

Travel futures: A pre-pay market mechanism for more effective demand responsive travel pricing

Nicholas Fournier, Ph.D.

Abstract

While transportation funding can be collected in a variety of direct (e.g., fares, tolls, and gas taxes) or indirect ways (e.g., property and sales tax), dynamic demand responsive pricing not only collects revenue but incentivizes travelers to avoid peak-demand periods, thus utilizing infrastructure capacity more efficiently. Unfortunately, the demand response to price changes, called price elasticity of demand, is generally greater for longer-term travel planning (e.g., air and rail travel) than for short-term (e.g., highway tolls and transit fares). While this is due to a plethora of factors (e.g., time flexibility, housing choice, automobile investment, etc.), a critical factor is that travelers simply lack sufficient information for future travel planning. For example, airline prices at different times can be easily compared, but a highway driver cannot accurately predict congestion pricing, thus causing price changes to have little effect on demand until after the fact. This leaves any congestion abatement up to inefficient trial and error, and anecdotal speculation. Moreover, dynamic pricing is a politically unsavory policy due to this price uncertainty and collateral equity concerns. This paper seeks to help remedy these concerns by proposing a simple “futures” market mechanism that can augment existing fare/toll collection technologies, providing travelers with sufficient pricing information and purchasing options to travel plan and avoid excessive prices. Users can pre-pay their future fares/tolls to lock in a lower price for expected trips, with prices rising as capacity fills up, thus encouraging good travel planning and efficient infrastructure utilization.

1 Introduction

Providing additional transportation capacity (e.g., more transit vehicles or road lanes) for only a very brief peak-demand period leaves infrastructure largely underutilized for the remaining time or simply cost prohibitive. The challenge is then how to more efficiently utilize existing capacity, rather than building oversized infrastructure solely for the peak periods. Tolls and increased fares are one tool used to raise revenue and mitigate congestion for decades, but conventional fixed prices fails to target the congested periods and merely apply a uniform downward disincentive on travel (see Figure 1). This is especially true when no travel alternative exists, such as with road tolls without a transit alternative. While it may reduce congestion by reducing overall travel demand, it does not necessarily encourage efficient use of infrastructure. Moreover, it is economically problematic by potentially dampening economic activity, as well as posing serious equity concerns to those who cannot afford costly travel.

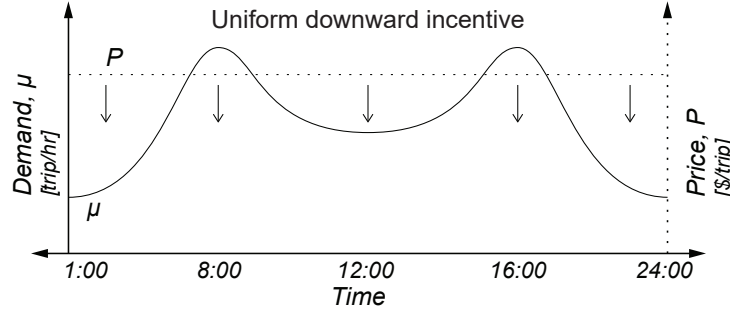


Figure 1: Fixed toll pricing incentive

The obvious solution to this problem is to dynamically adjust pricing based on demand to help discourage travel during targeted congested periods and shift it to periods of low demand (see Figure 2). Dynamic tolls have long been proposed as a solution to provide a more targeted congestion mitigating incentive, as has been done for decades in other transportation sectors (e.g., rail and air travel), but are far less common in more short-term high-frequency cases (e.g., highway tolls and transit fares). While dynamic pricing is less common for daily tolls and fares due to practical and logistical concerns, such as cost and complexity, it is largely a political barrier from public disdain for variable pricing. Understandably, travelers generally do not want to pay for travel, let alone be subject to unpredictable increased costs that they perceive as having little control over. Moreover, dynamic pricing creates equity concerns if it simply extorts travelers without any reasonable alternative travel options, such as highway tolls without a transit alternative.

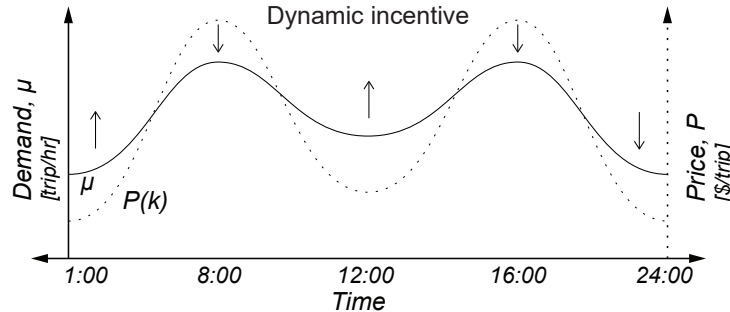


Figure 2: Dynamic toll pricing incentive

Fundamentally, dynamic pricing hinges on the notion that a change in price will result in some change in demand. To economists, this is called price elasticity of demand, where elasticity reflects the relative demand sensitivity to price change. This means that the aforementioned price uncertainty is a critical aspect

to dynamic pricing. For example, if the travel cost is uncertain, travelers will have a diminished ability to react, thus diminishing price elasticity of demand and having little impact on congestion. To explain using utility theory, where a decision making “agent” will choose the choice which provides the greatest “utility” ultimately assumes that the traveler has sufficient information about price and return (e.g., travel time) to make a decision about if and when to travel.

In practice, travelers may base their travel decisions off of anecdotal evidence of past experiences, speculating future traffic conditions and price in the case of dynamic pricing. For example, a daily commuter may have local knowledge of typical travel conditions and routing, enabling a reasonably informed decision after enough experience. However, this is hardly an efficient transmission of knowledge and only works in relatively steady-state conditions. Moreover, this would yield a delayed demand response to price changes, resulting in a very inelastic dynamic pricing system. Meaning, dynamic pricing would have little short-term effect on demand and that long-term demand change would only result after enough travelers were sufficiently disappointed by their previous choices. This then connects with the political aspect and public disdain for dynamic pricing schemes. Moreover, it is often an unfortunate truth that those most economically vulnerable often have the least travel flexibility (e.g., low income workers with strict working hours), and are often unfairly punished for this.

This paper seeks to help remedy these critical flaws of dynamic pricing through a pre-pay mechanism similar to a futures market. The proposed concept is simple, travelers are presented with the current price for future travel with the option to purchase trips ahead of time at that price. As more trips are booked by travelers, and demand approaches capacity, the price rises. This not only incentivizes travelers to avoid the peak congested periods, but offers an opportunity for regular commuters to minimize the risk of price-gouging from unexpected congestion by purchasing tolls in advance and compare prices. The proposed mechanism helps overcome some of the shortcomings of dynamic pricing, potentially yielding greater congestion mitigation benefits of dynamic pricing by effectively increasing price elasticity of demand.

