The Economics of Urban Road Pricing

Capita Selecta

Erik T Verhoef

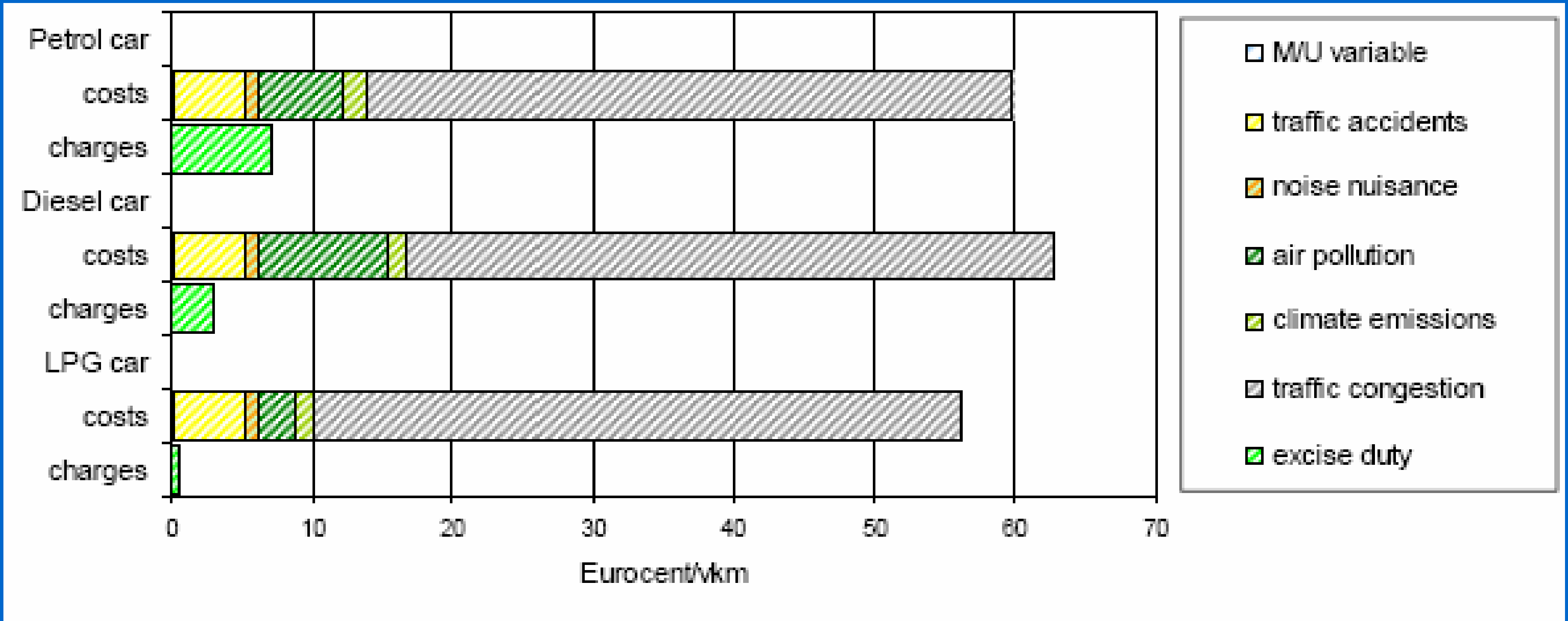
Free University Amsterdam

Outline

- Basic economic motivation
- Statics versus dynamics
- Second-best aspects
- Conclusions

Basic motivation road pricing

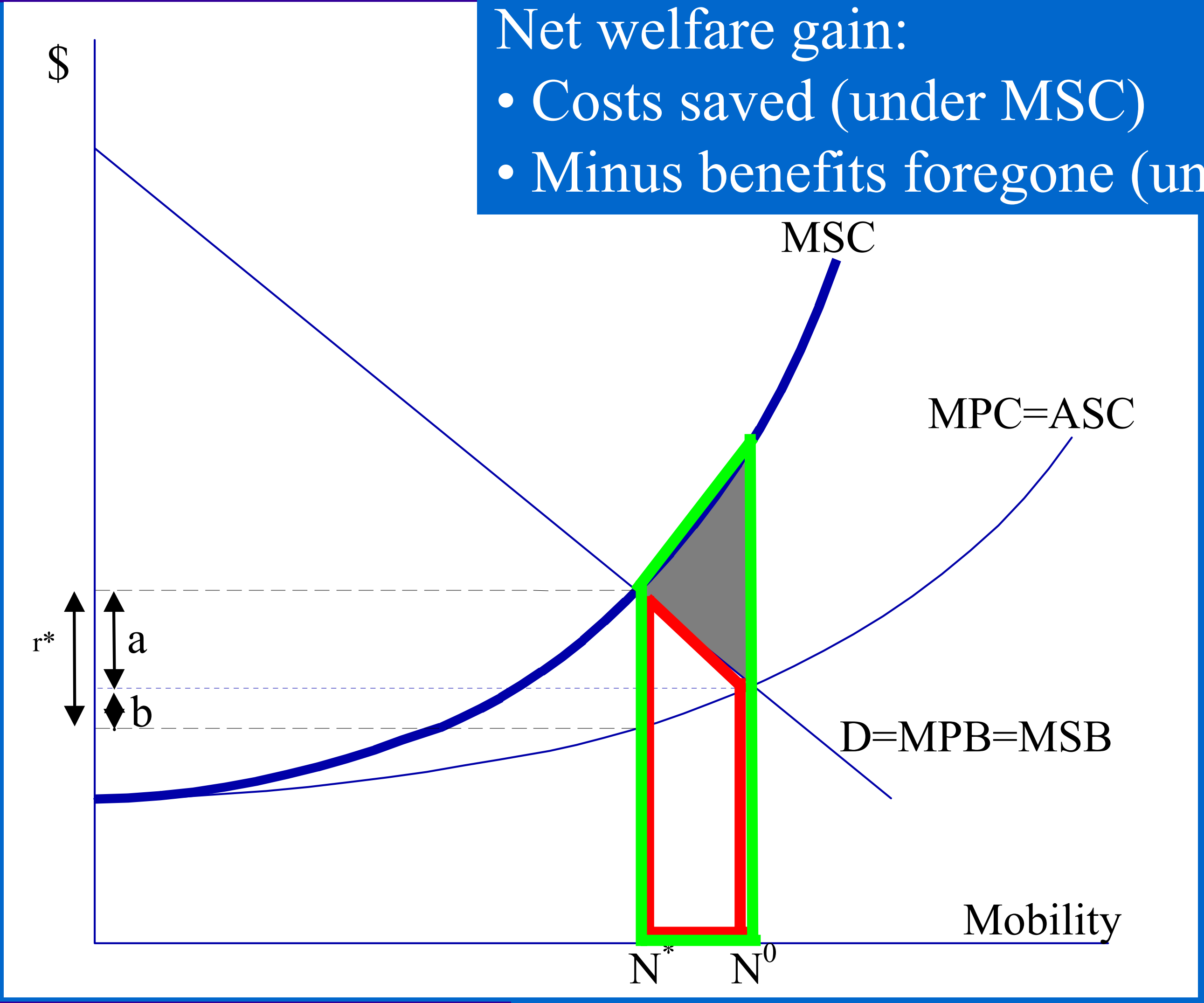
- Pigou (1920): external costs
  - At the margin, *mb* = *mpc* instead of *mb* = *msc*
  - Social welfare rises when discouraging traffic with *mb* < *msc*



Textbook basics

Net welfare gain:

