Optimal Congestion Pricing with Diverging Long-Run and Short-Run Scheduling Preferences

Erik T. Verhoef
Department of Spatial Economics, VU University Amsterdam
Introduction

- Values of travel delays and schedule delays are central concepts in transport economics.

- Recent evidence suggests that travellers decompose scheduling decisions into:
  - long-run choices of routines
  - short-run choices of departure times

- This paper: implications for optimal congestion pricing:
  - Do we need separate instruments to optimize both decisions?
Deterministic starting point
Vickrey 1968; Small 1982

Average cost

c^0 = p^0

c_S(t')

c_T^0(t')

c_S(t')

Departure rate first above then below capacity

Exit time t'

t_q

Congestion pricing with LR & SR scheduling
Two dimensions of SR vs LR

1. Different measures for preferred arrival time PAT
   - Long-run LRPAT ($t^*$): preferred arrival time if there were no congestion, ever
     - Interpretation in standard bottleneck model
   - Short-run SRPAT ($t^\#$): preferred arrival time given the expected pattern of travel times
     - Choice of ‘routines’ may make SRPAT deviate from LRPAT
     - With a LRPAT at 9:00, an SRPAT at 7:00, and a scheduled meeting at 7:30, an arrival time at 8:30 would bring cost of schedule delay late, not early
       - Evident: important to address in empirical modelling
Two dimensions of SR vs LR

2. Different values of time and schedule delay, depending on ‘degree of permanence’
   • A structural one-minute travel time gain brings more benefits *per day* than an incidental minute on a random day
   • An unanticipated schedule delay brings a greater disutility than schedule delays that are anticipated when forming routines
Empirical confirmation

- Peer, Verhoef, Koster, Knockaert (2015)
- Drivers plan their routines to avoid congestion
Empirical confirmation

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Long-Run</th>
<th></th>
<th>Short-Run</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Value</td>
<td>t-Statistic</td>
<td>Value</td>
<td>t-Statistic</td>
</tr>
<tr>
<td>$\beta_R$</td>
<td>0.22</td>
<td>4.87</td>
<td>0.13</td>
<td>5.78</td>
</tr>
<tr>
<td>$\beta_T$</td>
<td>-6.56</td>
<td>-7.31</td>
<td>-0.69</td>
<td>-1.45</td>
</tr>
<tr>
<td>$\beta_E$</td>
<td>-2.03</td>
<td>-13.28</td>
<td>-2.89</td>
<td>-18.38</td>
</tr>
<tr>
<td>$\beta_L$</td>
<td>-1.57</td>
<td>-13.90</td>
<td>-2.70</td>
<td>-20.34</td>
</tr>
<tr>
<td>$\theta$</td>
<td>-</td>
<td></td>
<td>0.43</td>
<td>6.25</td>
</tr>
<tr>
<td>VOT (Euro/h)</td>
<td>30.16</td>
<td></td>
<td>5.20</td>
<td></td>
</tr>
<tr>
<td>VSDE (Euro/h)</td>
<td>9.34</td>
<td></td>
<td>21.62</td>
<td></td>
</tr>
<tr>
<td>VSDL (Euro/h)</td>
<td>7.22</td>
<td></td>
<td>20.22</td>
<td></td>
</tr>
<tr>
<td>Nr. Obs.</td>
<td>1158</td>
<td></td>
<td>5965</td>
<td></td>
</tr>
<tr>
<td>LogLik.</td>
<td>-2681</td>
<td></td>
<td>-10550</td>
<td></td>
</tr>
<tr>
<td>Pseudo $R^2$</td>
<td>0.17</td>
<td></td>
<td>0.36</td>
<td></td>
</tr>
</tbody>
</table>
Implications for pricing?