2 Background

More recently, there has been some discussion surrounding the concept of travel credits, exploiting the aspect of finite roadway capacity as the basis for a “cap and trade” type model for mitigating congestion by crediting travelers for not traveling during the peak [1]. Another more market-based evolution of the approach has also been proposed, enabling travelers to “sell” their trip rights to the highest bidder, achieving a similar goal without the need for governments to credit travelers [2–5]. There has also been a great deal of hype around the concept of Mobility as a Service (MaaS) or Transportation as a Service (TaaS) where “travel brokers”, acting almost as an automated travel agency, provide integrated multi-modal trip packages to travelers. To date, few of these concepts have had much traction outside of academia in recent years, but some technology companies have implemented conceptually similar trip routing and payment integration features in their platforms (e.g., mobile payment systems with transit and ride hailing services), laying the groundwork for MaaS/TaaS concepts in the future.

The proposed futures mechanism is not necessarily a radical new concept, nor is it a technologically driven disruption, but merely an attempt to implement efficient congestion reducing market mechanisms by augmenting existing conventional revenue collection technologies (e.g., automated tolls and electronic fares). The goal of this proposed scheme is to enhance pricing information transmission and purchasing options for travelers, which can increase price elasticity of demand and improve congestion mitigation.

3 Conceptual framework

The basic concept for the system is shown in Figure 3, where travel demand (e.g., on transit, over a bridge, along a corridor, or in a zone) can be divided into discrete travel windows (e.g., between 8:00-8:15) where there is a finite quantity of capacity available (e.g., a road capacity of 2,000 vehicles per hour per lane). Travelers then have the option to pre-purchase a planned trip from the available capacity, with prices increasing as booked trips in that time-slot increase. This encourages travelers to avoid traveling during those peak periods as the price increases towards capacity, or to purchase trips in advance while the price

remains low or discounted. Otherwise, travelers that do not pre-purchase their trip are subject to the current “market rate” price determined by dynamic congestion pricing.

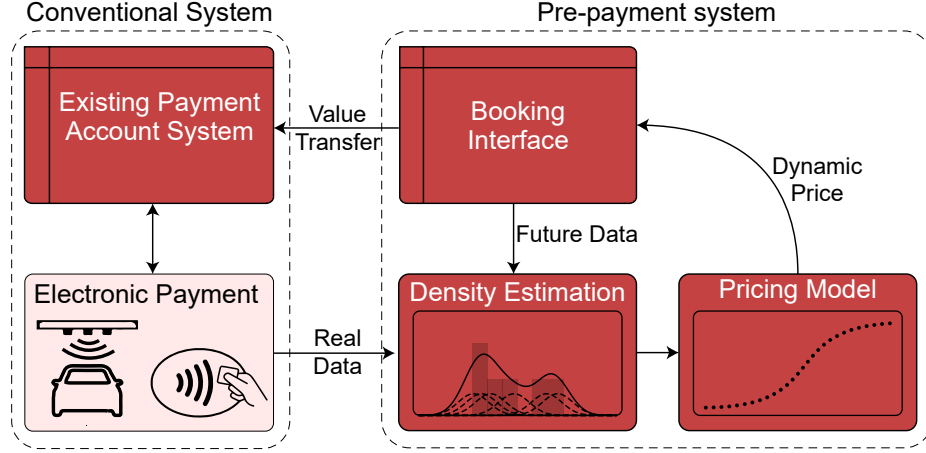


Figure 3: Conceptual system model

The overall process of booking, density estimation, and then pricing is continually updated until the horizon of the future trip time or some cut-off period is reached. At this point the sale of future trips in the “futures market” is closed to booking and any travel must be paid at the conventional real-time market price based on the actual observed traffic density. It is possible, and likely, that there will be a difference between the futures price and the real time market price. This difference is not a problem, and is a motivating force behind real futures markets; for buyers to try beat the market price, and for sellers to create stability of demand in exchange for a discount by locking buyers into a contract. In this case with transportation, as opposed to commodities like oil and grain, there are two possible outcomes for price differences:

Futures rate < Actual rate. This occurs when too few people booked their trips, having limited congestion mitigating effects until sufficient market penetration is reached. However, the booked travelers would yield a large discount against the conventional toll, thus incentivizing more travelers to adopt the booking system, eventually correcting this difference and increasing congestion mitigation effects.

Futures rate > Actual rate. Although highly unlikely, but still possible, this occurs when too many travelers failed to meet their target window or did not show up at all. The travelers would incur a loss on their booked trips, thus incentivizing travelers to be punctual or to sell unused trips back to the system, making that capacity available to others.

A more detailed description for each of the proposed mechanism’s components are contained in the following subsections.

3.1 Booking interface

A basic booking interface could be a website or app where users can pre-pay for their trip on a specified time slot for a given date period (e.g., hourly time slots for a given date). The same concept exists in ticketed transportation modes, such as airlines and rail travel. The difference in this case compared to conventionally ticketed modes (e.g., planes and trains) is that traditional ticketing is impractical due to the frequency and flexibility required for local travel. For these reasons it is essential that booking be made very convenient, such as buying transit or toll passes. For convenience, it would be wise to provide the option to buy trips in bulk for commuters (e.g., 9am-10am and 5pm-6pm time slots every Monday-Friday for the next month). The system can then be easily integrated with existing toll collection and payment systems where the pre-purchased trips are exchanged as a toll/fare waiver.

3.2 Demand for time and space

The overall objective is to increase prices when too many travelers try to occupy the same time and space in a transportation system. While such a system could be based on continuous time slots, for practical reasons it is useful to break time down into arbitrarily small quantum units of time, such as 1, 5, or 15-minute intervals, for every time increment until some rolling future time horizon (e.g., from now until one year). The increment depends on the desired resolution of the system and any computational limitations associated. Each time unit then becomes “tokenized” for purchase, with price and quantity of tokens being dependent upon some target capacity (e.g., road or transit capacity), see Figure 4.

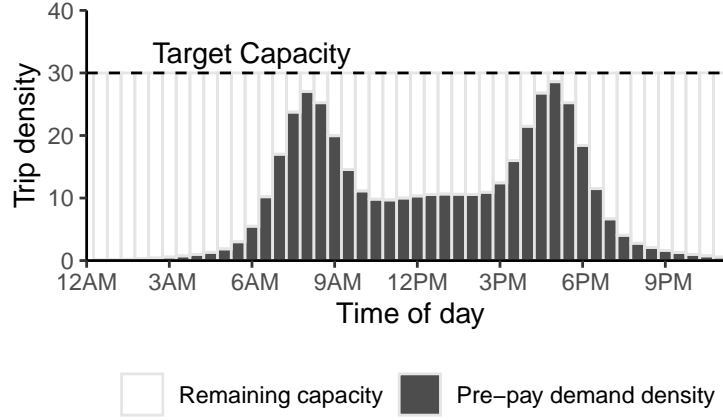


Figure 4: Capacity across descretized time

When travelers book their desired time window, they effectively purchase a sequential bundle of tokens. For example, a time window from 8:00am to 9:00am could have 60 tokens of 1 minute, or 12 tokens of 5 minutes, depending on the time increment size. The price of that window then depends upon the sum of token prices across that range, with individual token prices depending on the number of travelers that have purchased the same overlapping token. Depending on system design, the total cost of a typical time window (e.g., 30 minutes) should be calibrated to have a nominal total cost comparable to a typical toll/fare cost.