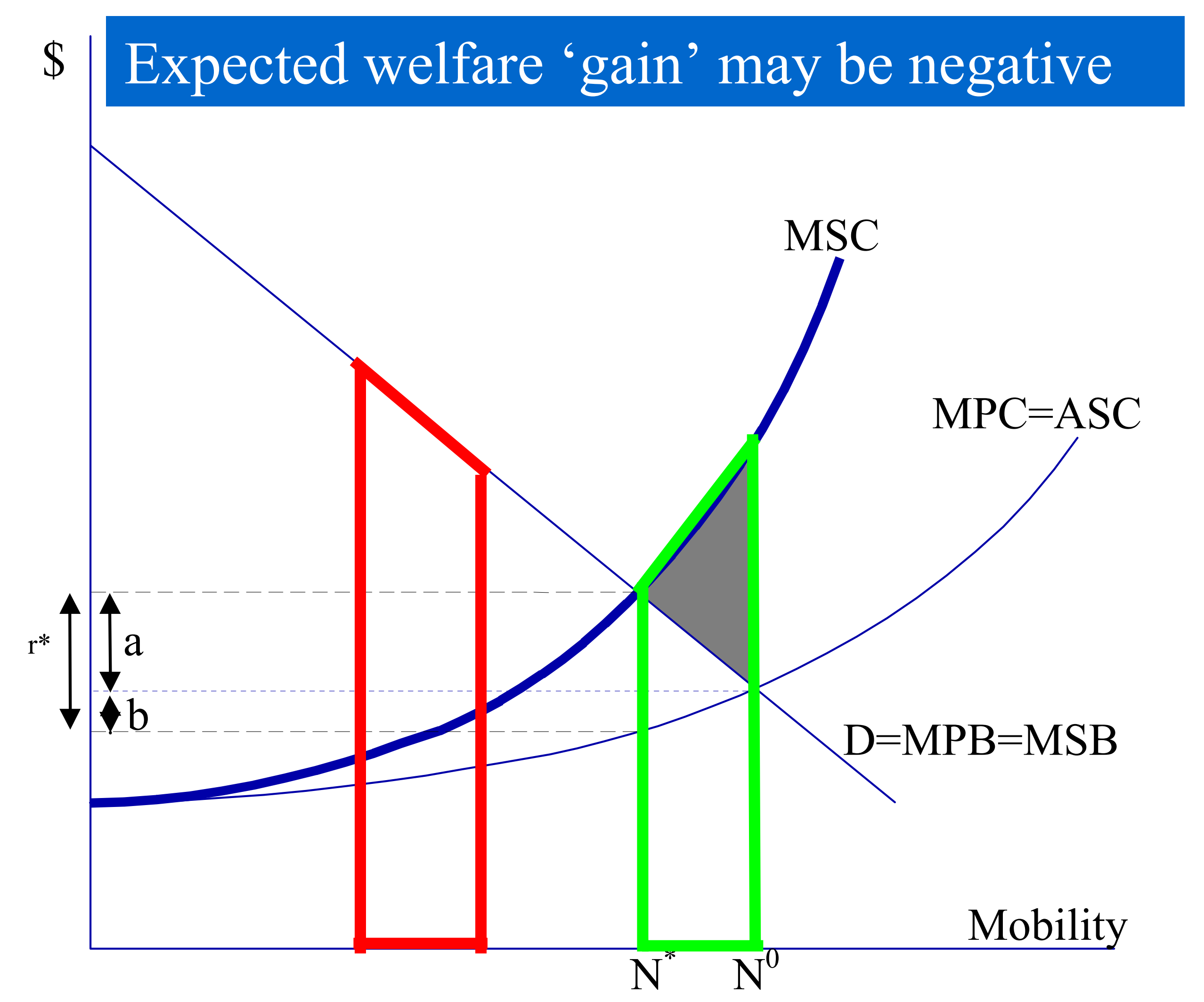
- Costs saved (under MSC)
- Minus benefits foregone (under MB)



Subtlety

- Pricing outperforms non-price regulation in terms of efficiency
  - Level and composition of road use matters
- Example: ‘Athens-type’ number plate policy
  - Does not discriminate according to WTP
  - Even if a clever design succeeds in achieving N^*, not (nearly) as efficient as pricing

Number plates vs pricing



Modelling of traffic congestion

- Advantages of the basic static model
  - transparent
  - basic economic principles
- Disadvantages: simplicity
  - dynamics
  - networks
  - technical, non-behavioural nature of congestion function
  - … basic model of little use in practice?

Dynamic modelling

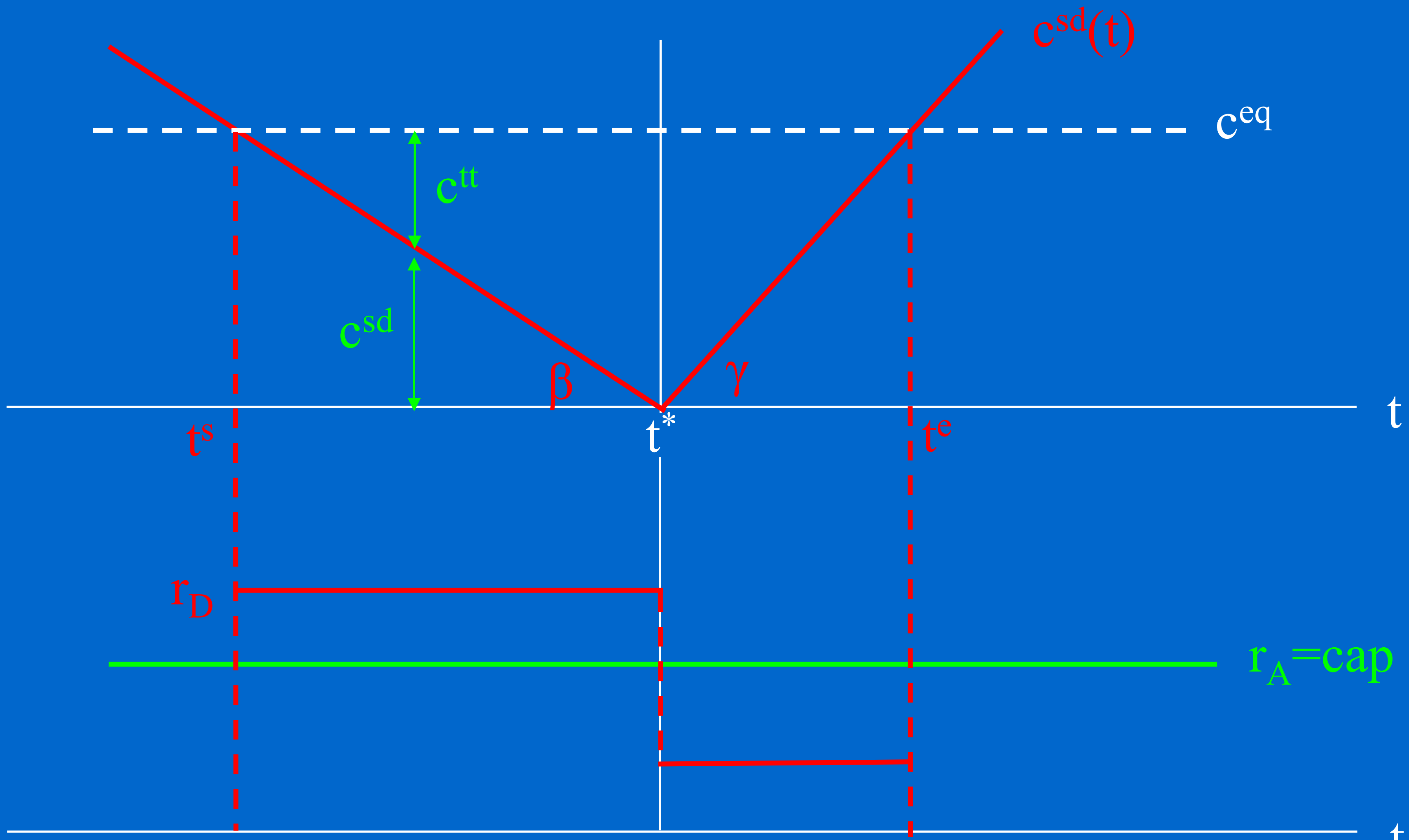
- Supply side: non-stationarity of traffic flows
- Demand side: dynamic equilibrium in terms of endogenized departure times
  - Generalized cost: schedule delay cost plus travel delay cost
  - Dynamic equilibrium: generalized cost constant over peak
- Important conclusions
  - No demand reduction needed to reduce congestion
  - Generalized price needs hardly rise with optimal tolling

Vickrey (1969)

