Mental Health Expenditure in England: a Spatial Panel Approach

Francesco Moscone, Martin Knapp, Elisa Tosetti

October 22, 2006

Abstract

We empirically investigate the determinants of local authority mental health expenditure in England. We adopt a reduced form demand and supply model, extended to incorporate possible interaction among authorities, as well as unobserved heterogeneity. The model is estimated using an annual panel dataset that allows us to explore both time-series and cross-municipality variation in mental health expenditure. Results are consistent with some degree of interdependence between neighbouring municipalities in spending decisions. This first attempt to apply spatial panels in investigating health expenditure offers insights and raises new questions.

Keywords: Mental Health, Social Care, Spatial Econometrics, Panels

JEL-Classification: I18, I38, C31, C33

1 Introduction

Recent years have witnessed a gradual rebalancing of responsibility from National Health Service (NHS) to local authority budgets in financing mental health services in the United Kingdom. Local authorities have assumed responsibility for funding placements, such as in nursing homes, residential care homes and supported accommodation, that previously would have
been paid for by the NHS or covered by social security transfers. As a consequence, in the last few years, expenditure on mental health services by municipalities has grown faster than NHS spending, accounting today for roughly one fifth of its size (Knapp et al., 2005). This proportion varies substantially across the country. For instance, social services spending on mental health by local authority in London ranges from 23% to 79% of health service spending (Aziz et al., 2003).

Two empirical strands of health economics literature have been developed to explore variations in mental health expenditure and costs. One strand uses data on individual people, perhaps collected in controlled experiments, to study how mental health costs vary in relation to characteristics such as diagnosis, illness severity, physical health, and socio-economic characteristics. A second strand focuses on area-level characteristics, including some that are recognized as proxies for mental health needs, such as ethnic mix and deprivation, in examining variations in mental health spending. The tendency to use risk factors rather than direct measures of needs follows from the lack of information on the prevalence of mental illness at area level, and from recognition that measures of treated prevalence depend on service supply as well as need. The work in this paper fits within the second strand of literature, taking as statistical unit the administrative area (local authority or council), and trying to explain the geographical disparity in mental health spending in relation to a set of risk factors. The approach taken is to extend the traditional demand and supply framework to allow for possible interaction among authorities as an additional source of expenditure variation. If it exists, interaction of this kind could have implications for the methods used by central government to monitor and influence local performance across the country and over time, and perhaps also to allocate resources across municipalities (cf. Carr-Hill et al., 1999; Jarman et al., 1992; Carr-Hill et al., 1994; Kelly and Jones, 1995; Smith et al., 1996; and Lelliot et al., 1996).

A substantial proportion of English local authorities’ revenue, roughly two-thirds, comes from central government grant and uniform business rates; the residual part is locally generated from council tax, sales, fees and charges. Central government grants are distributed via a formula that tries to identify the spending requirement of an authority if it were to adopt a common level of services, given differences in needs and exogenously determined input prices. The allocation is further adjusted by an equalization factor, to account for possible distortions generated by disparities in the council tax revenue. Central government assesses personal social services needs separately for each client group, and a number of formulae are used to estimate
the (adjusted) spending requirements. However, local authorities have con-
siderable autonomy in allocating these resources from central government,
both to and among the various local services (education, housing, leisure,
community resources as well as social services), prioritizing particular areas
and client groups in line with local interpretations of need, local preferences
and so on. Therefore, authorities’ actual spending will at least partly, and
perhaps predominantly, reflect local policies rather than the standard spend-
ing assessment (SSA) by central government. In particular, expenditure is
likely to depend on the risk factors perceived by the authority as well as lo-
cal preference and income, competition for other services, and local policies
(Carr-Hill et al., 1999).

In this study we adopt a general approach where expenditure choices
of municipalities are examined by reference not only to needs and supply
determinants, but also to policy interdependence and geographical concen-
tration of unobservable risk factors. A recent strand of literature in public
economics has analyzed the role of interaction among policy makers. In
particular, several studies from a range of countries have investigated in-
terdependency between local authorities in setting tax rates (Ladd, 1992;
Besley and Case, 1995; Revelli, 2001), in deciding standard and regulatory
measures (Brueckner, 2000; Fredriksson and Millimet, 2002), and in deter-
mining expenditure levels (Case et al., 1993; Kelejian and Robinson, 1993;
Bivand and Szymasky, 1997; Revelli, 2002). In contrast, the effect on local
decisions of expenditure behaviour in neighbouring areas has been almost
completely neglected in health economics. In the next section we therefore
provide a series of arguments that suggest why there might be strategic in-
teractions among municipalities. Our line of reasoning will build on recent
economic theories on interaction among individuals (Manski, 1993, 2000),
the literature on public economics (Brueckner, 2000), and intuition from
studying the English health and social care systems (e.g. Moscone and
Knapp, 2005). We provide an econometric framework that models mental
health expenditure as a function of risk factors, supply determinants, spend-
ing decisions of neighbouring municipalities, and time trends. This model
is then estimated using a yearly panel dataset that allows both time-series
and cross-municipality variation to be explored in per-capita mental health
expenditure by English local authorities.

The rest of the paper is organized as follows. Section 2 discusses some
possible reasons why local policy makers may interact. Section 3 synthesize
existing empirical evidences on interaction between municipalities provided
by a recent strand of literature on public economics. Section 4 introduces the
empirical models of health expenditure. Section 5 presents the data. Section
6 focuses on the estimation of our models, and emphasizes some limitations of the study. Finally, Section 7 closes with some concluding remarks.

2 Interaction and reaction in mental health policy

Local interaction models in economics can be defined as models in which each individual’s behaviour is affected by the characteristics or behaviour of other people (Manski, 1993). A major assumption of these models is that individuals interact locally with a finite subset of the population defined by a social, economic or geographical distance metric. Their interaction generates spillovers and externalities, and can lead to an emergent collective behaviour and aggregate pattern that empirically translates into a structure of correlation of data (Anselin, 2002).

The causes of interaction among economic agents are often discussed using a threefold categorization suggested by Manski (1993, 2000), Moffitt (2001), Brock and Durlauf (2000) and others: endogenous, exogenous and correlated effects. The endogenous effect hypothesizes that the propensity of an individual to behave in some way is causally influenced by the behaviour of other members of the group. According to the exogenous (or contextual) effect, individual action varies with observed attributes that define group membership. The correlated effect states that individuals in the same neighbourhood tend to behave similarly because they have similar characteristics or face similar opportunities and constraints. The microeconomic foundations underlying interdependence is that agents interact through their chosen action; an action chosen by one agent may influence the constraints, expectations, and/or preferences of his reference group (Manski, 2000).

These possible scenarios refer to what might be called within-neighbourhoods effects. In contrast, in this paper we explore associations among neighbourhoods, in particular whether there exist spillovers or other interactions among municipalities (Dietz, 2002). It is possible to suggest among-neighbourhood equivalents to the endogenous, exogenous and correlated effects, and certainly it is methodologically pertinent to be able to distinguish between them. In a recent study, Moscone and Knapp (2005) posited a number of reasons why there might be spatial dependence between local authorities in relation to mental health expenditure, labelling them the demonstrative, shared resource, directive, and correlated effects.