There are two benefits to having users purchase small tokenized units of time, rather than fixed large chunks of time (e.g., where users can only buy between 8am-9am, not 8:15am-8:45am). First is that it offers users greater flexibility in deciding when and how narrow their travel window will be. Second it further encourages users to be punctual and to shift their travel times away from congested periods. For example, 8:00am to 9:00am are congested and thus expensive, but users can still achieve a discount by shifting from 7:30am to 8:30am, which only overlaps halfway with the congested times.

3.3 Trip density

To estimate trip density (i.e., congestion) from discrete time increments requires a conversion to continuous density values. Furthermore, while flexible travel time window can be specified with very small time increments in order to approximate continuous values (e.g., $8:12 \pm 27$ mins), in practice it is likely that users will simply pick rounded numbers anyway (e.g. 8:00 to 8:30). In any case, this results in unrealistically concentrated expected trip densities that will eventually differ as individual traveler arrival times vary. Converting from discrete booked trips and into expected travel density can be achieved by smoothing the finite trips using Kernel Density Estimation (KDE).

$$\hat{k}(t) = \frac{1}{n} K_h(t - t_i) = \frac{1}{nh} \sum_{i=1}^n K \frac{t - t_i}{h} \quad (1)$$

where $\hat{k}(t)$ is the estimated density function for time t , K is the non-negative kernel function (e.g., Gaussian normal curve), and h is a smoothing bandwidth parameter that must be greater than 0. KDE essentially

functions by cumulatively applying a continuous density function K at each finite sample point, providing a smoothed probability distribution, such as shown in Figure 5.

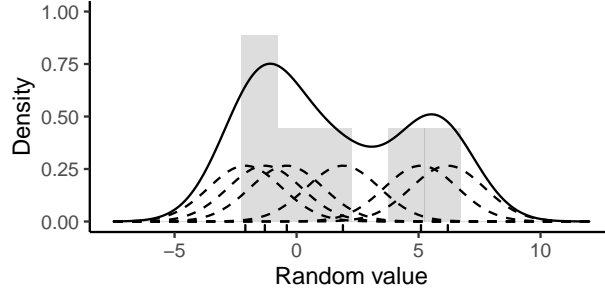


Figure 5: Kernel Density Estimation using Gaussian function

3.4 Pricing

Once a smooth and stable value of future traffic density $\hat{k}(t)$ is calculated in Equation (1), it can be easily inputted into a pricing function to calculate an estimated future price at that travel window. Although pricing could be determined by some free floating “free market” calibrated to match the limited “supply” of capacity, such an uncontrollable system may be unstable. Prices could fluctuate wildly due to extreme congestion scenarios or speculative trip booking.

Alternatively, an engineered market price could be determined using a parametric function designed to achieve the best possible congestion mitigating and revenue collecting results while remaining stable. A simple pricing function shown in Figure 6 uses an S-shaped sigmoid as a function of trip density k (e.g., traffic density in $veh/km/hr$).

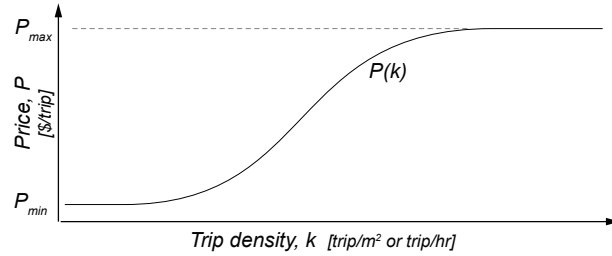


Figure 6: Pricing sigmoid function

Unlike a monotonic function that increases cost to infinity, a sigmoid allows for price minimum P_{min} and maximum P_{max} boundaries to be set, while remaining a smooth continuous function. A sigmoid price function could be expressed as

$$P(k) = P_{min} + \frac{P_{max} - P_{min}}{1 + e^{\alpha - \beta k}} \quad (2)$$

where α and β are tuning parameters to specify the horizontal shift and “steepness” of the function, respectively. A potentially useful feature of using a bounded sigmoid is that it can be presented to the public as a discount feature for booked and off-peak travelers, not a surcharge. For example, booked off-peak trips are discounted from the typical fixed toll/fare price. This may help make it politically more palatable. Of course, there is no technical difference and the upper boundary will need to be set high enough to collect sufficient net revenue.

While the objective in certain cases (i.e., roadways) is to maintain smooth traffic flow, a dynamic pricing function based directly on traffic speed would fail to capture the underlying demand. Traffic speed is not a reliable measure of demand as speed will remain relatively constant across a range of demand, and only

begins to break down when traffic density reaches a critical point. This makes it difficult to set prices for any speed below this critical point. However, given that the goal is to increase infrastructure utilization and efficiency, not simply maintain high speeds, basing pricing on density is a more robust and universal measure. Furthermore, density not only offers a direct input from the trip density estimation, but can be applied to broader range of applications where speed is less variable but density still affects operational efficiency (e.g., transit).

3.5 Implementation

A potential feature of the futures market concept is that it can be developed independently, simply augmenting an existing system rather than as a complete replacement. An existing dynamic toll system can be easily augmented with a futures market to further improve demand optimization with only software changes necessary. A fixed toll system can use a futures market to introduce dynamic pricing as an “opt in” program, incentivizing users with discounts for booking travel during periods of low demand. Infrastructure investment would be minimal, especially if electronic tolling is already in place. A futures market in this case would be a soft way to introduce dynamic pricing in an otherwise politically hostile environment. In this case the futures market may result in revenue loss if the fixed toll, which is used as P_{max} , is not sufficiently high enough. The loss could be considered acceptable if improving operational performance is priority, but it is also possible that with improved operations and reduced pricing overall demand could increase, offsetting the loss. However, this is difficult to speculate.

Although this paper presents the mechanism for a simple bridge toll, it can be easily expanded to other applications, such as corridor tolls, congestion pricing zones, transit fares, ride hailing, or applications beyond transportation. This could provide a basic framework for a broader, more mode agnostic, transportation system which is of growing importance as ride sharing, hailing, and other future mobility platforms emerge. However, when creating a new market, care must be taken to ensure operation is smooth and stable, but also equitable. The following subsections discuss several basic market safeguards and practical constraints.

