

[Recipient name redacted]

Please reply to : Sally McTernan  
Assistant Director  
Community Housing Services  
Civic Centre  
Silver Street  
Enfield

E-mail : [email redacted]

Phone : [number redacted]

Fax : [number redacted]

:  
:

Date : 28 April 2014

Dear [Recipient name redacted],

**Re: NKM Report: Understanding the relationship between private rented properties and anti-social behaviour in Enfield**

Thank you for your recent request for information which was received by Community Housing Services. My officers have been in communication with Professor Mayhew and organised a response on the points that you have raised with him.

In addition Professor Mayhew thanks you for your interest in the project and has provided the following responses to your queries set out in blue for ease of reference. The numbering follows the numbers that were used in your original correspondence, with the original query raised detailed below.

1. Does the NKM analysis conclusively prove that residents of private rented sector properties are the main cause of ASB in the London borough of Enfield?

**NKM analysis does not conclusively prove that residents of private rented properties are the main cause of anti-social behaviour in Enfield.**

2. What are the statistical results for the tests (e.g. R2 coefficient) with respect of ASB plotted against proportion of single family dwellings and HMOs? (These were not give in the report)

Assuming that this refers to Figure 6 in the report, this states that the correlation coefficient is 74% or 0.74. Since  $R^2$  is the square root of the correlation coefficient  $R^2$  is 0.55 in the case of high risk single family households. For high risk HMOs, there is no relationship (see later) i.e. the locations of single family households are correlated with the locations of ASB but high risk HMOs are not. [Added: The first "R<sup>2</sup>" should be just "R"]

3. What were the percentages of ASB attributable to higher risk rented households are what were they for other households? What were the confidence intervals of these estimates?

The data was not analysed on this basis because the great majority of ASB incidents could not be directly linked to individual households and only to areas or multi-story blocks as stated in the report. However, of the 1530 ASB incidents which could be linked to addresses, 590 were in the private sector. Analysis showed that average ASB rates among high risk HMOs were the same as for all private sector households but the rate for high risk single family households were double this.

4. What analysis did you do of a link between social deprivation and ASB:

Housing benefit, a measure of a low income household, was associated with single family properties as well as with elevated levels of ASB. Eligibility for means tested benefits is an often used proxy for social deprivation and so your question is addressed in this way. The Index of Multiple Deprivation (IMD), an often used measure of social deprivation, is unsuitable because it has not been updated for several years. It is also geographically inflexible and it is not clear how it could be applied convincingly in this context.

### **Methodology**

1. Newham (page 3) - What is the definition of a "benefit household"? Council tax and housing benefit are mentioned elsewhere but, "benefit household" is not formally defined.

A benefit household is one that received a means tested benefit such as Council Tax benefit, Housing Benefit or both.

2. Newham (page3) – the report states that officials made follow-up visits to properties. How many properties were visited and what proportion did this represent? What was the false positive rate for the modelling? Were you able to establish a false negative rate? Was any other kind of model validation done?

The work in Newham was done in stages. The first stage involved a simplistic model based on a range of factors. Newham then followed up with visits using a team of responsible officers. Over a period they visited some 370 properties to establish tenure and data retrieved from those visits were used to refine and re-calibrate the model resulting in a more accurate set of predictive risk factors. Local authority officers using the predictive data base take into account, not only the risk score but also ad hoc supplementary information before visiting suspected unlicensed properties. It is understood that that 90% or more of visits successfully predicted tenure based on the risk score and other prior information. Since then there have been many more visits and this information is being used to update and refine the analysis. On one group of visits to properties that was personally attended by Prof. Mayhew, the data base correctly predicted actual tenure in all cases. So the model achieves what it was intended to do.

3. (page 3) How were owner-occupiers with lodgers (aka) live in landlords classified?

They are classified as owner occupiers.

4. (page 4) How did you determine that 3 years of records was sufficient? Would this have picked up long term renters?

This was not determined at the outset, but from previous experience it was known that this would be sufficient for the purpose. The aim was to identify currently privately rented properties, not how long the household had been renting.