The demonstrative (or market leader) effects can be considered a characterization of the endogenous effect. One local authority’s good (or bad) performance may encourage its neighbours to mimic (or avoid) the activities
and expenditure patterns associated with such performance. Some municipalities may develop good reputations for their championing of particular causes or particular services, for example, providing better residential care for people with mental health problems, or more energetically developing procedures for user involvement in decision making, thus encouraging mimicking behaviour. For instance, an authority that has not previously faced high suicide rates, might respond to a significant increase in the number of suicide episodes by looking at the policies of local authorities that are (or that seem to be) more experienced in addressing this challenge. Put differently, local authorities might use information gathered from others’ actions to provide additional knowledge on a particular issue. We argue that this transfer of knowledge and information, which will then translate into a local spending decision, occurs more naturally between authorities that are neighbours for example in the geographical domain, although other domains are possible, as we explore later. For example, such transfer of knowledge might require face-to-face interaction, or a well developed sense of local issues (Rallet and Torre, 1999). Growing use of publicly announced and widely reported performance indicators, such as the performance ratings for social services (SSPR) introduced in 2001 by the central government, could encourage such information spillovers and diminish territorial inequality.

According to the shared resource hypothesis, which could also be labelled the contextual (or exogenous) effect, adjacent authorities share common observable characteristics that are independent of each authority’s attributes. We argue that changes in the organizational-financial structure of any given health or social care system might significantly affect surrounding systems. For instance, the closure of a large psychiatric hospital that accommodates people from a number of geographical areas (and hence a number of municipalities) might impact on the social care sector across a wide territory. A substantial increase in the need for social care services could then arise in a cluster of municipalities served by the (former) hospital. Other examples of the shared resource effect are environmental stressors, such as deprivation and poverty, or climate, or a shared gene pool, that pertain an area wider than a single authority. These factors play a role in explaining the prevalence of certain mental health disorders at a supra-authority level, which in turn is hypothesised to impact on spending of a single local authority.

The directive effect can be discussed under the exogenous effect heading, where local authorities react to common policy environments. For example, in the UK two-tiered structure of local government, lower-tier authority policy (districts) varies with upper-tier authority policies (counties), giving the misleading impression of horizontal interaction or competition (Revelli,
In addition, the regionally organized regulatory, inspection and auditing function of central government might offer common guidance across a region that may have influenced certain patterns of activity or expenditure across municipalities.

The correlated effects comprise those determinants of interdependence that are not observable to the econometrician. For instance, high psychiatric hospital admissions in two neighbouring authorities might be explained by aircraft noise, caused by the presence of an airport (Kryter, 1990).

From the above discussion the hypothesis emerges that one local authority’s expenditure choices are, in part, the result of strategic interaction with neighbouring municipalities, or respond to some either observable or unobservable common risk factors. We emphasize that contacts and interactions are likely to occur not only among geographically adjacent municipalities, but also among authorities that are close with respect to non-geographical metrics, such as the socio-economic space. These contacts and interactions are channeled by the use of new technologies, such as the internet.

In the following section we review some significant works that support the existence of local interdependence in public spending.

### 3 Evidence of interaction in public spending

There exists an ample literature studying cross-sectional variation of public spending in relation to various economic, political, and demographic factors. Differences in public expenditure across municipalities are traditionally explained by differences in local characteristics such as per-capita income, taxes, socio-economic attributes and political structure (Foster et al., 1980; Wildasin, 1986). Recent literature in public economics has pointed out that an important element explaining variations in per-capita public spending is represented by the spillover effect, that is expenditure on public services in a locality can have beneficial or harmful effects across a wider geographical area. Put differently, geographical proximity plays an important role in the activities of local authorities. Recalling the first law of geography (Tobler, 1979), the transfer of knowledge (or information), which will then translate into a local spending decision, spills over more naturally between neighbouring authorities than between those which are far apart. According to this argument, municipalities can be seen as concerned about how their actions, such as activities or expenditures or tax rates, compare with those of their neighbours.
A number of empirical studies support the existence of such phenomena. The earliest paper in the literature is attributed to Case et al. (1993), who used data on local spending in the United States to test for spending interdependence. The authors found evidence of positive correlation in states' per-capita expenditures in several different categories. They stressed the possibility that the sign as well as the magnitude of a locality's spending behaviour might be positive for some spending categories and negative for others. In general, the authors observed that a one dollar increase in a neighbouring locality increases a municipality's own level of expenditure by 70 cents. A similar line of reasoning is evident in the work by Kelejian and Robinson (1993). Using US county data they observed that police service expenditures in a given county are significantly and positively influenced by neighbouring county police expenditures. The authors argued that the decision of a county to spend more on police services generates a negative externality on its neighbours due to cross-overs between the borders. Murdoch et al. (1993) explored local authorities' expenditures on recreation in the Los Angeles area, detecting substantial positive correlations across territory, due to the fact that spending in neighbouring municipalities are complementary joint products. More recently, Lundberg (2001) observed that recreational and cultural expenditures at a municipality level in Sweden are concentrated and that these services are correlated with benefit spillovers between municipalities. Another recent investigation is that by Revelli (2002, 2006), who identified substantial spatial correlation in personal social service expenditure among English local authorities. Revelli (2006) observed a decline in local interaction in social spending after the introduction of the published performance rating system in England. Baicker (2005), exploring the extent to which health spending in one US state is influenced by the spending of neighbours, emphasized that states are most influenced by the states to and from which their residents are most likely to move. From this study emerges the result that, in response to a one dollar increase in these neighbours' spending, states raise their own spending by almost a full dollar.

The relevance of these results from an economic perspective is to suggest that spatial effects should be incorporated, or at least tested for, in analyses of health spending variations. To date, to our knowledge there is almost no work in this direction. We say ‘almost none’ because in the UK literature, for example, interesting work by Revelli (2002, 2006) detected substantial spatial patterns in overall social care expenditure among English local authorities, including apparent ‘mimicking behaviour’ in local property tax setting. In the following section we therefore suggest econometric models
that formalize the relationship between per-capita expenditure, risk factors and the average expenditure choices observed in neighbouring municipalities, and subsequently estimate them using approaches suggested by the panel regression literature and spatial econometrics.

4 Empirical models

We can suggest two econometric specifications that allow the representation and testing of the hypotheses introduced earlier.

The demonstrative or exemplar effects, explained by the benchmarking of each municipality performance against the performance of other authorities, might induce endogeneity of spending choices. The statistical framework for this effect is a generalization of models usually presented in the public economics literature, where the value of the dependent variable for one authority is simultaneously determined with that of neighbouring authorities (Brueckner, 2003; Anselin, 2002). Given $N$ municipalities observed over $T$ time periods, we assume that per-capita expenditure in mental health in the $i^{th}$ municipality at time $t$, namely $y_{it}$, is generated according to the following linear panel

$$y_{it} = \rho \sum_{j=1}^{N} w_{ij} y_{jt} + \beta' x_{it} + e_{it},$$

(1)

where $x_{it}$ is a $k \times 1$ vector of observed individual specific regressors on the $i^{th}$ cross-section unit at time $t$ ($i = 1, ..., N$ and $t = 1, ..., T$), $e_{it}$ is the error term, and $w_{ij}$ is the generic element of a positive, $N \times N$ matrix $W$, known as the spatial weights matrix. In a spatial weights matrix the rows and columns correspond to the cross-section observations, and the generic element $w_{ij}$ is usually interpreted as the strength of potential interaction between units $i$ and $j$ (Anselin, 2001, 2002). In our specific context, elements of $W$ can be regarded as how municipalities weight each others’ spending choices when deciding their own expenditure levels. The specification of the weights matrix is generally somewhat arbitrary, based on some measures of distance between units. A range of suggestions have been offered in the literature, based on geographical distance (Anselin, 1988), as well as more general metrics, such as economic proximity or similarity (Conley, 1999; Pesaran, Schuermann and Weiner, 2004) and social proximity (Conley and Topa, 2002). In this paper, we consider a number of alternative weights matrices, based on geographical contiguity as well as socio-demographic and political characteristics of local authorities. Specifically, we first adopt a
geographical contiguity criterion and assign $w_{ij} = 1$ when municipalities $i$ and $j$ share a common border or vertex, and $w_{ij} = 0$ otherwise. We then use geographical contiguity, weighted by socio-demographic characteristics. Indeed, it is possible that a municipality, when deciding its level of expenditure, assigns more weight to contiguous authorities that are larger in size, or that share similar characteristics. Let $x$ be the attribute on which we base our neighbourhood definition, we assign $w_{ij} = x_j$, if municipalities $i$ and $j$ share a common border or vertex, and zero otherwise (Baicker, 2005).