3.5.1 Market manipulators and exploitation

If unused trips are able to be “sold” back to the system, there is the potential for profiteering. This would only cause undesirable price fluctuations in the system that does not correspond to physical infrastructure capacity and demand. Thus to prevent users from inappropriately exploiting the system, restrictions should be set. For example, to prevent users from manipulating the market price, a practical restriction could limit users to only purchasing one trip per time slot (otherwise it would be a travel impossibility), or to limit the number of sales a user can make to ensure a net profit cannot be made. Ensuring zero profit would also help avoid any user tax complexities. Although restrictions may seem market prohibitive, the intent of the system to promote efficient travel and revenue collection, not market capitalization.

3.5.2 Pricing calibration

To avoid reaching outrageous prices and the public backlash that follows [cite Virginia example and uber surge pricing], reasonable upper limits on price should be set to avoid public backlash. Similarly, a lower price limit could also be set to provide a baseline revenue collection and travel disincentive. The pricing function will need to be calibrated to match demand sensitivity as well as meet revenue obligations, while remaining politically palatable. Ideally the upper bound will be kept as low as possible to avoid public dissatisfaction, but high enough to achieve the desired shift in travel behavior. The lower bound must also be kept low enough to encourage a shift in travel behavior, but high enough to collect sufficient revenue. Research will be needed to measure user arrival time demand elasticity, that is how sensitive are travelers to increases in price. From this, the function can be optimized to balance operations with revenue, and will likely need regular tuning to adjust for current conditions.

3.5.3 Time scale

Although theoretically possible to operate with narrow time intervals, such as <1-second, the computational limitations of accounting for that many tokens and transferring every second is likely intractable. Moreover,

travelers probably will not bother with such detail nor be held to such precision. It's likely that users will pick a target time \pm some buffer time (e.g., ± 30 minutes from expected arrival). However, it is uncertain whether travelers will be equally as likely to arrive early as late, which can be modeled with a Gaussian normal curve; or perhaps travelers trend on being late but not early, which can be modeled with a log-normal curve. Research is needed in this area to determine feasible parameters and the relative reliability of travelers. Alternatively, the system could be designed in such a way to allow for users to specify their desired travel window, offering greater discount to more punctual travelers and a penalty to those that fail to meet their target.

4 A simulation example

As an example application, suppose there are two adjacent urban centers in a metropolitan region as shown in Figure 7a, which are separated by a body of water. The two cities are connected by a bridge that carries 100,000 trips per day with an existing fixed toll. The daily travel demand possess a demand distribution with two severe peaks as shown in Figure 7b. For simplicity, assume the traffic flow is balanced in each direction and the bridge has three lanes in each direction.

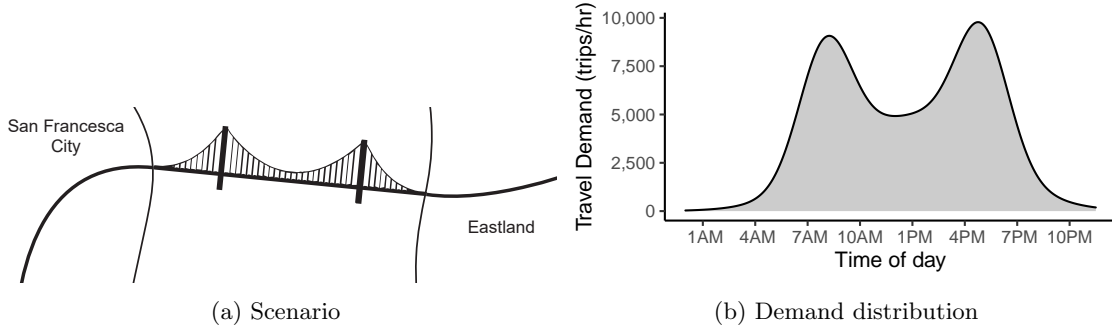


Figure 7: Toy simulation scenario

Suppose that the regional metropolitan transportation planning commission wishes alleviate congestion in the peaks with a dynamic tolling system. A study determines that the price elasticity of demand for a conventional dynamic tolling system without a pre-pay futures market is $\epsilon_0 = 0.1$, and that a futures market mechanism would boost elasticity to somewhere between $\epsilon_{low} = 0.3$ on the low end, and $\epsilon_{high} = 0.7$ on the high end. They determined that voters are willing to approve a dynamic toll system that is limited to between a 50% discount for off-peak travel, but no more than a 200% surcharge above existing fixed prices in the peak (See Table 1).

Table 1: Desired pricing parameters

Parameter	Value	Description
P_{min}	50%	Minimum dynamic price ratio (P_{min}/P_{fixed})
P_{max}	200%	Maximum dynamic price ratio (P_{max}/P_{fixed})
α	7	Price calibration parameter
β	0.3	Price calibration parameter

Calibration parameter values of 7 and 0.3 are set for α and β , respectively, which provides a smooth S-shaped sigmoid function with the upper and lower limits being approximately oriented on the traffic densities of 0 and 30 veh/km , as shown in Figure 8. A pricing function in this case should align the upper price limit to desired congestion values to incentivize demand to stay below those levels.

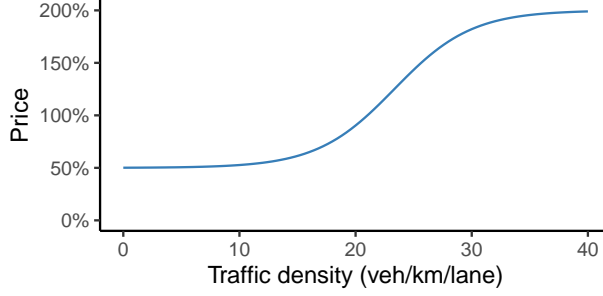


Figure 8: Pricing sigmoid function

4.1 Price elasticity of demand

To simulate the resulting demand shift from pricing, the simple micro-economic principle of price elasticity of demand can be utilized. Elasticity, ϵ , reflects the demand sensitivity to price, that is the percent change in demand resulting from a percent change in price. Recall that elasticities of $\epsilon < 1$ are considered “inelastic” where there is a small change in demand from a change in price, $\epsilon > 1$ are considered “elastic” where there is a greater change in demand from a change in price, and $\epsilon = 1$ are considered “unit elastic” where the change is equal. Price elasticity of demand in transportation tends to be fairly inelastic, typically in the range of 0.10 - 0.20 for shorter-term price changes and 0.20 - 0.80 for longer-term price changes, in absolute values, where short-term is considered price changes in a period of less than two years [6].