- Implications for pricing
  - Is a separate regulation of choice of $t^*$ desirable, above that of trip timing?
  - Peer and Verhoef 2012
    - Bottleneck model
    - Not conclusive on need for LR toll due to corner solutions
Henderson-Chu model

- Alternative to Vickrey – ADL bottleneck
  - Demand-side and scheduling behaviour identical
    - “α β γ” preferences
  - Congestion technology different
    - Vickrey: kinked performance function
    - Chu: smooth performance function
      - Delay is a function of outflow
      - E.g.: power function (“BPR”)
      - Optimal toll: instantaneous application of Pigouvian toll
  - Both have closed-form solutions
    - Also for equilibrium (time-independent) cost (c) and price (p)
Main ingredients

- $N$ identical travellers with “$\alpha \beta \gamma$” preferences
- LR VoSD fraction $g$ of SR VoSD
- LR VoT: relative premium of $a$ added to SR VoT
- SRPAT ($t^\#$) endogenous, LRPAT ($t^*$) identical and 0
- To avoid degenerate problem, we need variation between the days
  - Stochastic capacity $K$: $K_0 > K_1$
  - Probabilities: $(1-\pi)$ on state 0; $\pi$ on state 1
  - On the day itself, all travellers know the realization
Main ingredients con’d

- BPR travel time function
  - Ignore free-flow travel time
  - Delay: \( (r(t')/K)^x \)

- Equilibria
  - Short-run: equilibrium distribution of arrival times \( r(t') \)
    - … given the distribution of SRPATs \( z(t#) \) and given the realization of \( K \)
  - Long-run: equilibrium distribution of SRPATs \( z(t#) \)
    - … given that short-run equilibria as above will apply
LR equilibrium

- Three candidate types of LR equilibria
  - “Always Dispersed” (AD): values of $t^#$ are chosen so dispersed that all drivers arrive at $t^#$ in both states
  - “Sometimes Dispersed” (SD): density of $z(t^#)$ is so high that only state 0 is dispersed
    - State 1 is “condensed”: early drivers arrive before their $t^#$ and late drivers after their $t^#$
  - “Never Dispersed” (ND): both states “condensed”
    - ND is no equilibrium: it always pays off to widen $z(t^#)$ to save SR SDC and accept increased LR SDC ($g<1$)
Solution

- Important elements are the “reference arrival rate distributions” $r_0(t')$ and $r_1(t')$
  - These would apply in the basic Chu model with deterministic capacity $K_0$ or $K_1$, and with identical $t^*$

- Actual arrival pattern:
  - Condensed peak: reference arrival pattern
  - Dispersed peak: $r(t')$ equals $z(t^\#)$
    - Everybody arrives on time (at SRPAT)
Example: SD-NTE \( (\pi = 0.025) \)

- Actual arr rate 0

\[
\begin{align*}
r_0(t') \quad & \quad r_1(t') \quad & \quad z(t^#)
\end{align*}
\]

Congestion pricing with LR & SR scheduling
Expected SR cost and LR cost
Are long-run tolls needed? AD

\[ c^{LR}(t^\#) = \alpha \cdot (1 + a) \left( (1 - \pi) \cdot \left( \frac{z(t^\#)}{K_0} \right)^\chi + \pi \cdot \left( \frac{z(t^\#)}{K_1} \right)^\chi \right) + \left\{ \begin{array}{l} -\beta \cdot g \cdot t^\# \\ \gamma \cdot g \cdot t^\# \end{array} \right\} \]

\[ p^{LR}(t^\#) = c^{LR}(t^\#) + (1 - \pi) \cdot \tau_0^{SR}(t^\#) + \pi \cdot \tau_1^{SR}(t^\#) + \tau^{LR}(t^\#) \]

\[ mc^{LR}(t^\#) = c^{LR}(t^\#) + z(t^\#) \cdot (1 + a) \cdot \alpha \cdot \left( (1 - \pi) \cdot \frac{\partial T_0(z(t^\#))}{\partial z(t^\#)} + \pi \cdot \frac{\partial T_1(z(t^\#))}{\partial z(t^\#)} \right) \]