As a result, for each local authority, the spatial lag of the dependent variable is a weighted average of geographically adjacent municipalities, where weights are given by the selected socio-demographic characteristic. Finally, we depart from contiguity criterion, and introduce spatial matrices based only on socio-demographic distance. The rationale behind the use of non-geographical distance is that a local authority’s spending choices might be influenced by the actions of municipalities that are not geographically adjacent, but that are similar in terms of socio-demographic or other pertinent characteristics. Interactions can be eased by the adoption of new technologies, such as the internet. In such cases, physical distance ceases to be important in diffusing information. Following Case et al. (1991, 1993) and Baicker (2005), distance between any two local authorities $i$ and $j$ can be measured as

$$w_{ij} = \frac{1}{|x_i - x_j|}, \quad i \neq j = 1, ..., N.$$ 

The use of weights that are based on geographical distance, such as the contiguity criterion, ensures that they are exogenous to the model, a condition that is not guaranteed when adopting weights based on more general distance metrics (Anselin, 2002).

It is common practice to row-standardize the weighting matrix, so that the sum of the weights for each row is one. This ensures that all the weights are between 0 and 1, and that the weighting operation can be interpreted as an average of neighbouring values. While the standardization facilitates the interpretation of coefficients, a side effect is that the resulting matrix is likely to become not symmetric, thus entailing significant computational complexities in the estimation (Kelejian and Prucha, 1999).

In model (1), the coefficient $\rho$ measures how expenditure in one area is related to expenditure in neighbouring areas, conditional on the vector of explanatory variables. A positive, significant value for $\rho$ might indicate the existence of mimicking behaviour among neighbouring authorities, for example due to the influence of a market leader. However, we remark that a significant spatial coefficient $\rho$ could also be the result of forces within the
exogenous effect heading. Indeed, simultaneous models such as equation (1), attempting to capture interdependence among agent’s choices, lead to an identification problem — the inability to distinguish between behavioural and contextual factors. Such an issue is generally referred to by Manski (1993) as the reflection problem\(^1\). Hence, a significant spatial coefficient in our case might indicate endogeneity of spending choices or the effect of contextual characteristics and common policy on spending decisions of a set of local authorities.

An observed spatial pattern in expenditure choices is not necessarily due to interaction among local governments. Indeed, units sharing observable characteristics such as location or political party may also share unobservable characteristics that would lead the regression disturbances to be correlated. This conforms well with the existence of some environmental risks that are difficult to measure, such as air pollution, or exposure to acoustic pollution, which can be characterized by assuming that the errors are generated by a spatial process. Following most of the applied spatial econometric literature, we hypothesise that the error term of our regression model is generated by a spatial autoregressive (SAR) process

\[
y_{it} = \beta' x_{it} + e_{it},
\]

\[
e_{it} = \lambda \sum_{j=1}^{N} w_{ij} e_{jt} + \varepsilon_{it},
\]

for \(i = 1, ..., N\) and \(t = 1, ..., T\), where \(\varepsilon_{it}\) are IID random errors with zero mean and variance \(\sigma^2_{\varepsilon}\), and \(\lambda\) is a scalar parameter.

Equations (1) and (2)-(3) assume that the relationship between risk factors and mental health expenditure is homogeneous across municipalities. However, this assumption is rather restrictive in empirical applications and is unlikely to hold in this study. An earlier cross-sectional exploratory analysis of mental health expenditure suggested that substantial geographical heterogeneity of spending might not be captured by observable characteristics (Moscone and Knapp, 2005). This unobserved variability, if not properly incorporated in the model, may lead to incorrect conclusions of spatial correlation (McMillen, 2003). Thus, the empirical analysis of this paper is based on the estimation of a random effects panel data model, extended to include a spatially lagged dependent variable, common factors and spatial error correlation. The random effects specification allows us to capture time-invariant heterogeneity across municipalities through an individual authority-specific

---

\(^1\)See Rice and Sutton (1998) for an ample discussion of the identification problem.
error component, and thus may achieve a gain in information and efficiency when compared to a pooled regression (Hsiao, 2003). Note that we avoid the use of a fixed effects specification since our variables display little variation over time, and therefore might be highly correlated with the fixed effects. This would result in confidence intervals of estimates that are too large to be informative. Another reason for avoiding fixed effects is that data on per-capita expenditure cover a short time period. In fact, the reorganization of local government in England a few years ago changed some municipal boundaries, and we only have six years of data since the most recent boundary changes took effect.

The random effects model with spatially lagged dependent variable is

\[ y_{it} = \rho \sum_{j=1}^{N} w_{ij} y_{jt} + \beta' x_{it} + \epsilon_{it}, \]  

where \( \epsilon_{it} = \mu_i + \zeta_{it} \) is an IID error term, and \( \mu_i \) is a random effect associated to the \( i^{th} \) municipality, IID distributed with zero mean and variance \( \sigma^2_\mu \).

The random effects specification with spatial error correlation is (Baltagi et al., 2003)

\[ y_{it} = \beta' x_{it} + \epsilon_{it}, \] 

\[ \epsilon_{it} = \mu_i + v_{it}, \] 

\[ v_{it} = \lambda \sum_{j=1}^{N} w_{ij} v_{jt} + \zeta_{it}, \] 

where the notation is as above. In this case the error term is the orthogonal sum of two components: a time-invariant, municipality-specific random disturbance, and a spatial process.

A critical assumption in the random effects specifications is that \( E(\zeta_{it}\mu_i) = 0 \), and \( E(\mu_i x_{it}) = 0 \), for \( i = 1, ..., N \), and \( t = 1, ..., T \). Therefore, the individual-specific component is hypothesized to be orthogonal to the explanatory variables (Hsiao, 2003). If this hypothesis does not hold, estimates from the random effects model suffer from possible bias due to the correlation between the error term and the regressors. In the empirical section we will test this hypothesis using Hausman test statistics both in spatial and non spatial models.

In the econometric literature only a few studies have considered spatial random effects regression models. Anselin (1988) first tackled the specification and estimation of a panel with a random effects and spatial error
correlation, while later Case (1991) employed spatial random effects panels to analyse the household demand for rice in Indonesia. More recently, Elhorst (2003) discusses the random effects model extended to include a spatially lagged dependent variable and spatial error. To our knowledge, the approach employed here, estimating a spatial panel with a random error associated with each spatial unit, has not previously been explored in the field of health economics.

As for the estimation of the suggested models, we adopt a maximum likelihood approach. In the appendix, we briefly describe the likelihood functions of the two spatial versions of the random effects model, and provide some implementation details on the estimation procedure. Basically, estimates are obtained using an iterative two-stage estimation procedure that alternates the computation of GLS estimators and the use of search algorithms on the concentrated likelihood functions, until the convergence criterion is met. We fixed the convergence criterion to $10^{-8}$ (LeSage, 1999). Having obtained the estimates of the parameters, we constructed the asymptotic standard errors and the associated $z$-values using the numerical Hessian matrix of the full log-likelihood function. The numerical Hessian has been computed employing a routine provided by LeSage spatial econometric toolbox. Finally, we executed our routines for different initial values of parameters that appear in the likelihood functions in order to check the robustness of our results. The iterative procedure described above has been implemented using Matlab software from MathWorks Inc.