Although this temporal variation in transportation elasticity is due to a variety of complex socio-economic and behavioral factors, a simple explanation is that it takes time for people to adjust their behavior depending on conditions in their life. For example, housing choice and automobile ownership is a long-term decision. If fuel prices rise, they cannot easily move homes or purchase a more fuel efficient car. However, the objective of the proposed travel pricing futures market is to increase elasticity by giving travelers some of the advantages typically associated with longer-term travel behavior. That is, to be able to compare prices well in advanced, rather than speculating future traffic conditions based on anecdotal evidence. For this simulation, a constant isoelastic price-demand function was used, which can be expressed as

$$\Delta\mu = e^{-\epsilon \cdot \Delta P} - 1 \quad (3)$$

where $\Delta\mu$ is percent change in demand and ΔP is percent change in price, with the function centered on the origin, as shown in Figure 9. Since an improvement in price elasticity with a futures market is unknown, several different elasticities are explored in the following simulation. A range of elasticity points are used: a very low elasticity of $\epsilon_0 = 0.1$ to represent a conventional dynamic toll with no futures market mechanism, and then three increasing elasticities of $\epsilon_{low} = 0.3$, $\epsilon_{medium} = 0.5$, and $\epsilon_{high} = 0.7$ are used to represent dynamic pricing with a futures market mechanism.

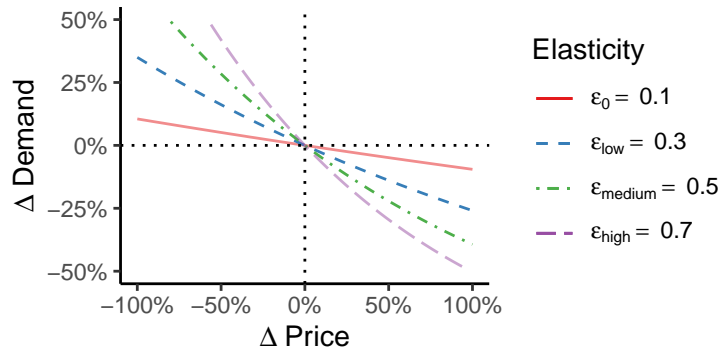


Figure 9: Price elasticity

4.2 Traffic flow

To account for congestion impacts on operation, the traffic flow across the bridge can be characterized by the macroscopic fundamental diagram using the following modified piece-wise flow-density function

$$q(k) = \begin{cases} k < k_c : & v_f k \left(1 - \frac{k}{2k_c}\right) \\ k \geq k_c : & \frac{k_c v_f}{2} \left(1 - \frac{(k-k_c)^2}{(k_j-k_c)^2}\right) \end{cases} \quad (4)$$

where k_c is the critical density when traffic flow is optimal, k_j is the jam density when traffic flow completely stops, and v_f is the free-flow traffic speed. For reference, an optimal traffic flow is often $\approx 20-40 \text{ veh/km}$, with traffic flow beginning to become unstable above this point until reaching the jam density at $\approx 100-150 \text{ veh/km}$. The modified piece-wise function is composed of two different parabolic functions in order to provide the more realistic asymmetric form often observed with a longer right-tail, as shown in Figure 10a.

Although a variety of more refined macroscopic models have been developed, many models are often very sophisticated requiring additional calibration parameters, or possess abrupt transitions that can create anomalies in analytical results. Two common classical models that require no additional calibration parameters are Greenshields' (1935) parabolic function and Daganzo's (1997) simple bi-linear model. Greenshields' seminal function is elegantly simple, but symmetric parabolic shape has since been proven a poor fit in reality, particularly when critical density, k_c , is exceeded. Daganzo's model in contrast provides a very simple parsimonious model with an asymmetric form, but its linearity assumes a constant free-flow speed until it abruptly transitions at critical density. The constant free-flow speed not only ignores minor delay caused by gradual slowing of traffic as density increases, but causes an abrupt transition that can cause erroneous results when results suddenly jump. The bi-parabolic function proposed is both parsimonious, requiring no additional calibration parameters, and contains no abrupt transitions. It is not the most advanced model, but satisfies the needs for this simple simulation.

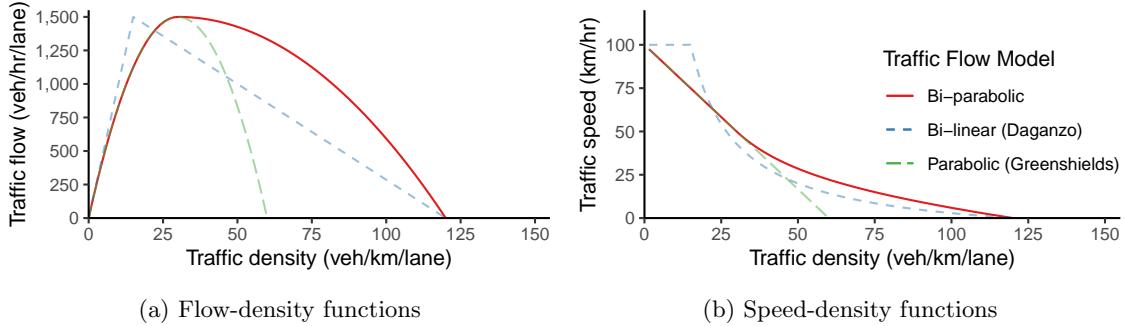


Figure 10: Macroscopic Fundamental Diagrams

The first piece when $k < k_c$ provides a steep curve that peaks at the critical density k_c , and the other case when $k \geq k_c$ sets the right-tail of the elongated parabola at jam density k_j . The critical and jam densities, k_c and k_j , are set to 30 veh/km , and 120 veh/km , respectively. This asymmetric shape also yields a more gradual decline in speed that is more similar to the bi-linear model (see Figure 10b, but without the abrupt speed transition).

4.3 Simulation results

Using the parameters in Table 1, the macroscopic traffic conditions can be simulated to compare results between a dynamic pricing scheme against a fixed toll. The price changes causes demand to shift from the peak to the off-peak, resulting in a decrease in congestion and an improvement in average travel speed as shown in Figure 11. Although there is a travel speed improvement in all cases, the magnitude of travel speed improvement depends upon the elasticity. The greater the elasticity, the greater the improvement.

