
- Holmes, Thomas J. 1998. The effect of state policies on the location of manufacturing: Evidence from state borders. *Journal of Political Economy* 106(4):667–705.
- Kahn, Matthew E. 2004. Domestic pollution havens: Evidence from cancer deaths in border counties. *Journal of Urban Economics* 56(1):51–69.
- Thompson, Robert and Sotiris Tsolacos. 2001. Industrial land values – a guide to future markets? *Journal of Real Estate Research* 21(1/2):55–76.
- Tiebout, Charles M. 1956. A pure theory of local expenditures. *Journal of Political Economy* 64(5):416–424.
- Wilson, John D. 1999. Theories of tax competition. *National Tax Journal* 52(2):269–304.

## **Appendix A. Correction of the standard errors**

When spatially differencing, an establishment  $i$  that has  $n$  neighbours will be in  $n$  pairs. This induces correlation in the error for all  $n$  of these pairs. The correlation arises because  $\epsilon_{it}$  (the error of establishment  $i$ ) enters the error of each pair. This imposes a particular structure to the covariance matrix which we use to correct the standard error. This appendix gives the details of that correction. Note that Bertrand, Duflo, and Mullainathan (2004) consider similar issues when proposing their correction for the standard errors of difference-in-difference estimators to measure treatment effects. There are several key differences between our correction and the one proposed there. First, their correction is difficult to apply to specifications with a large number of establishments and locations because they estimate the treatment effect directly without rewriting the model in difference-in-difference. Second, their correction requires the covariance matrix to be block diagonal. This means that it is not applicable to situations, like the one here, where there is no obvious way to construct closed sets of neighbours (because establishment A may be a neighbour to establishment B, who may be neighbour to C etc). In short we deal with an error structure which is considerably more complex. This comes at a cost: we ignore issues arising from serial correlation of the errors that are the key concern of that paper.

After spatial-differentiation and within-pair projection, the model can be re-written as follows:

$$W\Delta e = \alpha W\Delta r + W\Delta X\beta + W\Delta\epsilon \quad (\text{A } 1)$$

$$= Z\gamma + W\Delta\epsilon \quad (\text{A } 2)$$

where for any variable  $v_{it}$ , observations have been stacked in *pair* and time order,  $\gamma' = (\alpha, \beta')$  and  $Z = (W\Delta r, W\Delta X)$ . The OLS estimator is then:

$$\hat{\gamma} = (Z'Z)^{-1} Z' \Delta e = \gamma + (Z'Z)^{-1} Z' \Delta\epsilon \quad (\text{A } 3)$$

We suppose that the residuals  $\epsilon_{it}$  are *iid* with variance  $\sigma^2$ . The variance of the OLS estimator is:

$$V(\hat{\gamma}) = \sigma^2 ABA \quad (\text{A } 4)$$

where  $A = (Z'Z)^{-1}$  and  $B = Z'\Delta\Delta'Z$ . Matrix  $A$  is easy to compute since  $Z$  can be obtained after spatial-differentiation and a projection *within*-pair of the explanatory variables. Matrix  $B$  can also be computed after using an algorithm to obtain  $\Delta'Z$ . This algorithm relies on the fact that  $\Delta$  has a simple structure. Indeed, denote  $p \in \{1, \dots, P\}$  where  $P$  is the number of pairs and  $N_p$  the number of years that pair  $p$  appears in the data. We can decompose  $\Delta$  in blocks such that  $\Delta = (\Delta'_1, \dots, \Delta'_p)'$  where, for instance, block  $\Delta_p$  writes:

$$\begin{pmatrix} \dots & 0 & 1 & 0 & \dots & 0 & -1 & 0 & \dots & \dots \\ \dots & 0 & 0 & 1 & 0 & \dots & 0 & -1 & 0 & \dots \\ \dots & 0 & 0 & 0 & 1 & 0 & \dots & 0 & -1 & \dots \end{pmatrix} \quad (\text{A } 5)$$