In the rest of the paper we implement this procedure to estimate the relationship between mental health expenditure, risk factors, supply determinants, and spatial effects.

5 Data

Our empirical study follows the 150 English local authorities (also referred to as councils with social services responsibility) over a period of six years, from 1998 to 2003. The dataset is drawn from several sources. The dependent variable is net personal social services expenditure for adults aged under 65 years with mental health problems, published yearly by the Department of

\textsuperscript{2}www.spatial-econometrics.com

\textsuperscript{3}The expenditure measures covers the following services: assessment and care management, nursing home placements, residential care home placements, supported and other accommodation, direct payments (consumer-directed purchases of services), home care, day care, equipment and adaptations, meals, other services to adults with mental health
Health, standardized by the total population in each local authority.

We took information on socio-demographic attributes of English municipalities, such as population by age structure, and on median weekly earnings by local authority from the Office of National Statistics. We drew on data on average house prices in England from the Land Registry. Finally, we gathered information on local political control (Labour, Conservative, Liberal Democrats and others) from the House of Commons.

Two local authorities, City of London and Isles of Scilly, were excluded from the analysis, given their unusual socio-economic and demographic characteristics. Each is very small when compared to the size of the remaining 148 English authorities, thus generating possible distortions in the estimation. One (City of London) has an unusual population composition and both have atypical social services organisation. Almost all (non-spatial) studies of English local authorities exclude these same two areas because of these marked differences.

To induce normality, we chose the logarithmic transformation of the dependent variable. Table 2 reports some tests for normality by year and for the entire period, and Figure 1 displays the kernel density of the logarithm of PSS per-capita spending. The Shapiro-Wilks test appears to be not significant for the years 1998, 1999, 2001 and 2002 suggesting that for these time periods the variable conforms well to the normality hypothesis. Conversely, in the years 2000 and 2003 normality is not achieved. However, the Shapiro-Wilks test, when applied to the entire period, is not significant, supporting the normality assumption. This result is also confirmed by Figure 1, which visualizes the empirical kernel density against the normal curve.

As for the regression, we selected a set of explanatory variables suggested by the literature as area-level characteristics potentially linked to mental health needs (McCrone and Jacobson, 2003; Aziz et al., 2003; Glover et al., 1998). In particular, we explored the impact on per-capita spending of the variables density of population, percentage of population who are male, percentage of people aged under 14, standardised mortality ratio, number

---

4 http://www.dh.gov.uk/PublicationsAndStatistics/Statistics
5 http://www.statistics.gov.uk/
6 http://www.landreg.gov.uk/
7 http://www.parliament.uk/
8 Aziz et al. (2003) performed a factor analysis on a large number of variables linked to mental health needs, for Greater London local authorities. They found that the first four factors explain approximately the 79% of the variation and are mainly determined by the following variables: population density, population age and gender structures, crime, number of people living alone and number of people who are single, widowed or divorced.
of jobs\textsuperscript{9}, percentage of households headed by lone parent, number of unemployment claimant\textsuperscript{10}, a dummy variable indicating whether the authority is controlled by Labour and one for control by the Liberal Democrats. Furthermore, we included an average house price measure, and the median weekly wage to control for supply-side factors. Descriptive statistics for the selected variables are given in Table 1. In the regression equations, these variables appear in logarithms so that the coefficients can be interpreted as elasticities. Finally, in the rest of the paper, when discussing significance of the parameter estimates, we use the 95-percent level of significance.

Table 3 provides some information on the connectivity characteristics of the weights matrices employed in the spatial regression analysis and in some descriptive statistics. The first column shows the connectivity information of the pure contiguity matrix. The average number of links is 5, meaning that each local authority has on average 5 neighbours. Accordingly, the average link is approximately 0.2, given that the matrix has been row-standardized. Finally, although not shown in the table, we report that there are two most connected observations with 13 links, and 13 units with only one neighbour. The second, third and fourth columns summarize the connectivity characteristics of contiguity matrices weighted by, respectively, population, population density, and political party. The fifth column reports information on distance based on population and population density. Clearly, in this case the number on non-zero links is much higher when compared to the other matrices, due to the fact that nonzero weights are not limited to contiguous neighbours.

6 Empirical analysis

Figure 2 maps the geographical distribution of per-capita expenditure averaged over the six years. The higher level of mental health spending, indicated in the map by darker colour, shows important concentrations across space. It is evident that the variable tends to distribute in clusters, with the highest concentrations in metropolitan areas such as Greater London, Greater Manchester and Birmingham. Per-capita expenditure is also concentrated in the areas contiguous to these clusters, though with a lower intensity. Figure 3 plots per-capita expenditure, expressed in standardized form, against its

\textsuperscript{9}The number of jobs is is measured by the Labour Force Survey as the sum of employee jobs (as measured by surveys of employers); self-employment jobs, and government-supported trainees.

\textsuperscript{10}This variable measures the number of claimants of unemployment-related benefits on the Benefits Agency administrative system.
spatial lag, so as to show one commonly used measure of spatial correlation: the local Moran measure (Arbia, 2006). The majority of observations fall in the first and fourth quadrants, indicating that a high (low) level of expenditure in one authority is associated with a high (low) level of expenditure in contiguous authorities.

Table 4 reports for each year the estimates of models (1) and (2)-(3). The estimated spatial coefficient in the spatial lag specification is highly significant in the years 2001, 2002 and 2003. Further, we observe wide variation in the estimated endogenous effect over time, ranging in the interval 0.10-0.23, and reaching its maximum in the year 2003. The estimated spatial parameter for the spatial error model is highly significant in the years 1998, 2001, 2002 and 2003, taking values between 0.16 and 0.30, while it is not significant in the remaining time periods. These findings of spatial correlation during the years 2001-2003 seem to be in contrast with Revelli’s (2006) observation of a declining local interaction in social spending, after the introduction of the publicly released performance rating system in England in 2001. However, since Revelli considers the entire aggregate of personal social spending, such reduction in spatial correlation might be the product of different effects in the spending categories that overall cancel out (Case et al. 1993; Baicker, 2005). Further, as the first performance evaluation was released in 2002, we would expect it to be effective only after 2003, due also to some delay in conforming to the new rules. As Revelli’s analysis covers a temporal span from 2000 to 2003, we argue that whether the performance rating system has reduced local interaction among authorities is still an open issue.

As for the remaining regressors, it is interesting to note that in both spatial models there is likely to be some temporal instability for some determinants. For instance, the effect of population density on expenditure, which is positive and significant for each year, seems to be stronger in the spatial lag model during the years 2001 and 2002; the death rate displays in both models a significant effect only in the second half of the study period. While the temporal variation in the effect of risk factors on mental health spending is not the object of this work, we refer to Moscone et al. (2006) for an exploratory analysis of this issue.

Since these results derive from a separate regression analysis for each year, they are generally less informative than the panel estimation (for an ample discussion on the advantages of using panel data see Hsiao, 2006).

---

11 The performance star rating has been introduced in England in 2001 by the Secretary of State for Health.
Hence, in the following we focus on panel data regression models. Among the different possible ways to model unobserved heterogeneity, we selected the random effects model on the basis of a Hausman test, which has a value of 8.76 with a p-value of 0.92. However, we note that, since our variables display little variation over time, the within estimator would be imprecise, which in turn implies low power of the Hausman test.