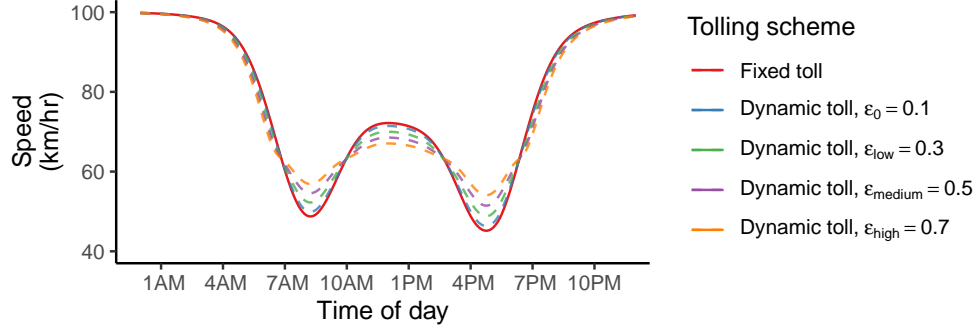


Figure 11: Average traffic speed by time of day

The elasticity will not only determine how sensitive travelers are to price changes, but also effects the aggregated total revenue collected. A revenue increase from a dynamic pricing scheme compared to a fixed price is not guaranteed and depends upon the parameters selected in the pricing function. Ultimately, the revenue collected from a dynamic pricing scheme primarily depends on the elasticity, ϵ , and the upper and lower limits of the pricing function, P_{max} and P_{min} . The net revenue resulting from the example actually yielded a slight increase in revenue in this case, as shown in Figure 12, despite offering a discounted P_{min} toll.

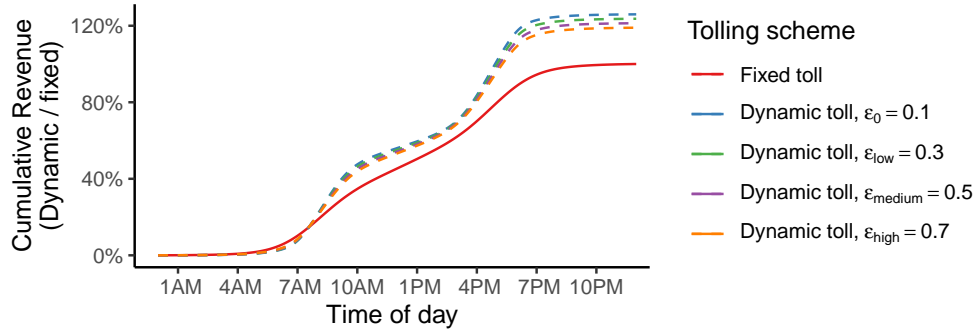


Figure 12: Cumulative revenue collection by time of day

This shows that a modest adjustment in price results in demand shift (i.e., commuters shifting behavior to avoid high tolls) without revenue loss. In fact, in this case it yields a revenue increase of 126% for the conventional dynamic pricing. However, as elasticity increases, it yields decreasing revenue of 124%, 121%, and 119% for low, medium, and high elasticity futures market cases, respectively. The reason for this is that if travelers are more sensitive to price changes, more will shift to cheaper off-peak times and the increased volume in the cheaper times does not outweigh the revenue loss in the peak. This highlights how critical the upper and lower price limits P_{max} and P_{min} are to both the congestion and revenue outcomes.

4.4 Optimal price limits

To explore the price and elasticity parameters effects, a matrix of simulation results can be computed and plotted. Figure 13 shows the percent change in revenue achievable by varying the combination of P_{min} and P_{max} simultaneously with a fixed elasticity of $\epsilon = 0.3$ used to form the surfaces.

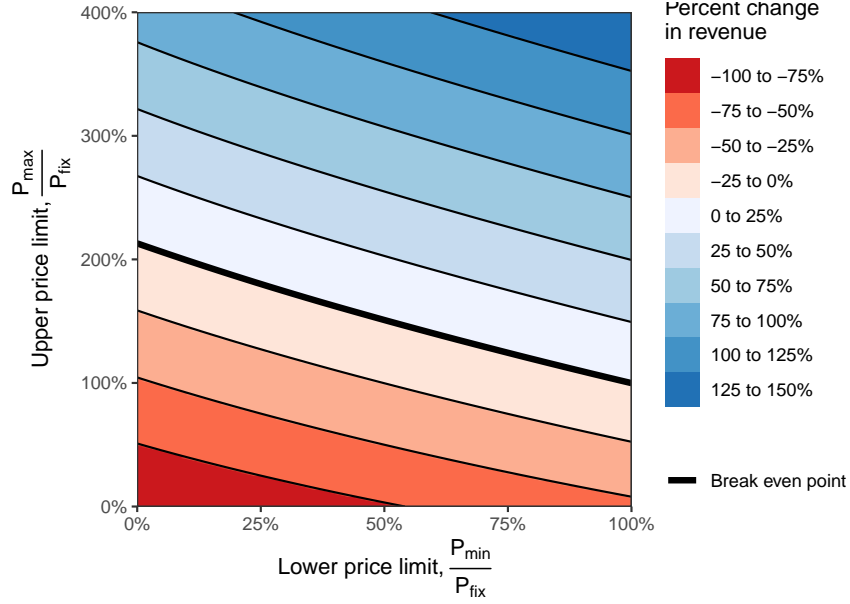


Figure 13: Revenue comparison between dynamic and fixed pricing varying P_{\max} and P_{\min} with constant elasticity of $\epsilon = 0.3$

Similar to revenue, operational performance can be evaluated in terms of delay. The total delay in each scheme can be compared as the percent change in aggregated total delay, calculated as

$$\Delta\% d = \frac{\sum_i d_{2i} - \sum_i d_{1i}}{\sum_i d_{1i}} = \frac{\sum_i \frac{n_{2i}}{v_{2i}} (v_f - v_{2i})}{\sum_i \frac{n_{1i}}{v_{1i}} (v_f - v_{1i})} \quad (5)$$

where delay is $d = n_i L (1/v_i - 1/v_f)$ for each time increment i . L is the distance traveled, which cancels out in percent change; n_{1i} and n_{2i} are the demand in number of trips, and v_{1i} and v_{2i} are the calculated traffic speeds for the fixed and dynamic tolls, respectively. Equation (5) effectively sums up the total delay experienced across the entire 24-hour period, and calculates the percentage change between the fixed and dynamic tolling schemes.