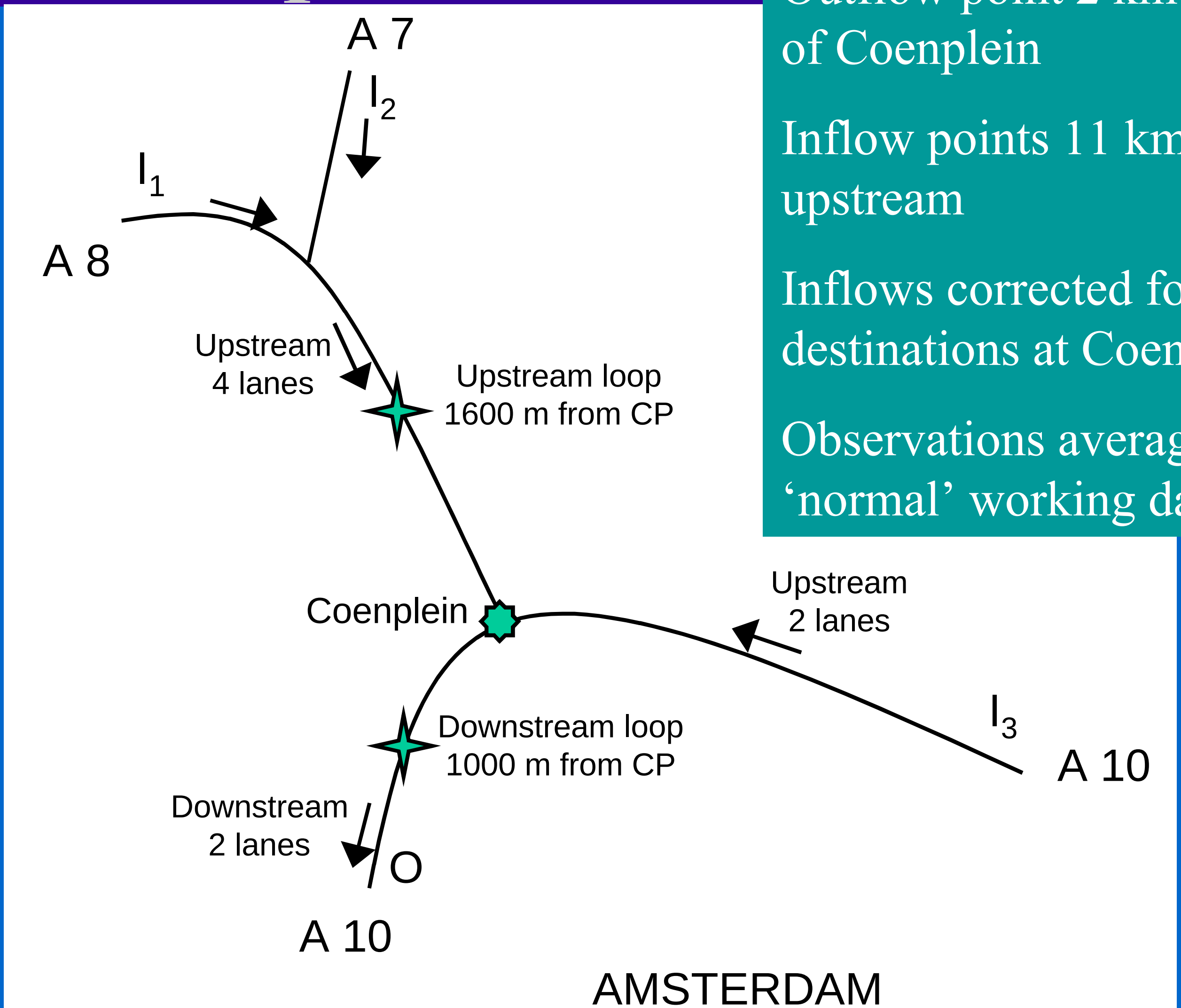
- Pure bottleneck congestion, for a single facility
  - no queue, inflow ≤ capacity:
    - outflow = inflow
  - else:
    - outflow = cap; growth of queue = inflow – outflow
- Dynamic equilibrium for homogeneous users with inelastic demand:
  - Early arrivals: inflow > capacity, queue grows over time
  - Late arrivals: inflow < capacity, queue shrinks over time

With linear SD-costs

$t^e - t^s = N/cap$

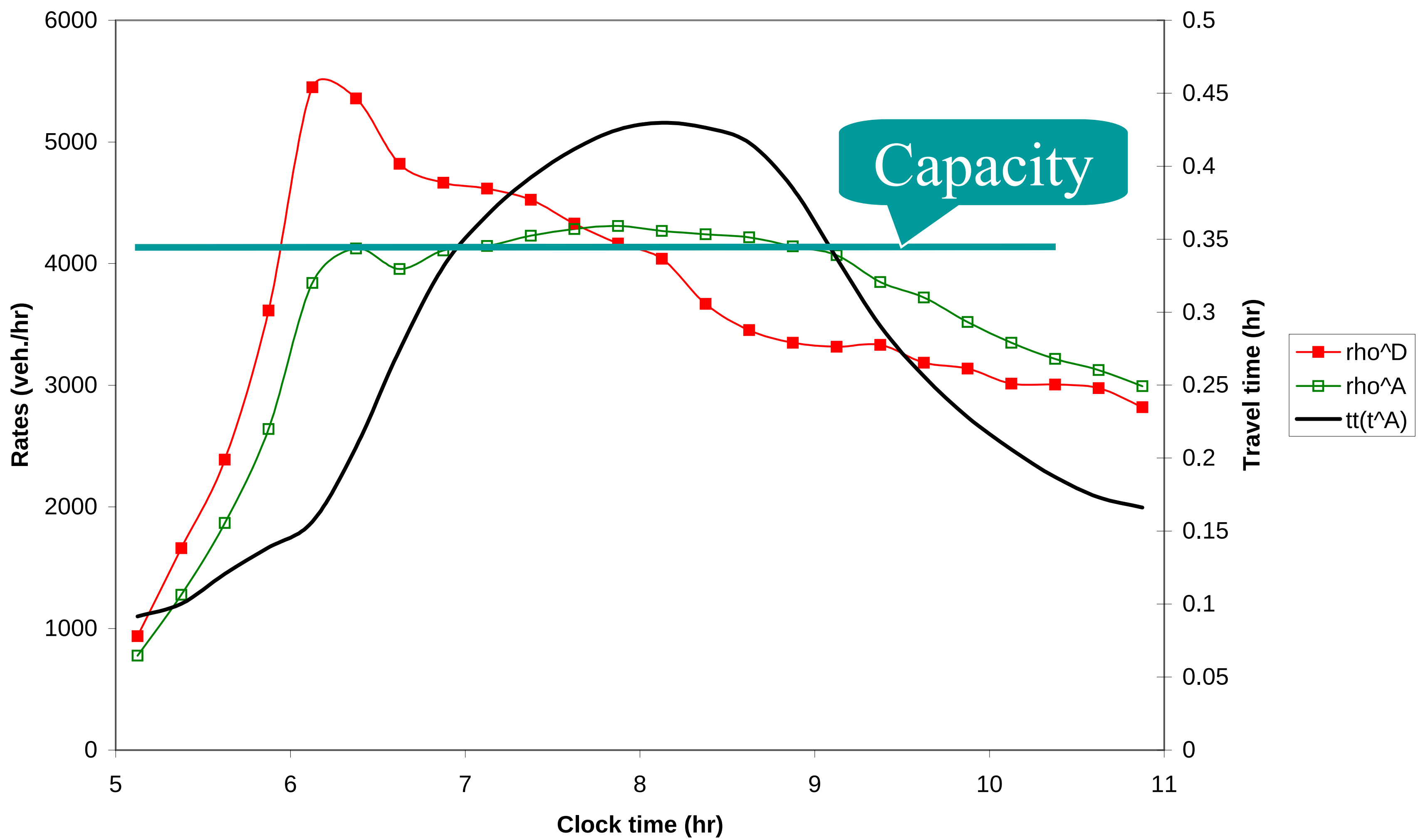


Empirical rele



- 1 downstream segment
- 3 'sub-queues'; weighted averaged travel times
- Outflow point 2 km downstream of Coenplein
- Inflow points 11 km (5.5 min.) upstream
- Inflows corrected for different destinations at Coenplein
- Observations averaged over 'normal' working days in 2000

