AN EXPERIMENTAL STUDY OF PIGOU-KNIGHT-DOWNS PARADOX FOR TRAFFIC CONGESTION UNDER A MODIFIED ROAD NETWORK

INTRODUCTION

This study aims to experimentally investigate the Pigou-Knight-Downs paradox in traffic congestion under a modified road system compared to the original version of the paradox.

Traffic congestion is a universal problem in urban areas, causing a waste of time and resources every year. Technological advancements allow easier road capacity increases, resulting in policies that try to remedy congestion by building more or widening existing roads. Yet, the effectiveness of these policies is questionable given the induced demand effect they have on public route choice behaviours (Dechenaux, Mago & Razzolini 2013).

Arnott and Small (1994) demonstrates three transportation paradoxes based on road users’ failure to internalise the costs of their travel imposed on others in congested conditions when making route-choice decisions and the effect of induced demand. As roads are expanded, commuters are induced to demand more by driving during hours or switching to routes they normally would not use. One of the paradoxes mentioned in Arnott and Small (1994) is the Pigou-Knight-Downs (PKD) paradox. Downs (1962) argues when commuters make their marginal route choices, travel time on each route will be eventually equalised at a stable equilibrium. Temporarily, if commuters find a faster route, the ‘explorers’ will switch to that route, leading to a disequilibrium. Over time, as more commuters move to the faster route, the travel time on that route will increase, while the travel time on the old route decreases. The switching continues until the equalisation of travel time on each route is reached. This gives rise
to the establishment of the PKD paradox: increasing road capacity does not reduce travel time, i.e. the benefit from increasing road capacity dissipates completely.

The PKD paradox is illustrated in a traffic network in which commuters, who want to travel to a destination, have to choose between a direct, congestible bridge and a circuitous, non-congestible road (Arnott & Small 1994). Travel time on the congestible bridge rises as more people choose it, while travel time on the other road is constant irrespective of the number of commuters on it. Suppose the bridge is widened, due to induced demand effect and commuters’ failure to consider the external cost they impose on others by switching to the bridge and slowing down others, the PKD paradox suggests the bridge expansion may result in no improvement to traffic congestion. To solve congestion, an emphasis is placed on the understanding of behavioural interactions of stakeholders in the transport problems.

Previous researches have derived choices in transportation networks from individual preferences and equilibrium behaviour under strategic interaction (Dixit et al 2015). This equilibrium behaviour has been mainly studied in a laboratory environment, since laboratory experiments allow more control than the field does. Anderson, Holt and Reiley (2008) and Hartman (2009) investigate subjects’ choice between a direct, but increasingly congested route and a non-congestible, indirect route, assuming time value is homogeneous. While Anderson, Holt and Reiley (2008) treats it as a market entry game, wherein subjects can choose to enter the risky route or stay out by adhering to the safe route, Hartman (2009) describes the congested situation as coordination problems that commuters fail to coordinate when making route-choice decisions simultaneously. Both papers reach similar conclusions. Although the average rate of entering the congestible route converges to the equilibrium prediction – where payoffs from choosing the two roads are equalised, variability still persists between rounds. The equilibrium
level is, however, not socially efficient. The socially optimal entry rate – where the total payoffs to all subjects are maximised – is found to be half of the equilibrium rate. When an entry fee or tolls are imposed on the congested route, the number of people on the route is reduced and approaches the efficient level, but this is not robust as fluctuations still exist.

Some other experimenters modified the congestible conditions to study equilibrium and efficiency. Selten et al (2007) design a route-choice experiment with two congestible roads for subjects to choose: a main road and a side road. They also report persistent oscillation although the mean numbers of subjects on both roads are close to the pure equilibrium. Using a partner-matching protocol, Gabuthy, Neveu and Denant-Boemont (2006) let subjects choose their departure times and a route to arrive at destination, either a city road or a tolled freeway. With low and high tolls as treatments, they find a low toll generates higher efficiency. Ziegelmeyer et al (2008) also allow subjects to choose departure times and show providing public retrospective information does not have a significant effect on mitigating congestion. Moreover, they find the commuter population size is independent of subject's coordination problems. This result contrasts with the conclusion made by Huyck, Battalio and Rankin (2007) and Dechenaux, Mago and Razzolini (2013) that subjects' capacity to coordinate falls with a larger population.

