



Entering the lists: what can be learned from local listings of third sector organisations? Results from a study of Northern England

John Mohan, David Kane, Julia Branson, Steve Barnard

UNIVERSITY OF
Southampton



www.nr-foundation.org.uk/thirdsectortrends

Entering the lists: what can be learned from local listings of third sector organisations?
Results from a study of Northern England

Acknowledgements: We are hugely grateful to a large number of organisations and individuals who supplied us with copies of their local listings of third sector organisations. We would also like to acknowledge the work of Mark Crowe and Ruth Beattie (Yorkshire and Humberside Regional Forum) who undertook the task of contacting these organisations and persuading them to share their information with us. Data on regulated third sector organisations was supplied by Guidestar Data Services. We would also like to acknowledge comments and suggestions made by participants in the Third Sector Trends study including Carol Candler, Tony Chapman, Mark Crowe, Cathy Pharaoh, Fred Robinson, Penny Wilkinson and Rob Williamson. John Mohan would also like to acknowledge the assistance of various colleagues in the Third Sector Research Centre, notably Andrew McCulloch, David Clifford, and Peter Backus. Finally, we are grateful for funding provided by the Northern Rock Foundation and by Yorkshire and Humberside Regional Forum.

Contents

1. Introduction

2. Sources of data

3. Analysis

3.1 Below-radar organisations

3.2 Classifying below-radar organisations

3.3 Differences between TSOs that appear on local listings, and those that do not

3.3.1 Organisations that appear on any local listings

3.3.2 Organisations that appear on their local CVS listing

4. Conclusions

5. Implications for policy

Appendix 1 – Examples of lists gathered

Appendix 2 - Techniques for excluding organisations, and deduplication of combined listings

Appendix 3 – Keywords used to classify organisations

1 Introduction

In our previous reports we presented a body of evidence about the characteristics of regulated third sector organisations, and of the associated workforce and volunteers, in Cumbria, the North East, and Yorkshire and Humberside, but there was also great interest in the characteristics of the population of organisations which is unregulated and sometimes termed below-radar in the sense that organisations have not needed to or have chosen not to adopt one or other of the range of available non-profit legal forms. We wanted to explore the evidence about how many such organisations existed locally in these regions, and to determine what differences there were, if any, between the characteristics and distribution of these organisations, and the characteristics and distribution of the regulated third sector.

In our interim report on this work¹ we described the methods we had used to identify duplicate entries in our dataset compiled from a large number of local and regional listings of third sector organisations (TSOs), gave some indications of initial findings, and outlined the way in which we were going to develop the analysis. The intention of this element of our work had been to attempt, as far as was possible from these sources, to quantify the size and characteristics of the below-radar population of third sector organisations in the northern regions and Cumbria. However, as we show in this paper, on the basis of an extensive analysis of the organisations represented on local listings from infrastructure bodies, much of the variation in the apparent size of that population is in fact due to the types of organisations captured by the compilers of those lists. Hence the emphasis shifts much more towards describing what can, and can't, be learned from these

¹ Mohan, J., Kane, D., Branson, J., and Owles, F. (2009) Beyond 'flat-earth' maps of the third sector. <http://www.nrf-foundation.org.uk/thirdsectortrends/downloads/NRF%20TST%20Report%20Beyond%20Flat%20Earth.pdf>

local listings. In a companion publication, we also report on the results of a survey of such organisations, which we compare with similar work done as part of the National Survey of Third Sector Organisations (NSTSO).

It may be worth briefly repeating why we took the approach that we did and how it compares with other ways of obtaining information about the organisations that do not appear on regulatory "radars" (there is a fuller discussion of this question in our earlier report). An alternative method would have been something along the lines of the LOVAS (Local Authority Voluntary Action Studies) research, in which lists of organisations are assembled from local respondents through a snowball sampling methodology. This method would work well for small areas -- most of the communities studied in the LOVAS programme were local authorities or even small areas within local authorities. It would simply not be feasible for an area the size of our combined regions. At the other end of the scale, the local area profiles methodology developed by NAVCA and the Audit Commission requires collation of every feasible listing for individual local authorities but envisages that a programme of work lasting up to a year would be involved. Our approach sits between these two - its advantage is that by attempting to gain a representative set of listings from a broad area we thought that a broadly consistent and comparable pattern would emerge which was representative of the kinds of organisation that had relationships with infrastructure bodies. Inevitably this means that very small and informal community groups will be missed out but extremely intensive methods are necessary to pick those up.

We believe that our approach generates useful information on the characteristics of organisations that appear on the listings of local infrastructure and other umbrella bodies. We analyse the geographical patterns of these organisations and their relationship with the distribution of the regulated third sector, and we classify them so that we have a better idea of the differences. However, if this information is to be of value to the sector then it is our responsibility to point out its strengths and weaknesses. One of the striking features of the work proved to

Entering the lists: what can be learned from local listings of third sector organisations?
Results from a study of Northern England

be the differences in the numbers and types of organisations captured by local listings. Hence the title of this report - our emphasis is very much on what useful information can be learned from these sources. We don't believe it would be defensible to claim that these sources give an authoritative and accurate portrayal of the numbers of below-radar organisations in an area but we do believe that they give us some useful information about broad characteristics of these organisations. But the use of that information - for example in discussions about the need for support for the voluntary sector - needs to be informed by a consideration of its limitations.

In the Middle Ages, "entering the lists" was what knights did to register that they were challenging a champion in a tournament. Although confronting the problematic nature of local listings of voluntary organisations does not compare with facing a fully-armoured opponent on horseback, there were times, as we explored our combined listings, when the complexity of deduplication routines and classification processes seemed equally daunting. The electronic equivalent of a ball of string was necessary at times; very meticulous recording of programming decisions is essential in this work. The complexity of classifying some 60,000 organisations from scratch has absorbed very considerable effort. Although some of this was described in our previous paper, we have developed our methodology since then and we therefore present some information on how we have improved it. However, we have tried to keep methodological discussion to a minimum; appendix 2, plus our interim report ("Beyond flat-earth maps") provide sufficient detail on the methods we have used.

We had a series of questions about the lists of local organisations gathered from a variety of sources:

- Which organisations feature on local lists and conversely, which organisations are not on them?
- How comprehensive are these listings? That is, to what extent does the information on these lists provide full coverage of the local third sector population?

- What variations are there between local listings in terms of the numbers and types of organisations on them? For example do some listings differ from others because they capture information about different entities (schools, churches, government bodies)?
- For registered third sector organisations, what differences are there (e.g. size, type) between those organisations that appear on local listings and those that do not?
- What are the key differences between the registered third sector population and the so-called below-radar population, e.g. in terms of geography and other characteristics?
- Is it possible to estimate the size of the below-radar population from these sources?

The exercise reported here is by far the largest of its kind conducted in England to date. There are academic studies of single cities which have relied almost entirely on one or two listings provided by the relevant local council for voluntary service (CVS) or by a local authority.² As we show, there are substantial variations between listings in the mix of organisations so we cannot assume that an individual listing is in any way typical. Local pilot exercises for the 2008 NSTSO study³ focused on the intensive examination of one local list maintained by a relevant infrastructure body in each of three localities (Bury, Gloucestershire, and Hackney) whereas, in contrast, we have over 20 listings from CVSs as well as a large number of other local listings of varying sizes, contents and quality.

² Fyfe, N. and Milligan, C. (2003) 'Space, citizenship and the 'shadow state': Exploring the voluntary welfare sector in Glasgow' *Environment and Planning C: Government and Policy*, 35

Maloney, W., Graham Smith and Gerry Stoker (2000) Social Capital and Urban Governance: Adding a More Contextualised 'Top-Down' Perspective. *Political Studies*, 48, 4, 802-820

Smith G, Stoker G, Maloney WA. [Building Social Capital in City Politics: Scope and Limitations at the Inter-Organisational Level](#). *Political Studies* 2004, 52(3), 508-530.

³ <http://www.nscsesurvey.com/pilots/>

Entering the lists: what can be learned from local listings of third sector organisations?
Results from a study of Northern England

In addition we pursue questions not attempted by the NSTSO study. Principal among these is an attempt to assess the extent to which registered third sector organisations featured on local listings: according to the NSTSO study, "a significant proportion" of organisations on local listings were in fact registered third sector organisations. That study argued that considerable effort would need to be made to achieve an assessment of this question, but we have successfully managed to match large numbers of organisations using automated methods.

A second original contribution of our work has been an attempt at classification of the unregistered third sector organisations; the NSTSO pilot used the responses to their questionnaire to classify organisations, whereas we used (in the absence of anything else) the names of the organisations, to classify over 30,000 entities. This allows for a more robust assessment of the characteristics of this population than has been possible in other studies to date. We are aware of no other study of this scale and complexity.

In previous work on this project we have produced separate reports for North East England, Cumbria and Yorkshire and Humberside. We have not followed that practice here because relatively small numbers of listings were supplied for some parts of these regions and therefore collation of the material into one report protects the anonymity of the organisations which supplied us with listings. It also gives us a much more robust basis for the conclusions we draw.

2 Sources of data

We received over 80 lists ranging in size from tens to several thousand entities. In comparison to previous report, here we give much more detail about the types of lists received and the characteristics of organisations which appeared on them.

Table 1: Summary of lists received

Type of list	Total records on list	Proportion of all records (%)	Number of lists	List size	
				Minimum	Maximum
CVS	18,394	31.2	30	137	2,136
Other local	306	0.5	7	18	85
Local Authorities	4,508	7.6	21	3	1,892
Community Foundations	3,118	5.3	2	911	2,207
National organisation	3,108	5.3	15	24	1,689
Regulator	20,450	34.7	2	5,365	15,085
Trade directory	1,194	2.0	2	206	1,228
Regional Infrastructure	3,388	5.7	9	3	1,228
Other – research	4,532	7.7	1	4,532	4,532
Total	58,998	100.0	89	3	15,085

We summarise the principal types of list in table 1 which also gives descriptive statistics for the number of entries on each listing and a headline summary of the coverage of the data. Numerically the largest sources were the Register of Charities and Companies House⁴. The largest number of listings - 30 - were supplied either directly from CVSs, or were extracted from larger databases which had been built up from local CVS listings and identified their initial source. These supplied over 18,000 records. A further 21 source lists were provided by local authorities or downloaded from publicly-available sources, and 15 more were

⁴ supplied by GuideStar Data Services

Entering the lists: what can be learned from local listings of third sector organisations?
Results from a study of Northern England

provided by national umbrella or membership organisations. The lists varied considerably in size, including 20 which contributed fewer than 50 entries to our combined database, and several of these only had a small number of entries within these regions.