\[ \tau_0^{SR}(t^\#) = z(t^\#) \cdot \alpha \cdot \frac{\partial T_0(z(t^\#))}{\partial z(t^\#)} \quad \quad \tau_1^{SR}(t^\#) = z(t^\#) \cdot \alpha \cdot \frac{\partial T_1(z(t^\#))}{\partial z(t^\#)} \]

\[ \tau^{LR}(t^\#) = z(t^\#) \cdot a \cdot \alpha \cdot \left( (1 - \pi) \cdot \frac{\partial T_0(z(t^\#))}{\partial z(t^\#)} + \pi \cdot \frac{\partial T_1(z(t^\#))}{\partial z(t^\#)} \right) \]

Congestion pricing with LR & SR scheduling
Intuition

- To establish short-run optimum in both states, short-run tolls must be based on short-run “α β γ”
  - Through Pigouvian form, α in particular

- But long-run expected travel times are proportional (probability-weighted) with short-run travel times
  - Same internalization argument applies
  - Must be a long-run toll in order not to distort short-run optima

- Value of long-run toll is simply a times the expected value of short-run tolls
Long-run tolls are less strongly needed in SD

\[ c^{LR}(t^#) = (1 - \pi) \cdot \alpha \cdot (1 + a) \cdot \left( \frac{z(t^#)}{K_0} \right)^\chi + \pi \cdot \left( c_{1}^{SR} + a \cdot \alpha \cdot \left( \frac{r_1(t(t^#))}{K_0} \right)^\chi \right) + \left\{ \begin{array}{l} \beta \cdot (g - \pi) \cdot -t^# \\ \gamma \cdot (g - \pi) \cdot t^# \end{array} \right\} \]

\[ p^{LR}(t^#) = \tau^{LR} + (1 - \pi) \cdot \left( \tau_{0}^{SR} + \alpha \cdot (1 + a) \cdot \left( \frac{z(t^#)}{K_0} \right)^\chi \right) + \pi \cdot \left( p_{1}^{SR} + a \cdot \alpha \cdot \left( \frac{r_1(t(t^#))}{K_0} \right)^\chi \right) + \left\{ \begin{array}{l} \beta \cdot (g - \pi) \cdot -t^# \\ \gamma \cdot (g - \pi) \cdot t^# \end{array} \right\} \]

\[ p^{LR}(t^#) - mc^{LR}(t^#) = \tau^{LR} + (1 - \pi) \cdot \left( \tau_{0}^{SR} -(1 + a) \cdot \alpha \cdot \frac{\partial T_0(z(t^#))}{\partial z(t^#)} \right) + \pi \cdot \left( p_{1}^{SR} - mc_{1}^{SR} + a \cdot \alpha \cdot T_1(t(t^#)) \right) \]

\[ \tau_{0}^{SR}(t^#) = z(t^#) \cdot \alpha \cdot \frac{\partial T_0(z(t^#))}{\partial z(t^#)} \]

\[ \tau_{1}^{SR}(t') = r_1(t') \cdot \alpha \cdot \frac{\partial T_1(r_1(t'))}{\partial r_1(t')} \]

\[ \tau^{LR}(t^#) = (1 - \pi) \cdot z(t^#) \cdot a \cdot \alpha \cdot \frac{\partial T_0(z(t^#))}{\partial z(t^#)} - \pi \cdot a \cdot \alpha \cdot T_1(t(t^#)) \]
Intuition

• For state 0, things work as in previous (AD) case
  • LR toll contains a factor \((1-\pi)\times a\) times SR toll in state 0

• But for state 1, the congestion externality is dropped
  • Marginal changes in \(z(t^\#)\) will not change traffic conditions in state 1: it is a condensed equilibrium
  • So no externality of that type enters the LR toll rule

• Instead, what is subtracted from the LR toll rule is the factor \(\pi \times a \times \text{(travel delay in state 1)}\)
  • It is part of the generalized price, but not of the marginal cost for \(z(t^\#)\)
  • A marginal change in \(z(t^\#)\) does not change these costs
Numerical illustration

