

## **Written evidence from the National Institute of Economic and Social Research**

The Committee's hearing on Wednesday 19<sup>th</sup> October 2016 discussed the National Impact Study of the Troubled Families Programme, published on the 17<sup>th</sup> October. There are number of points that NIESR would like to clarify.

### **The production of the report**

1. The data on families who met the eligibility criteria for the programme used in the report were collected and supplied by local authorities. As was discussed in the hearing, there were a number of issues relating to this data. This is not unusual in evaluations of this type using administrative data, but was compounded by the reliance on local authorities to collate information, sometimes from multiple sources and stored in different formats. Having received the data we required to commence data cleaning and analysis in January 2015, we supplied a full draft report to CLG on 28 July 2015. The main findings have not changed since this draft.
2. In response to a succession of queries raised by CLG over a period of some months, we reworked the analysis, omitting data from three local authorities (of 59) where the data appeared to be unreliable. This did not make a significant difference to any of the key findings. That is to say, the key finding of lack of significant or systematic impact was the case before and after this data was excluded. We also performed additional analyses. Again, there was no material change to any of the findings. This process is fully detailed in Annex A of the published NIS report.
3. In Spring 2016 the evaluation team suggested sending the report for external peer review to provide further reassurance to CLG that the approach take to the data and analysis was sound. The external peer reviewers, Professor Anna Vignoles (University of Cambridge) and Professor Jo Blanden (University of Surrey, were chosen and approached by CLG.
4. The review concluded that the methodology, approach and the quantitative analysis contained in the evaluation were sound. It also specifically praised the efforts of the evaluation team. It did, however, note the problematic nature of the underlying data. In particular, Professor Professor Vignoles concluded that "the underlying data was very messy (through no fault of the evaluators) with partial and incomplete data making an impact study of this kind always quite difficult." She stressed that the "main quality-assurance issues under discussion are rooted in this problem and challenges in the nature of the population of interest rather than the evaluation per se."
5. It is important to note that the review did not result in any changes to NIESR's analysis, nor to the key findings of the evaluation, although it did suggest some changes to the

presentation of those findings, which were incorporated into the final report, a near-final draft of which was supplied to CLG in June 2016. These changes highlighted further the problems with the quality of the data supplied by local authorities and their potential implications for the interpretation of the findings of the analysis. .

6. Both the “Key findings”, set out in the Executive Summary, and the caveats to those findings fully reflect our discussions with CLG and the input of Professor Vignoles. The final text was explicitly signed off by CLG. It is important to note that, as the text states, while the data was indeed imperfect, there is no reason to believe that our findings are biased (in the statistical sense); they are robust to a large number of checks, as detailed in the report and set out above; and they are consistent with the separate and independent analysis using survey data.

### **Public presentation of the report**

7. At Q52 (and again at Q127). Dame Casey states: “They [NIESR] had not, frankly, put any of the caveats in the public domain” and that “they have misrepresented their own research”. This is incorrect. [NIESR’s press release](#) (also attached as Annex A), issued immediately after the report was published, includes the precise text from the Executive Summary describing the data and the caveats; as noted above, this text had been explicitly agreed by CLG. We supplied a copy of our draft press release, including this text, to CLG on the morning of Friday 14 October.
8. Dame Casey also refers to a personal blog published by Jonathan Portes immediately after the publication of the report. As well as linking directly to the full evaluation reports and the NIESR press release, and quoting extensively from the press release, this blog states explicitly:

**It's worth noting the caveats in NIESR's evaluation, which state, carefully and correctly, that issues with data quality mean that “the results cannot be taken as conclusive evidence that the programme had no impact at all”.**

### **Interpreting the findings**

9. Ms Dawes and Dame Casey correctly made clear that the key finding of the evaluation was that there was no evidence that the programme had any significant or systematic (net) impact on the key outcomes for which we had data. As they also made clear, this did not mean that families who participated in the programme had not, on some measures, experienced improved outcomes: but rather that the evaluation found that very similar outcomes were experienced by a matched comparison group who did not participate in the programme (or who participated in the programme at a later date. In

other words, there is no evidence that the programme caused any of the changes (positive or negative) experienced by the families who participated in the programme.

10. It was precisely the objective of the evaluation to find such evidence, or its absence, and the evaluation achieved this objective, albeit in a negative sense. Ms Dawes stated that this was unusual or indeed unprecedented:

**[Q130] “It is important to draw out that although the Troubled Families programme evaluation is saying that there is no attributable impact from the programme when you do the detailed statistical analysis, that is not an analysis that the Dundee Project did. In fact, it is not an analysis that any other programme has ever done. So the Troubled Families programme tried to measure itself by a much higher bar than is usual in social policy.**

**[Q143] What we are not able to say definitively, with statistical certainty, is that it would not have happened anyway. No other programme has asked itself that question, of course; we haven’t been able to answer it either.**

This is incorrect. The question of whether a programme has achieved a (statistically robust and significant) net impact with respect to the counterfactual (of no programme, or of individuals not participating in the programme) is the standard question that almost any quantitative evaluation of a social policy programme asks. For example, the [early evaluation of the labour market impacts of Universal Credit](#) adopted a very similar (propensity score matching) methodology to that used in the TFP evaluation, and found a (small but statistically significant and robust) positive net, or “attributable”, impact. There are numerous other similar examples across the social policy field, showing both positive net impacts and, as with this evaluation, no net impact.