Table 2 presents a summary of the main types of organisation found in the listings and shows the effect of different criteria for selection; from a total of just under 59,000 entities in total, there are around 41,400 unique organisations within the region for which sufficiently detailed postcoded data are available for linkage, at a small-area level, to small-area census data.

Table 2: Entries by type of organisation

Organisation type	Total entries	...of which in study area	...of which unique	...of which with postcode
No legal form	26,830	25,091	22,297	18,677
Charities	18,723	16,739	14,784	14,334
Charitable Companies	5,044	4,875	2,992	2,853
CLG	5,038	4,546	3,860	3,679
IPS	2,144	2,046	1,883	1,849
Third Sector	57,779	53,297	45,816	41,392
Public sector	893	871	797	703
Individuals	154	142	94	80
Private sector	172	161	157	139
Total	58,998	54,471	46,864	42,314

The table in Appendix I gives some more detailed data for around 40 anonymised lists each of which originated within the region or which contained substantial numbers of organisations from it; the table shows those with at least 100 entries on their list and gives a classification by legal form. Even this tabulation begins to raise questions as it shows some clear points of difference between the lists. For example, there are a couple of local CVS lists which include substantial numbers of public sector organisations such as schools or Parish Councils, accounting for well over 10% of entries, and judging from the numbers in this table, these CVSs must

have made a conscious decision to include this information. Reflecting the fact that these are in effect the address books of local organisations, there are also numbers of public and private sector organisations and even individuals on the lists as well; examples would include local authorities and departments of local authorities, NHS bodies such as hospital trusts, and in some cases doctors' surgeries. We have classified entities by legal form where we were able to match them to organisations in the lists from Companies House and the Charity Commission.

In total 97% of records relate to third sector organisations, including 44% that we could not match successfully to the lists of regulated TSOs and which, in the absence of other information, we regard as our below-radar organisations, having excluded individuals, private companies and public sector bodies where it was possible to identify them. This varies between lists however, with below-radar organisations typically making up between 60% and 80% of the population of lists obtained from local and regional organisations (rather than regulators). Some example statistics from the largest lists can be found in Appendix I.

We gathered the lists in 2009 but we do not know exactly when they were compiled. Some could be several years old while others may have been of more recent provenance. This is potentially a source of confusion and could lead to the identification of organisations as being unique when, in fact, they are not. This arises because two or more different postcodes may be associated with the same organisation, meaning that it would not be identified as a duplicate. Some of the postcodes given for organisations are likely to be those of officials' homes, so the postcodes could change as and when the individuals involved changed, and if an entity was recorded on more than one list, it's possible that two separate postcodes might be recorded.

Entering the lists: what can be learned from local listings of third sector organisations?
Results from a study of Northern England

To reach these figures we needed to apply two cleaning techniques. Firstly, organisations that were not in scope needed to be excluded from the lists. Then the remaining organisations needed to be deduplicated so that only unique entities remained. We have previously described in outline the methods we adopted but in our more recent work we have refined them in several ways.

Organisations were excluded based on the following criteria:

- Geography – to exclude organisations from outside the study region;
- Type of entity – to exclude public and private sector organisations, and records relating to individuals.

Deduplication of the lists involved looking at the following issues:

- Matching using complete and partial postcodes,
- Matching based on organisation name,
- Organisations that are branches of national organisations.

Appendix 2 describes the criteria for exclusion and techniques for deduplication in more detail. We provide several tables which illustrate the process and demonstrate the extent to which duplicate entries were found, and we also provide examples of the kind of organisations that appeared frequently on our listings.

How confident can we be of the accuracy of the deduplication process? We believe it is recognising about five-sixths of the total number of duplicate entries in these data sources. We arrived at this figure by a manual check of the listings for one local authority which had a total of 1,584 entries. We identified 320 duplicates through our automated processes. A further 67 were identified through visual inspection, so the proportion of duplicates being identified by our process was 83%. The data for this area were particularly messy and complex in that

organisations were drawn from over 20 source lists. This increases the possibility of minor differences in the recording of information. In other parts of the region we sometimes only had a small number of lists to work from, such as a CVS listing and a local authority listing. The likelihood of duplication is therefore lower in such circumstances.

3 Analysis

3.1 Below-radar organisations

There is considerable interest in the question of how many organisations exist in the community that are not picked up on the "radars" of regulators such as the Charity Commission. If the upper-bound estimates of these are correct, suggesting somewhere between 600,000 and 900,000 community organisations, then we would expect to find in this region somewhere between three and five times as many community organisations as there are registered charities (based on the presence of around 180,000 charities on the register). This would imply something like 6,000 – 10,000 community organisations in Cumbria, up to 20,000 in the North East, and some 30,000 in Yorkshire and Humberside. To what extent do we find evidence of those sorts of numbers of organisations from the datasets on which we have drawn?

The first point to look at concerns the distribution of below-radar organisations and compares it to that of regulated third sector organisations. There are some minor changes here compared with our interim report on this work (2010), but the broad picture remains the same. Figure 1 compares the numbers of registered organisations in each local authority to the number of unregistered organisations. Note that we have included only 42 of the local government districts in existence at the time these lists were initially compiled, all of which had at least 100 below-radar organisations and registered TSOs. There are a small number of areas – mainly small rural district councils (this data refers to pre-2008 districts) – where the ratio of organisations to population is 6 per 1000 or above, and in one case 9 per 1000. This pattern is not altered significantly if one removes those organisations without postcodes. The reason for doing so is to increase the accuracy of the matching process; the more comprehensive the information used in matching, the more likely we are to find duplicate entries. However, removal of organisations for which no postcode information was supplied would mainly

affect a small number of urban areas in West Yorkshire as well as information from a larger number of listings with relatively few entries on them.

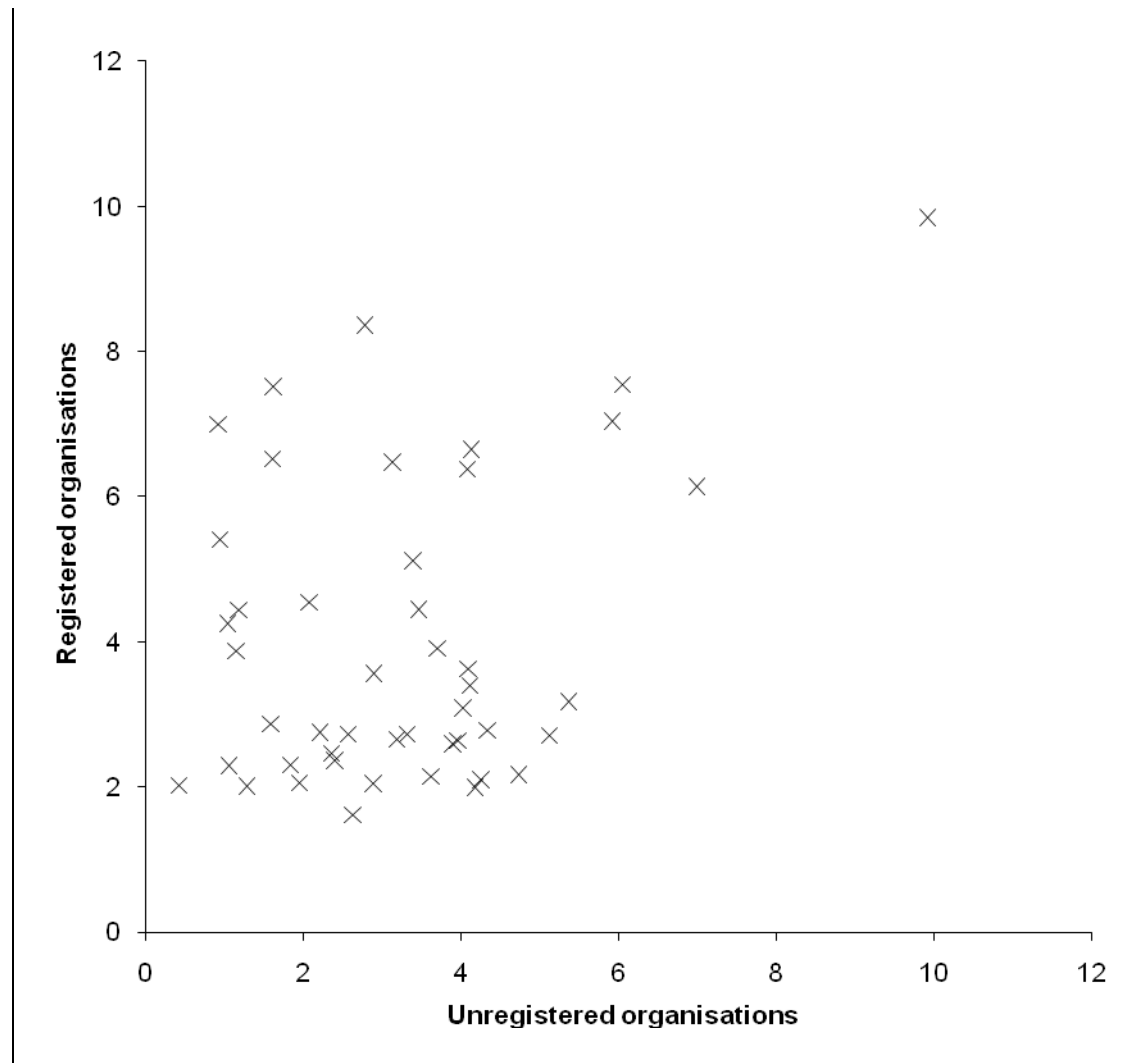


Figure 1: Ratios of registered and unregistered third sector organisations to population, by local authority (organisations per 1,000 population)

If extrapolated nationally the figures at the upper limit of our distribution would equate to something like 300,000 – 450,000 organisations, on top of the 180,000 registered charities. Note that these are well below the New Economics Foundation’s “Low Flying Heroes” estimates (600,000 to 900,000 organisations) and the typical figure for local authority districts in the three study regions is usually around 4 organisations per 1000. An obvious conclusion is that the figures

Entering the lists: what can be learned from local listings of third sector organisations?
Results from a study of Northern England

will depend on the sources from which they are drawn and on the numbers of sources available for any given local authority. At the top end of our distribution we found organisations from over 20 listings - in one area there were 1,584 organisations drawn from 28 listings, though only eight of those listings had more than 50 organisations on them. Later in this report we illustrate the degree of variation between sources in terms of the kinds of organisations listed there. To some extent this accounts for the variations in the ratios of organisations to population between local authorities: there are more lists in some areas, and some of the listings capture different types of organisation.

