THE LIKELY IMPACTS OF TARGET SETTING AND PERFORMANCE REWARDS IN LOCAL TRANSPORT

Dr Greg Marsden
Senior Lecturer Transport Policy & Strategy
Institute for Transport Studies, University of Leeds

Mrs Charlotte Kelly
Research Fellow
Institute for Transport Studies, University of Leeds

Mr John Nellthorp
Senior Research Fellow
Institute for Transport Studies, University of Leeds

Abstract

All local transport authorities in England have, since 2000, been obliged to submit five year plans for local transport. The plans set out the overall strategy, key policies that will be implemented and how the strategy will be resourced. The central government now adjusts the funding allocations up or down by up to 25% based on the quality of the plans and, on an on-going basis, achievement against the targets proposed in these plans. This paper presents a theoretical and practical assessment of the impacts of these incentives on local authority performance.

The research has employed a mixed methods approach with interviews, questionnaires, the development of a game theoretic representation of the process and a laboratory experiment. The findings have been discussed with practitioners. The research suggests that the presence of performance rewards, in a scheme where authorities believe they have a reasonable chance of being rewarded, leads to authorities setting more ambitious targets. Whilst it is not certain that these targets will be met it appears that the absolute outcomes achieved are likely to be better than they otherwise would have been. Generic conclusions are drawn about the conditions under which target-based performance reward schemes will work best.

1. Introduction

Targets set out the level of performance that an organization aims to achieve for a particular activity within a given timeframe. This might be for example a commitment to reduce fatalities on urban roads by 10% over the next five years. Managing public services through the use of targets is not new (Hood, 2006). However, it is becoming increasingly widespread globally (FHWA, 2004; Moynihan, 2006 and Hodgson et al., 2007) and in the field of transport. Despite this, a 2005 review of the use of targets in transport found that "there is little published evidence on the effect of targets on the performance of the transport sector." (Marsden and Bonsall, 2006, p191).

Since 2000, the Department for Transport (DfT) has required local authorities in England to prepare 5 year 'Local Transport Plans' (LTPs) in which local authorities set out the policies and expenditure required to make the integrated transport vision a reality (Kelly et al., 2006). Targets were also required in these first plans. However, a lack of experience in setting targets, an absence of good baseline data for many indicators and no clear guidance on how performance against these targets was to be assessed led to difficulties in comparing how well authorities had done (Atkins, 2007). Despite these difficulties, the annual assessment of the performance of local authorities has moved to one where authorities are ranked and rewarded according to their performance. The emphasis

on targets and monitoring performance has been taken through into the second round of LTPs (LTP2) and this initiative forms the basis of the research presented.

For LTP2 each authority is given an initial funding allocation for integrated transport measures for the period 2006/07 to 2010/11 using a formula allocation mechanism. The formula was developed based on four agreed priorities (congestion, accessibility, air quality and safety) and adapted on a needs basis (described in DfT, 2005). The funding allocation was to be varied up or down by up to 25% based on the quality of the LTP2 plan, past evidence of delivery, the ambition of the targets set and, over time, achievement against these targets (see DfT, 2004 for full details of the initial aims and requirements of the scheme). Kelly et al. (2006) suggest that most of the factors involved in varying the performance rewards were related to achievement of targets or "whether or not authorities are likely have set their targets wisely." (p2)

The goal of central government in developing this system appears to be a desire to maximise the effective outcomes of its local transport plan spending. That is to say, it wishes to reward those authorities that appear to deliver results most cost effectively. In so doing, it also has aspirations to raise the quality of planning and, through some form of competition, increase the net outcomes of its spending. Critics suggest that such processes are limiting in several ways. They may for example focus behaviours only around those aspects that can be measured, can encourage short-term decision making and create perverse incentives (Smith, 1995 and Hood; 2006)

A mixed-methods approach was required to research the likely outcomes of a performance-led rewards scheme in local transport planning. Underlying the central government's goals appears to be a belief that an incentive-led competition will drive up the net performance of all authorities. To assess whether there is an underlying theoretical basis for this, and under what conditions it was likely to hold true, a game-theoretical framework was established. Game theory can provide insights into the key drivers of outcomes in a situation where players interact but a number of simplifying assumptions are required to derive a model and these assumptions are critical to the usefulness of the model (Kreps, 1990). An initial investigation of the details of the performance regime and how local authorities were approaching the target setting task was undertaken and this is described in Section 2. Section 3 provides an outline of the game theoretic model and the key findings. To test the model assumptions further, a laboratory experiment was designed to test, through a simulated city environment, the understanding developed through the game theoretic framework and this is described and presented in Section 4. The implications of the findings from the three strands of work are then discussed in Section 5 before some general conclusions about the use of targets and performance rewards in transport are drawn.