5. (page 4) How did you test the validity of the key assumption that the Newham risk factors also exist in Enfield (statistically or physically)?

This was a plausible assumption which was discussed at the outset with officials. The Council needed a starting point to establish and confirm the size of the rented sector and types of property most at risk. In due course it was anticipated that the programme of visits would enable them to refine and verify their own risk factors, as has happened in Newham. Indication of the size of rented sector is given in the 2011 Census and this was later substantiated by the work undertaken by NKM.

6. (page 4) What validation work has been done in Enfield: Bullet 2 on this page suggests that visits have been made.

The visits refer to those made in Newham, there have been none so far in Enfield.

7. (page 5) How valid is the assumption that anti-social behaviour takes place very close to where people live?

The first law of geography states that everything is related to everything else, but phenomena that are proximal to each other are more likely to be related or associated in some way. The maps show that ASB is quite widespread but clearly concentrates in the east and south of the borough. There were two outliers **Added: should be "outliers"** mentioned in the report – one was ASB related to Chase Farm hospital and the other to an area of green space corresponding to Trent Park, neither places having many people living there. The reason for higher ASB in hospitals environs is not unusual (e.g. especially in A&E departments at certain times on certain days).

## Results

8. (page 5) Your analysis is that properties not receiving council tax benefit are 3.1 times more likely to be HMOs than properties that receive council tax benefits. This is 3.1 times more likely than what?

This means that the odds of being an HMO are 3.1 times higher than if they were receiving council tax benefit.

9. (page 5) Two or more changes in council tax liable surname in 36 months. How has have you removed possible misclassification caused by, for example, purchase under a maiden name and then change to a married name?

This cannot be verified from any local data source. Any effect is assumed to be trivially small. For example, analysis shows that the effect is not especially sensitive to whether 3 or more changes are stipulated instead.

10. (page 5) Electoral roll – at least one change in surname – how do you discount cohabiting couples with different couples? How do you classify live-in landlords taking in lodgers? How do you take account of children returning to a parental home after university (this would potentially add surnames)?

Surname changes are used as a proxy for turnover. The answer to the particular question cannot be ascertained from administrative sources and so may result in some misclassification between adjacent risk categories, but it must be emphasised that it is most unlikely to change the overall analysis or conclusions, since the risk factors have been shown to be highly predictive of rental status on the ground

11. (page 6) More than three surnames on electoral roll - is this three surnames or three changes in surnames? Again, how do you discount situations in which elderly relatives move in with couples who have different surnames, or where cultural reasons mean that surnames are based on first names between generations, or have different endings based on gender?

As previously mentioned, this variable is used as a proxy for turnover. It means someone new has come in or someone has left. The number of unique names is just the total registrants there are over that period. For example, Mr and Mrs Smith registered between 2011 and 2013 would

count as 2. If someone else extra was added onto the electoral roll in 2013 (as well as Mr and Mrs Smith) that would be 3 and so on.

12. Can you make available the list of properties classified in risk categories 1-3, or indeed in any of the other categories?

It is not possible to comply with this request for raw data.

13. (page 7) Why do risk factors drop in and out of the table? If the number of surnames is the strongest predictor of "HMO risk", then why is there a discontinuity at risk category 8? Similarly, the council tax benefit risk factor drops in and out at numerous points too.

The risk table enumerates sixteen possible risk outcomes based on which risk factors apply to which properties. If a risk factor applies it is denoted by 'Y'; otherwise it is left blank. Properties with the same combinations of risk factors are grouped and risk rated. The reason why the risk score drops after category 8 in the table is because the most influential risk factor drops out.

14. (page 6) What are the confidence intervals for the risk factor weightings? These are given for single family dwellings but not HMOs.

They are given between pages 5 and 8 of the report.

15. (page 7) There seems to be an implicit assumption that wealthier families do not rent in the private sector. Wouldn't this underestimate the number of properties in the single family rented classification?