Table 5 presents the estimation of the pooled panel by ordinary least squares (Column I), the maximum likelihood estimation of classic random effects (Column II), the maximum likelihood estimation of random effects models with spatially lagged dependent variable (Column III, equations (4)-(5)), and with spatial error correlation (Column IV, equations (6)-(8)), for the entire period 1998-2003, using the selected regressors. In all models, time dummies were included to account for the existence of global shocks that affect all municipalities simultaneously. Looking at the estimated coefficients and their significance for models in columns (I) and (II), the following findings emerge. The impact of "population density" on spending is positive and significant as expected in both specifications, since we expect higher mental health expenditure in inner-city areas, which are more densely populated. The estimation of the two models yields a non-significant effect for the variable "percentage of males". According to McCrone and Jackobson (2003), while gender differences in mental illness are known at the individual level, they do not appear to have an appreciable impact at the area level. The variable "people aged under 14" has a positive and sizeable effect only in the OLS specification, while it is not significant in the random effects. As for "mortality ratio", it has the hypothesized positive effect on spending in both models, since it can be considered an indicator of poorer mental health (Aziz et al., 2003). The variable "unemployment claimants" has the expected positive sign in both models. Indeed, high unemployment tends to be associated with higher crime rates together with a poorer physical, mental and social environment, thus leading to a rise in social expenditure on mental health (Robinson, 1988).\textsuperscript{12} The regressor "lone parents" has a positive influence in both specifications, since, like mortality, it is usually associated with poorer mental health (Aziz et al., 2003). The estimated coefficient for "number of jobs" is not significant in either specifications. Among the factors used to adjust for supply-side influences, namely "house price" and "median weekly earnings", in the OLS regression both appear to have

\textsuperscript{12}However, if unemployment is regarded as proxy of income, it is likely to be negatively related to the local tax rate, thus leading to a lower social spending. A negative impact of unemployment on English social expenditure has been found in some empirical works (Revelli, 2001, 2000).
an influence on the dependent variable. However, when we take account of heterogeneity among authorities, none of them shows a significant effect on spending. Finally, in both specifications, the estimation of political control dummies indicates that Labour-controlled municipalities on average do not allocate higher resources on mental health than Liberal Democrat or Conservative municipalities. This is somewhat surprising, since we would expect differences in social service provision policy due to political ideology. The usual argument is that local authorities of the left are likely to spend more on social care than Conservative authorities, though recent work has pointed out that the Left gains credibility in cutting expenditure (Tavares, 2004). Our result, however, corroborates Revelli’s (2006) finding that Labour authorities do not set social care spending that are higher than Conservative authorities. A similar result of no evidence of political effects is shown by Crivelli et al. (2006) in a study on the determinants of health expenditure in Switzerland. In the rest of the paper we will explore more closely the role of political control and its geographical concentration in determining expenditure.

In general, the results in columns (III) and (IV) of Table 5 are in line with those obtained for the classical random effects specification. We note, however, that the estimated effects of "median weekly earnings" and "people aged under 14" turn out to be significant with a positive sign in the spatial lag random effects and in the spatial error random effects models, respectively. There exists a positive relation between wage and spending, which seems reasonable since earnings are associated with service unit costs.

The spatial coefficients in both spatial models are positive and statistically significant, and therefore, the z-test rejects the null hypothesis of absence of spatial interactions. In particular, the spatial lag dependent variable is significant with parameter $\rho = 0.16$, that is a 1% increase in expenditure in neighbouring localities implies a rise of 0.16% in spending. Recalling the hypotheses discussed in Section 2, the estimated $\rho$ indicates possible mimicking behaviour, the so-called demonstrative effect, generated for example by learning and information spillovers among contiguous municipalities. The spatial error has a significant coefficient of $\lambda = 0.12$. We note that while the spatial coefficient $\rho$ has an intuitive and direct economic interpretation, the parameter $\lambda$ synthetizes the indirect effect on spending of non-observable variables or environmental risks that are difficult to measure.

Since the two spatial processes tend to be very similar, we distinguish between them by evaluating and comparing their log-likelihood values with that of the classic random effects model (Greene, 2003). Given that the
log-likelihood values of the spatial error random effects and of the classic random effects models are respectively 202.8 and 202.3, we conclude that they are insignificantly different, on the basis of the chi-squared test statistic with one degree of freedom. On the contrary, the log-likelihood of the spatial lag random effects is 206.3, which is significantly different from 202.3. We finally performed a Hausman test to contrast the fixed and random effects specification of the spatial lag model. As we obtain a value of 3.28, with a $p$-values of 0.99, we selected the random effects model with a spatial lag dependent variable as superior in explaining the process of mental health expenditure determination.

The size of the spatial effect in the dependent variable is smaller than the effects found by Moscone and Knapp (2005) and Revelli (2002, 2006). In the former, a cross-section study of log per-capita mental health social services expenditure by authorities in England detected $\rho = 0.25$. Similarly, Revelli (2002), by using English data on local authorities to explain total expenditure, observed $\rho = 0.24$. However, we note that, in contrast to these studies, we use a log-log model, where both the dependent variable and the regressors are expressed in logarithms.

Finally, we observe that the size of the spatial effect in the random effect spatial lag specification is very similar to that in the spatial model estimated year by year (Table 4).

We now turn to the estimation of spatial models, using a wide range of weights matrices. We focus our attention only on the random effects model with spatially lagged dependent variable, since this was selected in the previous analysis as the most powerful model explaining mental health expenditure. The aim is to explore whether there is evidence of spatial patterns which depart from the simple geographical contiguity assumption.

The first three columns of Table 6 present the estimation of spatial random effects models, where we employed spatial weights matrices based on geographical contiguity weighted by population, population density, and political characteristics (see Section 4). Accordingly, for each local authority, more weight is attached to geographically contiguous municipalities with larger size, which is approximated by population and population density, or similar ideology. To avoid potential endogeneity in the contruction of weights matrices, we used information in the first period of our time span, namely the year 1998, rather in the current year.

The use of these spatial weights corroborates our previous findings of a significant and positive spatial effect between municipalites. It is interesting to observe that the size of the spatial effect remains significant, whether we take a straight average as in the pure contiguity case, or an average
weighted by a set of characteristics. However, we should emphasise that this result does not contradict the finding of an insignificant political effect on spending. Indeed, the spillover effect in this case is likely to measure local interaction among contiguous municipalities that belong to the same political party. This can be consistent with the evidence of an overall, non-spatial effect of ideology on spending. The remaining columns of Table 6 show the estimation of spatial random effects models, where we employed non-geographical distance on the basis of population and population density (see Section 4). It emerges that, when space is intended in a general sense, based on demographic characteristics, there is no evidence of spillover effects.

Results are consistent with some degree of policy interdependence between neighbouring municipalities, corroborating previous findings in public economics, synthetized in Section 3. However, given Manski’s reflection problem, our empirical evidence might also indicate the effect of contextual characteristics on spending decisions of a local authority, or what we have previously defined as a directive effect.

It is important to stress some limitations to our empirical study. First, given that the statistical unit of our analysis is the council, we would expect variations within each administrative area. Therefore, subsequent analysis would benefit from more disaggregated data, and the accompanying use of "spatial" multilevel techniques. Second, we observed temporal variation in the impact of some determinants, considering therefore time-specific effects in our work. Future work should be directed to the use of models that incorporate dynamics in spending, and control for unobserved heterogeneity.

Third, though we have used different measures of neighbourliness, these are certainly not exhaustive. Therefore, we cannot rule out the existence of alternative specifications of potential interaction based on economic and policy distances.

