

# Evaluation of the Phase 1 Troubled Families programme

1

Susan Purdon

**bpsr** Bryson Purdon  
Social Research

# Overview

2

- Aims of the TF programme & the claims made
- The evaluation
  - The various elements
  - What each added to the overall evidence
- The impact studies
  - What they show ('no discernible impact')
  - The data issues/problems
  - Could we have missed genuine impacts?

# The TF programme (2012-2015)

3

- 120,000 'troubled families' were targeted
- Criteria were:
  - Involved in crime or anti-social behaviour
  - Children not in school (exclusion or truancy)
  - Adult on out-of-work benefits
  - Cause high costs to the public purse

# The TF programme (2012-2015)

4

- LAs were given flexibility, but guidance from CLG noted five key factors:
  - A dedicated worker, dedicated to a family
  - Practical 'hands-on' support
  - A persistent, assertive and challenging approach
  - Considering the family as a whole – gathering the evidence
  - Common purpose and agreed action
- Intended to generate a move from reactive service provision to earlier intervention

# The TF programme (2012-2015)

5

- DCLG operated a Payment-By-Results (PBR) model: payments being made for families ‘turned around’
- At May 2015, CLG statistics showed that 117,910 families had been identified and 116,654 (99%) had been ‘turned around’:
  - 90% because of improvements in crime/ASB or truancy
  - 10% because of entering continuous employment

# The evaluation

6

- Commissioned in Jan 2013, to a consortium led by Ecorys
- Five key elements (ordered by approx. timing of findings):
  - Family monitoring data (tracking of families by LAs)
  - A process evaluation
  - An impact evaluation based on a family survey
  - An impact evaluation based on administrative data
  - An economic assessment (planned as a CBA)

# Family Monitoring Data (FMD)

7

- Evidence from the FMD was extremely positive
- Not strictly part of the evaluation, but data recorded by LAs on a random sample of their families:
  - Three points in time: at start; at time of PBR claim; at case closure
  - Gathered detailed profile data on families, including the range of presenting problems and how these progressed
  - Data on around 17,000 families

# % families with reduced problems at case closure

8

Problem at baseline	% of families improved
Unauthorised absences	74%
Workless household	17%
Risk of eviction	61%
Rent arrears	40%
Parenting issues	38%
Adult with proven offence in last 6 months	61%



# But:

9

- Data collected and recorded by LA staff
- No information about sustainability of any improvements

# Process evaluation

10

- Evidence from the process evaluation was more mixed but still broadly positive
- Qualitative research involving:
  - professionals from 20 case study LAs, tracked over three years
  - interviews with a further 50 TF co-ordinators
  - two interviews with each of 22 families (start; 12-18 months later)

# Process evaluation

11

- Widespread evidence of transformation in local services and systems;
- The appointment of a TF co-ordinator helped to generate buy-in;
- Encouraged data-sharing across the various agencies;
- Families were generally positive;
- The PBR funding allowed for expansion of the key-worker model that was firmly believed to be a better approach
- Inevitably, some LAs more successful than others

# Impact study 1: survey

12

- Two groups of families interviewed:
  - 495 families who had started the programme around 9 months earlier (intervention group);
  - 314 families that were just about to start (comparison group)
- Difference in outcomes for the two groups (after matching) gives an estimate of impact
- Purpose was to generate estimates of impact across a very broad range of outcomes

# Number of problems though:

13

- Comparison group was intended to be made up of any non-TF families who were TF eligible 9 months earlier
  - But, proved impossible to identify and recruit these
- Instead, comparison group was made up of those just about to start
  - Implication is that this is closer to a pre-post design than a formal comparison group design

# Biggest problem:

14

- Interviews with the TF group took place around 9 months after programme start
- Cross-checking with the FMD suggested that around 70% were still on the programme at the time of interview
- So the survey did not collect post-intervention outcomes for most