In our initial findings we argued that there was a clear inverse gradient with a higher presence of regulated third sector organisations in the most prosperous parts of the region being mirrored by a bigger presence of below-radar organisations in the more deprived areas. However, at that time we did not standardise for population; if more people lived in the less prosperous parts of the region than in the more prosperous ones, the ratio of organisations to population might not differ very much. In the next figure we divide the combined regions into 10 categories according to relative levels of deprivation, and we calculate the share of the population living in each of these; it ranges from about 10.6% to 9.8%. We then compared this with the share of organisations occurring in each area, shown in figure 2.

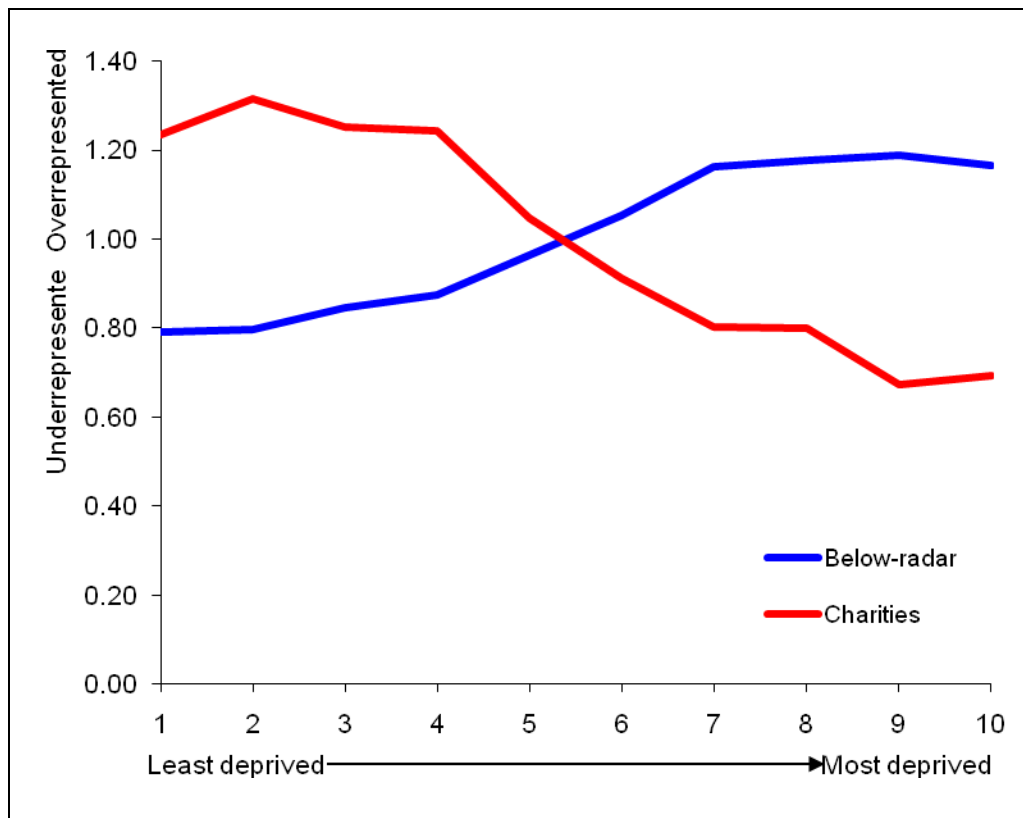


Figure 2: Distribution of charities and below-radar organisations, by level of area deprivation

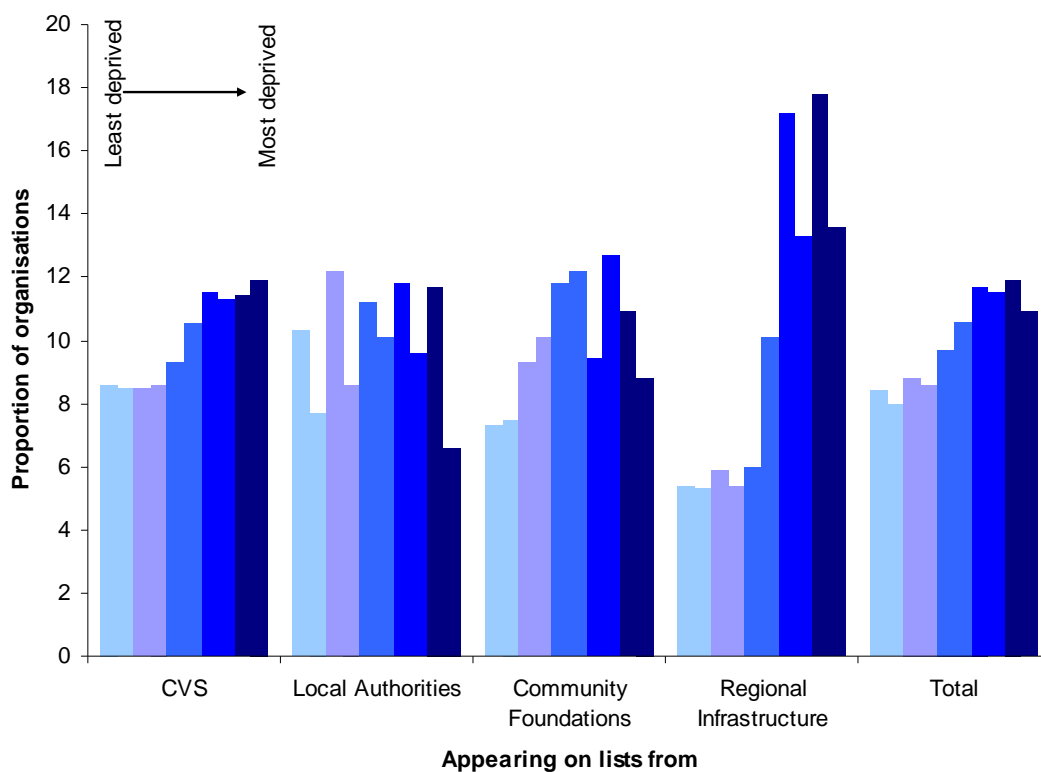
Areas are ranked from left to right according to increasing levels of deprivation for Census output areas. A score of one indicates that the proportion of charities, or below-radar organisations, in that area is equivalent to its share of population. Thus the figure shows that areas classified in the least deprived areas have around 8.3% of the BTR organisations, but 12% of charities, compared to 10.5% of the population. Essentially the graph indicates that, allowing for differences in population, charities are overrepresented in the more prosperous parts of these regions, and underrepresented in the less prosperous areas; the gradient for the below-radar organisations is less steep, and goes in the opposite direction. Note that postcodes are missing for far more below-radar organisations than for charities, which makes it impossible to link them to the census data. Nearly half of the below-radar organisations with missing postcodes are in five authorities in former metropolitan counties within Yorkshire and it is plausible to assume (if the

Entering the lists: what can be learned from local listings of third sector organisations?
Results from a study of Northern England

distribution of organisations in these areas that do have postcodes is any guide) that these are likely to fall in some of the more disadvantaged areas of the region.

Since we have a range of types of list a reasonable question to ask is whether some of these are more likely to be recording organisations in disadvantaged areas than in more prosperous areas, or vice versa. In figure 3, therefore, we present figures by decile of the index of material deprivation (in other words, one 10th of all observations are in each category) firstly for the listings as a whole, and secondly we break it down by different types of lists – CVS listings, Community Foundation listings (we only have two of these), local authority listings, and data provided by regional infrastructure bodies.

Figure 3: Distribution of below-radar organisations by index of multiple deprivation



On the basis of these figures it looks as if the gradient remains in place for CVS listings - in that there are about 1.5 times as many below-radar organisations in

the most deprived areas as there are in the least deprived areas. The gradient is less clear for other types of listing, but we have many fewer observations of below-radar organisations from such sources (at most 2000 or so observations with postcodes, compared to over 10,000 from CVSs). Much of the gradient is accounted for by particular types of organisation - the gradient is particularly strong for organisations in the social services and "development" categories, where, other things being equal, one would expect a greater presence in the more disadvantaged parts of the community.

On the basis of this information the CVS listings do pick out many organisations operating in deprived areas and to that extent it can be argued that the "below-radar" organisations complemented the distribution of the regulated third sector. But it is possible that more extensive listings from elsewhere - e.g. local authority listings and community foundation listings - would modify this pattern to a degree.

3.2 Classifying below-radar organisations

The next stage is to look at what kinds of organisations have been identified as below-radar and to describe their characteristics. This stage of research involved an exhaustive two-stage process. Firstly, using the titles of organisations, keyword searches were run, which generated a listing of around 740 keywords or phrases. Secondly, these were aggregated up to the ICNPO classification⁵. A summary of the process is provided in appendix 3, which sorts types of organisation by frequency, and cross references them against the ICNPO. Around 36,000 organisations were classified in this way, including many not recognised as below-radar organisational forms. The first 70 or so words or phrases accounted for more than half of these but thereafter diminishing marginal returns set in.