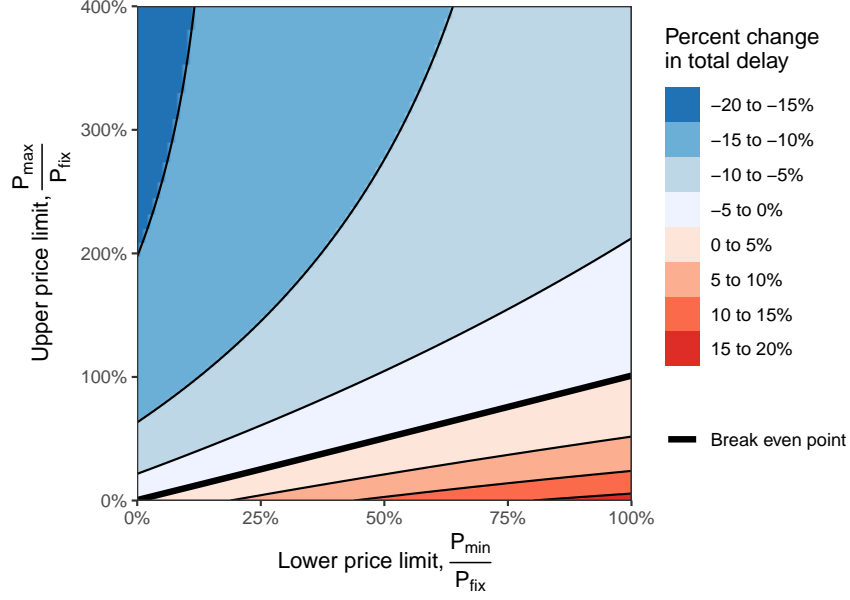


Figure 14: Delay comparison between dynamic and fixed pricing varying P_{max} and P_{min} with constant elasticity of $\epsilon = 0.3$

328 The dark solid line in Figures 13 and 14 are the break even point where the upper and lower price limits
 329 balance out, resulting in 0% change in the objective of either revenue or delay. Extracting the break even
 330 lines and combining the two, the general regions of comparative gains can be illustrated in Figure 15. It is
 331 apparent is that congestion can be mitigated as long the upper price limit P_{max} is greater than the lower
 332 price limit P_{min} . This makes intuitive sense, any amount of dynamic pricing where $P_{max} > P_{min}$ will yield
 333 delay improvement, but may come at the cost of revenue loss.

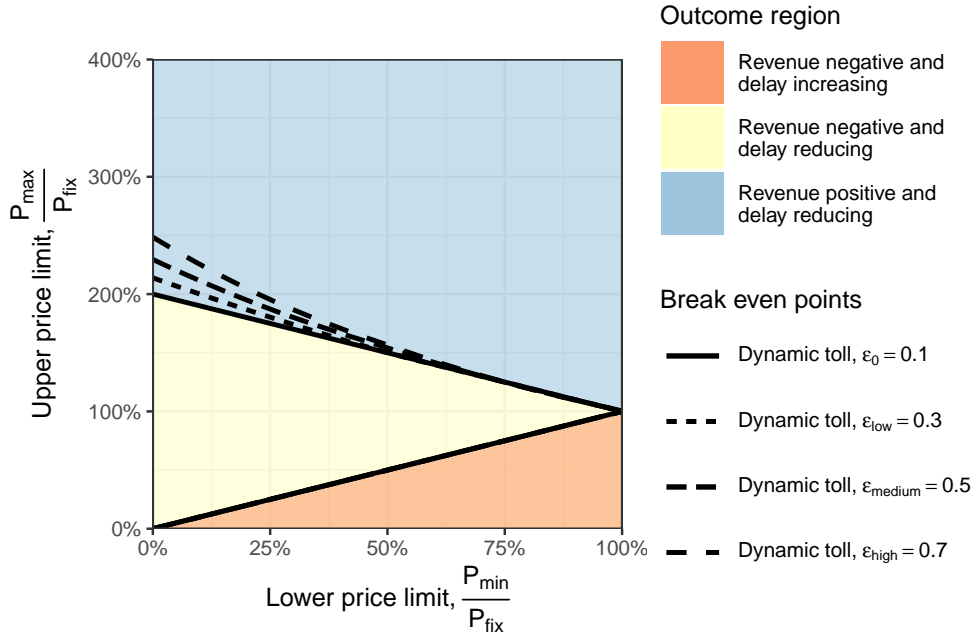


Figure 15: Break even points between dynamic and fixed pricing with constant elasticity of $\epsilon = 0.3$

To avoid revenue loss, the upper pricing limit must be approximately double the amount lost from P_{min} compared to the fixed toll, or $P_{max} \approx P_{fixed}(2 - \frac{P_{min}}{P_{fixed}})$. However, this varies depending on the elasticity, which causes the revenue break even line to curve slightly upward, increasing the upper price limit P_{max} . In contrast to the revenue break even point, the delay break even point appears unchanged by elasticity. This is merely because the break even point shifts proportionally with elasticity from this particular perspective shown in Figure 15, but in reality elasticity does affect demand. The following analysis will explore how elasticity affects both revenue and delay in more detail.

4.5 Optimal elasticity

Exploring how elasticity effects the revenue and delay outcomes, a set of plots can be similarly created by varying the maximum price and elasticity, P_{max} and ϵ . For simplicity, a lower price limit P_{min} of \$0 was chosen as the extreme floor for dynamic pricing. While it is technically possible to go below \$0, effectively offering a subsidy for off-peak times, for simplicity this theoretical scenario is not explored. As reflected in the break even region plot, Figure 16 shows that as elasticity increases, an exponentially higher P_{max} is required to maintain revenue levels.

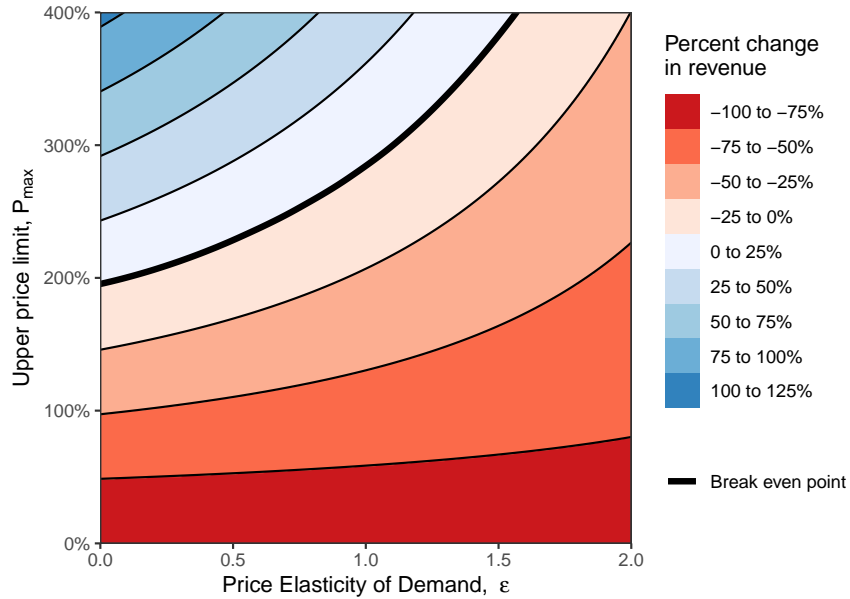


Figure 16: Revenue comparison between dynamic and fixed pricing varying elasticity ϵ and P_{max} with constant $P_{min} = 0$

Unlike with revenue, delay as elasticity varies provides a more interesting outcome in Figure 17. The maximum performance benefits appear to be saddled around an elasticity of 1.0 and a maximum price above 100%. The saddle shaped region is an interesting outcome, showing that there is diminishing returns as elasticity moves away from unit elasticity. In the most extreme case, where $\epsilon > 1$ and $P_{max} > 100\%$ travelers shift demand such that overall traffic congestion can worsen. However, this is a very unrealistic scenario since transportation demand is generally very inelastic, especially with short-term price changes. In practice it is likely that the feasible range exists where $\epsilon < 1$ and delay is improving.