# The two groups had different services

15

	Troubled Families Group	Comparison group (unmatched)
	%	%
<b>Whether received support in the last year regarding:</b>		
Getting on better as a family	40	15
Getting children to school each day	29	11
Managing money/debts	21	8
Getting job/training	18	6
Finding things to do in community	17	8
Accept responsibilities better	15	3
Daily routines	14	4
Looking after children better	14	9
Stopping getting involved in crime	12	5
Feeling safer at home	11	4
Making home nicer (eg housework)	10	3
Keeping children living at home	8	4
Reduce drugs	6	4
Reduce alcohol	4	1
Health issues	1	1
School issues	-	-
None of these	23	56

# But very little evidence of impact

16

	Troubled Families Group	Matched comparison group	Impact
	%	%	
Main carer in work or job-seeking	43	41	2
No threat of eviction	93	88	4
Kept up with bills in last 3 months	79	79	0
No ASB actions in last 3 months	84	86	-2
Scored less than 8 on Malaise scale (no sign of depression)	45	43	2
Scored 23 or more on SWEMWBS (higher well-being)	48	54	-5
No domestic abuse in family in previous three months	58	61	-3
Child attended at least 85% time	70	65	5
Not been called in to school about child's behaviour	55	53	2
<i>'Managing well financially'</i>	69	59	10*
<i>'Managing better than a year ago'</i>	26	19	7*
<i>Feel in control</i>	69	60	9*
<i>Feel positive about the future</i>	69	61	8*
<i>Confident worst problems behind them</i>	68	52	17*



# Comments

17

- If impacts occur, they occur after more than 9 months;
- If survey design mimics a pre-post, little evidence of any net change in family outcomes over a nine month period

# Impact study 2: administrative data

18

- Based on data from 56 LAs
- Families (and individuals) going through TF were matched to national admin data and tracked for up to 18 months:
  - DWP data (benefits and employment)
  - National Pupil Database (truancy)
  - Police National Computer (cautions and convictions)
- Total sample=c.25,000 families

# Comparison group

19

- Families identified as meeting two of the three criteria, but not entering the programme before Autumn 2014
  - c.10,000 families

# Matching to admin data

20

- Inevitably messy (matching done on name, address, gender, postcode, dob)
- Assumptions needed for when no match found:
  - If not matched to employment then assume not-employed
  - If not matched to PNC then assume no contact with police
- Inevitably some degree of measurement error & misclassification

# Matching TF group to comparison group

21

- Propensity score matching
- Matching variables being:
  - Household characteristics (age, gender, ethnic group)
  - Recent history on outcomes (employment, benefits, truancy, crime)
  - No match possible on ASB
- Assumption: if control for recent history, *if* TF is effective then outcomes will be better in TF group; and
- Without TF, the two groups (TF and comparison) would have been on the same trajectory

# No evidence of a TF impact

22

	Troubled Families Group	Matched comparison group	Impact
	%	%	
% claiming out of work benefits at 18 months	45.3	45.3	0
% employed at 18 months	40.7	43.6	-2.9
Adults with caution or conviction in 7-18 month period	8.7	7.8	1.0
Children with caution or conviction in 7-18 month period	5.8	5.7	0.1
Children absent for 15%+ of time	20.3	21.2	-0.9

# And not a great deal of change over time

23

	Troubled Families at start	Troubled Families after 18 months
	%	%
<b>% claiming out of work benefits</b>	43.9	45.3
<b>% employed</b>	28.5	40.7
<b>Adults with caution or conviction in 12 month period</b>	10.0	8.7
<b>Children with caution or conviction in 12 month period</b>	5.6	5.8
<b>Mean % time absent</b>	11.2	9.6

# How to reconcile the PBR/FMD data with the impact study data

24

- PBR/FMD both show large amount of improvement for families
- Admin and survey data show much smaller levels of change
- But, PBR/FMD measured at time point chosen by LA teams;
- Admin and survey data collected at fixed point in time
- Most plausible explanation is that families cycle in and out of problems