In addition, there has been experimental evidence of other transportation paradoxes. The Braess paradox, stating adding an extra link to a congested traffic network can raise travel time for commuters, is experimentally corroborated by Rapoport et al (2006; 2009); Morgan, Orzen and Sefton (2009). Laboratory research (Denant-Boemont & Hammiche 2009; Dechenaux, Mago & Razzolini 2013) has also confirmed the Downs-Thomson paradox, which states increasing the road capacity causes total travel time to rise for commuters who have to choose between public transport and private vehicle. In contrast, the PKD paradox, despite being widely accepted in
theory, has been little empirically supported. In an experiment, by enhancing the congestible route in a network configuration which also has a non-congestible route, Morgan, Orzen and Sefton (2009) examines the PKD paradox and ‘the least congestible route principle’ that an expansion should be made on the least congestible route. They find the PKD paradox holds but not strongly as the dissipation in benefit from having the congestible route improved is not complete. Travel times in the enhanced network are 4% lower.

The want of experimental investigation into the PKD paradox has motivated my paper. The paper’s objective is to validate the PKD in a modified road network, in which two roads available for commuters are both congestible: a circuitous, large road and a direct, small road. Employing the PKD paradox and ‘the least congestible route principle’, I aim to test the question whether an improvement made to the circuitous, large road will paradoxically lower travel cost more than to the direct, small road.

THEORETICAL BACKGROUND

Consider a basic traffic system based on Arnott and Small (1994) with n commuters who have to travel to a destination by choosing between a direct, congestible road and an indirect, non-congestible road. Under the assumption of homogenous time value across commuters, travel cost on the non-congestible road is constant and independent of the number of people on it, while cost on the congestible road increases on the number of people on it. A simple linear cost function is to be used for the congestible road as literature has shown equilibrium and convergence behaviours do not significantly differ between linear and non-linear cost functions (Denant-Boemont & Fortat 2013).

Travel time cost incurred by each commuter on the congestible road is expressed as follows:
\[ T_C = a + bx \]

where \( x \) is the number of people choosing the congestible road, \( a > 0 \) is the minimum time to arrive at the destination without congestion, \( b > 0 \) is the marginal cost for an additional commuter on the road.

Travel time cost on the non-congestible road is fixed at \( T_N \).

There is a user equilibrium, where travel time costs on both roads are equalised. This is also a Nash equilibrium:

\[ T_C = T_N \Rightarrow a + bx = T_N \Rightarrow x_e = \frac{T_N - a}{b} \]

Since there are \( n \) commuters in the system, \( \binom{n}{x_e} \) combinations of \((x_e, n - x_e)\) constitute a Nash equilibrium.

Nevertheless, \((x_e, n - x_e)\) is not socially efficient, for each commuter fails to internalise the external cost they impose on the others by switching to the congestible road. The efficient number of people on the congestible road minimises the aggregate travel cost of the system, which is \( x(a + bx) + (n - x)T_N \). By applying the first-order conditions, the efficient \( x \) happens to be \( \frac{x_e}{2} \) in this system.

Furthermore, there is uncertainty in commuters’ coordination, which leads road users to use mixed strategies. Therefore, mixed-strategy Nash equilibria exist, but they are not to be the focus of this paper since according to Arnott, De Palma and Lindsey (1993), mixed strategies might not be as realistic as pure strategies because commuters prefer a routine travel pattern.
The Pigou-Knight-Downs paradox postulates an increase in the congestible road capacity does not decrease total travel cost for all commuters. As the travel cost on the non-congestible road remains unchanged, at the user equilibrium, the equalisation of travel costs on both roads will result in no decrease in total travel cost for all commuters in the system. This paradox motivates the first hypothesis of this paper.

**Hypothesis 1:** Increasing the congestible road capacity does not decrease total travel cost, while increasing the non-congestible road capacity does.

Similarly, in a modified network with two congestible road, the travel cost incurred by each individual on the small, direct road is: $T_{SD} = a_{SD} + b_{SD}x$, while the travel cost on the large, indirect road is: $T_{LI} = a_{LI} + b_{LI}(n - x)$. In the above cost functions, $x$ is the number of people choosing the small, direct road, $a_{LI} > a_{SD} > 0$ is the minimum time to travel to the destination without congestion, $b_{SD} > b_{LI} > 0$ is the marginal time cost for an addition commuter on the road.