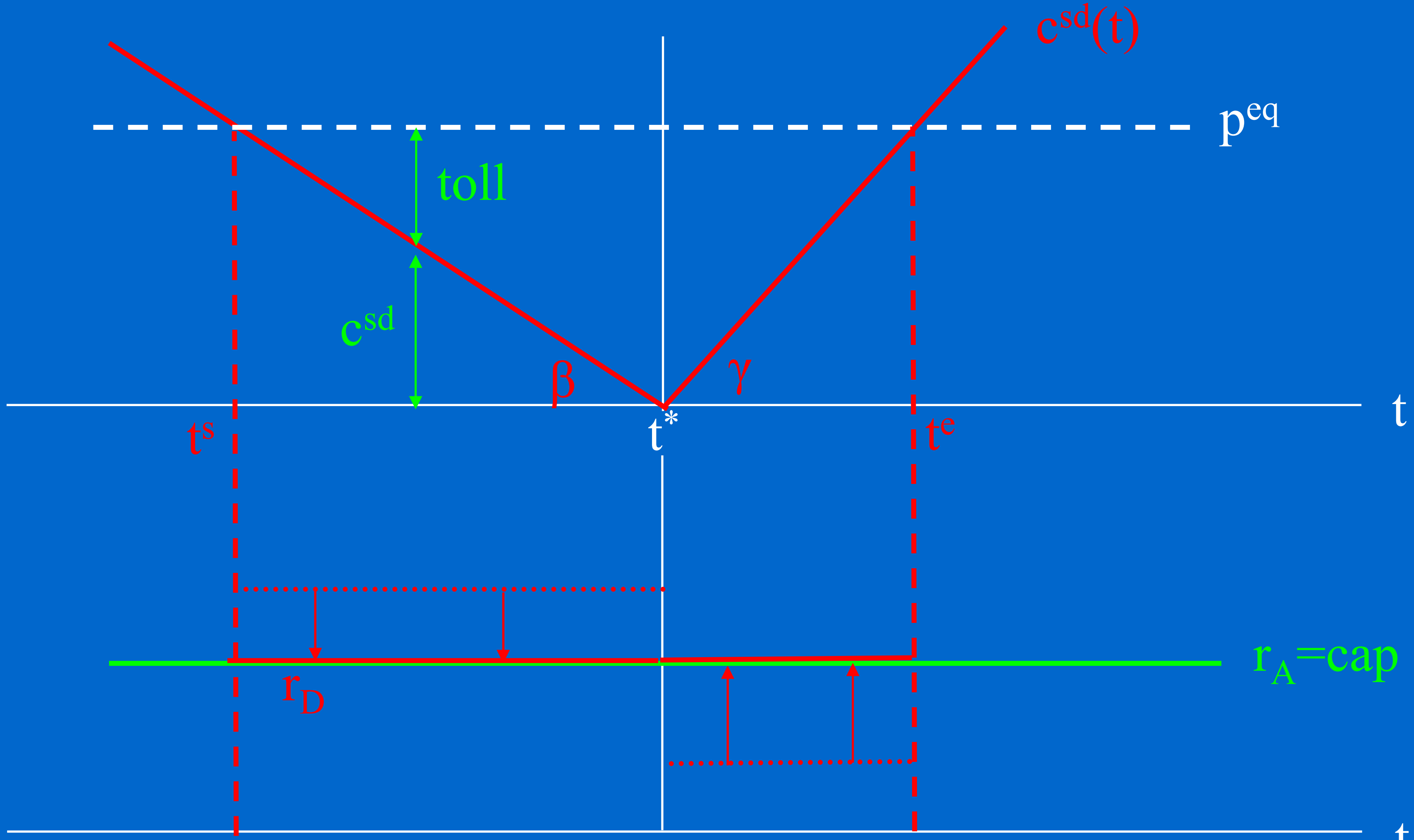
Dep. & Arr. rates and travel times



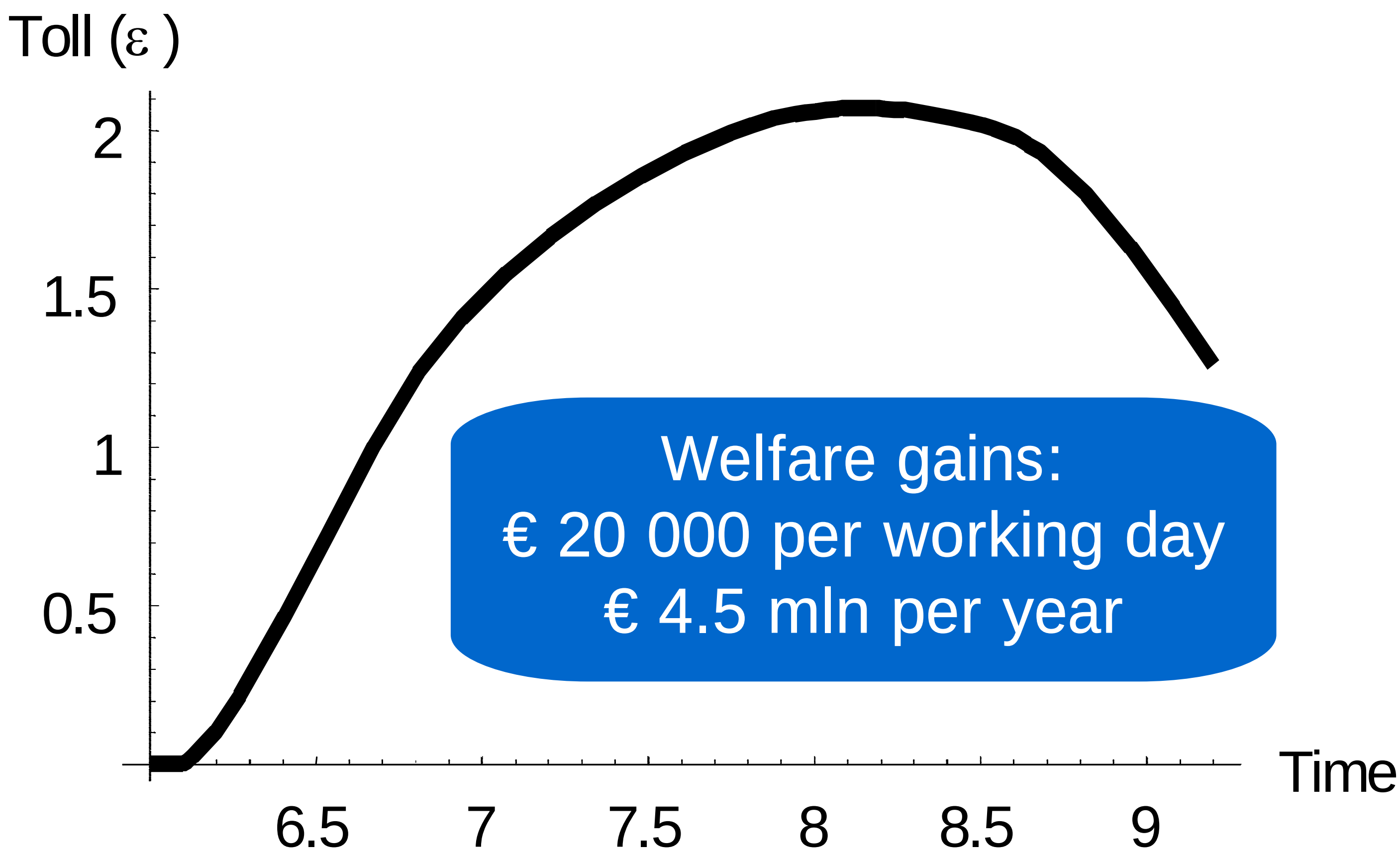
Optimality for a bottleneck

- Time spent queuing is a pure waste, but needed to achieve a dynamic equilibrium
  - Avoidance of queues, while keeping throughput at capacity, would eliminate travel delay cost without raising schedule delay cost
- Dynamic tolls
  - Purpose: inflow = capacity = outflow throughout peak as a ‘decentralized optimum’
    - Avoid wasteful queuing
  - Needed: time-varying tolls that replicate the dynamic equilibrium pattern of travel delay cost

With linear SD-costs



Prescription for ‘Coenplein’



Vickrey vs ‘standard’: surprises

- Congestion eliminated without demand reduction
  - No need to change mode, give up job, carpool, *etc*.
- Same arrival flow over the same time span
  - No need to arrive earlier or later at work
  - Only departure times are adjusted: everybody departs later than without tolling
- Acceptability of road pricing should be no problem with optimal time differentiation
  - Generalized equilibrium costs remain unchanged