(supposing  $N_p = 3$ ). The first line corresponds to the first year that the pair is in the data, the second line to the second year, etc... Each column of  $\Delta$  can contain the values 1 and  $-1$  several times depending on the number of times an establishment has been matched with neighbours in the corresponding year. For a column  $i$  of  $\Delta$ , denote  $j_i$  the identifier of the corresponding establishment and  $t_i$  the corresponding date. If *establishment* is the vector containing all the establishment identifiers and *year* is the vector containing the years,  $j_i$  and  $t_i$  can be retrieved from the first element in column  $i$  that

takes a value different from 0. For example, in the GAUSS language, column  $i$  of  $\Delta$  is then of the form  $(establishment .eq establishment [j_i]). * (year .eq year[t_i]) - (establishment_n .eq establishment [j_i]). * (year_n .eq year[t_i])$  where  $establishment_n$  contains the identifier of the neighbouring establishments and  $year_n$  their years. Element  $(n,k)$  of  $\Delta'Z$  can be computed using column  $n$  of  $\Delta$  and column  $k$  of  $Z$ . The whole matrix  $\Delta'Z$  is obtained from a loop over  $n$  and  $k$ .

We now propose an estimator of  $\sigma^2$ . Denote  $\widehat{W\Delta\epsilon}$  the vector of residuals from the OLS estimation. We have:

$$\widehat{W\Delta\epsilon} = W\Delta\epsilon - Z\hat{\gamma} = M_Z W\Delta\epsilon \quad (\text{A } 6)$$

where  $M_Z$  is the projector in the dimension orthogonal to  $Z$ . We then get:

$$\widehat{W\Delta\epsilon}'\widehat{W\Delta\epsilon} = \epsilon'\Delta'WM_ZW\Delta\epsilon \quad (\text{A } 7)$$

From this formula, we obtain:

$$E\left(\widehat{W\Delta\epsilon}'\widehat{W\Delta\epsilon}\right) = Etr\left(\widehat{W\Delta\epsilon}'\widehat{W\Delta\epsilon}\right) \quad (\text{A } 8)$$

$$= \sigma^2 tr\left(\Delta'W\Delta\right) - \sigma^2 tr\left[Z\left(Z'Z\right)^{-1}Z'\Delta\Delta'W\right] \quad (\text{A } 9)$$

$$= \sigma^2 tr\left(W\Delta\Delta'\right) - \sigma^2 tr\left(AB\right) \quad (\text{A } 10)$$

We can recover  $tr(AB)$  very easily from  $A$  and  $B$ . It is also possible to simplify the expression:  $tr(W\Delta\Delta')$ . We can write  $\Delta\Delta'$  in blocks corresponding to pairs. Indeed, the  $(p,q)$ -block writes:  $\Delta_p\Delta'_q$ .  $W$  is block diagonal. Thus, the  $(p,q)$ -block of  $W\Delta\Delta'$  writes  $W_p\Delta_p\Delta'_q$  where  $W_p$  is the  $(p,p)$ -block of  $W$ . Hence, we get:  $tr(W\Delta\Delta') = \sum_p tr\left(W_p\Delta_p\Delta'_p\right)$ . As we have  $\Delta_p\Delta'_p = 2I_{T_p}$  and  $tr(W_p) = T_p - 1$  (where  $T_p$  is the number of years that pair  $p$  appears in the data), we finally get:  $tr(W\Delta\Delta') = 2(N - P)$ . An unbiased (and consistent) estimator of  $\sigma^2$  is then:

$$\hat{\sigma}^2 = \frac{1}{2(N - P) - tr(AB)}\widehat{W\Delta\epsilon}'\widehat{W\Delta\epsilon} \quad (\text{A } 11)$$