2. The Local Transport Plan game in practice

Five year LTPs were introduced in 2000 and replaced a previous system of annual funding settlements. An evaluation of the first LTP period (2000/01 to 2005/06) has revealed overwhelming support for the development of such strategies amongst local authority transport practitioners both in terms of the ability to plan in a more strategic manner but also to raise the profile of transport within local government (Atkins, 2007). Concerns were raised however about the degree to which progress was being made against targets that had been set although this was, in part, due to a lack of expertise in and guidance on target setting when the plans were submitted. The second round of LTPs introduced two key changes of relevance to this research:

1. The first round of LTPs had been a bid for capital funds. Rewards were linked to the ambition of the bid and, subsequently, achievement against ambition (although this was only weakly signaled in the 2000 guidance). The second round of LTPs was to be based on a formula allocation which would then be adjusted up or down (as described in Section 1). The Department for Transport stated that no extra funds would be available for performance. The implication of the zero sum game is that there would be winners and losers and this could affect what was delivered in any authority.
2. Clearer guidance was given on what had to be measured and how it would be assessed. 15 mandatory indicators were identified (Table 1) and a maximum of 40 (local and mandatory) indicators was recommended (DfT, 2004). Thresholds which would denote satisfactory and stretching performance requirements were to be published for some of the 15 mandatory indicators with assessment of the ambition of local indicators to be made on a case by case basis for which the authorities were to provide supporting evidence.

Table 1: Mandatory and Best Value Performance Indicators

Mandatory LTP Indicators	Mandatory Best Value Performance Indicators
LTP1: Accessibility target	BVPI 223: Principal road condition*
LTP2: Change in area wide road traffic mileage	BVPI 224: Unclassified road condition*
LTP3: Cycling trips	BVPI 99x: Total killed and seriously injured casualties (KSI)*
LTP5: Bus punctuality indicator*	BVPI 99y: Child killed and seriously injured casualties *
LTP6: Changes in peak period traffic flows to urban centres^*	BVPI 99z: Total slight casualties*
LTP7: Congestion^	BVPI 102: Public transport patronage*
LTP8: An air quality target^	BVPI 104: Bus satisfaction*
	BVPI 187: Footway condition*

Key ^only a requirement for certain authorities
*Indicators where stretching/ satisfactory thresholds were set.
Source: DfT(2004)

The Department for Transport had therefore established a system whereby the local authorities bid against each other for performance reward based on the targets they set and the apparent (and then actual) ability to deliver them. The key guidance given to local authorities relating to how they would be assessed was that targets should be based on outcomes (i.e. authorities were competing on the number of fatalities they reduce not the number of street lamps that are working) and that targets should be "challenging but realistic".

To obtain an in-depth understanding of how local authorities were approaching target setting interviews were conducted with a range of people from six authorities involved in LTP2 submission ranging from those responsible for individual targets, heads of LTP, regional government officers, and consultants to those with political control (e.g. councillors). These interviews were used to help provide a perspective on the process of LTP2 target setting and to develop the questionnaire that was sent to all LTP submitting authorities in March 2006 investigating how they chose their targets for LTP2. The questionnaires were sent to all 82 LTP2 submitting authorities of which 31 responded (a 38% response rate). The responses from the questionnaire were then compared with how the submitted targets had actually been assessed by the DfT (DfT, 2006a, 2006b). A full account of this process is available in Kelly et al. (2008).

The majority of authorities agreed with the statement that they were more likely to increase/maintain their future funding levels if they achieved the targets they had set. In addition 75% stated that if they did not receive their full allocation of funding (100% formula) that they would not be able to achieve their targets in 2010/11, no authorities disagreed with this statement. Authorities were therefore aware that they need to set targets that they could meet, as this would impact on their funding levels.

The interpretation of challenging was more varied across authorities. Authorities were asked to rate the likelihood of meeting the targets they had set. The average values for all authorities are shown in Table 2.

Table 2: Risk Assessment by target classification (satisfactory, stretching)

	Mean % chance of meeting target		
Indicator	**All Targets**	**Satisfactory Targets**	**Stretching Targets**
Killed and Seriously Injured	79.2	80.4	78.2
Child KSI	75.8	76.2	74.6
Slight accident rate	76.3	73.9	78.4
Bus Satisfaction	72.6	71.9	77.5
Bus Patronage	72.2	63.0	76.8
Bus Punctuality	70.2	68.3	71.7
Unclassified Road Condition	74.1	72.3	79.3
Footway condition	74.1	69.3	80.0

Across all authorities the average expectation of meeting the targets is 70 to 80% suggesting that authorities are indeed challenging themselves and taking some risks. Interestingly, with the exception of Killed and Serious Injury accidents (where the differences are small), authorities are more certain of achieving targets which are assessed as stretching. This is indicative of authorities playing strategies where they set stretching targets in areas they know they will perform better in.

Authorities were asked to make an assessment of how they thought the Department for Transport would assess their targets and how they would assess their own targets for the mandatory indicators. In reality, very few unsatisfactory targets were set across all local authorities so the authorities were marked with a 2 if they set (or were assessed as having) stretching targets and 1 if they were not stretching. A comparison of the forecast and actual assessments (DfT, 2006a; DfT, 2006b) is shown in Figure 1.