- Parameters:
  - $N = 10\,000$
  - $K_0=10\,000; K_1=5\,000$
  - $\pi=0.25$
  - $\chi=4$
  - $\alpha=10$
  - $\beta=5$
  - $\gamma=20$
  - $\delta=4$
  - $a=3$
  - $g=0.5$
Still, modest cost reduction compared to QFB
- QFB realizes 77% of FB cost reduction (SD)
- Absence of LR toll makes SR tolls higher; $E$ peaks near 4
Relative efficiency $QFB: a$

Note: NTE is AD throughout; QFB is SD for $a=\{1,2,3\}$ and AD for $a=\{4,5\}$; FB is SD for $a=1$ and AD for $a=\{2,3,4,5\}$

Figure 3. Varying $a$: relative efficiency $\omega$

LR toll more important if $a$ higher
Relative efficiency QFB: $g$

Note: NTE is AD throughout; QFB is AD for $g=\{0.3, 0.4\}$ and SD for $g=\{0.5, 0.6, 0.7\}$; FB is AD for $g=\{0.3, 0.4, 0.5, 0.6\}$ and SD for $a=0.7$

Figure 4. Varying $g$: relative efficiency $\omega$
Relative efficiency QFB: $\pi$

Note: NTE is SD for $\pi=0.05$ and AD for $\pi=\{0.15,0.25,0.3,0.45\}$; QFB is SD for $\pi=\{0.05,0.15,0.25\}$ and AD for $\pi=\{0.35,0.45\}$; FB is SD for $\pi=0.05$ and AD for $\pi=\{0.15,0.25,0.3,0.45\}$.

Figure 5. Varying $\pi$: relative efficiency $\omega$
Relative efficiency QFB: $\chi$

Note: NTE is AD throughout; QFB is AD for $\chi=\{1,2.5\}$ and SD for $\chi=\{4,5.5,7\}$; FB is AD for $\chi=\{1,2.5,4,5.5\}$ and SD for $\chi=7$

*Figure 6. Varying $\chi$: relative efficiency $\omega$*

High chi: only small changes in flow are needed to optimize. Once that is done, little gain from LR policy
Conclusion

- Long-run toll is needed when short-run and long-run valuations of time diverge
  - Surprisingly, the need is larger for instances if the first-best is more lightly congested
  - Reason: in a condensed equilibrium, arrival pattern becomes insensitive to marginal changes in desired arrival times

- QFB has relative efficiency that may falls as low as 0.6 in the numerical example used
Solution

- Solution proceeds technically in the same way for the three pricing regimes
  - NT = No Toll
  - QFB = Quasi First-Best: short-run Chu-tolls only
  - FB = First-Best: QFB plus possibly a long-run toll to optimize the choice of $t^*$
- Solution differs between AD and SD
- Main steps:
  - Solve the partial differential equation for $z(t^*)$ that makes the long-run generalized price constant over time, given the short-run equilibria (and toll rules)
  - Solve for $t_f$ and $t_{f'}$ that guarantee $N$ drivers and equalized generalized prices at those two moments
Solution Chu model

Generalized price always takes the form \( (X \text{ and } d \text{ depend on AD vs SD and on pricing regime}):\)

\[
p^{LR}(t^#) = X \cdot z(t^#)^x + Cons + \left\{ \begin{array}{c} -d \cdot \beta \cdot t^# \\ d \cdot \gamma \cdot t^# \end{array} \right\}
\]

\[
\dot{p}^{LR}(t^#) = X \cdot \chi \cdot z(t^#)^{x-1} \cdot \dot{z}(t^#) + \left\{ \begin{array}{c} -d \cdot \beta \\ d \cdot \gamma \end{array} \right\} = 0
\]

