

Summary of PPP Performance Regime

The PPP performance regime taken as a whole is staggering in its complexity. And deploying it—placing the very future of the Underground and indeed London’s lifeblood and economy at its mercy—is a risky endeavour and a dangerous game.

Consider the historic models of how other aging metro systems have been turned around. Putting it simply, problem areas were identified by management, various capital works and maintenance solutions were identified and prioritised by management, and were then implemented, often by private contractors. The contractors would be paid as they delivered desired products, and would be responsible if the products didn’t work as promised (warranty and reliability requirements). Given stable funding, the record is clear. These programmes work.

BACKGROUND

Under PPP, the Infracos will be paid by LUL for maintaining the trains, stations, track and signals under contract for 30 years. LUL will continue to staff the stations and operate the trains.

However, the PPP contract does not provide for payment to the Infracos based on concrete deliverables such as a new fleet of trains or carrying out specific and routine track maintenance.

Instead, LUL is forced to pay the Infracos on the basis of how the Underground is performing, and whether the Infracos are responsible for the improvement or deterioration of that performance. How the Infracos attempt to deliver the performance specified in the contracts is almost entirely up to the Infraco.

Because there is no agreed checklist of work that the Infracos have to deliver, LUL has had to develop a ‘performance regime’ to judge how well or badly the Infracos are doing their job. This regime attempts to measure thousands of minute aspects of the Underground’s operation that affect journey times and ambience and other elements that comprise the “Underground Experience”. Once the data is collected and agreement is reached on who is responsible for what “incidents” arising on the Underground, the relevant data is plugged into numerous complex formulae to determine whether the Infracos’ performance meets a set of benchmarks laid out in the contract documentation. This regime will determine the payments (both penalties and bonuses) that LUL makes to the Infracos.

Thus, in order to quantify the Infracos’ performance, LUL has had to create and incorporate into the contracts:

- A very complex mathematical model of the entire Underground system;
- and a system for attributing responsibility for “incidents” on the system.

PERFORMANCE REGIME

The complexity of creating a mathematical model of the entire Underground system and all its parts has led to the creation of an equally complex and unwieldy system of gauging the performance of the Underground based on that model.

“The idea of incentivising Underground improvements almost purely through operation of a mathematical model that captures thousands and thousands of minute elements of Underground performance is fascinating to think about – in the abstract. However, it is reckless to experiment with implementing such an untested idea – and to bind yourself to the experiment for thirty years.”

PERFORMANCE REGIME – SIZE

The explanation of the performance regime, and how it is to be implemented and administered is not kept in one volume, and is to be found variously in:

- the 200+ page Performance Measurement Code;
- the Service Output Schedules;
- Line Specific Performance Measurements Schedules;
- Nominally Accumulated Customer Hours (NACH) Tables; [which TfL has not seen]
- Capability Model; [which TfL has not seen]
- Performance Payment Mechanism Schedules.

Determining key elements of the performance regime also requires extensive cross-referencing to other lengthy and equally complex documents.

PERFORMANCE REGIME - COMPLEXITY

The mathematical model LUL intends to use is far from simple, and “ ‘Very complex’ is an understatement”.

For example, the formula to determine the ‘Ambience’ or the state of the Underground environment is:

$$A = \frac{\sum_{js} QASAS_{js} \times W_{js}^S + \sum_{jL} QATAS_{jL} \times W_{jL}^T + \sum_{kL} QATAS_{kL} \times W_{kL}^T}{\sum_{js} W_{js}^S + \sum_{jL} W_{jL}^T + \sum_{kL} W_{kL}^T}$$

Similarly, the equation to determine the state of a train, or the ‘Train Ambience’ is no less complicated:

$$TMS = \sum_j \left(\frac{QATAS_j \times W_j^T}{\sum_j W_j^T + \sum_k W_k^T} \right) + \sum_k \left(\frac{QATAS_k \times W_k^T}{\sum_j W_j^T + \sum_k W_k^T} \right)$$

There appear to be a large number of these formulae, with a vast array of inputs needed to perform the calculations ranging from litter to lighting to quality of staff facilities.

On top of this “there will be thousands and thousands of incidents to evaluate every month, just to execute one facet of the multi-faceted performance regime.” These range from an out of service escalator to whether the Infraco is responsible if a staff toilet becomes blocked due to inappropriate materials being flushed!

PPP has even attempted to quantify and qualify the state of litter and cleanliness on the Underground in the minutest detail. There are even instances where litter is not litter. If it’s less than the width of a Tube ticket it does not count. The amount of chewing gum on station walls must be measured. Then the relative importance of the dirt given its location and its degree is measured by weighting formulae.

“Is there a document that explains in plain English the relative weightings of ambience factors? For example, how will ‘very clean train seats with only minor areas of dirt or ‘dull’ appearance’ affect payments to the Infracos relative to ‘some areas of dust, dirt or staining, e.g., dirt build up in corners and crevices’ in the station ticket hall? What is the relative weight of graffiti in the ticket hall to non-scratched graffiti in the trains?”