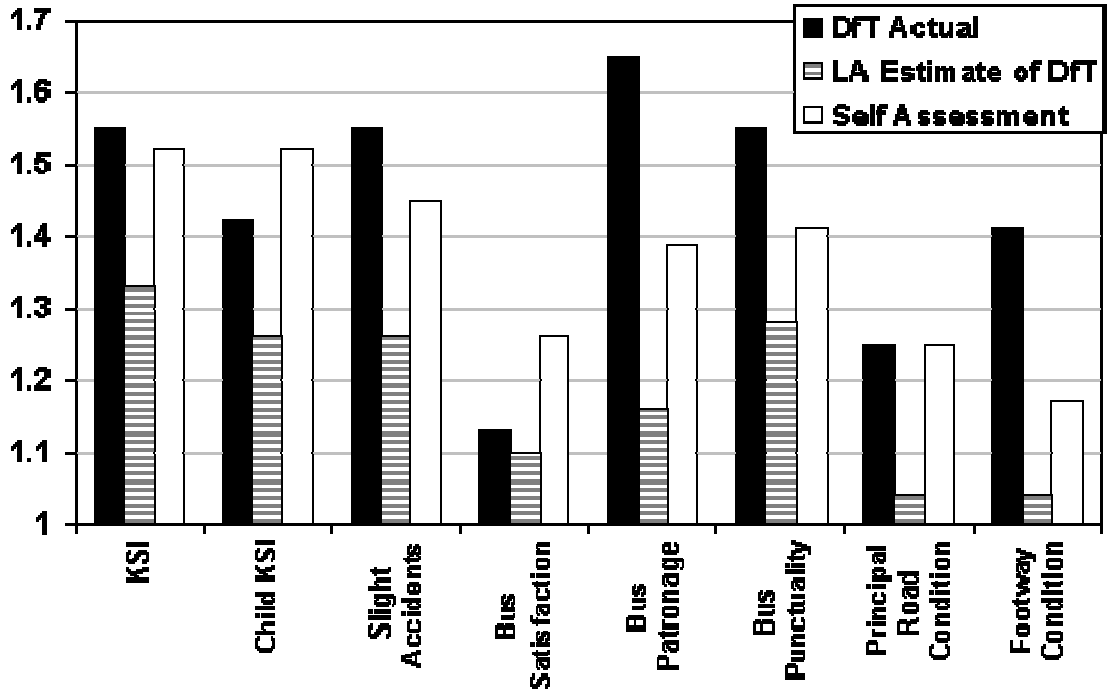

Figure 1: Average Scores for each target for the three assessments (Source: Kelly et al., 2008)

It is clear that the local authorities took a pessimistic view of how they would be assessed by the Department for Transport. The actual assessments are more closely in line with the authorities' own perceptions of the degree of stretch that the targets represented to them given their own local circumstances. This implies that the DfT has invested substantial effort to create a level playing field where it assesses each target on its merit and not just according to deterministic rules.

Evidence from the questionnaire and our interviews suggested that whilst the exact process from turning their LTP submission into a performance reward was not clear, the authorities were competing with each other, the performance rewards (or penalties) were important to them and they were submitting plans which would demonstrate the use of targets which they felt were challenging but achievable. The relative pessimism over how they would be assessed suggests that some concerns existed over the fairness of the assessment process but that there was little complacency in the approach to setting targets.

Limited evidence of authorities gaming the system was found. For example, some authorities chose to set targets for 0% growth in bus patronage even where the evidence suggested a decline. This may have been due to the local political implications of 'aiming for decline' or due to the potential funding implications. E.g. "targets were set to at least meet the minimum criteria of the LTP guidance even if this seemed ambitious" (LTP officer).

We conclude that, overall, the authorities were competing with each other and were using their targets as the primary signal for doing so. They understood the difficulties of meeting targets in

different areas and prioritized the stretch they put forward in their plans accordingly. The funding reward was important to them for delivery although it was also important for other reasons such as their position within wider corporate assessments.

3. The Local Transport Plan Game in Theory

To inform our understanding of the likely impacts of the performance reward scheme for LTP2 a theoretical model was developed. Game theory was developed for studying strategic interactions between small numbers of participants whose actions impinge on each other and has been suggested as a "core approach to the analysis of institutional relationships" (De Palma and Lindsey, 2003, p1). The official rules as well as the self-reported interviews and questionnaire findings suggest that the LTP2 contest can be characterized as a competition between authorities for funds from the Department for Transport. The game has been represented as a rent seeking contest (Clark and Riis; 1998; Blavatskyy, 2004). The model description and development was described in a UTSG paper from 2007 (Nellthorp, 2007) so this section focuses on the key results emerging from the model:

1. Whether or not it is worth a player (authority) actively playing the game by putting in some positive effort, depends on how deterministic the game is and on the total effort exerted by the other players
2. For an equilibrium solution, the power function which defines the level of determinism within the game should not exceed N/(N-1) (= 82/81 in this case) and the number of prizes should be below 0.63N (approximately).
3. As the number of players increases (and 82 is a large number of players in this context) the proportion of the prize fund that must be allocated to the first few prizes in order to secure a symmetrical pure strategies equilibrium increases. If such an equilibrium is not achieved then either an asymmetric equilibrium may be achieved or no equilibrium may be possible.
4. The symmetric equilibrium has the desirable property that it maximises aggregate effort across the players.