⁵ The International Classification of Non-profit Organisations is designed for the classification of voluntary organisations and other non-profit groups.

Entering the lists: what can be learned from local listings of third sector organisations?
Results from a study of Northern England

Approximately 3,600 below-radar organisations were not classified and this reflects the limits to keyword-based classifications drawing on the titles of organisations. Whether these are substantively different from those for which classification was possible is difficult to say. Some will undoubtedly be organisations which are branches of registered charities not based in the region but which happen to feature on local lists. Others will be names of individuals and potentially government departments. A difficulty is that names of organisations in languages other than English are unclassifiable – we did not have the resources to translate these items. However, the ICNPO classification does not distinguish beneficiary groups, but the main areas of activity carried out by organisations.

Table 3: Contribution of different sources to estimates

	CVS (%)	Local Authority (%)	Community Foundation (%)	Regional Infrastructure (%)	Total (%)
Culture/Recreation	29	28	47	25	30
Education	10	12	11	14	11
Health	3	3	1	4	3
Social Services	23	23	19	23	23
Environment	4	4	3	3	4
Community Development	18	21	14	17	18
Law, advocacy, politics	3	2	1	4	3
Infrastructure	1	1	1	3	1
International	1	0	0	1	1
Religion	7	6	3	6	7
Other	0	0	0	0	0
Total	100	100	100	100	100

Table 3 includes a breakdown of the ICNPO classification of local listings. The main difference between list types appears to be the large numbers of organisations on the community foundation lists that are from the broad “recreation and culture” category – nearly half the below-radar organisations on the two community foundation listings we obtained are of that kind. This clearly illustrates the way in which different listings capture different segments of the below-radar population.

In the areas for which we had community foundation listings, these two listings alone accounted for nearly 25% of the below-radar organisations.

Figure 4: Distribution of organisation types by ICNPO classification

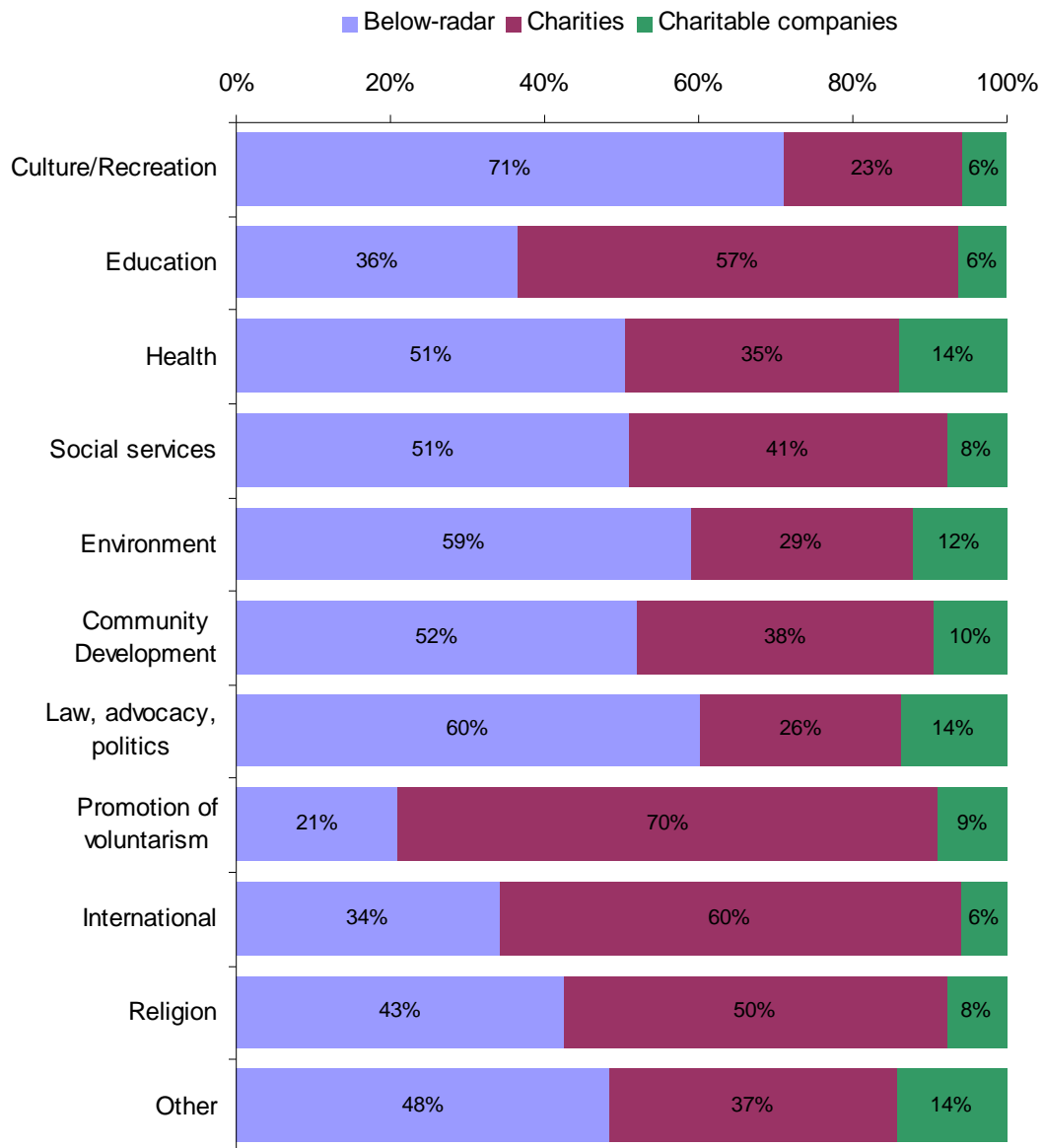


Figure 4 shows the broad pattern, and compares only below-radar organisations with charities and charitable companies because of the difficulties of classifying other legal forms such as Companies Limited by Guarantee (CLGs) and Industrial

Entering the lists: what can be learned from local listings of third sector organisations?
Results from a study of Northern England

and Provident Societies (IPs) against the ICNPO. The key point of difference seems to be the substantial proportion of below-radar organisations that are in the culture and recreation sector. Nearly 30% of the below-radar organisations are classed in this way, compared to under 13% for charities. In turn, this reflects the very substantial numbers of sports clubs which feature on local listings. These accounted for nearly 15% of the total of below-radar organisations in our database. Sport per se is not a charitable purpose and consequently sports clubs are likely either to take other legal forms or none at all.

Another way of looking at this is to look at the distribution across legal forms of different types of organisations; again, because of low numbers, we confine this to below-radar, charities, and charitable companies. The most striking feature of figure 4 is that of the 7,600 or so organisations in the culture and recreation sector which we have found in these regions, approximately 70% have no legal form. Other sectors in which more than half of entities are classed as below-radar include the categories of law, advocacy and politics, and community development. The former include numerous organisations which may, or may not, be projects run by regulated third sector organisations or even potentially the public sector such as support for victims of abuse or violence, and advice of various kinds. The latter includes local project shops run by community development charities (many of the BTR organisations with "Oxfam" in their title will probably feature here) and it also incorporates entities whose names appear to suggest an international orientation, but this is inevitably subjective. Nearly 60% of organisations classified under the heading of "environment" in the ICNPO appeared to have no legal form either. This includes a large number of conservation and environmental groups sometimes associated with particular properties or sites (e.g. allotments, canals, parks and open spaces).

At this level of the ICNPO, other differences seem less apparent but when we break down the numbers in more detail we find that there are some interesting contrasts. For example, using this categorisation of organisations, there are

approximately 4,000 organisations in the broad social care classification for both below-radar organisations and registered charities. However, some 900 of those refer to organisations in the Scout and Guide movements, large numbers of which are unregistered; without these, the proportion of below-radar organisations in this classification of the ICNPO would be much lower. And while there are comparable numbers of organisations classified as religious, several hundred of the below-radar organisations are churches and other places of worship which are not registered charities. Some 17% of the below-radar organisations fall within a general category of “community development”. This includes large numbers of community associations, tenants associations, and residents associations, some of which are registered and some of which are not. The “community development” category includes several hundred village halls, though around 80% of these are registered charities.

3.3 Differences between TSOs that appear on listings, and those that do not

So far we have shown that there are differences in the types of organisations that feature on the various local listings. Another perspective on this can be provided if we look at the extent to which different sorts of regulated third sector organisations appear on local listings, and at whether there are any contrasts between the kinds of regulated third sector organisations that do appear on such lists and those that do not.

For this part of the analysis we concentrated on CVS listings since this is our largest single category of listings (30 were supplied) and the organisations also share a common purpose. These listings had a combined total of 17,413 unique entities (excluding individuals, public sector organisations and private companies) of which 13,239 were below-radar, 1,861 registered charities and 1,582 charitable companies; a further 553 CLGs and 178 IPSs appeared on these listings. Overall, therefore, 20% of organisations featuring on local lists were registered

Entering the lists: what can be learned from local listings of third sector organisations?
Results from a study of Northern England

charities, with a further 4% being either CLGs or IPSs. However, the proportion of registered third sector organisations which appeared on CVS lists varied considerably: only 12% of charities, 11% of CLGs and 8% of IPSs did so, although 35% of charitable companies in the region featured on the listings of CVSs. Appendix I shows how the proportion varies between individual (anonymised) lists.

So what kinds of regulated third sector organisations do appear on local lists? And what proportion of regulated third sector organisations in a given local authority are recorded on their local listing?

3.3.1 Organisations appearing on any local listings

Where we have at least 50 local third sector organisations appearing on a local listing, the proportion of third sector organisations in an area which are on ANY local list varies between about 40% and 5%, but there is great variation by legal form. Very few IPSs appear on local listings – rarely more than 10% in any one local authority, and these are a mixture of housing associations, credit unions and assorted others. The same is true for CLGs – though the proportions are slightly higher. For charities as a whole, the proportions appearing on some local list or other are higher than for CLGs or IPS, but not as high as for charitable companies. The figures are largest for charitable companies where in numerous local authorities at least 50% of these appear on one or more of the listings compiled by local infrastructure bodies and this presumably reflects infrastructure bodies’ awareness of the major players in their area. Organisations are more likely to take this form if they are employing people and if they have substantial assets, both of which mean they are likely to be larger than average.