PERFORMANCE REGIME - LOOPHOLES

The PPP contracts, it is claimed, are designed to incentivise the Infraco to be rewarded for good performance and be penalised for poor performance. However, “the performance regime is also rife with opportunities for the Infracos to play the system.”

The system of attribution “provides a disagreeable Infraco ample opportunity to challenge and/or delay being held accountable for any incident” – a potentially lengthy ten step process, eventually ending up in a Court of Law, is laid out in the contracts – “and, as we read in the contracts, the incident doesn’t even go into the hopper for payment purposes until any dispute over responsibility for the incident is settled. Very realistically, this could be years.”

LUL’s own presentation documents on the performance code admit that there “will be [a] degree of ‘horse trading’ during attribution” – meaning it is acknowledged that an Infraco can try to mitigate its financial responsibility for individual incidents on the Underground – therefore breaking the very link that meant an Infraco would be penalised financially for poor performance.

-- Contrary to statements from DTLR this October in their document ‘Your Tube: Publicly Run, Privately Built’, the Infracos will not be charged for each train delay or escalator failure – instead the specific performance of an Infraco is averaged out over a period of 3 (or for speed restrictions 6) months. Even if there has been a significant failure attributable to the Infracos, they may escape penalty if their average performance over the time meets the specified benchmark for overall performance.

“Only if the three month average exceeds the benchmark level of permitted customer delays will the Infraco begin to pay anything for this type of problem in performance. Indeed, if they are above ‘benchmark’ (which we understand will be 5% below current performance), they will get a bonus. Imagine that. Service reliability deteriorates below what it is today – and a bonus is paid?”

-- “The Infracos will control much of the information that is necessary to evaluate the cause of incidents that are to be input in to the various formulae.”

For example “the most recent draft of the documents provided that the Infraco identifies a pool of trains from which test trains may be selected. Fully 40% of the trains in a fleet can be kept out of the pool – allowing the Infraco to have only its better trains count for this critical measure driving their own compensation.”

-- “Infracos have far too much control over the test process. They design the test programme, and they can retreat to the lengthy dispute resolution procedure if they are unhappy with any LUL objections to that test programme.”

PERFORMANCE REGIME: WEAKENING OF THE CONTRACT

The performance regime is replete with contract terms that are generous to the Infracos, some of which appeared in the documents only recently, after the Best and Final Offers and selection of preferred bidders. In earlier correspondence Mr. Kiley warned that the loss of leverage by naming bidders prematurely could lead to precisely these types of outcomes.

For example, originally LUL provided that it would have the right to choose the trains that would undergo testing as part of journey time capability measurement. The selection of trains can have an enormous influence on test results. Now the contracts allow the Infracos to exclude up to 40% of the trains from the tests—providing an opportunity to avoid testing the weakest performing lot.

In another section of the contract, we found a change to the formula for adding up service points for fault rectification delays that appears to give each Infraco a sizeable allowance of Service Points for

such delays (which the Infracos accrue for poor performance) – over 40,000 per four week period in the case of one Infraco and approximately 30,000 per four week period for each of the other Infracos. This appears to add up to a waiver of well over £2 million per month for one Infraco alone, and over £1.5 million for each of the other Infracos. It also appears that instead of only docking Infracos for poor service, the Infracos are now allowed to build up credits against service failures if they manage to rectify other faults ahead of contractually specified deadlines.

In yet another section, the Infracos appear to have won relief on many targets for overall condition of assets.

At the time of BAFO 50% of open section rail on one line was permitted to be in “D or worse condition” at the 7-½ year mark. Only E is worse. Now, fully 100% of open section rail on that same line is permitted to be in “D or worse condition” at the 7-½ year mark.

BAFO requirement: 46% of “sleepers and base concrete” in tube sections on one line could be in D or worse condition at 7 ½ years. Preferred Bidder revised requirement: 92.5%.

BAFO requirement: 0% of viaducts on one line could be in E or worse condition at 7 ½ years. Preferred Bidder revised requirement: 7%.

What persuaded LUL to adopt these and other reductions in Infraco obligations so late in the game? Will the ongoing negotiations result in additional degradation?

CONCLUSION

- Is this any way to run a railway? How effective can a system like the performance measurement regime be, when the process for imposing penalties for shoddy Infraco performance is so convoluted and so attenuated from day to day work?
- “If PPP is concluded and TfL inherits the Tube, I will be explaining to delayed Tube passengers that their experience was or was not attributable to the recently privatised Infracos based on the application of these formulae!”
- “Words simply cannot begin to accurately convey the complexity of this mathematical modeling of the Underground, and the inherent risks it represents. I am deeply disturbed by the prospect that the riding public in London, indeed the very economy of London, will be held hostage to this model and the inevitably continuous battles that must be fought as the private sector attempts to eke advantage from it.”