The sensitivity of the results was assessed using a range of prize structures, varying both the number of prizes and the prize gradient (the ratio of the first prize to the k^{th} prize). The results of a sensitivity test run are shown in Figure 2.[1]

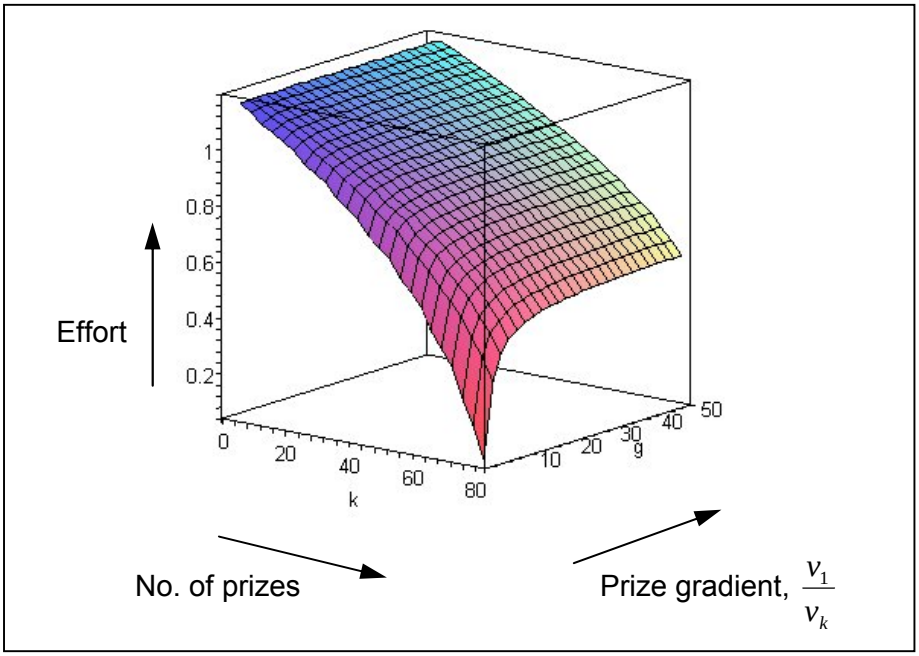

Figure 2: Effort responses in the 82 player game

Figure 2 suggests that the greatest effort will be exerted when a single prize is offered. Intuitively, this is because for a given prize mass, the single prize creates the greatest competitive force

[1] Sensitivity test for a prize fund of V = 100 with a power function of r = 1

between the players who, by assumption, are all equally capable of winning the prize. However, if the designer does choose to offer a large number of prizes, the prize gradient begins to play a major role. For example, "when there are in excess of 60 prizes, moving from a prize gradient of 2 to a prize gradient of 10 has the effect of roughly doubling the amount of effort exerted, for the same prize fund" (Nellthorp and Marsden, 2008, p13).

Like the single prize game the effect of the power function, r, is to increase the players' effort for a given prize structure. Increasing the prize fund, V, has the effect of proportionately increasing effort. That is in line with the finding in much of the rent-seeking literature that the players will exert more effort as the aggregate rewards increase.

The Department for Transport has, by allowing authorities to choose which targets to stretch and by taking account of local circumstances in making their assessment of what is stretching, attempted to create a level playing field where authorities are rewarded for spending resources 'effectively'. It is by no means certain that a level playing field has been created however. Blavatsky (2004) examines the impacts of asymmetric ability between players for a 3 player situation and conclude that this creates a case for a wider distribution of the prize mass. This is intuitively sensible as weaker players who are unlikely to win the 1st prize still have an incentive to compete (invest positive effort). This creates a countervailing force to the competitive force created by the single prize. The ideal prize structure therefore has to balance the competitive force that fewer prizes offer with the desire to motivate a broader range of authorities by offering more prizes. The LTP2 game is characterized by a high potential number of prizes which would be sub-optimal if all players have an equal chance of submitting the best plan but preferable if there is substantial asymmetry of players abilities. The prize gradient is, if exercised in full (±25%), quite steep.

The game theoretic model therefore allows us to establish some testable hypotheses:

H1: The effort exerted by authorities when faced with performance incentives related to targets will be greater than that without performance incentives;

H2: Where the authorities are broadly homogeneous in their ability to compete for the prizes on offer, fewer prizes will encourage greater effort than many prizes.

To test whether these hypotheses might stand up in the complex arena of transport strategy development and delivery a laboratory experiment was designed and implemented as set out in Section 4. The laboratory experiment also provides an opportunity to understand further interesting behavioural dynamics between competing players and over the course of the game. The experiment therefore also allows an assessment to be made of whether the simplifications required for the game theoretic model undermine its explanatory capabilities.

4. The Local Transport Plan game in the laboratory

To bridge the gap between the simplified approach required in investigating the underlying behaviour predicted by game theory and the realities of decision-making in a real-world environment an experimental approach to understanding the behaviour of local transport planners under incentives has been adopted. Funke (2001) reviews the ability of a variety of experimental decision-making literature to capture the impacts of behaviour in complex environments and we discuss some of the limitations of the approach adopted below before presenting the results

4.1 Experimental Description

A simulation of the Local Transport Plan system was developed. The experiment involved groups of players using the PLUTO land-use transport software model (Bonsall, 1994). PLUTO is a tool developed at the Institute for Transport Studies for assessing transport strategies in an idealised city. It is already used as a training tool for professional transport planners. The model currently places the user in the role of the local authority with decision making powers on infrastructure investment, public transport subsidy, traffic restraint, land use policy and safety measures, the tools commonly available to local authorities. Policy decisions are taken annually and budgets are also automatically updated annually. The city is a stand-alone city with no transport interaction with adjacent cities. The software was modified to allow the research team to play the role of the Department for Transport (and allocate funding annually) whilst players simultaneously play the role of the local authorities.