By equating the travel cost on both roads, the pure-strategy Nash equilibrium $x$ can be found:

$$x_e = \frac{a_{LI} - a_{SD} + b_{LI}n}{b_{SD} + b_{LI}}$$

Morgan, Orzen and Sefton (2009) posits ‘the least congestible road principle’ such that an expansion is more effective when made on the route that is least sensitive to congestion. This gives rise to the following hypothesis for the modified network, which aims to test the robustness of the Pigou-Knight-Downs paradox in a less extreme congestible conditions.

**Hypothesis 2:** An improvement to the large, indirect road lowers total travel cost more than that to the small, direct road.
EXPERIMENT DESIGN

The experiment is designed, with a number of modifications, based on the travel cost structures in Dechenaux, Mago and Razzolini (2013), in whose experiment subjects have to choose between a Road and the Metro. In my experiment, there are three treatments: a baseline network, a network with an improvement made to the more congestible road, and a network with an improvement made to the less congestible road. These treatments are performed under two levels: an unmodified network (one direct, congestible road and one circuitous, non-congestible road – the original version of the PKD paradox) and a modified network (one direct, small road and one circuitous, large road; both congestible). The following table denotes all cases:

<table>
<thead>
<tr>
<th></th>
<th>Baseline Network</th>
<th>Expansion made to the more congestible road</th>
<th>Expansion made to the less congestible road</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Unmodified</strong></td>
<td>UnBase</td>
<td>UnExM</td>
<td>UnExL</td>
</tr>
<tr>
<td><strong>Modified</strong></td>
<td>ModBase</td>
<td>ModExM</td>
<td>ModExL</td>
</tr>
</tbody>
</table>

The experiment comprises 32 sessions, using students at the University of X as subjects. In each sessions, 16 inexperienced subjects are required and they can participate in the experiment only once. On arrival, subjects are paid a show-up fee of $10.

During the experiment, they are instructed to play a route-choice game that has 40 rounds. In every round, they must arrive at a destination D to be rewarded 275 points. In order to reach D, they have to choose between Road A and Road B. The two options are given to them with no reference to their congestibility or distances to the destination to avoid personal biases subjects hold in real life. However, the information of travel cost structures on both roads is common knowledge to all subjects, which is displayed on the screen every round. Travel cost on a
congestible road is an increasing linear function of the number of people on that road, while travel cost on a non-congestible road is a constant and independent of the number of players on it. If more players choose the congestible road, travel cost for each player on the road will rise, and vice versa. Each subject’s payoff is the difference between the reward they receive for arriving at D and the cost they incur when choosing a particular road. Accordingly, a subject’s payoff depends not only the decision they make but also others’ decisions. To facilitate subjects’ decision making and learning of others’ behaviours, at the end of each round, they will be informed of their payoff for the round and the numbers of players on each road.

For the first set of 16 sessions, each cohort of 16 subjects begins every session with 20 rounds of a baseline network, either unmodified or modified. Then the next 20 rounds, they will play an improved network, which could be an expansion to the more congestible road or to a less congestible one. For the remaining 16 sessions, the sequence will be reversed, i.e. subjects will begin the first 20 rounds with an improved network, then play the next 20 rounds with a baseline one. The purpose of reversing sequences is to control for sequencing effects as subjects’ behaviours in the ensuing 20 rounds might be affected by the experience gained from the first 20 rounds. Moreover, letting subjects participate in two treatments at a time can reduce subject variability.

The allocation of sessions is as follows:

<table>
<thead>
<tr>
<th></th>
<th>Unmodified network</th>
<th>Modified network</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Expand the more congestible road</strong></td>
<td>Sessions 1-4: 20 rounds of UnBase, 20 rounds of UnExM (UnBase, UnExM)</td>
<td>Sessions 9-12: ModBase, ModExM</td>
</tr>
<tr>
<td></td>
<td><strong>Reversed sequences:</strong></td>
<td><strong>Reversed sequences:</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sessions 13-16: ModExM, ModBase</td>
</tr>
</tbody>
</table>
Expanding the less congestible road

Sessions 5-8: UnExM, UnBase

Sessions 17-20: UnBase, UnExL

Reversed sequences:

Sessions 21-24: UnExL, UnBase

Sessions 25-28: ModBase, ModExL

Reversed sequences:

Sessions 29-32: ModExL, ModBase

Cost structures for the networks

The following section gives a functional form of cost on each road in all treatments, whereas participants will be provided both functional and tabular forms in the experiment when making route-choice decisions.

Road A refers to the more congestible road, while Road B is the less congestible one of the two.