Dynamic congestion technologies

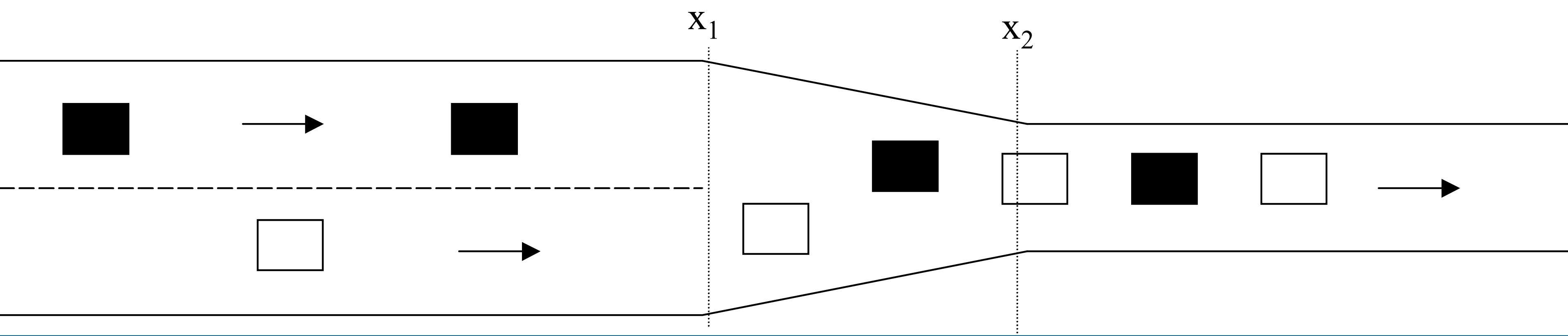
- Alternative flow-based representations
  - 'Instantaneous propagation' (Agnew, 1977)
    - Speeds along the road equal at every instant
  - 'No propagation' (Chu, 1995)
    - Drivers have constant speed over their entire trip, depending on arrival rate at instance of departure or arrival
  - 'Hybrid' (Mun, 1999)
    - Chu + basic bottleneck
  - 'Finite propagation': car-following modelling (Verhoef, 2001, 2003, 2004)

Which insights survive?

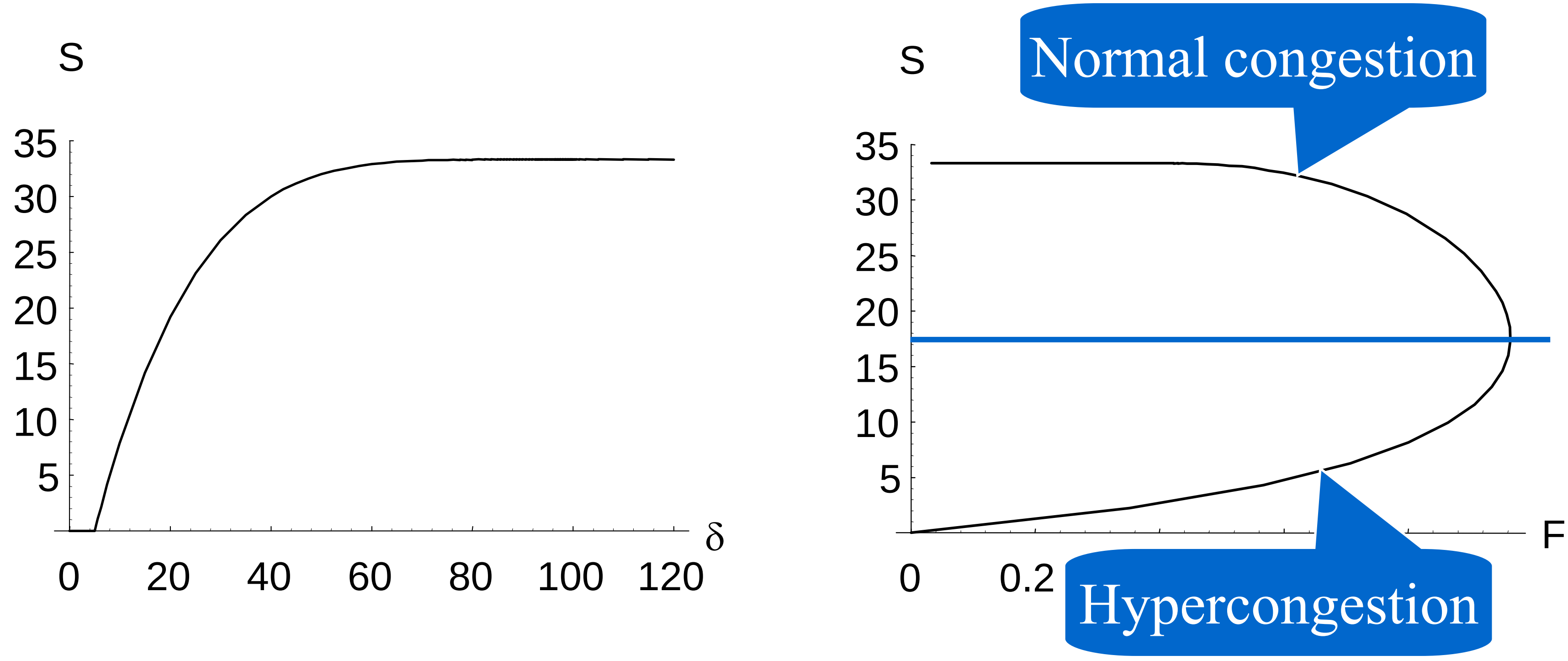
- Importance of rescheduling of departures for optimality
  - Need for continuous toll differentiation over time
- Modest increase in generalized price with optimal tolling; more optimisitic view on acceptability
  - Especially if the congestion technology allows for / incorporates some form of ‘hypercongested’ queuing
  - In practice: difference between ‘flowing traffic’ and ‘jammed’ traffic
  - Therefore: relevant for the most visible type of traffic congestion

Example from Verhoef (2003)

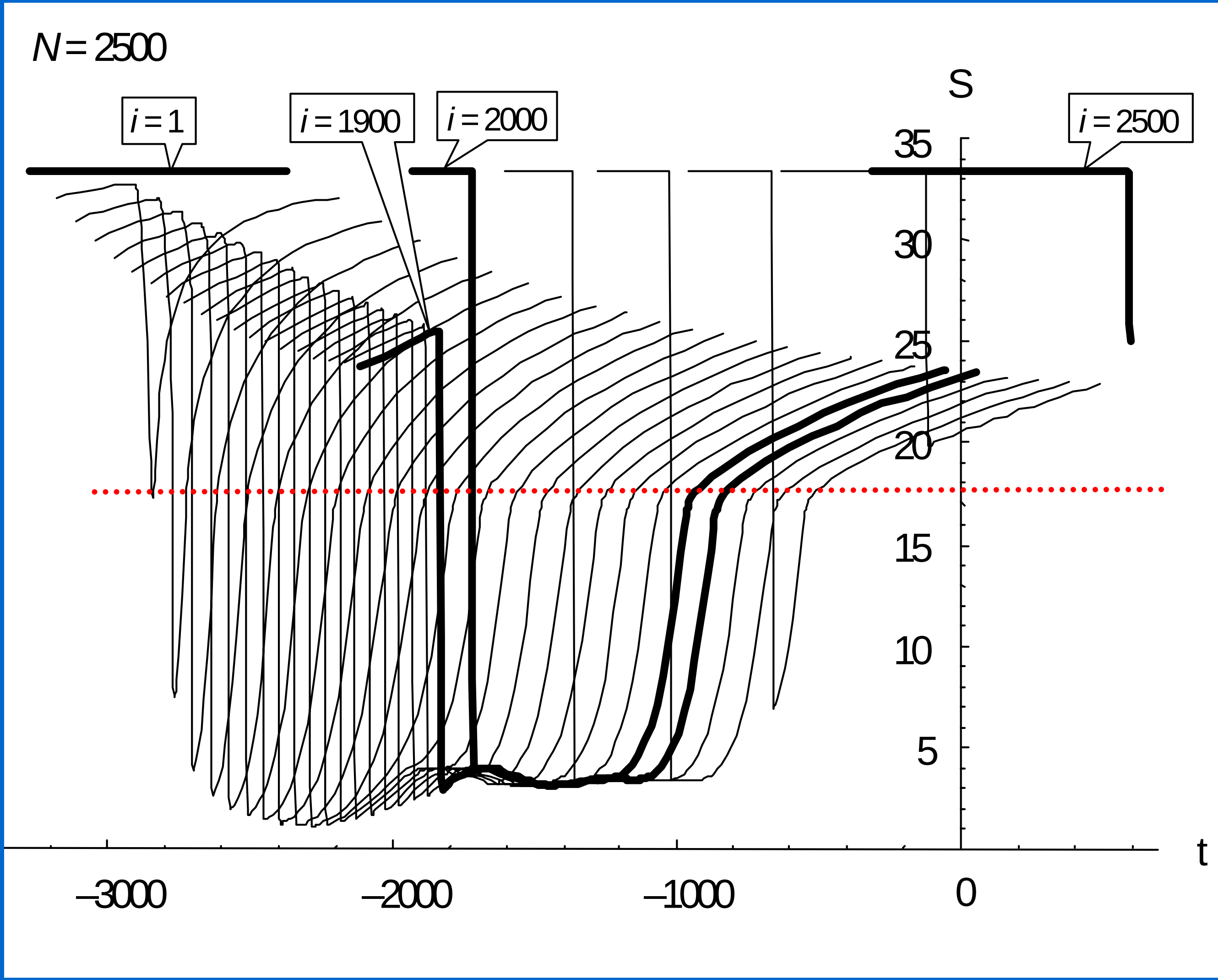
- Single origin and destination, one road
- Car-following congestion technology
- Numerical solutions only
- Bottleneck due to lane-merging
- ‘Loops’ to ‘monitor’ traffic dynamics