The experiment placed eight groups of five players in the local transport plan game environment. Within each group the five players were to compete against each other – each developing and submitting a plan with seven targets. Rewards and penalties were applied to the capital funding settlement according to, the experimental condition they faced and their ranking in the competition between players. The city model which each player has to use is identical. The eight groups were exposed to each of three different experimental scenarios:

- E_0 – no incentives ($10m per annum irrespective of performance)
- E_1 – one prize ($20m for best performer $7.5m for others)
- E_2 – multiple prizes ($12.5m, $11.25m, $10m, $8.75m, $7.5m)

The hypotheses set out at the end of Section 3 are therefore tested as follows:

- H1: True if effort in E_1 and $E_2 > E_0$
- H2: Effort in $E_1 > E_2$ if players are homogenous

One of the main limitations of the experimental approach is of replicating the motivations of the real game players (ecological validity). The use of LTP practitioners would clearly be preferable to reduce the impact of mismatched motivations of players. It was however deemed impractical to gain participation from around one half of all authorities for a staff member to attend a four day experiment. We therefore selected 40 students that came almost entirely from the faculties of engineering, mathematics, sciences, computing and environment but who had not had previous experience of transport planning. A series of tests of the degree to which their self reported motivations differ from the practitioners was therefore an important part of the study design. A further challenge is the establishment of the relative capabilities of the players (homogeneity). The degree to which the selection of students from the listed disciplines delivered homogeneity is discussed in the results. Other limitations in the representation of reality include:

- The simulated city does not have any neighbouring authorities which might in reality act as a brake on certain policies (particularly restraint) due to spatial competition (Marsden and May, 2006).
- There is greater flexibility to move money around between different local authority functions in reality than in the simulation. A poorly performing authority could, for example, invest in new personnel or in infrastructure to overcome a poor league table ranking. This is not possible within the game which, we anticipate, might lead to stronger spirals of success and failure than would exist in reality.
- Game playing such as falsification of results and the strategic selection of easy to achieve local indicators cannot occur.

Funke's work (2001) stresses the importance of a good degree of training to avoid the learning process unduly influencing the experiment. A one-day training session was therefore provided as part of the experiment and the order with which groups took the experiment was randomised to remove any learning effect impacts from the comparison. For each experiment participants visited the computing laboratory for a maximum of three hours to set their targets (of which seven were required) for the experimental condition they were presented with. The funding rules were signaled clearly to the players. A questionnaire on the effort exerted setting targets was completed using the NASA RTLX scales. They returned on a subsequent day to receive their initial settlement (based on target ambition) and then to play the five year strategy through the software simultaneously with the competing players. One hour was allowed for each simulated year of play which proved generous. A spreadsheet model was written to calculate the performance of players against their stated targets each year for each of the seven targets and this performance was summed to give the ranking. The ranking of players was not announced in condition E_0 until after year 5, was only indicated through the settlement to the best player in E_1 and was presented to all players in E_2 as this is implicit in the five settlements. Participants were asked to complete a questionnaire on effort exerted during the task similar to the target setting exit questionnaire. They were also asked to complete a more detailed questionnaire on the whole experiment after the last session.

Participants were paid a fee based on attendance, not performance. That is to say that there was no financial incentive to the players to compete. This is a contentious decision in the literature although meta-analyses suggest that "the presence or absence of financial incentives was not a crucial factor in discouraging or encouraging expected utility violations" (Wickham, 2007, p6; Camerer, 1995); Beattie & Loomes, 1997). In this experiment, the decision was motivated by the lack of direct financial reward to players in the real game (i.e. rewarding the experimental players

would not provide a good map of the real game). If anything, the decision is conservative in the degree to which it motivates players of the game to participate.

Assessment of the impacts of the incentives provided a challenge. The theoretical framework suggests that effort should be captured. The Department for Transport appears to focus on achievement of outcomes as its key measure. Funke (2001) suggests that both results and process oriented measures should be taken. Table 3 shows the measures adopted.

Table 3: Effort Measures

Process oriented effort measure		**Outcome oriented effort measure**	
Measure	Description	Measure	Description
Level of stretch set	The ambition of the targets set	Indicator levels vs. stretch set	The extent to which targets are met
Planning time invested	The amount of time invested in deciding what to set	Indicator levels vs. no incentive scenario	The change over a system with targets but no incentive
Number of scenarios run	The number of times participants tried different strategies each year		
Self reported input effort	The reported effort on the task		

Ultimately, the focus of the experiments is to see if there is a positive impact from a well-designed incentive regime on transport outcomes. The result oriented measures therefore appear of greatest use. However, the theory also suggests that to achieve the outcomes more input effort will be invested.