1. Unmodified Network (Hypothesis 1)

Baseline:

Travel cost on Road B is constant at $T_B = 175$ points for each commuter on the road, and per-commuter travel cost on Road A follows the formula: $T_A = 100 + 12.5x_A$ with $x_A$ being the number of commuters on this road. A subject’s payoff for each round is $275 - T_A$ if the subject chose Road A or $275 - T_B$ if they chose Road B.

Treatment UnExM:

The capacity of Road A is expanded, thus reducing travel cost on Road A. Its new cost function is as follows: $T_A = 100 + 6.25x_A$.

Travel cost on Road B is unchanged, $T_B = 175$ points for each commuter.
Treatment UnExL:

Holding the capacity of Road A unchanged with the cost function \( T_A = 100 + 12.5x_A \), an improvement is instead made on Road B, resulting in a reduction in travel cost to \( T_B = 150 \).

2. Modified Network (Hypothesis 2)

Baseline:

Travel cost on Road A is \( T_A = 50 + 18.75x_A \), while travel cost on Road B is \( T_B = 100 + 12.5x_B \), both depending on the number of commuters on the chosen road and subject to \( x_A + x_B = 16 \). The travel cost functions indicate that in the absence of congestion, it takes longer to reach the destination by choosing Road B, meaning Road B is circuitous, and Road A is direct. However, higher marginal cost for an additional commuter on Road A than on Road B specifies Road A is more congestible. In sum, Road A is the direct, small road, while Road B is the circuitous, large road.

Treatment ModExM:

When an improvement is made on Road A, its cost function becomes \( T_A = 50 + 12.5x_A \), which has a lower marginal cost per additional commuter. As the travel distance is unchanged, time cost in the absence of congestion is assumed to be fixed at 50 on Road A. Travel cost on Road B remains \( T_B = 100 + 12.5x_B \).

Treatment ModExL:

If Road B is expanded, its cost function is \( T_B = 100 + 6.25x_B \), while travel cost on Road A is unchanged and determined by \( T_A = 50 + 18.75x_A \).

EXPECTED RESULTS
The measures the study would use to examine subjects’ route-choice behaviours are the mean numbers of players on each road, which will be compared with the equilibrium levels, and the total travel cost for all players each round. Comparisons can be made between the two treatments, expanding the more congestible road and the less congestible one to see the difference in effectiveness of each improvement; as well as between the two levels, modified and unmodified networks, to test the robustness of the PKD paradox when loosening the non-congestible condition.

Predictions of equilibrium numbers of commuters on each road and total travel cost per round:

1. **Unmodified Network (Hypothesis 1):**

   **Baseline:**

   The pure equilibrium level is reached when \((x_A, x_B) = (6, 10)\), such that costs are equalised on both roads. Total travel cost of the system at equilibrium is 2800 points, calculated by
   \[
   TC = T_A x_A + T_B x_B.
   \]

   **Treatment UnExM:**

   When \((x_A, x_B) = (12, 4)\), the pure equilibrium is achieved. Total equilibrium travel cost does not decrease after expansion made to Road A.

   **Treatment UnExL:**

   The pure equilibrium is at \((x_A, x_B) = (4, 12)\), and total travel cost is predicted to decline to 2400 if the equilibrium is reached.

   If the results are as above, Hypothesis 1 holds true, thus supporting the PKD paradox.
2. **Modified Network (Hypothesis 2):**

**Baseline:**

By equating the two cost functions, the pure equilibrium is found at \((x_A, x_B) = (8, 8)\). Total travel cost at equilibrium is 3200.

**Treatment ModExM:**

The pure equilibrium is at \((x_A, x_B) = (10, 6)\), where total cost is 2800.

**Treatment ModExL:**

The pure equilibrium is at \((x_A, x_B) = (6, 10)\), where total cost is 2600, lower than expanding Road A.

If this occurs, it can be concluded that expanding the less congestible road is more efficient.

**CONCLUSION**

The research empirically investigates the Pigou-Knight-Downs paradox that states increasing road capacity does not reduce travel time, which questions the common remedy of expanding networks to reduce congestion. Albeit theoretically fundamental in transportation economics, the paradox has yet to be empirically supported. The research also generalises the paradox by relaxing its extreme network condition with one route being non-congestible. Due to the difficulty of studying route-choice behaviours in the field, this research is conducted in a laboratory setting and potentially provides an experimental evidence for the paradox that can inform policy-makers of the effectiveness of road expansion in mitigating congestion.
References