The aim of the project was to establish the probable scale and location of the private rented sector and to establish the strength of any links to ASB. A 'wealthy' family would still show up in the table even if it would be 'lower risk'. It must be understood that the method profiles all private sector properties and establishes the risk of private rented status. It turns out that that risk is higher, if the household receives housing benefit

among other risk factors. Wealthier renters are harder to detect and are fewer in number, but it is suspected that the landlords of these properties are more likely to comply.

16. (page 7) Why would three or more adults at an address make it less likely to be a single family rented property?

The main reason is that in a simple nuclear family there are two adults with none or more children. This is an assumption and obviously does not cover all contingencies or family or household types (e.g. single adult family, three generation households with grown up children etc.) but it is predictive as a proxy. It is known from work in other places that this type of household is dominant.

17. (page 8) Is the Housing Benefit system sufficiently robust that this can be used as a predictor of tenure type? What checks are required before this benefit is granted and who receives the payment?

Housing benefit data is as accurate as it can be since involves the disbursement of public money and is strictly administered. The detailed benefit rules are extensive and complex, however the key point is that statistically HB status predicts tenure status very well and so the rules themselves are not of relevance in this application.

18. (page 8) Given the uncertainty in your analysis, what validation work did you undertake on these estimates?

Simple plausibility checks were carried out by comparing numbers in each risk category with Newham and this is given in the report. Detailed validation checks will be possible once licensing starts e.g. landlords will self-identify as they apply for licenses and their application can be checked against the data base. Non-applicants that were previously identified as high risk would presumably receive a visit at some point whereupon tenancy could be validated.

19. (page 10) Your analysis has included all anti-social behaviour reports. What have you done to exclude potential duplicates or to

remove things like malicious communications (are they nuisance phone calls or swearing?), which account for over 7% of instances and which may not emanate from within the borough?

ASB data would have been prepared in accordance with agreed criteria as to what constitutes ASB and how it should be collated.

20. (page 10) What would error bars around those data points look like?  
It is difficult to tell whether any rise or fall is statistically significant.

The data are a 'complete' record and not therefore subject to sampling error unless there has been a change in collection methods. In all 29k incidents were analysed; even if this excluded some unreported incidents it is still a very large number. For these reasons, confidence intervals would not add any useful information and could be misleading. Better to look at deviation from the average or speak to the collators to ascertain why the pattern is like it is. The key point to take away from the data is that there is a significant and persistent ongoing ASB problem.

21. (page 10) Your report explicitly states that the number of uniquely linkable addresses was very low compared with the total number of ASB incidents [page 5 suggests 5%]. You then seek to establish a correlation anyway. How valid is this approach?

The possibility of exact address referencing was limited and grid references could refer to open spaces or to blocks of buildings such as flats. It is therefore assumed association by proximity, but this is not the same as cause and effect. In the 1530 cases where ASB could be linked exactly to an address rather than by proximity, 590 of these were in the private sector. The rate among high risk HMOs among these properties was no different from the Enfield average, but for high risk single family properties it was double.

22. (page 11) You have used a 1km x 1km grid, which crossed features such as major A roads [looking at page 12]. Did you then carry out any grid dependence tests to check whether your results still held? What was the grid size for the chart of p14?



In our experience of using spatial analysis re-basing the grid would make no difference to the conclusions. An alternative approach such as imposing a radius around each high risk property and ascertaining whether or not there were more incidents than in low risk properties could provide another comparator if necessary, but one would need to avoid the possibility of 'double counting'.

23. (page 11) Why did you divide number of ASB incidents by number of households and not number of people living in that area? Does your approach assume that entire households perform anti-social behaviour?

The question is not relevant since the aim was to find whether a link existed between ASB and households, not individuals living in households. Note: Census data, the main source of local population data, is now 3 years old. It does not apply to or provide information at a household level and would not therefore add any useful information.

24. (page 12) Your conclusion that "It shows that the pattern of single family privately rented properties tracks ASB incidents quite closely suggesting that there is a relationship" is not obvious to me, both for those your model has classified as single family and those classified as HMO dwellings. (page 13) What do the cumulative distributions prove?