4.2 Experimental Results

This section begins by discussing how the participants engaged with the experiment as this is essential to interpreting the outcome results. We begin by assessing the external validity of the game through responses to the end of experiment questionnaire shown in Table 4.

Table 4: End of experiment questionnaire results

	1 = Disagree to 5 = Agree	
Statement	**Mean**	**σ**
Setting ambitious targets led to greater rewards	**3.7**	**1.1**
Setting ambitious targets was more risky	**3.6**	**1.0**
I understood how the assessments were made each year	**4.0**	**1.1**
Performance rewards helped me to achieve more	**3.5**	**0.9**
Performance penalties stopped me achieving my goals	**2.8**	**1.1**
I was penalised if I did not achieve the targets I set	**2.5**	**1.0**
I tried hard harder at the start of each five years than the end	**2.5**	**1.3**
I changed my strategy a lot for each experiment	**2.8**	**1.2**
The strategies I developed would be acceptable to the public	**2.8**	**1.2**
I competed against the other players	**3.6**	**1.2**
I found the PLUTO model difficult to use	**1.6**	**0.7**
The players in my group were of an equal ability	**3.0**	**1.1**

The results suggest that the experiment reproduced the key features of the LTP system and that participants understood the relationship between stretching targets, risk and performance rewards. The findings are less conclusive on performance penalties where it appears that the penalties did not have as large an impact on participants' ability to achieve their plans. This may be, as in the real game, due to the ability to offset losses through other measures such as slightly increased parking fees. Encouragingly participants, on average, reported competing with each other, understood the software tool and did not appear to just put in effort at the start of the game. Participants were also quizzed about the ease of meeting the satisfactory and stretching thresholds given for each of the seven targets and there was evidence that participants understood which were easiest and could focus their efforts on those (as the interviews and questionnaire suggested

happens in the real game). The motivations of the participants in respect of the key features of the game therefore map reasonably well to practitioners suggesting that the ecological validity is not substantially compromised and that the results will therefore offer some useful insights.

Despite the provision of a training day, there was very strong evidence of learning effects across the three experimental rounds which underlined the importance of randomized play. Figure 3 shows the results of planning time invested in the first, second and third rounds of play. Whilst the time spent setting targets for all 40 players followed the order $E_1 > E_2 > E_0$ the differences were not significant statistically and likely to be somewhat influenced by learning.

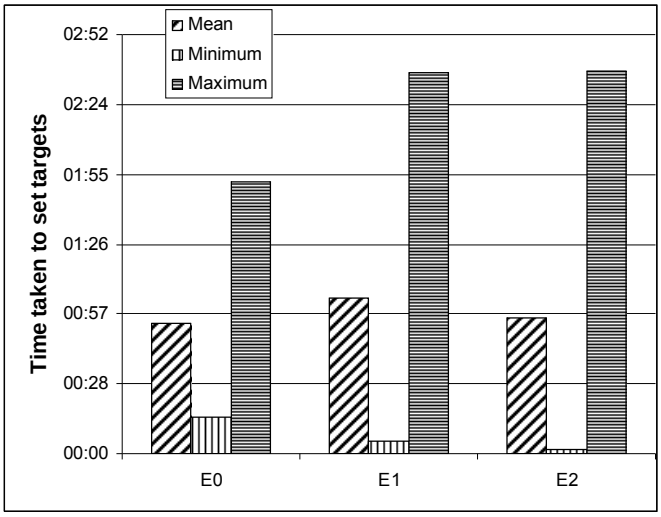

Figure 3: Time spent setting targets for the three experiments

Despite the selection of students in an attempt to attain homogenous players it appears that capabilities differed between players within groups and also between different groups. Figure 4 shows the target setting and performance over five years for two of the groups playing the same experimental condition.

It can be seen that there is substantial variation between the target ambition set and achieved by some players in group Y and little in group X. This is found throughout the data set so there is heterogeneity to varying degrees in all groups. When asked in the exit questionnaire however, respondents (on average) did not appear to perceive a difference in abilities of players in their groups.

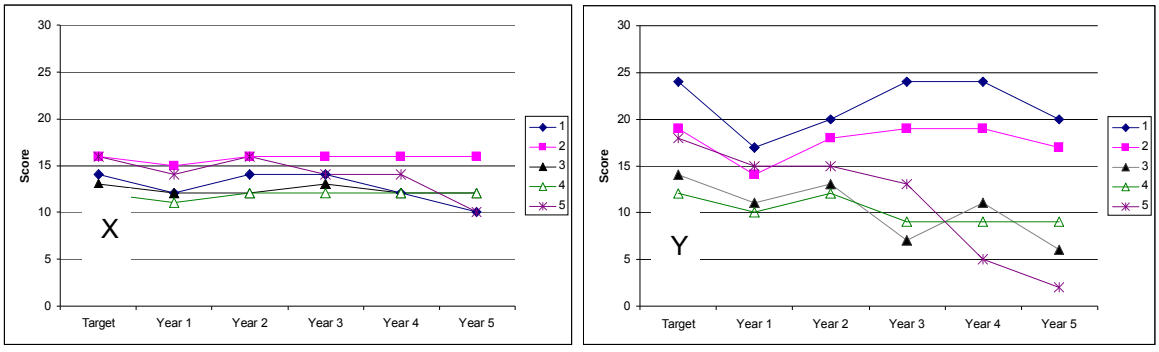

Figure 4: Target setting and performance scores over the 5 years for the five individuals in groups X and Y for experimental condition E2