We also find differences in the type of organisations that appear on local listings vis a vis those that don’t. For charities, firstly, the great majority of charities appearing on local lists (96%) fall within the NCVO “General Charities” definition.

Those charities in our regions which appeared on local listings and were not general charities were mainly religious bodies of one sort or another. Secondly, there are differences in terms of the type of charities that appear on the lists of infrastructure bodies. Compared to those charities which did not appear on local listings, those which featured on the lists were more likely to be in the area of community development (these accounted for 27% of charities on local listings, compared to 15% of those which were not) and social services (27% compared to 22%). Conversely those charities which did not feature on local listings were more likely to have been classified as education (22% compared to 8%) and religion (12% compared to 5%). Discussion of this question is less straightforward for IPSs and CLGs where our classification of third sector organisations is not comprehensive, but in any case the proportions appearing on the lists are much smaller.

Finally there are variations by size. Those charities which appeared on local listings were typically substantially larger than those which did not; differences in the mean expenditure were statistically significant (£320,000 for those organisations appearing on local lists compared to £190,000 for those that did not). However a more appropriate comparison is the median since the mean is skewed by large values: the median expenditure for those charities appearing on local listings is £39,000 compared to £7,000 for those that do not appear on these lists.

3.3.2 Organisations that appear on their local CVS listing

It is possible to work out which organisations appear on the listings of their own local CVS and this gives us some indication of the "reach" of these infrastructure bodies. We can ask, for example, what proportion of the organisations in their area they have a relationship with. These figures suggest that typically a single-district CVS (i.e. one which just works within its immediate local authority, rather than a county-wide-body) will have a relationship with between 10 and 30% of

Entering the lists: what can be learned from local listings of third sector organisations?
Results from a study of Northern England

charities in its area although exceptionally the proportion rises to over 40%. The figures are much higher for charitable companies though, where it is not unusual for over half, and in some cases three-fifths or more, of these bodies to feature on the list of their local CVS.

This isn't necessarily to be interpreted as a comment on how well infrastructure bodies such as CVSs are 'performing'. In some of the larger urban authorities in these regions there are upwards of 500 charities and over 1,000 in each of Leeds and Sheffield; in some rural districts such as North Yorkshire there will be several hundred spread over a very large area. Keeping in touch with all of these would be a formidable undertaking. We should also consider the kind of services and functions for which organisations might want to have advice from an infrastructure body – large numbers of small local charities might not be likely to require the services of, or a relationship with, the CVS.

There are also variations in the kinds of community-based organisations which appear on local listings, as can be seen by comparing the proportion of organisations in various ICNPO categories across CVS lists. One CVS has 35% of entries in the broad social services category compared to 16% in another, yet they are both urban CVSs in the same region with broadly comparable social conditions and populations. There is an even greater range for the "culture/recreation" category. In one listing, for a CVS in a fairly prosperous rural area, over 50% of entries are in this category while at the lower end of the range only 11% of below-radar organisations are in this category in two contrasting locations (one a large city, the other a largely rural area with a previous history of industrial development). There are CVSs serving adjacent districts in former metropolitan areas which have over 40% and under 15% respectively of their below-radar organisations in the culture/recreation category.

Without further information on how these lists came to be compiled it is difficult to go further with this analysis but it certainly appears likely that much of the apparent variation in the ratio of organisations to population is going to reflect the priority or otherwise attached to particular types of organisation. As an example, on the basis of these listings, there are more sport and recreation below-radar entities in some rural districts with populations of below 100,000 than there are in some of the larger urban centres, which is implausible, unless one believes that for some reason the likelihood of participating in certain types of activity is greater in rural areas than in urban ones. Another illustration is provided by Scout and Guide groups. There is one CVS which records over 50 of these, accounting for nearly 10% of the entries on their list, and three others which record 20 or more whereas several CVSs have hardly any such groups on their lists. Depending on the size of the population of the local authority in question, these alone could increase the ratio of below-radar organisations substantially. It also means that our ratios are likely to be underestimates, assuming that is that we agree that Scout and Guide groups are part of the population in which we are interested in.

A further point which is relevant here is that there could be variations in the likelihood that organisations of a particular type (as opposed to the ICNPO classification) are of interest to particular sources. To explore this we look at some types of organisation which are easily identified from keywords in their titles, such as Scout groups, community centres, social clubs, sports clubs and so on. We then assess the extent to which it was possible to match these to lists provided by regulators. As can be seen in table 4, whereas around 40% of sports, community centres, church, scout groups and organisations for young people could not be matched, and therefore appear not to be registered, the proportion is much higher for Women's Institutes, village halls, and social clubs (in each case over 50% of the organisations of that kind which we found only appeared on lists provided by regulators). If one is interested in certain kinds of organisation, then, some lists will provide a more comprehensive source of information than others and researchers would have to ensure they obtained information from the relevant sources. Most obviously, if seeking information on social clubs, the best

Entering the lists: what can be learned from local listings of third sector organisations?
Results from a study of Northern England

place to start would clearly be with a listing from Companies House, whereas CVS and community foundation lists would provide a better basis for a search for sports clubs or various kinds of community organisation.

Table 4: Which organisations are found in which lists

Type of Organisation	List source						Total (%)
	CVS (%)	Local Authority (%)	Community Foundation (%)	Regulator (%)	Regional Infrastructure (%)	Other – research (%)	
Social Clubs	10	2	2	83	1	2	100
Sports	46	5	21	19	3	7	100
Community Centres	49	3	6	22	6	11	100
Community – various	46	2	6	23	6	12	100
Tenants/Residents	48	20	13	6	5	8	100
Village Halls	32	4	6	53	0	4	100
Women’s Institutes	14	1	0	77	2	6	100
Church	45	4	3	27	5	13	100
Scouts	39	6	7	34	3	10	100
Young people	44	4	16	22	4	8	100
Total	37	5	9	37	3	8	100

4 Conclusions

There is clearly considerable interest in the scale of the third sector that exists below the radar of regulatory bodies and this project represents probably the largest single effort to explore the characteristics of that part of the sector using listings provided by local infrastructure bodies. We argue that this is important for at least two reasons. Firstly, as public spending declines and there is a sense amongst some policymakers that the voluntary sector is too dependent upon the state, there is more interest in understanding the capacity and resilience of communities. In other words, what are the structures and institutions that will sustain them when funding is withdrawn? As such, there is more interest in informal voluntary action and not just in big charities or more traditional organisations. Secondly, if we are to understand such community capacity, and its relationship with informal clubs, societies and groups, we need to come up with a stronger evidence base with regard to their number, distribution, contribution and leverage role. Moreover, we need to be clear that being 'below the radar' is in itself not an indicator of weakness or a lack of capacity; rather, an indicator of a different sort of capacity. We acknowledge that enumerating below-radar groups is a difficult, time consuming business, but nevertheless one that is important if we are to make good decisions about the ability of communities to self-organise and ultimately provide more support to themselves.

There has been work of this kind previously but our work is distinctive because of its scale. In contrast with the NSTSO pilot studies it covers many more local authorities and a wider range of sources, and has also relied on automated matching rather than upon visual methods. This imposes consistency on a very large data set, but perhaps lacks some flexibility. There is inevitably a trade-off. The NAVCA / Audit Commission profiles "toolkit" is an impossible counsel of perfection - it details a substantial range of possible listings and a work programme for one local authority alone which lasts a year. We are not aware of a single local authority in which such work has been done exhaustively. We have traded off local detail (and therefore probably not identified as many duplicate

Entering the lists: what can be learned from local listings of third sector organisations?
Results from a study of Northern England

entries as there possibly are) for a methodology which we think will identify the great bulk of duplicates in a set of matched lists.

It will be clear that we do not believe that this work provides an unambiguous answer to the question of how many below-radar organisations there are. Previous work suggests that estimates of below-radar organisations vary substantially - at the high end, reports such as *Low Flying Heroes* by the New Economics Foundation (NEF) (which produced an estimate of between 600,000 and 900,000 community organisations) suggest a great unrecognised resource that, if true, implies a massively latent Big Society that policy-makers simply need to harness. But we do not know whether such optimistic appraisals reflect the particular communities chosen, or the research method deployed, or the definitions used. In areas where we have a large number of listings the upper limit of the numbers of organisations we have found is around six organisations per thousand population although in one local authority the figure was approximately nine per thousand. On the one hand this could well be an over-estimate because of the complexity of listings in those areas and we suggested in the case of the local authority at the top end of the distribution that a more realistic figure might be in the range of seven to eight per thousand because of the likelihood of false positives being recognised.

This is clearly well below the NEF figure which is widely cited in debate about the scale of the sector. On the other hand, there is evidence of variation in the numbers and types of organisations captured on local listings (something which was also pointed out in the NSTSO pilot studies) with some infrastructure bodies having details of large numbers of sports groups, Scout and Guide groups, etc, while others eschew them. Given this, if one were to have greater time and if a larger number of lists were obtained, it is probable that higher figures would be found. Even so we don't think they would approach the levels estimated by NEF, and we argue that variations in estimates from existing studies have as much to

do with the sources from which they are drawn as with the characteristics of the underlying population of below-radar entities.

Can we nevertheless provide a basis for estimation of numbers using these listings which might, in turn, give guidance for future exercises? One approach to this is to consider only those parts of the region for which we have a representative set of different types of listings. We do this in two parts.

First, we have chosen four contiguous local authorities where we have the listing from each of the local CVSs plus a listing from a community foundation covering the area. There are just over 2,500 regulated TSOs in this area although, as might be expected, they don't all appear on local listings. The first row of table 5 therefore shows the maximum number of organisations in each category that we have identified in these four local authorities – a total of 5,437 distinct entities. The next row gives the numbers which we have identified from the relevant local CVS listings combined with the list from regulatory bodies. As can be seen it is approximately two-thirds of the total. We then add to this several hundred organisations from a listing supplied by the local community foundation. This contributes nearly 1,000 extra below-radar organisations, and the total now accounts for over four-fifths (84%) of those entities recorded in these local authorities. So the inclusion of the community foundation is, as we would expect given earlier analysis of the contents of listings, adding an extra layer which is not picked up elsewhere.