# Are our estimates correct?

25

- Evaluation team concluded the impact estimates were close to correct;
- Govt took the view that the data problems undermined that
- But sensitivity analysis did not change the results;
- No obvious explanation as to how the data problems could have obscured positive impacts
- It is *possible* that the TF and matched comparison groups *were* on different trajectories, but unlikely that the difference was very large

# Conclusions

26

- The Phase 1 TF programme did not change outcomes for families, at least within 18 months of start dates;
- But there is still uncertainty as to why;
- Although the evaluation has shed a lot of light, the optimal large-scale approach remains elusive



# Troubled Families Evaluation: Discussion

**Paul Gregg**  
University of Bath

# The Policy

Parenting/Family Support Interventions have been widely discussed and advocated but few substantial government initiated interventions

Of these few – interventions have not been sustained and have poor evaluator record. Family Nurse Partnership did not replicate US success. Sure Start changed parenting and some child behaviours but not cognitive attainment.

# The Policy

3. Initially 120k families but expanded to 400k of 8m families in UK or 5%. This not a target programme on extreme minority but a moderately broad target – larger than Family Nurse Partnership for instance. Hence a major addition to the Social Work / Family Support policy agenda.
4. On range of widespread prevention programme and more targeted after problems revealed (Cure) spectrum, it sits in the middle.

# The Policy

1. Origins of Troubled Families policy lie in **Family Intervention Projects** – hailed as effective policy in Labour Era. As far as I can see these were not subject to rigorous evaluation with clear building of a comparison group or even clear data tracking of family improvements
2. In contrast there are evaluated positive programmes in this area. Such as Pre-natal Early Infancy Project PEIP or more closely in age Strengthening Families Programme

# The Evaluation

A good evaluation rests on a number of key ingredients:

# The Evaluation

Clear Policy Objectives: Here there are a wide range  
Evaluation gives little guidance on importance  
of outcomes

In terms of costs to government there is broad  
hierarchy, for families may be similar ranking

1. Children in Care  
Child Protection Plan  
Child in Need Plan
2. Crime:- especially serious crime
3. Health behaviours:- Substance abuse, drugs,  
alcohol, smoking
4. Education of child and parental work/benefits  
receipt



# The Evaluation

## Treatment and counterfactual

We need to know both what the treatment was but also what the alternative was.

Here the treatment was fairly well documented by the Process Evaluation – the alternative was less clear.

Produced three delivery models:

1. Dedicated team
2. Hybrid
3. Embedded in delivery teams also potentially doing whole family intervention

Policy for others than Troubled Families target group may well have changed – do areas in group 1. produce a cleaner counterfactual

# The Evaluation

## Treatment and Counterfactual Populations

The evaluation methods is a delayed treatment model with Propensity matching to balance the two samples

1. Treated group includes a lot of referral from other services and within LA team
2. Are the delayed group in receipt of any services? If they weren't, why weren't they :- lower need?
3. Is there possible unobserved differences between those referred on subjective criteria
4. Need more on pre-programme trends/data of the two groups – matching on cross-sectional data plus trend data?
5. Propensity matching does not guarantee all characteristics are the same, just the basket of predictors. Do populations look different at all.

# The Payment System

Payment By Results creates powerful incentives to claim success – Individual Learner Accounts

There are large potential problems where agency receiving payment can influence the measure used for payment without real success being achieved.

You need either: Administrative data which is accurate and timely or

Independent assessment of outcomes

Sustained outcomes: Used in the Work Programme but not here where it seems crucial

Study should be able to simulate PBR triggers being met but not sustained.

# Conclusions

- Clear pre- and post- programme tracking of objectives
- Funding on sustained improvements on clear set of independently verifiable objectives
- A range of support programmes for specific needs of family (e.g. mental health, anger management or family functioning)
- Explore an earlier in childhood intervention version