Table 5 shows the effort committed by players in setting targets and playing the games across the three scenarios. There is no significant difference between the self-reported effort invested in target setting across the three experimental conditions but self reported effort in playing the game in both E_1 and E_2 are statistically significantly greater than for E_0. The t – test results found that for $E_1 > E_0$, $t = -2.411$, $p = 0.021$ and $E_2 > E_0$, $T = -2.208$, $P = 0.034$. Page's L test was run on the eight groups (each group was treated as single data point) and found that there was a significant level of agreement between the predicted ranking (E1>E2>E0) and the experimental ranking ($L = 104$ $\alpha = 0.05$, one tailed test) for self reported effort in playing the three games.

Table 5: Average Self Reported Effort in Setting Targets and Playing the Game[2] (the maximum effort score = 100)

	E0 No prizes	E1 1 prize	E2 Multiple prizes
Target Score	37.52	39.33	37.48
Experiment Score	34.00	40.12	39.34

This next section reviews the transport outcomes of the different incentive regimes. Table 6 shows the summary results for target ambition (set at the start of the experiment), target attainment (were the targets met in year five) and absolute outcomes (how much was achieved ignoring whether or not the target was met).

Table 6: Targets and Outcome Results

	Number of Indicators E1 > E2 > E0	Number of Indicators (E1 + E2) > E0
Target Ambition	3/7	5/7
Target Attainment	1/7	3/7
Actual Attainment	3/7	6/7

The first row of Table 6 suggests that for 3 of the 7 indicators the level of target ambition followed the expected assumptions for homogenous players that $E_1 > E_2 > E_0$ and that this increases to five out of seven indicators if the conditions are relaxed to E_1 and E_2 both being greater than E_0 but in any order. This suggests that the targets set are more ambitious with some form of incentives in place than without. On target attainment it appears that setting more ambitious targets does not mean that targets will be achieved with two of the five indicators for which more ambitious targets were set in both E_1 and E_2 not being met (5/7 reducing to 3/7). Politically this may be important if ambitious targets cannot be delivered. It is however necessary to look at the absolute level of achievement to see if the system has delivered its goals of incentivising better performance. In this regard, 3 of the seven indicators were higher strictly in the order $E_1 > E_2 > E_0$ and that this increases to six of the seven indicators if the conditions are relaxed to E_1 and E_2 both being greater than E_0 but in any order. This implies that although the targets might not all be met, the absolute outcomes are better when there is some form of incentive than when there is not. The Page L-test was conducted to determine whether the differences are statistically significant. Again because the actions of players within groups are not independent (within E_1 and E_2) each group was treated as a data point. It has not been possible to establish statistical significant within such a small sample. Despite this limitation we can tentatively conclude that the experiment supports H_1 and that the presence of incentives leads to the creation of greater levels of progress towards the identified indicators than when targets are set but no incentives are provided. We are unable to conclude on H_2 (extent to which one prize or multiple prizes are preferable) due to the heterogeneity of the players. It seems unlikely that the conditions of homogeneity necessary for a one prize game could exist in reality despite the careful system design.

5. Conclusions

This research set out to determine the likely impacts of a performance reward funding regime for local transport planning which is linked to achievement against targets. It is without doubt a complex topic and difficult to unpick. None of the methods adopted alone would have provided a picture robust enough for us to draw conclusions from. Taken together however, there are some key themes and conclusions.

First, the theory and experiment taken together suggest that linking performance rewards to target setting will lead to competition between authorities and, designed right, this will lead to greater levels of achievement against the key metrics in the system than if no performance rewards are available.

[2] Self reported effort is calculated using the NASA RTLX scales

Secondly, in designing a performance reward regime the prize fund, number of prizes and gradient of value of prizes are all important design factors. It is likely, in any real local transport situation, that players will have differing abilities as well as differing starting points. The case for multiple prizes increases as the difference between players increases. In a game of 82 players it is also politically more attractive to spread the prize fund round to avoid difficult knife-edge decisions. There are however also lessons to be drawn on running competitions for clearly defined funding pots (e.g. to tackle urban congestion). Here, only those self-selecting authorities that believe they have a good chance of winning will compete. Smaller numbers of prizes are likely to be more appropriate in such situations.

Thirdly, the system has a substantial management cost. In the UK, significant effort has been made by the managing organization (Department for Transport) to try to establish a level playing field between authorities. However, this has only been possible (and even then not uniformly accepted) for a narrow sub-set of indicators. Whilst the assessments of central and local authorities appear more consistent than was anticipated there is still an apparent knowledge asymmetry between central and local government on likely performance for some indicators. Our work suggests that splitting the contest into several smaller pools of more evenly-matched authorities would have only a small negative impact on the incentive to effort, so could offer net gain from central government's point of view.