Entering the lists: what can be learned from local listings of third sector organisations?
Results from a study of Northern England

Table 5: Contribution of different sources to estimates

1. Four contiguous urban authorities	Below-radar	Charities	Charitable Companies	CLG	IPS	Total	Percentage of total (%)
Total	2,883	1,475	298	509	272	5,437	
...from CVS and regulators	1,266	1,373	245	465	267	3,616	67
...from CVS, regulators and community foundations	2,207	1,388	252	481	268	4,596	85
2. Two rural districts							
Total	874	667	141	128	35	1,845	
...from CVS and regulators	344	598	96	107	30	976	53
...from CVS, regulators and community foundations	488	631	120	119	33	1,391	75
...from CVS, regulators and local authority	602	622	112	112	31	1,479	80

Secondly, in these areas we do not have any local authority listings to complement those from the relevant Community Foundation, but in two rural districts elsewhere in the region we do. We identified a total of 1,845 organisations in these areas, of which 874 do not appear to be on the lists of regulators. Of these, 344 are on a listing supplied by the relevant CVS and regulators. A very detailed listing from a local authority supplies a further 258 entries and the local community foundation provides another 140 or so. The combined listings from the local authorities, CVSs, the community foundation and the regulators gives a total of just under 1,500, i.e. four-fifths of the total number of organisations we have identified. In this area we had several other local, regional and national listings but few of them contributed more than 10 entries to the total.

Depending on the purposes of these exercises, therefore, and given that there is inevitably going to be limited resources and compromises, a good starting point would appear to be a combination of CVS, Community Foundations (where they are in operation) and local authority grant listings. It seems to us that these are picking up the great majority of organisations in our area. Where we have such listings the ratio of below-radar organisations to registered charities rarely exceeds 1.25 so a generous estimate might be that there are 125 such organisations for every 100 charities. We could add perhaps a further 20% to these figures on the grounds that additional listings revealed the presence of more organisations, and work with a figure of around 150 organisations per 100 charities.

As there are approximately 150,000 general charities in England and Wales, this would imply an upper limit of around 225,000 organisations for England and Wales. This is relevant for policy because it highlights that the vast majority of the third sector is unregulated and as such we may wish to recalibrate our concerns that voluntary action is stymied by over-burdensome regulation. In these regions, such an estimate would imply something of the order of 3,000 – 3,500 organisations in Cumbria, 9,000 in the North East, and 15,000 in Yorkshire and Humberside. These could be generous upper estimates depending on the extent to which further duplicate entries were found.

It is also a matter of judgement whether the so-called below-radar organisations constitute a distinct segment of the third sector population. When we carried out a small-scale survey we found numbers of organisations which did not differ hugely in terms of size, paid and unpaid staff and funding mix from smaller scale charities identified in the NSTSO. There is also a degree of consistency between the local pilot surveys of the below-radar organisations in three NSTSO areas (and it is reassuring in some ways to note that despite the very intensive work they did on individual listings, their work and ours both encountered the problem that some large organisations had slipped through the net). The most substantial

Entering the lists: what can be learned from local listings of third sector organisations?
Results from a study of Northern England

difference appears to be in terms of the classification of the below-radar population where there is a general over-representation of the areas of recreation and sport.

5 Implications for policy and practice

This study (and our previous work on the paid and unpaid workforce) highlights that regulatory returns are inevitably limited in assessing the totality of volunteers and active in the UK. Whilst this is unsurprising, it perhaps should give policymakers more confidence in the much higher estimates of volunteering obtained from surveys of individuals. It also highlights that much voluntary action takes place within an informal, unregulated setting - so, again, we perhaps need to take care that we do not overestimate the potential impact of regulatory barriers upon voluntary action.

It is worth noting the implications of these estimates for the forthcoming Community Organisers programme. These numbers imply that the 500 senior community organisers will potentially be working with 900 groups each, though of course there are significant variations between localities.

Being able to monitor and in turn report on the full range of community action is an important role for infrastructure bodies that is recognised by NAVCA in its quality standards for such organisations. Our research estimates that organisations hold contact information for between 10-30% of the **charities** in their geographical remit. In other words, they are missing organisations that are **above** their local radar. **Conversely, information on the majority of the organisations on the lists held by local infrastructure organisations (65%) is not held on national lists.** This is, we believe, a significant finding that stresses the importance of local bodies in maintaining our understanding of local voluntary action.

We believe it would strengthen local infrastructure organisations if it was easier to maintain a dynamic link with the Charity Commission register so as to benefit from updates such as new organisations or changes in contact details. The sector as a whole should consider the open data agenda and in particular encourage the Commission to develop web technologies which allow data to be gathered and

Entering the lists: what can be learned from local listings of third sector organisations?
Results from a study of Northern England

disseminated quickly and cheaply. Ultimately we believe strategy and policy development will benefit from local infrastructure organisations having as complete a list as possible.

The Charity Commission will almost certainly focus its effort as a regulator on fewer, larger charities in the future. Registration thresholds will be raised. Therefore, maintaining local listings will become an even more important role for local infrastructure organisations, for which they may need support. Funders and government should better support local infrastructure bodies to hold and maintain such data, where possible agreeing to share organisational contact details.

There is significant variation between local areas in terms of number of organisations per 1,000 population. This potentially highlights variation in community capacity, but in some cases it may reflect local infrastructure organisations' focus, development work or capacity to maintain listings. This is not a recommendation for uniformity, but rather a suggestion that peer comparison may highlight differences in practice and opportunities for improvement.

Almost one-third of the below-radar organisations in our combined listings were arts and sports organisations. These are typically peripheral in discussions of the third sector, yet often central to discussions of participation and volunteer involving organisations.

Appendix I – Examples of lists gathered

Table A1: illustrative data from listings used to compile database

Type of source	Proportion of organisations				Total number of organisations
	Not regulated (%)	Charities (%)	Charitable companies (%)	CLGs (%)	
CVS	88.5	4.1	2.7	2.5	489
CVS	83.6	8.8	3.2	1.9	532
CVS	83.5	6.8	6.4	1.7	468
Other Local	82.0	9.6	6.3	1.7	460
CVS	81.4	9.5	4.0	0.9	349
CVS	81.1	7.0	5.6	3.3	640
Local authority	80.9	6.4	5.5	0.9	110
CVS	80.5	8.7	7.4	1.3	747
Local authority	79.1	6.6	1.0	0.0	196
CVS	78.6	8.1	8.6	3.7	1,776
Other Local	78.4	8.9	7.3	3.6	740
CVS	77.5	7.4	9.0	3.5	311
Local authority	76.4	10.9	6.5	2.3	432
CVS	76.3	12.4	3.9	4.9	865
CVS	76.2	11.8	7.2	2.5	692
CVS	76.0	9.7	6.0	3.4	1,004
CVS	75.6	8.8	9.6	2.7	2,071
Local authority	75.0	8.0	8.9	3.6	112
Local authority	74.6	13.2	5.5	3.7	669
CVS	73.1	8.6	12.4	3.8	372
National	72.9	12.9	8.3	3.6	387
Local authority	71.1	6.9	8.0	6.5	1,803
National	70.1	12.2	12.7	4.1	221
CVS	69.9	14.7	14.0	0.7	136
Local authority	69.5	8.9	7.9	0.5	190
CVS	69.3	14.4	10.4	4.5	374
CVS	68.4	11.7	13.7	4.0	351
CVS	68.3	11.9	15.1	2.8	218
National	68.0	13.0	14.8	4.1	169
CVS	67.6	18.6	10.9	1.6	515
Local authority	66.1	13.4	8.1	2.7	186
CVS	65.0	19.1	12.0	3.3	183
Local authority	63.0	26.5	1.6	2.1	189
Other Local	61.9	13.4	13.2	9.7	893
CVS	59.8	15.2	19.6	2.2	224
CVS	58.3	26.4	11.9	2.1	571
CVS	57.5	9.4	6.6	2.5	863
CVS	57.1	15.6	20.0	5.9	506
CVS	56.8	13.1	11.8	3.1	1,055
CVS	53.3	21.3	20.7	0.7	150
Other Local	53.1	9.4	18.6	9.8	967
CVS	49.9	18.2	23.3	5.4	429
National	37.6	19.7	33.6	6.6	1,540

Figures refer only to entities within the NE, Cumbria or Yorkshire and Humber

Entering the lists: what can be learned from local listings of third sector organisations?
Results from a study of Northern England

Appendix 2 – Techniques for excluding organisations and deduplication of lists

To reach the figures in this document we needed to apply two cleaning techniques. Firstly, organisations that were not in scope needed to be excluded from the lists. Then, the remaining organisations needed to be deduplicated so that only unique entities remained. We have previously described in outline the methods we adopted but in our more recent work we have refined them in several ways.

1.1 Excluding organisations

1.1.1 Geography

Our combined listings totalled 58,998 entries, but for analytical purposes we restrict ourselves to 54,471 entities for which we had enough information to allocate to a local authority within the North-east, Cumbria, and Yorkshire and Humberside. (1,077 organisations fell outside this area of interest and a further 2,802 were in parts of the region from which we were unable to obtain local listings. There were also 252 entries in lists which had insufficient information to enable us to allocate them geographically to local authorities).

Linkage to socio-economic statistics was carried out by using the postcode. Our information on postcodes was incomplete (table A2) - partly because of incomplete data in the original source, and partly because of the reluctance of some suppliers of lists to provide it, though some provided partial postcodes; we had complete postcode information for over 53,000 entities but there were around 4,350 for which no postcode data were supplied and a further 700 for which partial data were available. We therefore had around 41,400 third sector organisations within our region for which we have a valid index of deprivation, representing coverage of over 90%. Some organisations preferred to provide us with no postcode information or only minimal postcode data rather than the full postcode, in some cases citing data protection legislation. Because of this, in the areas affected, mainly parts of urban Yorkshire, some of the numbers of unique entities may be overestimated.