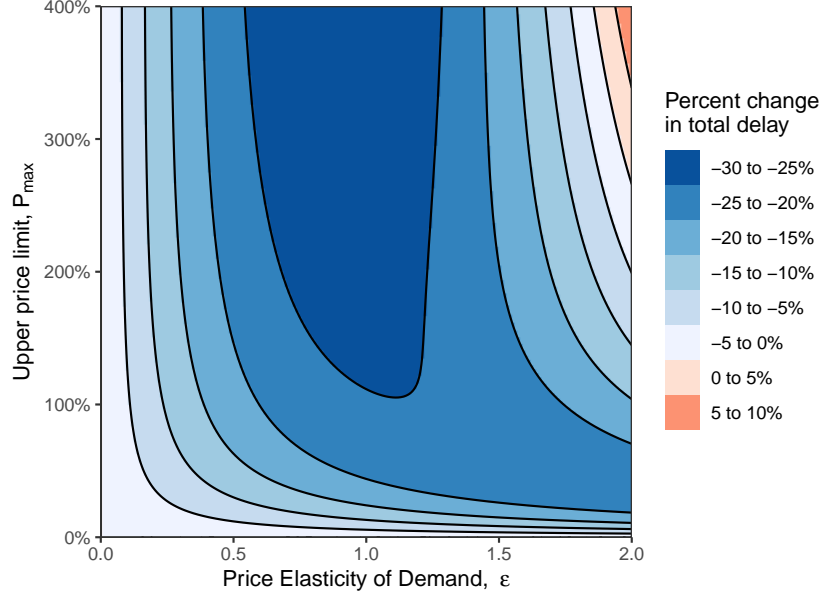


Figure 17: Delay comparison between dynamic and fixed pricing varying elasticity ϵ and P_{max} with constant $P_{min} = 0$

5 Conclusions

This toy example yielded interesting results, showing strong potential improvements in both revenue and performance, but it is important to point out that this is a very simplified case. It assumes a closed system of trips, meaning that people merely adjust their trip schedule and do not change modes or choose to not travel entirely. Moreover, elasticity may also vary depending on a variety of factors, such as on time, purpose, and individual preferences and flexibility. This is an important reality to consider, particularly in regards to transportation equity when considering increasing transportation costs for individuals.

An important element not considered is an overall demand suppression from pricing. While dynamic pricing is intended to more efficiently utilize infrastructure capacity, it is possible that an existing fixed toll had suppressed demand overall. Replacing a fixed toll with a dynamic toll may actually cause an increase in overall trips due to the newly available peak-hour capacity and the reduced priced off-peak trips. This could cause both positive and negative outcomes in regards to revenue and delay, as well as other endogenous outcomes, such as emissions and fuel consumption.

While simple micro-economic elasticities may often be too theoretical for practical use, they do offer general insight for approximate optimization. It is clear from this simulation that price elasticity of demand is critical to both congestion and revenue. A greater elasticity means travelers will more easily change their trip time, having a greater effect on delay, but it also means that more travelers will choose to travel at lower priced times, thus decreasing total revenue. Intuitively, this means that high inelasticity is good from a revenue perspective, making it easy to price-gouge travelers unwilling to change their behavior, but does little to mitigate congestion, and vice versa. In practice, simple price elasticity of demand may be too crude a value for day-to-day optimization. A more refined approach might be to calibrate the parameters with a bounded bi-criterion objective that seeks to optimize for both revenue and/or congestion in a bounded region, such as in Figure 15 but also accounting for variable elasticity. Further, the pricing function itself could be modified from the symmetric sigmoid function to some custom optimized function using empirical data (e.g., use artificial intelligence to set prices to maximize the objective). The booking apparatus lends itself to such an optimization, providing data of future demand to help best predict optimal prices; a feature unavailable in current dynamic pricing systems. Moreover, integration and revenue sharing policies with other modes, such as public transit, could yield wider benefits.

References

- [1] Adriaan Perrels. User response and equity considerations regarding emission cap-and-trade schemes for travel. *Energy Efficiency*, 3(2):149–165, 2010. ISSN 1570646X. doi: 10.1007/s12053-009-9067-5.
- [2] Takashi Akamatsu and Kentaro Wada. Tradable network permits: A new scheme for the most efficient use of network capacity. *Transportation Research Part C: Emerging Technologies*, 79:178–195, jun 2017. ISSN 0968090X. doi: 10.1016/j.trc.2017.03.009. URL <http://dx.doi.org/10.1016/j.trc.2017.03.009> <https://linkinghub.elsevier.com/retrieve/pii/S0968090X17300839>.
- [3] Wenbo Fan and Xinguo Jiang. Tradable mobility permits in roadway capacity allocation: Review and appraisal. *Transport Policy*, 30:132–142, nov 2013. ISSN 0967070X. doi: 10.1016/j.tranpol.2013.09.002. URL <http://dx.doi.org/10.1016/j.tranpol.2013.09.002> <https://linkinghub.elsevier.com/retrieve/pii/S0967070X13001352>.
- [4] Xiaolei Wang, Hai Yang, Daoli Zhu, and Changmin Li. Tradable travel credits for congestion management with heterogeneous users. *Transportation Research Part E: Logistics and Transportation Review*, 48(2):426–437, mar 2012. ISSN 13665545. doi: 10.1016/j.tre.2011.10.007. URL <http://dx.doi.org/10.1016/j.tre.2011.10.007> <https://linkinghub.elsevier.com/retrieve/pii/S1366554511001323>.
- [5] Hai Yang and Xiaolei Wang. Managing network mobility with tradable credits. *Transportation Research Part B: Methodological*, 45(3):580–594, mar 2011. ISSN 01912615. doi: 10.1016/j.trb.2010.10.002. URL <http://dx.doi.org/10.1016/j.trb.2010.10.002> <https://linkinghub.elsevier.com/retrieve/pii/S0191261510001220>.
- [6] Todd Litman. Understanding transport demands and elasticities. *Victoria Transport Policy Institute*, pages 1–76, 2019.
- [7] BD Greenshields, JR Bibbins, WS Channing, and HH Miller. A study of traffic capacity. In *Highway research board proceedings*, volume 1935. National Research Council (USA), Highway Research Board, 1935.
- [8] Carlos Daganzo. *Fundamentals of transportation and traffic operations*, volume 30. Pergamon Oxford, 1997. ISBN 9780080427850.