We can finally deduce an estimator of the variance of  $\hat{\gamma}$ :

$$\widehat{V}(\hat{\gamma}) = \hat{\sigma}^2 ABA \quad (\text{A } 12)$$

We now compute the standard errors when instrumenting. The model is:

$$W\Delta e = \alpha W\Delta r + W\Delta X\beta + W\Delta\epsilon \quad (\text{A } 13)$$

$$W\Delta r = Y\delta + \eta \quad (\text{A } 14)$$

with  $Y = W\Delta P$  where  $P$  are some political variables, and  $cov(Y, \eta) = cov(Y, W\Delta\epsilon) = 0$  by assumption. Denote  $\hat{\delta}$  the OLS estimator of  $\delta$  obtained from equation (A 14) and  $V = \hat{V}(\hat{\delta}|X)$  an estimator of its covariance. This covariance estimator may simply be the usual OLS estimator. It may also take into account clusters at the jurisdiction level. Equation (A 13) rewrites:

$$W\Delta e = \alpha Y\hat{\delta} + W\Delta X\beta + W\Delta\epsilon + \phi \quad (\text{A } 15)$$

$$= \tilde{Z}\gamma + W\Delta\epsilon + \phi \quad (\text{A } 16)$$

with  $\tilde{Z} = (Y\hat{\delta}, W\Delta X)$  and  $\phi = \alpha(Y\delta - Y\hat{\delta})$ .  $\phi$  is such that  $E(\phi|X, Y) = 0$  and  $V(\phi|X, Y) = \alpha^2 YVY'$  with  $V = V(\hat{\delta}|Y)$ . The IV estimator is:

$$\hat{\gamma}_{IV} = (\tilde{Z}'\tilde{Z})^{-1} \tilde{Z}'W\Delta e = \gamma + (\tilde{Z}'\tilde{Z})^{-1} \tilde{Z}'(W\Delta\epsilon + \phi) \quad (\text{A } 17)$$

Assuming that  $\hat{\delta}$  is known, the variance of the IV estimator can be approximated by:

$$V(\hat{\gamma}_{IV}) \approx (\tilde{Z}'\tilde{Z})^{-1} \tilde{Z}'(\sigma^2 W\Delta\Delta'W + \alpha^2 YVY') \tilde{Z} (\tilde{Z}'\tilde{Z})^{-1} \quad (\text{A } 18)$$

$$\approx \tilde{A}(\sigma^2 \tilde{B} + \alpha^2 \tilde{C}) \tilde{A} \quad (\text{A } 19)$$

with  $\tilde{A} = (\tilde{Z}'\tilde{Z})^{-1}$ ,  $\tilde{B} = \tilde{Z}'W\Delta\Delta'W\tilde{Z}$  and  $\tilde{C} = \tilde{Z}'YVY'\tilde{Z}$ .  $\tilde{A}$  and  $\tilde{B}$  are easy to compute (see the OLS case).  $\tilde{C}$  is also easy to compute since we can first compute  $Y'\tilde{Z}$ , which has a small dimension. We now propose an estimator of  $\sigma^2$ . Denote  $\widehat{W\Delta\epsilon + \phi}$  the vector of residuals from the IV second-stage estimation. We have:

$$\widehat{W\Delta\epsilon + \phi}' \widehat{W\Delta\epsilon + \phi} = (W\Delta\epsilon + \phi)' M_{\tilde{Z}} (W\Delta\epsilon + \phi) \quad (\text{A } 20)$$

Consider for a while that  $Z$  is non random (i.e.,  $\hat{\delta}$  is non random). We have:

$$E(\widehat{W\Delta\epsilon + \phi}' \widehat{W\Delta\epsilon + \phi}) = trE(\epsilon'\Delta'WM_{\tilde{Z}}W\Delta\epsilon) + trE(\phi'M_{\tilde{Z}}\phi) \quad (\text{A } 21)$$

$$= \sigma^2 [2(N - P) - tr(\tilde{A}\tilde{B})] + \alpha^2 [tr(VY'Y) - tr(\tilde{A}\tilde{C})] \quad (\text{A } 22)$$