Table A2: Criteria used to identify duplicate entries

Criterion	Frequency	Percent (%)
Name, local authority, PCU2	2,115	27
Name, local authority, PCU2 (sometimes missing)	1,000	13
Nested name, PCU1	745	10
Name, local authority, PCU2, address	696	9
Charity Commission number, name, local authority, PCU2	658	9
Charity Commission number, company number, name, local authority, PCU2	414	5
Name, local authority, Pcsect	414	5
Charity Commission number	336	4
Name, pcdist, provided that two of that name only in whole file	199	3
Name, local authority, PCU1	187	2
Charity commission number, name, local authority, PCU2, address	148	2
Name, local authority, PCU2 plus nested name and PCU1	126	2
Charity Commission and company number, name, local authority, PCU2, address	105	1
Charity Commission number, name, LA, postcode sector	77	1
Charity Commission number, name, LA, PCU2 (sometimes missing)	75	1
Other	414	5
Total	7,709	100

PCU2 – full postcode (e.g. NE1 7RU)

PCU1 – first half of postcode plus first two characters of unit postcode (e.g. NE1 7R)

Pcsect – postcode sector – e.g. NE1 7

Pcdist – postal district – e.g. NE1

1.1.2 Type of entity

These listings contain many entries for individuals, government departments or agencies, public sector bodies, and private firms. Indeed it is not possible unambiguously to classify every single entity on a list. We developed a number of

Entering the lists: what can be learned from local listings of third sector organisations?
Results from a study of Northern England

exclusion criteria which account for several hundred entities. Note that we do not automatically exclude them from analysis. For example it may be useful to know, for comparison, that some local listings capture much detail about organisations which might be regarded as part of government, such as parish councils, or about regional and indeed national public sector bodies. However, there are circumstances where we would wish to exclude such organisations; for example when we are trying to contrast the distribution of below-radar organisations with that of registered charities, it would clearly be undesirable to include the 750 or so government/public sector organisations that we have identified.

1.2 Deduplicating the lists

Firstly we would emphasise again that in contrast to the NSTSO studies, which appeared to be based on manual cleaning of individual records, we automate the process of name cleaning and of ensuring consistency in the recording of particular terms (eg standardising "Scout" or "Scout Group" simply as "Scouts" (thereby also limiting the risk of identifying "false positives" – eg the "Gosforth Scout Group" and the "Gosforth Scouts" being treated as separate entities), or expanding / contracting appropriate terms (eg "Assoc" to "Association").

Secondly, we had concerns about the accuracy of our matching procedures. In this regard it is important to consider the kind of information that we have from our listings. This is in two parts. There is usually something which identifies the kind of organisation under consideration (youth club, Scout troop, community centre) and there is generally, though not always, some kind of information about the location (e.g. "Blaydon" youth club, "Scotswood Road" Scout troop, "Bamburgh" Community Centre). The parts are not always in a specific order, and they are often interleaved with one another. Differences may be very minor - "St Mark's Church Hall" and "St Mary's Church Hall" differ by only one letter – or quite substantial ("Guides" versus "Girl Guiding Association"). In our initial attempts at removing duplicates we used automated routines which required electronic sorting of combinations of variables. A challenge here is that minor variations in the recording of information could lead to a failure to recognise duplicates. This might be because of the addition of place names into the name of the

organisation, spelling errors, or component of names appearing in a different order (e.g. "League of Friends of the Wigton Hospital" vs "Wigton Hospital League of Friends"). Even sophisticated methods like splitting names of organisations into their component words and then re-sorting the words alphabetically have their limitations. Thus the hypothetical illustrations of the Parkinson's Lane Guide Association and the Guide Lane Parkinson's Association would be rendered as the same entity if reordered alphabetically ("Association Guide Lane Parkinson's") but are clearly separate entities. A more amusing example is from the *Life of Brian*: the "People's Front of Judea" and the "Judea People's Front" would be rendered by this method as the "Front Judea People", but devotees of the film will recall that the two organisations had very different aims!

To minimise error it is not sufficient simply to compare organisations ordered alphabetically in sequence. That might be feasible for individual listings such as those used in the NSTSO pilot studies but it is not practicable for a database of this size. Instead one needs to automate the process of comparison of the name and geographical identifiers of every organisation against every other organisation. This is time-consuming, but manageable, and that is what our deduplication processes do. An additional feature is that, in the case of duplicates, we identify a "parent" and transfer any relevant regulatory (e.g. registration number) classificatory (e.g. ICNPO) or financial information derived from other sources to its "children" – ie those organisations which are shown to be duplicates of it. This allows a more robust analysis particularly of differences between listings.

A further feature of these lists is the large number of entries which seem to be local projects of national organisations. The listings include 43 entries including the term "Salvation Army", 43 with "Oxfam" in their name, 26 each containing "British Red Cross", "Barnardo's" and "Help The Aged", 23 "Scope", 19 "MENCAP" and so on. Some of these are charity shops; others are undoubtedly local projects. There is no way of distinguishing which is which, though arguably it doesn't

Entering the lists: what can be learned from local listings of third sector organisations?
Results from a study of Northern England

matter: each serves its own purpose, whether this be delivering services or raising money.

Using a number of methods we were able to identify approximately 7,700 third sector organisations as duplicates on our lists. Ideally this was based on a match of the complete name and the full unit postcode, but in many instances we had partial postcodes or incomplete names and had to decide whether or not to accept minor variations of postcodes or spellings of names as plausible matches. We have several different elements which are used in varying combinations because of the likelihood that some sources will provide more complete and accurate information than others.

The method used treats the list of regulated third sector organisations as the "master" list and seeks to match everything else against it. It therefore begins by taking registration numbers of regulated third sector organisations and ensuring that there are no duplicates. As it happens there are very small numbers of duplicate registration numbers which arise principally in the case of registered charities where we had a registration number from a local listing. The process then works through a series of steps in decreasing order of accuracy, beginning with identical matches on name, local authority and full postcode, establishing matches where possible, and then working through the rest of the list to identify plausible matches. Thus, for example, 414 duplicates were identified using name, local authority and postcode sector, it being a reasonable assumption that two different organisations with identical names would not be found in a relatively small area.

Of the 45,000 organisations uniquely identified, it can be seen in table A3 (below) that about 11% have duplicates, but this proportion rises to 35% in the case of charitable companies. These, of course, tend to be larger organisations, often delivering substantial services over a wide area. The total number of duplicate

entries is 7,480, representing approximately 14% of the 53,000 organisations identified here (this table ignores individuals, public sector and private sector bodies). We also show that there are significant differences between the types of organisations that appear on local CVS listings and the types that don't. We generated a statistic to show the number of times organisations appeared in the database (shown in table A4); this is set to 1 if the organisation only appeared once, so if every organisation in a local authority was found only once in our database then that would be the average value for this statistic. In 11 local authorities, however, the average value for this was at least 1.2, indicating that up to 20% of entries for organisations located in the local authority proved to be duplicates.

Table A3: Distribution of duplicate entries by legal form

Number of Duplicates	No Legal Form	Charities	Charitable Companies	CLGs	IPs	Total
1	20,044	13,453	1,937	3,396	1,765	40,595
2	1,859	965	621	322	87	3,854
3	308	227	227	93	20	875
4	55	77	113	30	8	283
5	18	34	46	12	3	113
6	6	13	28	4	0	51
7	4	5	8	0	0	17
8	1	7	2	3	0	13
9	1	2	8	0	0	11
10	0	1	1	0	0	2
11	0	0	1	0	0	1
12	1	0	0	0	0	1
Total	22,297	14,784	2,992	3,860	1,883	45,816
Proportion with at least one duplicate (%)	10	9	35	12	6	11

Entering the lists: what can be learned from local listings of third sector organisations?
Results from a study of Northern England

Table A4: Extent of duplication, by local authority

Local authority	Duplication
Sedgefield	1.36
Newcastle upon Tyne	1.28
Allerdale	1.25
Eden	1.25
Blyth Valley	1.25
Leeds	1.24
Carlisle	1.23
Easington	1.22
Chester-le-Street	1.21
Gateshead	1.21
Durham	1.20
Sheffield	1.18
Darlington	1.17
Wear Valley	1.16
Sunderland	1.16
Kirklees	1.16
Bradford	1.16
Wakefield	1.15
Middlesbrough	1.15
Richmondshire	1.14
Wansbeck	1.14
Teesdale	1.14
Copeland	1.14
Redcar and Cleveland	1.13
Hambleton	1.12
Doncaster	1.12
Castle Morpeth	1.12
Barnsley	1.12
Calderdale	1.11
York	1.11
Derwentside	1.11
North Tyneside	1.11
Rotherham	1.10
Tynedale	1.10
Barrow-in-Furness	1.10
Alnwick	1.10
South Lakeland	1.09
Berwick-upon-Tweed	1.09
Stockton-on-Tees	1.07
Scarborough	1.06
Selby	1.06
Harrogate	1.06
South Tyneside	1.06
Ryedale	1.05
Hartlepool	1.05
Craven	1.03

This was especially noticeable in parts of the region where we had multiple listings. It could also reflect the presence of bodies with a regional remit, as was the case for numerous bodies located in Newcastle, Sheffield or Leeds. To illustrate the extent to which organisations appeared frequently, 48 organisations are listed which appeared more than 5 times (table A5). This backs up the point in the previous paragraph - these tend to be larger organisations and many will be recognised as significant parts of the infrastructure in these regions.