At this stage in our work we have not been able to explore the underlying reasons for the spatial interdependence. Some of the hypothesised influences or pathways of influence might be amenable to statistical testing if appropriate data could be located (for example on psychiatric hospital closures or regional directive effects embodied in specific service targets). Others would need a more qualitative form of interrogation, understanding the behaviour of key actors and the decision-making processes of elected councils. These further developments are beyond the scope of the present paper and may be the subject of future work.
7 Conclusions

In this paper we have explored the possibility that policy makers’ choices concerning mental health expenditure are interdependent. In particular, we offered a range of motivations underlying possible interactions between local authorities when deciding their own level of social care spending on mental health. This led us to suggest two econometric specifications to reflect these hypothesis. Using data on English personal social service mental health expenditure by local authority, we explored a number of the sources of spatial autocorrelation in local government spending decisions. The results of the estimated models indicate that spatial autocorrelation characterises local expenditure decisions, and support both hypothesis of mimicking behaviour among policy makers and contextual effects. These results still hold when considering patterns of spatial interaction that are more complex (and more interesting) than simple contiguity. Due to Manski’s reflection problem we are not able to provide a firm answer on which among these is the cause of the underlying spatial pattern in spending.

However, these results may help central government to gain a better understanding of the factors that influence local spending levels, including variations over time and between municipalities in their achievement of expenditure-related and perhaps other performance targets. There seems little doubt that positive interdependence is an important feature of decision making.

7.1 Appendix

Random effects model with spatially lagged dependent variable.

Consider (4)-(5), expressed in matrix form

\[ y = \rho (I_T \otimes W) y + X \beta + e, \quad (9) \]
\[ e = (i_T \otimes \mu) + \varepsilon, \quad (10) \]

where \( y = (y_1', \ldots, y_T')' \), \( X = (X_1', \ldots, X_T')' \), \( \varepsilon = (\varepsilon_1', \ldots, \varepsilon_T')' \), \( i_T \) is a \( T \times 1 \) vector of one, and \( \otimes \) is the Kronecker product. Assume \( \varepsilon \sim N(0, \sigma_\varepsilon^2 I_{NT}) \), and define \( A = I_N - \rho W \) and \( \theta = \frac{\sigma^2}{\sigma_\varepsilon^2 + \sigma^2} \), the log-likelihood function of model (9)-(10) is

\[ L = -\frac{NT}{2} \ln (2\pi \sigma^2_\varepsilon) - \frac{N}{2} \ln \theta + T \ln |A| - \frac{1}{2\sigma^2_\varepsilon} e' [I_{NT} - (1 - \theta) \frac{1}{T} (i_T i_T') \otimes I_N] e, \]
where \( e = y - \rho (I_T \otimes W) y - X \beta \). The parameters \( \beta \) and \( \sigma^2_\varepsilon \) can be solved from their first order maximizing conditions (Elhorst, 2003)

\[
\hat{\beta} = (X'^*X^*)^{-1} X'^*y^*,
\]

(11)

\[
\hat{\sigma}^2_\varepsilon = \frac{e^*e^*}{NT},
\]

(12)

where \( e^* = y^* - X^*\hat{\beta} \), and

\[
y^* = [I_T \otimes A] \left[ y - \left( 1 - \sqrt{\theta} \right) (I_T \otimes \bar{y}) \right],
\]

\[
X^* = X - \left( 1 - \sqrt{\theta} \right) (I_T \otimes \bar{X}).
\]

with \( \bar{y} = (\bar{y}_1, \ldots, \bar{y}_N)' \), \( \bar{X} = (\bar{x}_1, \ldots, \bar{x}_N)' \). For a given value of \( \rho \) and \( \theta \), \( \hat{\beta} \) is a GLS estimator and can be computed by running the standard OLS regression on the transformed variables \( y^* \) and \( X^* \). From the above equations, note that as \( \theta \to 0 \), \( \hat{\beta} \) uses only the temporal variation of each cross-section unit, and therefore is identical to the Least Square Dummy Variable estimator in a spatial panel. Conversely, as \( \theta \to 1 \), \( \hat{\beta} \) only considers the variation in the cross-section, and hence converges to the between-municipalities estimator.

The estimators for \( \beta \) and \( \sigma^2_\varepsilon \) are both functions of \( \rho \) and \( \theta \), and can be obtained applying the following two-stage iterative procedure. We fix an initial value for \( \rho \) and \( \theta \) and compute \( \hat{\beta} \) and the corresponding residuals \( e^* \). Hence, in a second stage we find \( \rho \) and \( \theta \) that maximize the concentrated likelihood \( L_C \), which is, ignoring the constant terms

\[
L_C = -\frac{NT}{2} \log \left( e^*e^* \right) + \frac{N}{2} \ln \theta + T \ln |A|.
\]

(13)

Given \( \hat{\rho} \) and \( \hat{\theta} \), we can compute \( \hat{\beta} \) and a new set of residuals \( e^* = y^* - X^*\hat{\beta} \). We can alternate back and forth between the estimation of \( \rho \) and \( \theta \) conditional upon a vector of residuals \( e^* \) (generated for a value of \( \hat{\beta} \)) and the estimation of \( \beta \) and \( \sigma^2_\varepsilon \) conditional upon a value for \( \rho \) and \( \theta \), until convergence is obtained.

**Random effects model with spatial error correlation.**

Consider (6)-(8) expressed in matrix form

\[
y = X\beta + e,
\]

(14)

\[
e = (I_T \otimes \mu) + [I_T \otimes (I_N - \lambda W)^{-1}] \varepsilon,
\]

(15)

21
where the notation is as above. Defining $\mathbf{B} = \mathbf{I}_N - \lambda \mathbf{W}$, and $\eta = \sigma^2_{\mu} / \sigma^2_{\varepsilon}$, the log-likelihood function of a random effects model is

$$L = -\frac{NT}{2} \ln \left(2\pi \sigma^2_{\varepsilon}\right) + (T - 1) \ln |\mathbf{B}| - \frac{1}{2} \ln \left|T\eta \mathbf{I}_N + (\mathbf{B}' \mathbf{B})^{-1}\right| \quad (16)$$

$$-\frac{1}{2\sigma^2_{\varepsilon}} \mathbf{e}' \left[\frac{1}{T} \mathbf{i}_T \mathbf{v}_T \otimes \left[T\eta \mathbf{I}_N + (\mathbf{B}' \mathbf{B})^{-1}\right]^{-1}\right] \mathbf{e}$$

$$-\frac{1}{2\sigma^2_{\varepsilon}} \mathbf{e}' \left[(\mathbf{I}_T - \frac{1}{T} \mathbf{i}_T \mathbf{v}_T) \otimes (\mathbf{B}' \mathbf{B})\right] \mathbf{e},$$

where $\mathbf{e} = \mathbf{y} - \mathbf{X} \hat{\mathbf{b}}$, and $\mathbf{B} = \mathbf{I}_N - \lambda \mathbf{W}$. Again, the maximum likelihood estimates can be obtained using a two stage iterative procedure that alternates the estimate of $\beta$ and $\sigma^2_{\varepsilon}$ on one side, and $\lambda, \eta$ on the other. The first-order maximizing conditions are

$$\hat{\mathbf{b}} = (\mathbf{X}' \mathbf{X}^*)^{-1} \mathbf{X}' \mathbf{y}^*, \quad (17)$$

$$\hat{\sigma}^2_{\varepsilon} = \frac{\mathbf{e}' \mathbf{e}^*}{NT}, \quad (18)$$

where $\mathbf{e}^* = \mathbf{y}^* - \mathbf{X}^* \hat{\mathbf{b}}$, and

$$\mathbf{y}^* = (\mathbf{I}_T \otimes \mathbf{B}) \mathbf{y} - \mathbf{I}_T \otimes (\mathbf{P} - \mathbf{B}) \bar{\mathbf{y}},$$

$$\mathbf{X}^* = (\mathbf{I}_T \otimes \mathbf{B}) \mathbf{X} - \mathbf{I}_T \otimes (\mathbf{P} - \mathbf{B}) \bar{\mathbf{X}},$$

with $\mathbf{P}$ being the Cholesky decomposition of $[T\eta \mathbf{I}_N + (\mathbf{B}' \mathbf{B})^{-1}]^{-1}$. Upon substituting $\hat{\mathbf{b}}$ and $\hat{\sigma}^2_{\varepsilon}$ in the log-likelihood function, the concentrated log-likelihood function for $\lambda$ and $\eta$ is, ignoring the constants

$$L_C = -\frac{NT}{2} \log \left(\mathbf{e}' \mathbf{e}^*\right) + (T - 1) \ln |\mathbf{B}| - \frac{1}{2} \ln \left|T\eta \mathbf{I}_N + (\mathbf{B}' \mathbf{B})^{-1}\right|. \quad (19)$$

Acknowledgment: We would like to thank participants at the 14th Workshop in Econometrics and Health Economics in Dublin, Mario Tosetti and two anonymous referees for valuable comments and suggestions.