A (consistent) estimator of  $\sigma^2$  is then:

$$\hat{\sigma}_{IV}^2 = \frac{\widehat{W\Delta\epsilon} + \phi' \widehat{W\Delta\epsilon} + \phi - \hat{\alpha}_{IV}^2 \left[ \text{tr}(\widehat{V}Y'Y) - \text{tr}(\widehat{A}\widehat{C}) \right]}{2(N - P) - \text{tr}(\widehat{A}\widehat{B})} \quad (\text{A } 23)$$

with  $\widehat{V}$  an estimator of  $V$  obtained from the first-stage equation,  $\widehat{C} = \widetilde{Z}'Y\widehat{V}Y'\widetilde{Z}$ . Note that when the residuals  $\eta_{it}$  are *iid* with variance  $\theta^2$ , we have  $\text{tr}(\widehat{V}Y'Y) = N\theta^2$  and  $\text{tr}(\widehat{A}\widehat{C}) = \widetilde{\theta}^2 \text{tr}(P_{\widetilde{Z}}P_Y)$ . The estimator of  $\sigma^2$  becomes:

$$\hat{\sigma}_{IV}^2 = \frac{\widehat{W\Delta\epsilon} + \phi' \widehat{W\Delta\epsilon} + \phi - \hat{\alpha}_{IV}^2 \theta^2 [N - \text{tr}(P_{\widetilde{Z}}P_Y)]}{2(N - P) - \text{tr}(\widehat{A}\widehat{B})} \quad (\text{A } 24)$$

Finally, the variance of the IV estimator can be approximated by:

$$\widehat{V}(\hat{\gamma}_{IV}) = \widetilde{A} \left( \hat{\sigma}_{IV}^2 \widetilde{B} + \hat{\alpha}_{IV}^2 \widetilde{C} \right) \widetilde{A} \quad (\text{A } 25)$$

## Appendix B. Non-spatial results for different samples of establishments

Tables 4 and 5 present results for the non-spatial employment specifications using less restrictive samples than those used in the text. These results should be compared to those in table 1, where establishments must be (i) in a pair in which (ii) both establishments simultaneously report employment in at least two years (the sample used for our preferred specification). Table 4 uses the largest possible sample for each specification imposing neither restriction (i) or (ii). Table 5 uses the largest possible sample of establishments that are part of a pair. That is, we impose restriction (i) but not restriction (ii). As can be seen from the comparison of tables 1, 4 and 5, all point estimates are of the same sign and do not significantly differ from each other.

Table 6 present results for the non-spatial entry specifications using all entrants. These results should be compared to those in table 3, where establishments must enter within 1 kilometre of an LA boundary to be part of the sample. All point estimates are of the same sign and do not significantly differ from each other.

	OLS	WITHIN	WITHIN IV
(log) tax rate	0.222 <sup>a</sup> (0.029)	0.131 <sup>a</sup> (0.016)	0.481 <sup>a</sup> (0.062)
age censored dummy	0.652 <sup>a</sup> (0.033)	0.290 <sup>a</sup> (0.016)	0.294 <sup>a</sup> (0.015)
age	-0.015 (0.018)	0.053 <sup>a</sup> (0.005)	0.055 <sup>a</sup> (0.005)
age squared	0.011 <sup>a</sup> (0.002)	-0.002 <sup>a</sup> (0.001)	-0.003 <sup>a</sup> (0.001)
Adjusted R-squared	0.13		
Number of observations	61785	53684	53684
Number of establishments	21813	13875	13875

Notes: Standard errors under coefficients. <sup>a</sup>, <sup>b</sup> and <sup>c</sup> denote significance at the 1%, 5% and 10% level respectively. Largest possible sample. First column (OLS) only restricted by data availability. Second and third column require at least two observations per establishment. Compare to table 1 where establishments must be (i) in a pair in which (ii) both establishments simultaneously report employment in at least two years.