Table A5: Examples of duplicates

Name	Numbers duplicates	Legal Form	2007 Expenditure (£)
VONNE – Voluntary Organisations Network North East	11	Charitable Companies	505,148
ME North East	9	Charitable Companies	109,753
Voluntary Action – Leeds	9	Charitable Companies	757,031
Victim Support County Durham	9	Charitable Companies	332,164
Leeds Jewish Welfare Board	9	Charitable Companies	1,061,640
Children North East	9	Charitable Companies	1,574,404
Voluntary Action Sheffield	9	Companies	5,102,363
Fit4Funding (Charities Information Bureau)	8	Charities	1,060,739
Churches Regional Commission in the North East	8	Charitable Companies	192,579
Black Ethnic Minority Community Organisations Network (BECON)	8	Charities	315,220
Northern Rock Foundation	8	Charities	24,213,000
Nepacs – Building Bridges for Prisoners and their Families	8	Charities	655,101
The Miscarriage Association	7	Charitable Companies	190,768
Finchale Training College	7	Charitable Companies	3,336,478
Age Concern York	7	Charitable Companies	1,002,695
Yorkshire Rural Community Council	7	Charitable Companies	3,114,699
Relate North East	7	Companies	330,979
Voluntary Action Calderdale	7	Charities	535,626
Advocacy Network Leeds	7	Charitable	255,897

Entering the lists: what can be learned from local listings of third sector organisations?
Results from a study of Northern England

		Companies	
East Leeds Health for All	7	Charities	388,732
Paddock Community Forum	7	Charities	312,104
		Charitable	
Leeds Involvement Project	6	Companies	447,369
County Durham Society for the Blind and Partially Sighted	6	Charities	93,844
		Charitable	
Leeds Church Institute	6	Companies	412,787
		Charitable	
Community Links Ltd	6	Companies	5,178,717
		Charitable	
Artlink West Yorkshire	6	Companies	102,798
St Oswalds Hospice Ltd	6	CLG	6,184,740
		Charitable	
Cloverleaf Advocacy 2000 Ltd	6	Companies	784,696
Durham Association of Clubs for Young People	6	Charitable	
		Companies	38,992
		Charitable	
Hamara Healthy Living Centre	6	Companies	513,771
		Charitable	
Heeley City Farm	6	Companies	965,564
		Charitable	
Reading Matters	6	Companies	206,929
Age Concern Metropolitan	6	Charitable	
Gateshead Ltd	6	Companies	561,683
Leeds Connecting Communities	6	Charities	152,823
Lowedges Community and Safety Forum	6	Charitable	
		Companies	383,693
		Charitable	
DISC (Sedgefield)	6	Companies	9,569,859
Leeds Girls' High School	6	Charities	278,222
		Charitable	
Northern Pinetree Trust	6	Companies	433,067
Roshni - Sheffield Asian Womens Centre	6	Charities	250,985
Mental Health Matters - Sunderland	6	Charitable	
		Companies	6,226,383
		Charitable	
Burley Lodge Centre	6	Companies	560,955
Community Action Northumberland	6	Charitable	
		Companies	898,608
Great North Air Ambulance Service	6	Charitable	
		Companies	3,440,708
		Charitable	
Leeds Mind	6	Companies	2,151,541
		Charitable	
Yorkshire Wildlife Trust	6	Companies	1,851,784
Step Development Trust	6	Charities	835,727
Age Concern Durham County	6	Charities	960,012

Appendix 3 - Keywords used to classify below-radar organisations into ICNPO categories

This table illustrates the process used for classification into ICNPO categories. For brevity we have listed only those words or phrases which appeared on more than 25 occasions.

Table A6: Keywords used in classification exercise

Type of Organisation	N	ICNPO
Religious group	1401	10100
Village/community hall	1147	6110
Youth club/project/forum	1036	4110
Playgroup/pre-school	960	2130
Tenant/ resident assoc	940	6200
Music	723	1100
Community Association	677	6100
Sport – Football	643	1200
Social club	603	1300
Art/craft	599	1100
Scout Association	562	4110
Sport – Bowling	371	1200
Friends of Cemetery/Park	366	1100
Religious groups	360	10100
Parish Council	341	Government
Guide Association	339	4110
Sport – Cricket	324	1200
History/arch. Society	313	1100
Womens group/club	298	4100
School	292	2100
Community Partnership	284	6100
Sport – dance	253	1200
PTA	243	2110
Theatre	234	1100
Charity	221	Generic
Cultural Soc./Assoc	219	1100
No Organisation Name	206	Generic
Sport – Rugby	199	1200
Women’s Institute	197	4100
Festival/carnival	183	1100
Allotment	170	5100
Conservation/Environment	167	5100
Housing project	162	6200
Sport	154	1200
Lunch club	153	4100
Community/Resource Centre	151	6100
Parent advice/support	151	4100
Blind/ visual impairment	146	4100
Community project	146	6100

Entering the lists: what can be learned from local listings of third sector organisations?
Results from a study of Northern England

Mental Health Group	143	3300
Community group	140	6100
Education Project/Service	139	2300
Nursing home	138	3200
After school/Out of School	136	2100
Medical	135	2410
Transport	134	6100
Age Concern	132	4100
Drug/Alcohol project/group	132	4100
Sport – martial arts	129	1200
Over 50/60	127	4100
Family support	126	4100
Drama	123	1100
Recreation/Play Ground	122	1100
Carers group/forum	120	4100
Environmental group	119	5100
Citizens Advice Bureau	114	7100
Young Farmers	109	4110
Development Trust/agency	108	8100
Family/Children's Centre	108	4100
Hospice	108	3100
Horticultural Society/Garden club	107	5100
Community forum	105	6100
Action group	104	6300
CVS	104	8200
Museum/archive	104	1100
Cadets	101	4110
Health Promotion		
Club/Forum/Network	101	3400
Day Centre	100	4100
Deaf	100	4100
Sport – Swimming	100	1200
Volunteering	99	8200
Salvation Army	98	4100
Sure start	90	2130
Disability group/association	85	4100
Advocacy	84	7100
Asylum/Refugee Council/Forum	84	7100
Credit union	83	6100
Townswomen	83	6100
Charity shop	82	Generic
Housing Association	82	6200
MIND	80	3300
Campaign group	78	7100
Housing	77	6200
Rotary	76	1300
Autism/ADHD	74	2410
Mencap	74	4100
Community Action Group	71	6100
Barnardo's	67	4100
Council	67	Government
Stroke	67	2410

Animal Sanctuary/Rescue	66	5200
Regeneration	66	6100
Sport – angling	66	1200
Golf	65	1200
Community Garden	64	5100
Training	64	6300
Educational Trust/Foundation	63	2120
Elderly group	63	4100
Overseas	63	9100
University	63	2200
Girls/ boys club	62	4110
Support Group	62	4100
Bereavement support	61	3300
Councillor	61	Government
Activity group/centre	60	6100
Community services/support	60	6100
Enterprise Agency/Community		
Enterprise	60	6100
Sport – keep fit	60	1200
Victim support	60	7200
Domestic violence	59	7100
Homeless Project/Support	59	6200
Alzheimers	58	2410
Badminton	58	1200
Government department	58	Government
Voluntary Action	58	8200
Charitable trust	57	8100
Oxfam	57	9100
YMCA	57	4100
St John Ambulance	56	3400
Business network/partnership	55	6100
Community/Hospital Radio	55	6100
Community/social trust/fund	55	8100
Employment/benefits advice	55	6300
Boys’/Girls’ brigade	54	4110
Cancer Support	54	2410
Sports Club	54	1200
Conservation Volunteers	53	5100
Disability	53	4100
In Bloom	53	5100
British Red Cross Society	52	4100
U3A	52	2300
Borough/District Council	51	Government
Education centre	51	2300
Riding	51	5200
School fund	50	2100
Sport – Boxing	50	1200
Walking Club	50	1200
Multiple Sclerosis	47	2410
Disability support	46	4100
Ethnic minority group	46	7100
Individual	46	Individual

Entering the lists: what can be learned from local listings of third sector organisations?
Results from a study of Northern England

Community development	45	6100
Language/Cultural School	45	2300
Neighbourhood Watch	45	6100
PCT	45	3100
Safety group/accident prevention	45	7100
Tennis	45	1200
Arthritis	44	2410
Disability information/advocacy	44	7100
Fire service/police	44	Government
Counselling/mentoring	43	3300
Political party	43	7300
Racial Equality	43	7100
Royal British Legion	43	4100
Scope	43	4100
Wildlife Trust	43	5200
Animal Protection	42	5200
Armed Forces	42	4100
Home start	42	4100
Millennium Green Trust	42	5100
Fostering/adoption	41	4100
Furniture	41	6100
Recreation/Leisure Group	41	1100
Social enterprise	41	Generic
Writers group	41	1100
Disability group	40	4100
First responders	40	3400
Library	40	2300
League of Friends (Hospital)	38	3100
Sport - Netball	38	1200
Childcare	37	4100
Lions	37	1300
Mediation	37	7200
Medical support	37	4100
Model club	37	1100
Cancer Research	36	2410
Childminding	36	4100
Elderly support	36	4100
Jehovah's Witnesses	36	10100
Marriage guidance/relate	36	1100
Play network/partnership/association	36	4100
Crime prevention	35	7200
Samaritans	35	3300
Buddhist	34	10100
Civic society	34	6100
Media group	34	Private sector
Women's Aid	34	3300
NCH	33	4100
Pigeon racing	33	1100
Town Council	33	Government
Advice Centre	32	7100

Community Trust	32	8100
Productions	32	1100
Rape/Sexual Abuse		
Crisis/Counselling	32	3300
Adult education	31	2300
Community Network	31	6100
Epilepsy	31	2410
NSPCC	31	4100
Parkinsons Disease	31	2410
Playscheme	31	4100
Preschool Learning Alliance	31	2130
Sport – hockey	31	1200
Sport – sailing	31	1200
Voluntary Sector Support	31	8200
Athletics	30	1200
Business advice/centre	30	6100
Volunteer Centre	30	8200
WRVS	30	3400
Community care	29	4100
Community Council	29	6100
NHS/NHS trust	29	3100
Surgery	29	3100
Dyslexia	28	4100
Family/Children's Project	28	4100
Sport – Basketball	28	1200
Camera club	27	1100
Help the aged	27	2410
Miners Welfare	27	6100
Sport – Table tennis	27	1200
Agricultural Society	26	5100
Community Farm	26	6100
Live at home	26	4100
Mothers Union	26	6100
Voluntary/community transport service	26	6100