\[
z(t^#) = \left( \frac{1}{X} \cdot \left\{ \begin{array}{c} d \cdot \beta \cdot (t^# - t_l) \\ d \cdot \gamma \cdot (t_l - t^#) \end{array} \right\} \right)^{\frac{1}{x}}
\]

\[
t_l = -\left( N \cdot \frac{\gamma}{\beta + \gamma} \cdot \frac{1 + \chi}{\chi} \cdot \left( \frac{X(\phi)}{d \cdot \beta} \right)^{\frac{1}{x}} \right)^{\frac{x}{1 + \chi}}
\]

\[
t_l' = \left( N \cdot \frac{\beta}{\beta + \gamma} \cdot \frac{1 + \chi}{\chi} \cdot \left( \frac{X(\phi)}{d \cdot \gamma} \right)^{\frac{1}{x}} \right)^{\frac{x}{1 + \chi}}
\]
<table>
<thead>
<tr>
<th>No-toll equilibrium</th>
<th>First-best optimum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short-hand $\Psi_i$</td>
<td>$\Psi_i = \left( \frac{N}{K_i} \cdot \frac{1 + \chi \cdot \delta}{\chi \cdot \alpha} \right)^{\frac{1}{\beta}}$</td>
</tr>
<tr>
<td>Arrival rate $r(t')$ early ($t' \leq t^*$)</td>
<td>$r(t') = K_i \cdot \left( \frac{\beta}{\alpha} \cdot (t' - t_q) \right)^{1 - \frac{1}{\beta}}$</td>
</tr>
<tr>
<td>Arrival rate $r(t')$ late ($t' &gt; t^*$)</td>
<td>$r(t') = K_i \cdot \left( \frac{\gamma}{\alpha} \cdot (t_q - t') \right)^{1 - \frac{1}{\gamma}}$</td>
</tr>
<tr>
<td>Early interval: $t^* - t_q$</td>
<td>$t^* - t_q = \Psi_i \cdot \frac{\alpha}{\beta}$</td>
</tr>
<tr>
<td>Late interval: $t_q - t^*$</td>
<td>$t_q - t^* = \Psi_i \cdot \frac{\alpha}{\gamma}$</td>
</tr>
<tr>
<td>Generalized price $p$</td>
<td>$p = \Psi_i \cdot \alpha$</td>
</tr>
<tr>
<td>Average generalized cost $\overline{c}$</td>
<td>$\overline{c} = \Psi_i \cdot \alpha$</td>
</tr>
<tr>
<td>Total travel delay cost $TDC$</td>
<td>$TDC = \frac{1 + \chi}{1 + 2 \cdot \chi} \cdot \Psi_i \cdot \alpha \cdot N$</td>
</tr>
<tr>
<td>Total schedule delay cost $SDC$</td>
<td>$SDC = \frac{\chi}{1 + 2 \cdot \chi} \cdot \Psi_i \cdot \alpha \cdot N$</td>
</tr>
<tr>
<td>Total toll revenue $TR$</td>
<td>$TR = 0$</td>
</tr>
<tr>
<td>Total social cost $C$</td>
<td>$C = \Psi_i \cdot \alpha \cdot N$</td>
</tr>
</tbody>
</table>

| Toll $\pi(t')$ | $\pi(t') = 0$ | $\pi(t') = \alpha \cdot \chi \cdot \mathbf{a(t')} \left( \frac{a(t')}{} \right)^{\chi}$ |
| | | $= \alpha \cdot \chi \cdot (T - T_i)$ |
| | | $= \frac{\chi}{1 + \chi} \cdot \beta \cdot (t' - t_q)$ if $t' \leq t^*$ |
| | | $= \frac{\chi}{1 + \chi} \cdot \gamma \cdot (t_q - t')$ if $t' > t^*$ |

**Note:** Costs and prices are net of free-flow travel times $T_i$. Inclusion would require adding $\alpha T_i$ for average cost and generalized price measures, and $N \alpha T_i$ for inclusion in total costs measures.

*Table 1. Equilibrium and first-best optimum*