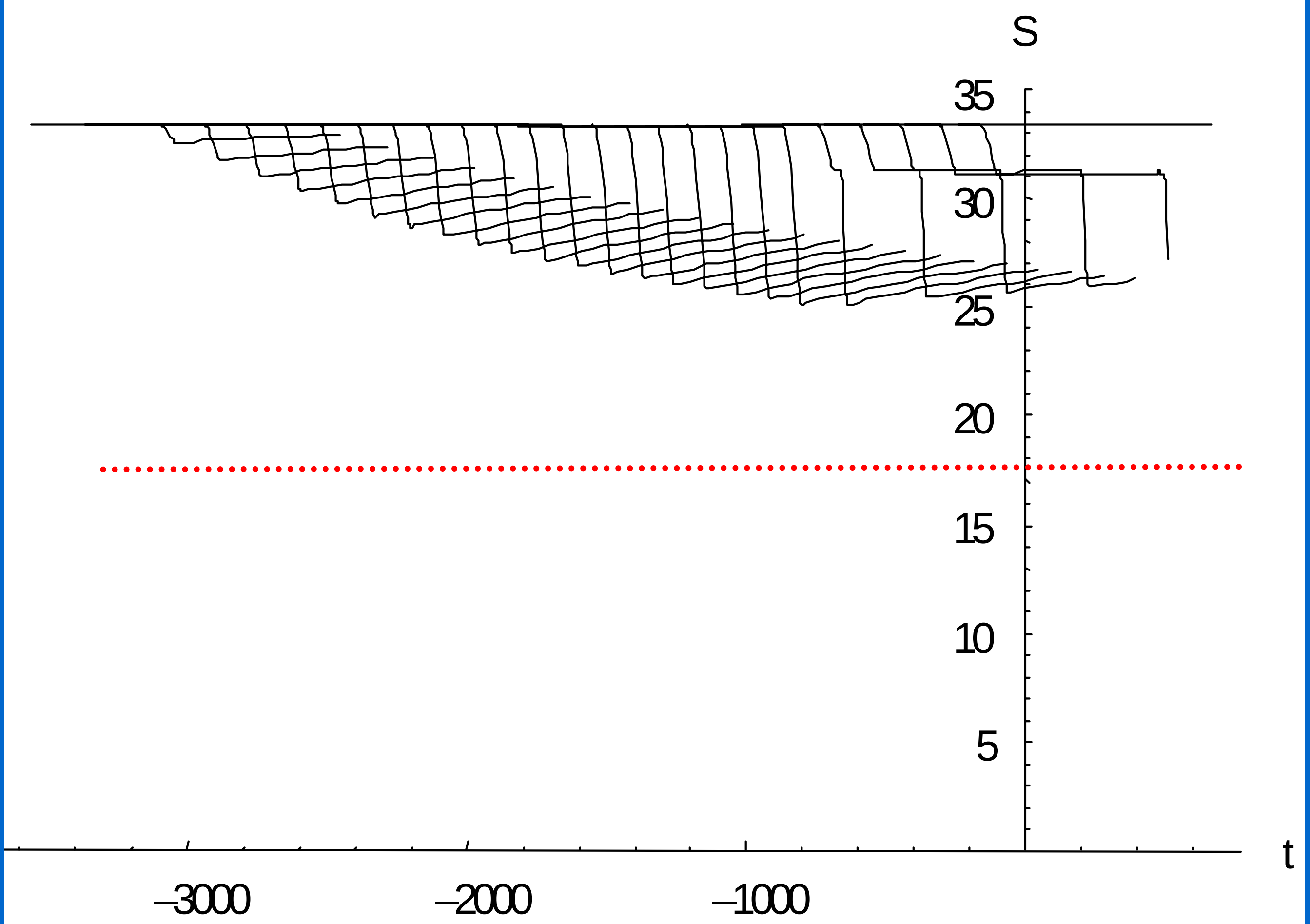
Assumed car-following equation



Clock-time speed functions: no tolls



Clock-time speed functions: optimum



Comparison with basic bottleneck model

Optimum *vs* equilibrium:	Bottleneck	Car-following (N = 2500)
Duration peak	+ 0%	+ 12%
Total variable cost	– 50%	– 40%
Total variable travel time cost	– 100%	– 85%
Total schedule delay cost	+ 0%	+ 10%
Generalized price (net of free-flow travel time)	+ 0%	+ 12%

Therefore:

- Dynamic models
  - Endogenize scheduling decisions
  - Importance of toll differentiation over time
  - Departure time adjustments may yield considerable gains even with perfectly inelastic demand
  - Generalized price needs not rise by much due to optimal tolling, especially with initial hypercongested queuing

Second-best pricing

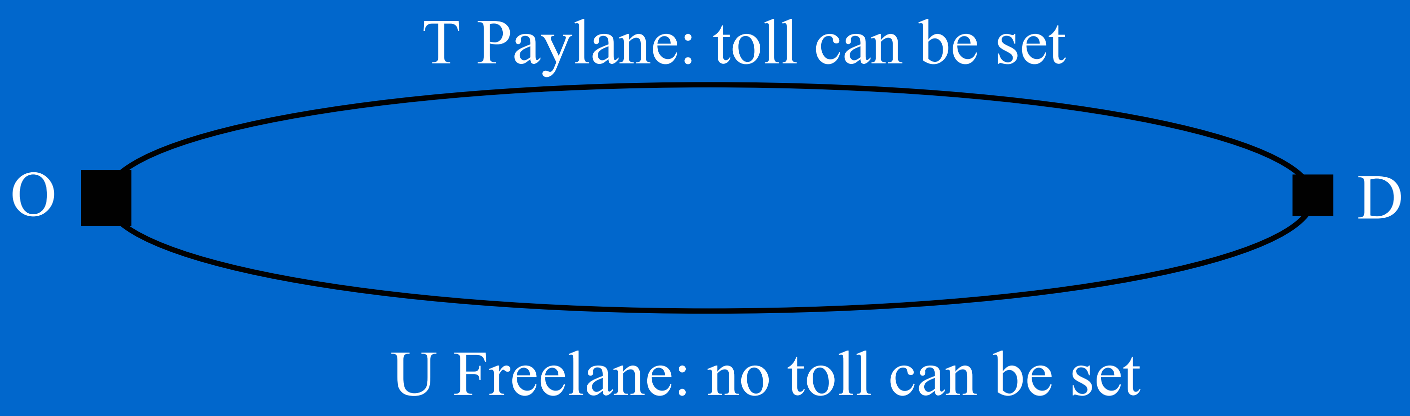
- Taxes as discussed up to here assume
  - No distortions in the economy but the externality under consideration
    - But: environmental pollution, market power, distortive labour taxes, etc.
  - Taxes can be differentiated perfectly over users
    - Time of day
    - Route followed
    - Vehicle used & maintenance
    - Driving style
- When violated: ‘Second-best pricing’

Therefore…

- Second-best pricing will be the rule rather than the exception
- Substantial literature on second-best pricing has recently emerged
- General issues best illustrated using an example

‘Two-route problem’

– Marchand and Levy-Lambert (1968)
– Typical of pay-lanes
– What is the optimal toll, which are the impacts?