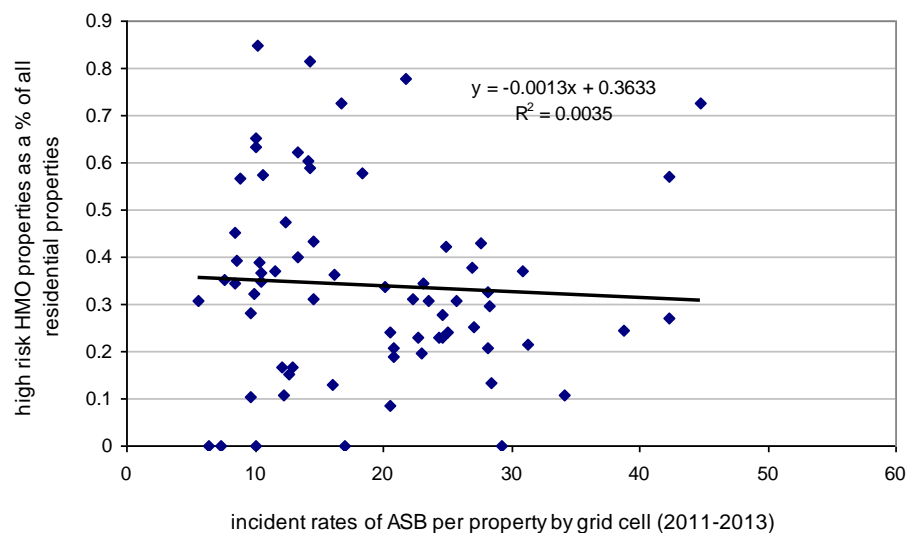
The cumulative distributions show that the pattern of ASB moving east to west relative to the pattern of HMOs differs. Using the K-S test which tests for differences in cumulative distributions the results indicate that the west-east pattern of high risk single family households and west-east pattern of ASB are not dissimilar but that the west-east distribution of high risk HMO households and ASB are significantly different from each other. In plain terms HMOs tend to be spread throughout the borough whereas ASB tends to concentrate in the east and south; single families also tend to be more concentrated in the east and south and are geographically co-incident with where most of the ASB occurs.

25. (page 14) Can I have the backing data for this chart please? What is the R2 coefficient for the straight line? The line of best fit doesn't go through the origin. What does that mean?

Coefficients of regression lines are subject to statistical uncertainty. In this case the regression constant is not statistically significant from zero. So the argument about it not going through the origin is not sustained. The actual equation is:  $Y = 0.2439 + 0.1076 X$

26. (page 14) Have you produced a similar chart for HMOs? What would the R2 coefficient be?

See below. Unlike the single family case, HMOs show no relationship with rates of ASB



27. (page 14) You state that "because we have not uniquely ascribed ASB to individual households (for reasons previously given) this does not necessarily demonstrate cause and effect". You further state (at the foot of page 14) that "evidence suggests that there is a correlation between rented properties and incident rates of ASB although statistically the correlation is weak". This seems a contradiction in terms. Either there is a statistical correlation, or any

possible idea that there is a correlation is not valid? Which is it please?

No, it is not a contradiction. The results demonstrate association by proximity, but not by cause and effect. The evidence for a direct correlation is weaker, but as previously stated we found that ASB rates in high risk single family households for which data could be directly linked was double the rate for other household types.

28. What analysis have you done of a possible link between the Enfield Homes properties (social rented sector, not leaseholders) and ASB?

No work was done on the social housing sector because that was not the aim of the project, although it would be possible to do so for comparative purposes, in addition the identity of social landlords is already known and social housing is already subject to regulation.

This concludes our response to the questions that you have raised thus far, I hope this information has been helpful. If you have any further queries, please do not hesitate to contact me.

Yours sincerely,

**Sally McTernan**

**AD Community Housing Services**