Policy Discussion:
Peak-load pricing

In class we talked about the role of “peak-load pricing” in determining the optimal provision of a commodity whose demand varies by time of day, month, or year. Typical examples include roads and public transit systems (for which the demand is highest at rush hours), electricity (demand is highest in early evening during winter time), but also lineups at the bank, etc.

Since demand varies over time then so does the cost of supplying the commodity – with costs being greatest when demand is highest, at which time the system is subject to problems of congestion/queueing. Therefore prices should vary over time as well. Moreover, since optimal prices will be highest at peak-load times, when capacity may be insufficient to satisfy demand, it is possible to see the pricing system as a way of requiring demanders at peak times to pay for the additional units of capacity that must be provided to satisfy demands.

In summary, peak-load pricing is both a way to ration demand at times when capacity is insufficient to satisfy everyone, and it is also a form of benefit taxation that charges a higher price to those who create the need for additional capacity costs. The following worked example illustrates these ideas.

The capacity of a subway system is measured by the maximum number of passengers per direction per hour (ppdph) it can carry. Suppose that the capacity of the TTC subway system is fixed at $K$ ppdph. The capital cost of the subway is $cK$, and operating costs are zero. The TTC has estimated that the inverse demand curve for total trips $T$ on the subway at rush hour is given by the inverse demand function

$$w_r(T) = A_r - T$$

whereas at non-rush hour times it is given by

$$w_n(T) = A_n - T$$

Since operating costs are zero, our previous discussion suggests that social surplus is maximized by setting the price of a ride (the “fare”) to zero. But suppose at first that capacity $K$ has been chosen so that $A_r > K > A_n$: if the fare were zero, then there is excess demand for rides at rush, but excess
supply of rides outside rush hours. In this case, it is optimal to set the fare to zero outside rush hours – \( f_n = 0 \) – but the fare at rush hour should be increased until demand is equal to supply. (Why use fares instead of just making people wait their turn, by the way?) This occurs when customers are on their inverse demand curve, i.e.

\[
    f_r = w_r(K) = A_r - K
\]

This is illustrated in Figure 1: up to \( T = K \), the marginal cost of a ride is zero, so \( f_n = 0 \) is optimal, since \( A_n < K \). But it is not possible to satisfy all demand at rush hour, since \( A_r > K \), which means that the marginal cost curve become vertical at \( T = K \), as shown by the dashed line. If a “peak-load fare” of \( f_r = w_r(K) = A - K \) is chosen, then supply equals demand at rush

Figure 1: Peak load pricing and optimal capacity.
hour.

This analysis takes the capacity as given at some level $K$. But $K$ could be increased above the level shown in the figure (by building more subway cars); what is the optimal level $K^*$? The cost to society of serving one more consumer at rush hour is the marginal capacity cost $c$ (since we know that the subway is operating at capacity during rush hour). The marginal social benefit of one more unit of capacity is the marginal willingness to pay for a ride at rush hour, which is $w_r(K)$. It is therefore optimal to increase capacity to the point $K^*$ at which

$$c = w_r(K^*) = A_r - K^*$$

or $K^* = A_r - c$ as shown in the Figure. At this point, the optimal peak-load fare is

$$f_r(K^*) = w_r(K^*) = c$$

therefore the revenues of the transit system

$$R = f_rT_r + f_nT_n = cK + 0 = cK$$

are exactly equal to the total costs of building the system. (Recall that operating cost is zero.) Therefore the optimal peak-load pricing scheme is also a budget balancing pricing system in this case.