References


22


[26] Elhorst J.P. (2003), Specification and Estimation of Spatial Panel Data Models, International Regional Science Review, 26, 244-268


[38] Kelejian, H.H., Robinson D. (1993), A suggested method of estimation for spatial interdependence models with autocorrelated errors, and an application to a county expenditure model, Papers in Regional Science, 72, 297-312


[42] Kryter K.D. (1990), Association of Heathrow Airport Noise with psychiatric admissions, Psychological Medicine, 20, 1022

[44] LeSage J.P. (1999), The Theory and Practice of Spatial Econometrics, Department of Economics, University of Toledo


[61] Revelli F. (2001), Spatial Patterns in local Taxation: Tax Mimicking or Error Mimicking?, Applied Economics 33, 1101-1107


Table 1: Descriptive statistics on the selected variables (years 1998-2003)(*).

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. Err.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSS per-capita exp. (£)</td>
<td>13.33</td>
<td>6.02</td>
<td>3.10</td>
<td>53.42</td>
</tr>
<tr>
<td>Population Density</td>
<td>2.43</td>
<td>2.67</td>
<td>0.06</td>
<td>14.53</td>
</tr>
<tr>
<td>% of Males</td>
<td>48.88</td>
<td>0.72</td>
<td>46.61</td>
<td>51.45</td>
</tr>
<tr>
<td>% People&lt;14</td>
<td>17.88</td>
<td>3.11</td>
<td>9.05</td>
<td>26.51</td>
</tr>
<tr>
<td>Mortality Rate</td>
<td>1.01</td>
<td>0.18</td>
<td>0.57</td>
<td>1.54</td>
</tr>
<tr>
<td>Unemployment Claimants</td>
<td>1.58</td>
<td>0.74</td>
<td>0.30</td>
<td>5.53</td>
</tr>
<tr>
<td>% Lone Parents</td>
<td>6.36</td>
<td>2.31</td>
<td>2.90</td>
<td>14.80</td>
</tr>
<tr>
<td>N. of Jobs</td>
<td>122</td>
<td>93</td>
<td>11</td>
<td>458</td>
</tr>
<tr>
<td>House Price</td>
<td>119,269</td>
<td>72,935</td>
<td>36,870</td>
<td>673,092</td>
</tr>
<tr>
<td>Median Weekly Earnings</td>
<td>310</td>
<td>60</td>
<td>181</td>
<td>586</td>
</tr>
</tbody>
</table>

(*): N=888

Table 2: Shapiro-Wilks tests for normality of PSS per-capita mental health expenditure (in logs).

| Year | N. obs | W   | V   | z    | Pr>|z|
|------|--------|-----|-----|------|-----|
| 1998 | 148    | 0.993 | 0.840 | -0.396 | 0.654 |
| 1999 | 148    | 0.990 | 1.157 | 0.330 | 0.371 |
| 2000 | 148    | 0.976 | 2.786 | 2.321 | 0.010 |
| 2001 | 148    | 0.994 | 0.694 | -0.829 | 0.796 |
| 2002 | 148    | 0.996 | 0.430 | -1.913 | 0.972 |
| 2003 | 148    | 0.980 | 2.278 | 1.865 | 0.031 |
| All years | 888 | 0.998 | 1.173 | 0.392 | 0.347 |
Figure 1: Kernel density of PSS per-capita mental health expenditure (in logs, years 1998-2003).

Table 3: Connectivity characteristics of the spatial weights matrix.

<table>
<thead>
<tr>
<th></th>
<th>Contig. by pop.</th>
<th>Contig. by pop. dens.</th>
<th>Contig. by pol. party</th>
<th>Pop./Pop. dens.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimension</td>
<td>148</td>
<td>148</td>
<td>148</td>
<td>148</td>
</tr>
<tr>
<td>% non-zero weights</td>
<td>3.44</td>
<td>3.44</td>
<td>3.44</td>
<td>2.00</td>
</tr>
<tr>
<td>N. non-zero links</td>
<td>748</td>
<td>748</td>
<td>748</td>
<td>436</td>
</tr>
<tr>
<td>Max. n. of links</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Average n. links</td>
<td>5.05</td>
<td>5.05</td>
<td>5.05</td>
<td>2.95</td>
</tr>
<tr>
<td>Average link</td>
<td>0.198</td>
<td>0.197</td>
<td>0.197</td>
<td>0.300</td>
</tr>
</tbody>
</table>
Figure 2: Quantile distribution of PSS per-capita mental health expenditure(*) (in logs, years 1998-2003)

(*) : GeoDa was used to generate the mapping (free download at http://sal.geoda.uiuc.edu/default).
Figure 3: Plot of the dependent variable against its spatial lag
<table>
<thead>
<tr>
<th></th>
<th>1998</th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff.</td>
<td>z-value</td>
<td>Coeff.</td>
<td>z-value</td>
<td>Coeff.</td>
<td>z-value</td>
</tr>
<tr>
<td>Population Density</td>
<td>0.1156</td>
<td>4.0618</td>
<td>0.1312</td>
<td>4.7466</td>
<td>0.0920</td>
<td>3.5684</td>
</tr>
<tr>
<td>% of Males</td>
<td>1.0251</td>
<td>0.3733</td>
<td>-2.0837</td>
<td>-0.8042</td>
<td>1.2536</td>
<td>0.5889</td>
</tr>
<tr>
<td>% People&lt;14</td>
<td>0.8807</td>
<td>1.8273</td>
<td>0.6100</td>
<td>1.4236</td>
<td>0.2487</td>
<td>0.6744</td>
</tr>
<tr>
<td>Mortality Rate</td>
<td>0.5641</td>
<td>1.7640</td>
<td>0.4227</td>
<td>1.2898</td>
<td>0.3528</td>
<td>1.2882</td>
</tr>
<tr>
<td>Unemployment Claim.</td>
<td>0.0544</td>
<td>0.5764</td>
<td>0.0393</td>
<td>0.4196</td>
<td>0.0790</td>
<td>0.9166</td>
</tr>
<tr>
<td>% Lone Parents</td>
<td>0.2407</td>
<td>1.5130</td>
<td>0.2474</td>
<td>1.6380</td>
<td>0.3142</td>
<td>2.2371</td>
</tr>
<tr>
<td>N. of Jobs</td>
<td>-0.2024</td>
<td>-0.6902</td>
<td>-0.0095</td>
<td>-0.0346</td>
<td>-0.2540</td>
<td>-1.0146</td>
</tr>
<tr>
<td>House Price</td>
<td>0.2510</td>
<td>1.5782</td>
<td>0.1331</td>
<td>0.9733</td>
<td>0.1708</td>
<td>1.6743</td>
</tr>
<tr>
<td>Median Weekly Earn.</td>
<td>-0.0056</td>
<td>-0.1356</td>
<td>-0.0196</td>
<td>-0.5066</td>
<td>-0.0976</td>
<td>-2.6354</td>
</tr>
<tr>
<td>Labour</td>
<td>-0.1500</td>
<td>-0.2086</td>
<td>-0.0329</td>
<td>-0.4739</td>
<td>-0.0521</td>
<td>-0.7919</td>
</tr>
<tr>
<td>Liberal Democrat</td>
<td>0.0828</td>
<td>0.9659</td>
<td>0.0550</td>
<td>0.0661</td>
<td>-0.0143</td>
<td>-0.1854</td>
</tr>
<tr>
<td>Intercept</td>
<td>-7.4221</td>
<td>-0.6902</td>
<td>-0.0095</td>
<td>-0.0346</td>
<td>0.0503</td>
<td>0.0991</td>
</tr>
<tr>
<td>( \rho )</td>
<td>0.1878</td>
<td>1.9202</td>
<td>0.1096</td>
<td>1.1266</td>
<td>0.1508</td>
<td>1.5694</td>
</tr>
</tbody>
</table>