**Table 4.** Non-spatial regression results for largest possible samples

	OLS	WITHIN	WITHIN IV
(log) tax rate	0.156 <sup>a</sup> (0.042)	0.105 <sup>a</sup> (0.023)	0.472 <sup>a</sup> (0.080)
age censored dummy	0.662 <sup>a</sup> (0.054)	0.208 <sup>a</sup> (0.025)	0.211 <sup>a</sup> (0.025)
age	0.004 (0.030)	0.046 <sup>a</sup> (0.008)	0.047 <sup>a</sup> (0.008)
age squared	0.008 <sup>b</sup> (0.003)	-0.002 <sup>a</sup> (0.001)	-0.003 <sup>a</sup> (0.001)
Adjusted R-squared	0.13		
Number of observations	25579	22803	22803
Number of establishments	5564	5852	5852

Notes: Standard errors under coefficients. <sup>a</sup>, <sup>b</sup> and <sup>c</sup> denote significance at the 1%, 5% and 10% level respectively. Sample restricted to establishments that are part of a pair. First column (OLS) only restricted by data availability. Second and third column require at least two observations per establishment. Compare to table 1 where establishments must be (i) in a pair in which (ii) both establishments simultaneously report employment in at least two years.

**Table 5.** Non-spatial regression results for largest possible samples of establishments in pairs

	CL	CL IV
(log) tax rate	0.633 <sup>a</sup> (0.070)	0.267 (0.244)
Number of establishments	81,042	81,042

Notes: Number of entrants as a function of local tax rates. First column (CL) reports results for conditional logit, second column (CL IV) instruments local taxes using political variables. Largest possible sample. Standard errors under coefficients. <sup>a</sup>, denotes significance at the 1% level. Estimates are from a Poisson regression using the equivalence result from Guimaraes *et al.* (2003).

**Table 6.** Non-spatial regression results for largest possible sample

## Appendix C. First stage regression

Table 7 presents the results from the first stage regression of spatially difference (log) tax rates on the exogenous variables and spatially differenced instruments. We present the *within* version that is used for instrumenting the spatially differenced specification in the text and report both corrected and uncorrected standard errors.

spatial difference of	(1)	(2)
age censored dummy	-0.0029 (0.0036)	-0.0029 (0.0036)
age	-0.0006 (0.0011)	-0.0006 (0.0011)
age squared	0.0001 (0.0001)	0.0001 (0.0001)
share Conservative	0.1279 <sup>a</sup> (0.0436)	0.1279 <sup>a</sup> (0.0482)
share Labour	0.2466 <sup>a</sup> (0.0429)	0.2466 <sup>a</sup> (0.0470)
share Liberals	0.1236 <sup>a</sup> (0.0394)	0.1236 <sup>a</sup> (0.0426)
Conservative controlled	0.0813 <sup>a</sup> (0.0178)	0.0813 <sup>a</sup> (0.0226)
Labour controlled	-0.1715 <sup>a</sup> (0.0154)	-0.1715 <sup>a</sup> (0.0173)
Liberal controlled	-0.1376 <sup>a</sup> (0.0244)	-0.1376 <sup>a</sup> (0.0281)
share Conservative (if control)	-0.2278 <sup>a</sup> (0.0325)	-0.2278 <sup>a</sup> (0.0431)
share Labour (if control)	0.2645 <sup>a</sup> (0.0317)	0.2645 <sup>a</sup> (0.0355)
share Liberal (if control)	0.2659 <sup>a</sup> (0.0428)	0.2659 <sup>a</sup> (0.0497)
Number of observations	18370	18370
Number of establishments	6087	6087

Notes: Standard errors under coefficients. <sup>a</sup>, <sup>b</sup> and <sup>c</sup> denote significance at the 1%, 5% and 10% level respectively. Results from first stage regression of spatial difference of tax rates on establishment fixed effects, exogenous variables and political instruments. Column (1) presents results with uncorrected standard errors. Column (2) presents results with standard errors corrected according to Appendix A.

**Table 7.** First stage regression results