The second-best optimal toll

- Trade off:
  - Good news: reduction of congestion on pay-lane
  - Bad news: increase in congestion on free-lane
- Constrained optimization:

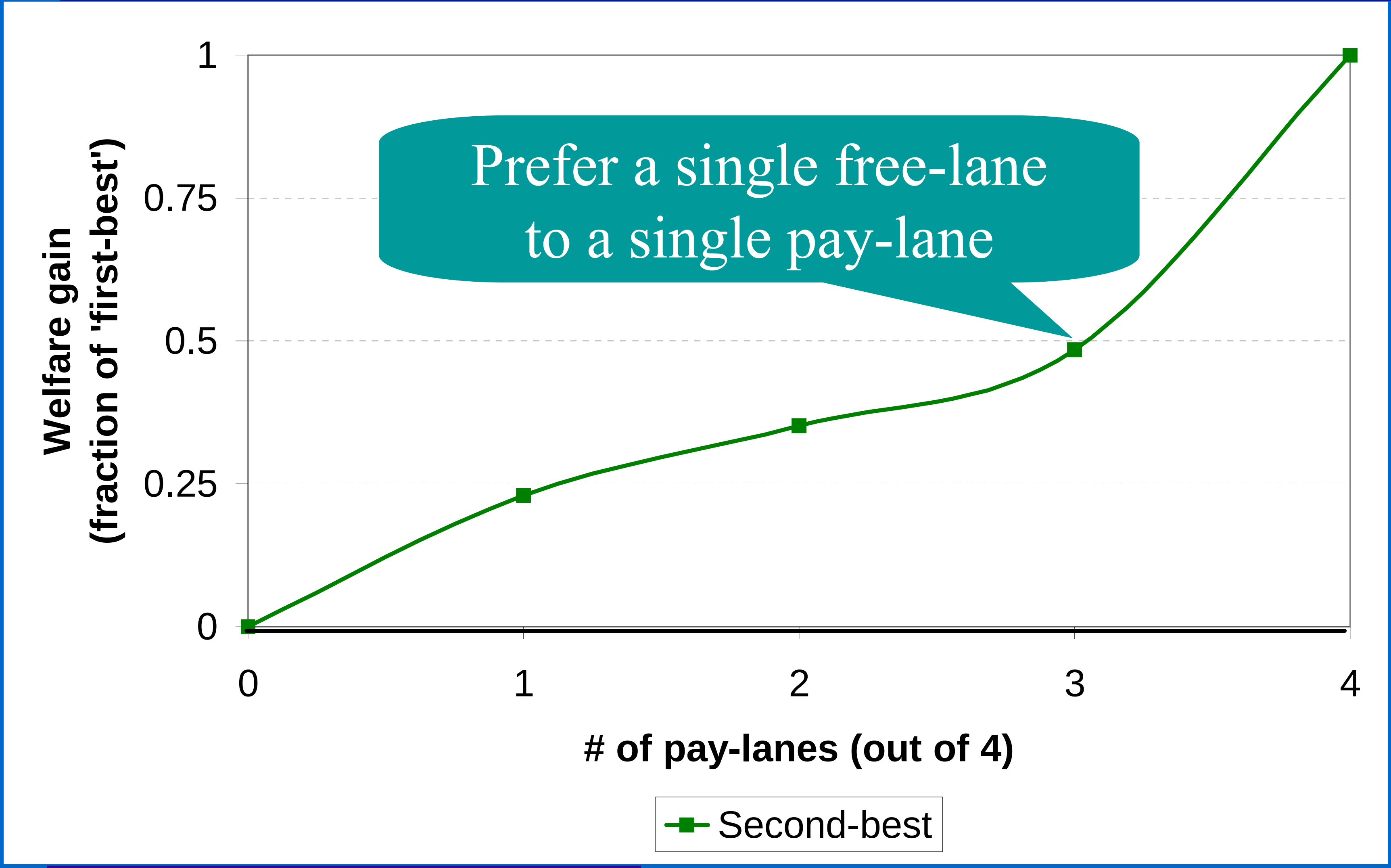
$$\tau = mec_T - mec_U \cdot \frac{-D'}{c'_U - D'}$$

- Two special cases:
  - Perfectly inelastic demand: s.b. toll equal to mec-difference
  - Perfectly elastic demand: s.b. toll ignores route U

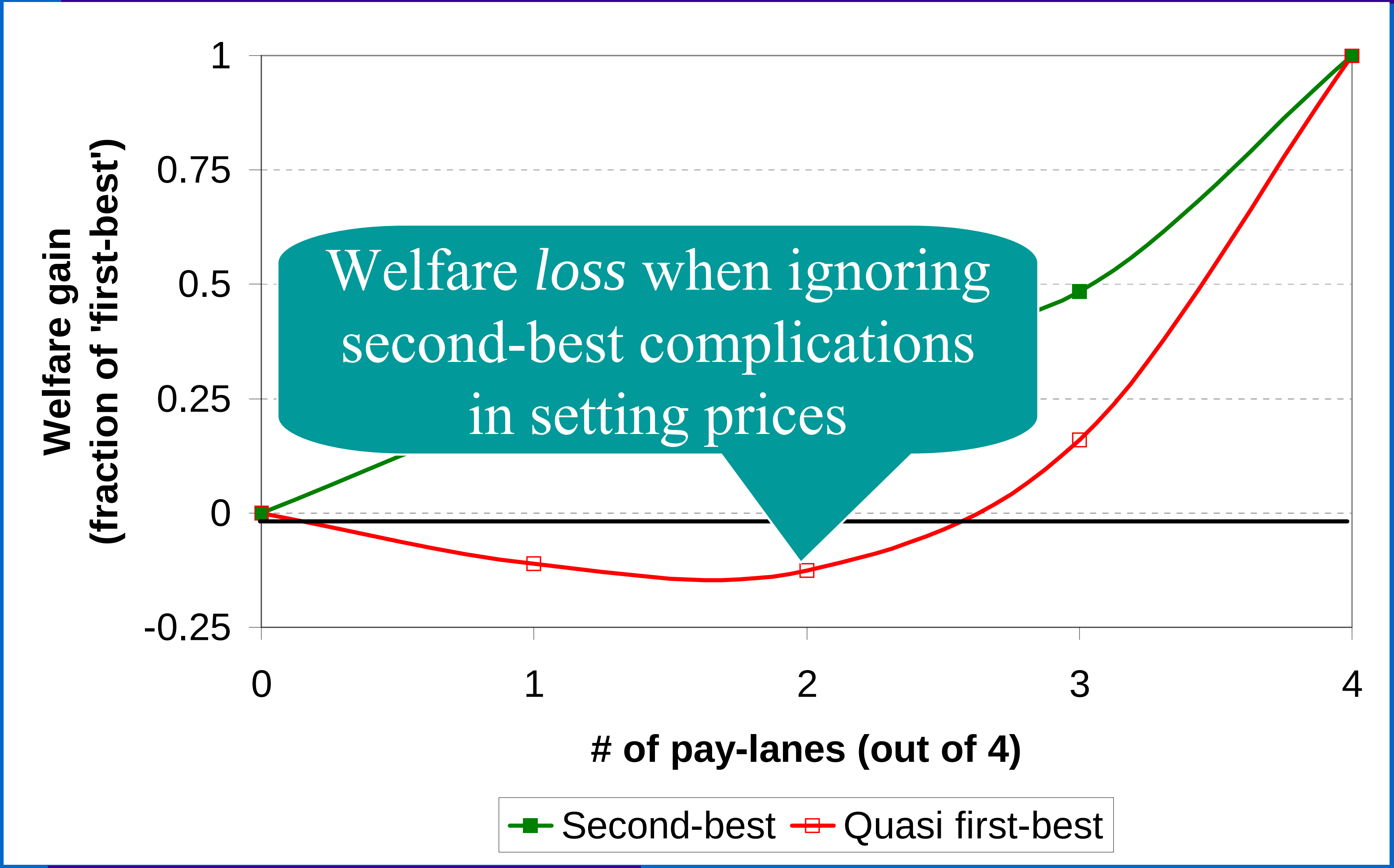
Illustration: extended version

– Account for heterogeneity of users (value of time)
– 4-lane highway
– A third, serial link where users from both routes interact
– Numerical model: calibrated so as to replicate Dutch peak conditions
– Results from Verhoef and Small (2004)