**Table 4:** Maximum likelihood estimation of spatial lag and spatial error coefficients by year (dep. var. PSS per-capita expenditure).
Table 5: OLS, random effects, spatial lag and spatial error random effects regressions (*) (dep. var. PSS per-capita expenditure).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff.</td>
<td>t-value</td>
<td>Coeff.</td>
<td>z-value</td>
</tr>
<tr>
<td>Population Density</td>
<td>0.1209</td>
<td>11.0623</td>
<td>0.1261</td>
<td>5.5046</td>
</tr>
<tr>
<td>% of Males</td>
<td>-0.2996</td>
<td>-0.3297</td>
<td>-0.6456</td>
<td>-0.5767</td>
</tr>
<tr>
<td>% People&lt;14</td>
<td>0.4195</td>
<td>2.8096</td>
<td>0.2653</td>
<td>1.2701</td>
</tr>
<tr>
<td>Mortality Rate</td>
<td>0.5334</td>
<td>4.5723</td>
<td>0.3550</td>
<td>2.1614</td>
</tr>
<tr>
<td>Unemployment Claim.</td>
<td>0.1042</td>
<td>2.8592</td>
<td>0.0795</td>
<td>1.7520</td>
</tr>
<tr>
<td>% Lone Parents</td>
<td>0.2188</td>
<td>3.7608</td>
<td>0.2044</td>
<td>1.9976</td>
</tr>
<tr>
<td>N. of Jobs</td>
<td>0.0118</td>
<td>0.1179</td>
<td>0.1980</td>
<td>1.8325</td>
</tr>
<tr>
<td>House Price</td>
<td>0.1822</td>
<td>4.0153</td>
<td>0.0337</td>
<td>0.5762</td>
</tr>
<tr>
<td>Median Weekly Earn.</td>
<td>-0.0448</td>
<td>-2.8493</td>
<td>-0.0414</td>
<td>-1.3547</td>
</tr>
<tr>
<td>Labour</td>
<td>-0.0353</td>
<td>-1.0885</td>
<td>-0.0469</td>
<td>-1.2024</td>
</tr>
<tr>
<td>Liberal Democrat</td>
<td>-0.0202</td>
<td>-0.6267</td>
<td>-0.0633</td>
<td>1.3390</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.1544</td>
<td>1.0412</td>
<td>2.5264</td>
<td>0.5681</td>
</tr>
<tr>
<td>$\rho$</td>
<td>-</td>
<td>-</td>
<td>0.1579</td>
<td>2.6049</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>-</td>
<td>-</td>
<td>0.2455</td>
<td>0.2569</td>
</tr>
<tr>
<td>$\theta$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$\eta$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3.0962</td>
</tr>
<tr>
<td>Log-Lik</td>
<td>-211.1</td>
<td>202.3</td>
<td>206.3</td>
<td>202.8</td>
</tr>
</tbody>
</table>

(*) In all regressions we included time dummies.
Table 6: Spatial lag random effects regressions with different weights matrices (*)

(dep. var. PSS per-capita expenditure).

<table>
<thead>
<tr>
<th>Neigh. def. by:</th>
<th>Contiguity by population</th>
<th>Contiguity by population density</th>
<th>Contiguity by political party</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff.</td>
<td>z-value</td>
<td>Coeff.</td>
</tr>
<tr>
<td>Population Density</td>
<td>0.1197</td>
<td>5.4884</td>
<td>0.1167</td>
</tr>
<tr>
<td>% of Males</td>
<td>-0.2250</td>
<td>-0.2043</td>
<td>-0.2482</td>
</tr>
<tr>
<td>% People&lt;14</td>
<td>0.2680</td>
<td>1.1355</td>
<td>0.2165</td>
</tr>
<tr>
<td>Mortality Rate</td>
<td>0.3724</td>
<td>2.9784</td>
<td>0.3501</td>
</tr>
<tr>
<td>Unemployment Claim.</td>
<td>0.0754</td>
<td>1.6902</td>
<td>0.0767</td>
</tr>
<tr>
<td>% Lone Parents</td>
<td>0.1820</td>
<td>1.9773</td>
<td>0.2171</td>
</tr>
<tr>
<td>N. of Jobs</td>
<td>-0.0392</td>
<td>-1.3792</td>
<td>-0.0406</td>
</tr>
<tr>
<td>House Price</td>
<td>0.0154</td>
<td>0.7607</td>
<td>0.0110</td>
</tr>
<tr>
<td>Median Weekly Earn.</td>
<td>0.1820</td>
<td>2.1759</td>
<td>0.1874</td>
</tr>
<tr>
<td>Labour</td>
<td>-0.0466</td>
<td>-1.2444</td>
<td>-0.0525</td>
</tr>
<tr>
<td>Liberal Democrat</td>
<td>0.0655</td>
<td>1.4138</td>
<td>0.0609</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.8444</td>
<td>0.1909</td>
<td>1.1236</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.1463</td>
<td>2.4197</td>
<td>0.1168</td>
</tr>
<tr>
<td>$\theta$</td>
<td>0.2546</td>
<td></td>
<td>0.2557</td>
</tr>
<tr>
<td>Log-Lik</td>
<td>205.1</td>
<td></td>
<td>205.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Neigh. def. by:</th>
<th>Population</th>
<th>Population density</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff.</td>
<td>z-value</td>
</tr>
<tr>
<td>Population Density</td>
<td>0.123</td>
<td>5.419</td>
</tr>
<tr>
<td>% of Males</td>
<td>-0.612</td>
<td>-0.861</td>
</tr>
<tr>
<td>% People&lt;14</td>
<td>0.266</td>
<td>1.177</td>
</tr>
<tr>
<td>Mortality Rate</td>
<td>0.362</td>
<td>2.770</td>
</tr>
<tr>
<td>Unemployment Claim.</td>
<td>0.080</td>
<td>1.770</td>
</tr>
<tr>
<td>% Lone Parents</td>
<td>0.209</td>
<td>2.053</td>
</tr>
<tr>
<td>N. of Jobs</td>
<td>-0.037</td>
<td>-1.271</td>
</tr>
<tr>
<td>House Price</td>
<td>0.040</td>
<td>0.756</td>
</tr>
<tr>
<td>Median Weekly Earn.</td>
<td>0.198</td>
<td>1.838</td>
</tr>
<tr>
<td>Labour</td>
<td>-0.046</td>
<td>-1.178</td>
</tr>
<tr>
<td>Liberal Democrat</td>
<td>-0.046</td>
<td>-1.178</td>
</tr>
<tr>
<td>Intercept</td>
<td>1.963</td>
<td>0.542</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.1178</td>
<td>1.1535</td>
</tr>
<tr>
<td>$\theta$</td>
<td>0.2539</td>
<td></td>
</tr>
<tr>
<td>Log-Lik</td>
<td>200.1</td>
<td></td>
</tr>
</tbody>
</table>

(*): In all regressions we included time dummies