11. It should also be noted that it was originally planned that the evaluation would include a cost-benefit analysis of the programme. However, since the impact evaluation was unable to attribute any quantified benefits to the programme, this analysis was not produced. Obviously, had CLG believed that it was extremely unlikely that the evaluation would show any quantified benefits, they would not have commissioned a cost-benefit analysis in the first place.

### **The NIESR evaluation team**

12. The NIESR team was led by Dr Helen Bewley from the inception of the project until June 2016 (when Dr Bewley left NIESR). Anitha George and Cinzia Rienzo also played an important role in the analysis and reporting from 2014 onwards. Mr Portes joined the

research team in April 2015 and oversaw the final stages of agreeing the report with DCLG following Dr Bewley's departure from NIESR.

13. Mr Portes had, prior to joining the team, commented publicly on a number of design aspects of the programme (in particular the origins and use of the 120,000 number); and his personal blog revisits these broader issues in the light of the findings of the evaluation. Since the blog, and Mr Portes' other public commentary on the programme, ranges considerably wider than the findings of the evaluation, he has done so in a personal capacity. His views are therefore his own and not those of NIESR, the NIESR evaluation team nor of the evaluation consortium.

*26 October 2016*

**Annex A: Text of NIESR press release issued October 17, 2016 [at approximately 6.30 pm, subsequent to the official publication of the evaluation on the CLG website]**

**Key findings**

NIESR's analysis, published today, forms part of the wider evaluation of the Troubled Families Programme, commissioned by the Department of Communities and Local Government from a consortium led by Ecorys. The key finding is that across a wide range of outcomes, covering the key objectives of the Troubled Families Programme - employment, benefit receipt, school attendance, safeguarding and child welfare - we were unable to find consistent evidence that the programme had any significant or systematic impact. The vast majority of impact estimates were statistically insignificant, with a very small number of positive or negative results. These results are consistent with those found by the separate and independent impact analysis using survey data, also published today, which also found no significant or systemic impact on outcomes related to employment, job seeking, school attendance, or anti-social behaviour.

**Methodology**

The purpose of the analysis was to estimate the impact of participation in the initial phase of the Troubled Families programme on a range of outcomes encompassing benefit receipt, employment, educational participation, child welfare and offending. The data used were compiled from information provided by local authorities and from national administrative datasets covering tax and benefits receipt, offending, educational attainment, schooling and child social care. Two methods were used to estimate the causal impact of the Troubled Families programme. The main focus in the report is on the analysis which uses a technique known as propensity score matching (PSM). PSM matches those subject to the programme to a comparison group of families that have a similar propensity to be treated, based on their observed characteristics, but who do not in fact receive assistance. Provided the two groups are matched on all characteristics which determine both whether the family participates in the programme and which impact on the outcomes that they experience as a result of participation, the impact estimate should provide an accurate and unbiased estimate of the true impact of the intervention. A supplementary analysis exploited the fact that families started on the programme at different points in time. This was known as the waiting list analysis. Those who started on the programme at a later date were used as a comparison group for families that started earlier, with outcomes for the comparison group observed in the period before they started to receive support.

**Data issues and caveats**

Fifty-six local authorities provided the data used in this study between October and November 2014. These data were then matched to national level administrative data sets. Data was obtained on approximately 25 percent of the 120 000 families that participated in

the programme, representing a large sample, and enabling us to undertake detailed analysis. However, the data supplied was of variable quality. As a result, some important data was missing, and it was necessary to make certain assumptions in assigning individuals to treatment and control groups. In addition, a significant number of individuals were not matched to certain of the administrative data sets, and this necessitated further assumptions – Press Release for example, we assume that individuals not matched to employment records were not employed. It is possible that in some cases these assumptions did not reflect the true circumstances of particular families and as a result, the findings may be subject to measurement error. However, as long as the prevalence of missing or incomplete data is random and/or does not differ systematically between the treated and control groups, the conclusions will remain unbiased. None of the available evidence suggests that such systematic differences exist, and the possibility of significant bias therefore appears unlikely. The fact that separate analysis, using survey data, produces results consistent with our analysis is further evidence that any bias that would impact our results to a significant degree is unlikely. However, given the data issues this possibility cannot be entirely excluded. It is not possible to say with certainty how this might affect the reported results.