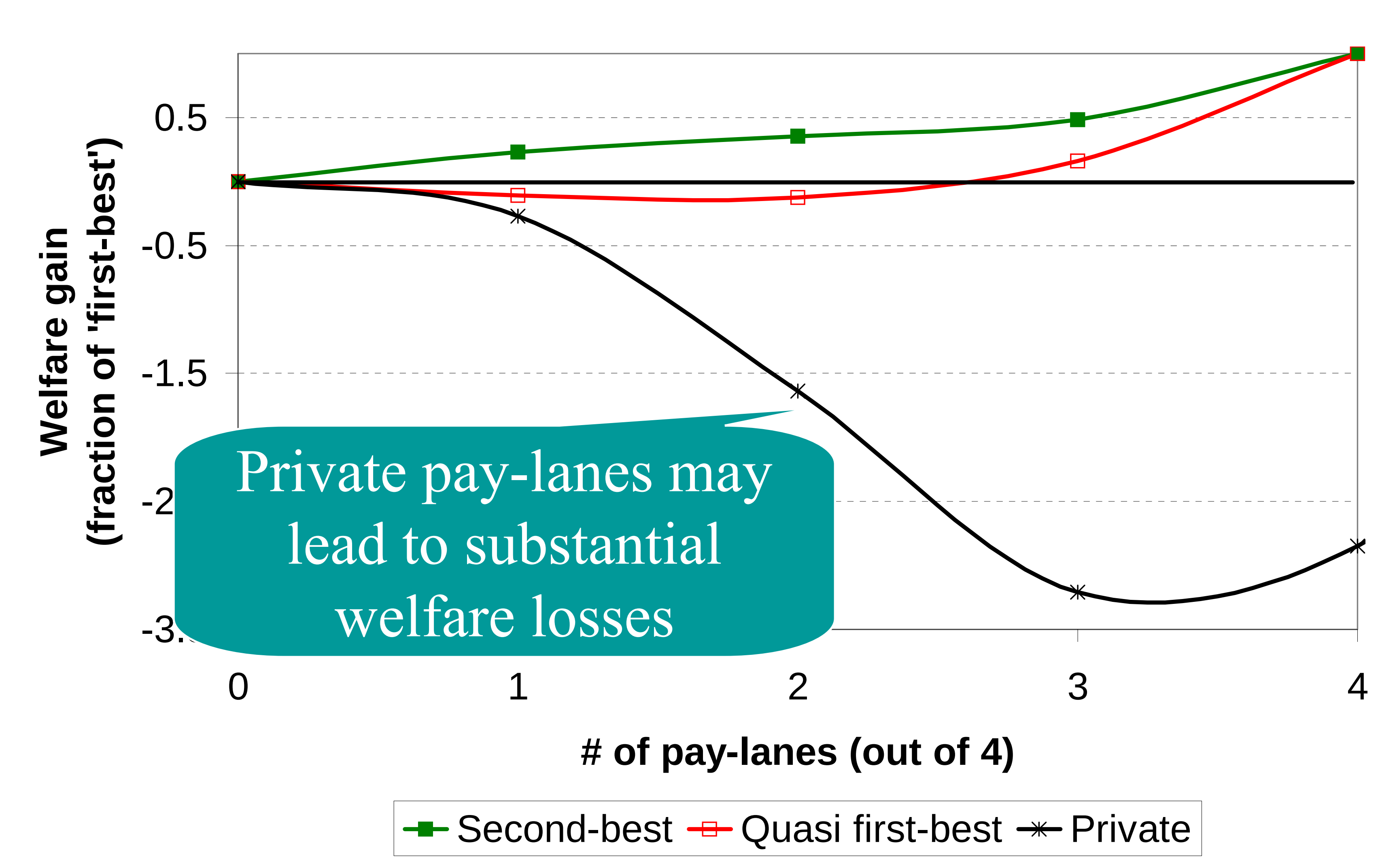
Relative efficiency



‘Quasi first-best’: $\tau = mec_T$



Private pay-lane



Generalization to larger networks?

- Pay-lane toll can be shown to be a special case of

$$f_j = \frac{\sum_{p=1}^{P} \delta_{jp} \cdot \frac{\sum_{m=1}^{J} \delta_{mp} \cdot \left(\sum_{q=1}^{P} \delta_{mq} \cdot N_q \cdot c'_m \right) - \sum_{q=1, q \neq p}^{P} \lambda_q \cdot \left(\sum_{m=1}^{J} \delta_{mp} \cdot \delta_{mq} \cdot c'_m \right) + \sum_{i=1}^{I} \sum_{q=1, q \neq p}^{P} \delta_{ip} \cdot \delta_{iq} \cdot \lambda_q \cdot D'_i - \sum_{m=1, m \neq j}^{J} \delta_{mp} \cdot \delta_m \cdot f_m}{\sum_{m=1}^{J} \delta_{mp} \cdot c'_m - \sum_{i=1}^{I} \delta_{ip} \cdot D'_i}}{\sum_{p=1}^{P} \frac{\delta_{jp}}{\sum_{m=1}^{J} \delta_{mp} \cdot c'_m - \sum_{i=1}^{I} \delta_{ip} \cdot D'_i}}$$

$$\forall\ j \text{ with } \delta_j = 1 \quad \text{and} \quad \forall\ p \text{ with } \delta_{ip} = 1 \quad \text{and} \quad \forall\ q \text{ with } \delta_{iq} = 1$$

- So: theoretically possible, but notationally cumbersome

One other example

- Distortions on labour market
  - Mayeres & Proost (2001), Parry & Bento (2001)
    - General equilibrium, endogenous labour supply
    - Distortive labour taxes
  - Conclusions:
    - Congestion charges may aggrevate distortions on labour market
    - Eventual welfare effects may depend strongly on use of revenues
    - Hence: not just a ‘tool to buy acceptance’

Main lessons from s.b. literature (1)

- Tax ‘rules’ become much more complicated than the simple “tax = m.e.c.” rule, to reflect indirect effects
- Regulator, in addition, needs more information to set prices optimally
- The risk of ‘government failures’ thus increases
- Potential efficiency gains of second-best pricing may be well below, or close to, those from first-best pricing, depending on the circumstances

Main lessons from s.b. literature (2)

- Naïve use of taxes - ignoring the second-best nature of the tax - will lead to even smaller efficiency gains; or even losses
- Second-best pricing lacks the property of giving optimal incentives for all behavioural dimensions
- In a second-best world, the use of tax revenues is not just an issue affecting the distributive effects of pricing, but also directly affects its efficiency

Alarming message?

- MC-based pricing in realistic second-best situations
  - risk of doing it ‘wrong’ is not insignificant
  - careful study of actual application and an identification of the relevant second-best aspects is necessary before implementing

To conclude

- MC pricing appears straightforward as a concept
- Intruiging / important aspects arise when looking at actual implementation
  - Acceptability
  - Dynamics
  - Second-best issues
  - … and more…
- Challenges for the design of pricing policies, as well as for further research

‘Acknowledegement’

- This presentation uses material from

- Verhoef, E.T. (2001) “An integrated dynamic model of road traffic congestion based on simple car-following theory: exploring hypercongestion” *Journal of Urban Economics* **49** 505-542.
- Verhoef, E.T. (2002) “Second-best congestion pricing in general static transportation networks with elastic demands” *Regional Science and Urban Economics* **32** 281-310.
- Verhoef, E.T. (2002) “Second-best congestion pricing in general networks: heuristic algorithms for finding second-best optimal toll levels and toll points” *Transportation Research* **36B** 707-729.
- Verhoef, E.T. (2003) “Inside the queue: hypercongestion and road pricing in a continuous time – continuous place model of traffic congestion” *Journal of Urban Economics* **54** 531-565.
- Verhoef, E.T. (2003) “Speed-flow relations and cost functions for congested traffic: theory and empirical analyses” Discussion paper TI 2003-064/3, Tinbergen Institute, Amsterdam-Rotterdam.
- Verhoef, E.T. and K.A. Small (2004) “Product differentiation on roads: second-best congestion pricing with heterogeneity under public and private ownership” *Journal of Transport Economics and Policy* **38** (1) 127-156.