Fourthly, it has been suggested that performance rewards are inappropriate for essential public services as penalizing poor performing authorities will lead to unacceptable spirals of decline. Although limited by the experimental environment we did not see substantial evidence of spirals of decline, perhaps because (as in the real game), the baseline funding of 75% is still sufficient to avoid 'spirals'. In reality, Atkins (2004) also found that poor performance acts as a trigger within an authority to put further management action into supporting transport.

Finally, the new approach to target setting appears to have had the desired effect of focusing local authority transport planning departments on the quality of their planning. Whilst this is welcome and acknowledged there is still a huge tension between their support for the five year planning process and the feeling that too much is required of local authorities by central government.

This study has provided some empirical evidence on the potential impacts of targets and performance rewards on transport outcomes and on the behaviour of participants under any such regime. Targets and performance rewards are becoming more widespread globally throughout public services. Whether they will prove to be a good thing for transport remains to be seen. Whilst the system appears to offer benefits of enhanced performance on defined indicators this may lead to imbalances in priorities between those things for which performance is rewarded and those for which it is not. In this respect, as with the use of multiple prizes, we find the UK system robustly designed with a mixture of national and local indicators. Interestingly, at the time of writing, the whole basis of the UK system is being reviewed so we may not get to see the practical validation of these findings in future years.

Acknowledgements

The research was funded by EPSRC under grant number EP/C002350/1 'Optimal Incentive Structures for Integrated Transport Strategies'. The authors are grateful to the Department for Transport and the authorities that participated in the research. Responsibility for the findings nevertheless rests with the authors.

References

Atkins (2004) Working with weaker local authorities. Final summary report on findings, Report to Department for Transport, London

Atkins (2007) Long term process and impact evaluation of the Local Transport Plan policy. Final Report to Department for Transport, London, www.dft.gov.uk/

Beattie, J and G. Loomes (1997) "The impact of incentives upon risky choice experiments", *Journal of Risk and Uncertainty*, 14:155-168.

Blavatskyy, P. (2004) *Why the Olympics have three prizes and not just one*. Working Paper No. 200. Institute for Empirical Research in Economics, University of Zurich.

Bonsall, P. (1994) PLUTO: computer assisted training for transport planners. *Proc. Conference on computer aided learning in engineering*, University of Sheffield.

Camerer, C. F. (1995) "Individual decision making", in J. H. Kagel and A. E. Roth (Eds) *Handbook of Experimental Economics*. Princeton, N. J.: Princeton University Press.

Clark, D.J. and Riis, C. (1998) Influence and the discretionary allocation of several prizes, *European Journal of Political Economy*, **14**, 605-625.

De Palma, A. and Lindsey, R. (2003) *Approaches to Institutional Issues*, Deliverable 2.2 to the Transport Institutions in the Policy Process project, EU 5th Framework Project Contract No: GMA2/2001/52043 S07.15419 TIPP, DG TREN.

Department for Transport (2004) Full Guidance on local transport plans: second edition www.dft.gov.uk/

Department for Transport (2005) Financial Planning guidelines for local transport plans. Consultation paper, July 2005 www.dft.gov.uk/

Department for Transport (2006a) Local transport plans settlement letters. www.dft.gov.uk

Department for Transport (2006b) Technical Note about the Assessment of Final Local transport plans. www.dft.gov.uk

FHWA (2004) Transportation Performance Measures in Australia, Canada, Japan, and New Zealand, International Scanning Study Team Report, US Federal Highway Administration, http://international.fhwa.dot.gov

Funke, J. (2001) 'Dynamic systems as tools for analyzing human judgement', *Thinking and Reasoning*, **7** (1), pp69-90

Hodgson, L., Farrell, C.M. and Connolly, M. (2007) Improving UK Public Services: A Review of the Evidence, *Public Administration*, **85**(2), 355-382

Hood, C. (2006) Gaming in Targetworld: The Targets Approach to Managing British Public Services, *Public Administration Review*, **66**(4), 515-521

Kelly, C., Nellthorp, J. and Marsden, G. (2006) *How do local authorities set their LTP targets?* Proc. Transport Practitioners Annual Meeting, Manchester, 25-26 July

Kelly, C., Marsden, G. and Nellthorp, J. and (2008) An assessment of target setting in local transport plans in England. *Submitted to Local Government Studies* (September 2007)

Kreps, D. (1990), Game Theory and Economic Modeling

Marsden, G. and Bonsall, P. (2006) Understanding the role of performance targets in transport policy, *Transport Policy* 13 (3), 191-203

Marsden, G. and May, A.D. (2006) Do institutional arrangements make a difference to transport policy and implementation? Lessons for Great Britain. *Environment and Planning C: Government and Policy*, **24**(5), 771-790

Moynihan, D.P. (2006) Managing for results in state government: Evaluating a decade of reform. *Public Administration Review* **66**, 77-89

Nellthorp, J. (2007)

Smith, P. (1995). On the Unintended consequences of publishing data in the public sector, *International Journal of Public Administration* **18**(2/3): 277-310.

Whickham, P. (2007) *Financial Incentives in Economic Experimentation*, ITS Research Discussion note, Institute for Transport Studies, Leeds, February 2007.