Reconciling the Off -Net Cost Pricing Principle with Efficient Network Utilization

Patrick DeGraba
Federal Trade Commission
600 Pennsylvania Ave.
Washington, DC
pdegraba@ftc.gov

The off-net-cost pricing principle argues that under a broad range of environments a positive "access" charge paid by originating networks to interconnected terminating networks would cause networks to set on-net usage charges equal to off-net rates, and that these charges would fully reflect the access charge. However, other results in the literature provide reasonable conditions under which on-net usage charges will not reflect access charges, but would be set to induce the social surplus maximizing level of on-net usage. This paper harmonizes these two apparently opposing results by showing that retail usage charges depend on two effects. One is a rent seeking effect on the part of networks and the other is an efficient utilization effect. In models in which the rent seeking effect is more important, on-net usage charges will tend to equal their off-net usage charges and incorporate the access charge. In models in which the efficient utilization effect matters more, off-net usage charges will reflect access charges, but on-net usage charges will not be affected by the level of access charges, but

1. Introduction

Recent deregulation of telecommunications has eliminated many monopoly franchises held by incumbent local phone companies, creating competition between multiple networks in the same local calling area. Federa on-net costs of providing service, while the rates for off-net service depend both on the costs of providing the service and the access fee.

The purpose of this paper is to reconcile these two seemingly opposing results. I show that these different results depend on different assumptions regarding how usage rates affect consumer usage, and about whether subscribers to a telephone network both initiate and receive calls, or whether one set of customers only originates calls while a separate set only receives calls. a teli]TJarrates paysntsmcesy e calls, 8(er)-9(s e pluc (leonsis)sn bs fee.)pitiay a tele turns out to be remarkably robust to generalizations of the model: mixed traffic, variable demand…" [5]

The intuition behind this result is that the resource costs of serving the originator of an off-net call is the same as the resource costs of terminating an off-net call. However, the network that terminates a call receives a payment, while the network that originates the call makes a payment. Thus, an interconnection regime that involves access charges endows a customer that terminates more traffic than he originates with a rent. Bertrand competition to serve such customers allows these customers to capture all of the rents, and this results in on-net usage rates equal to their off-net counterparts.

To see why this can happen, suppose initially there is a single network serving all customers, customer usage is unaffected by the per minute retail rates they are charged, all marginal costs are zero, and the access rate for off-net calls is set at 1 penny per minute. Suppose there were a customer that only received phone calls. Clearly a competing network (with zero marginal costs) would be willing to pay this customer up to 1 penny per minute for each off-net minute he terminated to subscribe to the competing network. Since the network serving all other customers would have to pay 1 penny per minute if this customer left, it too would be willing, and would have to pay him 1 penny per minute for each on-net minute he terminated to keep him on the network. Thus, competition in this example causes off-net termination and on-net termination to be priced at the same rate. (A similar argument explains why on-net origination rates must equal off-net origination rates.)

In contrast to the LMRT result, research by DeGraba (2003)[6] presents a model in which

---

[5] See LMRT pg 2. The intuition behind this claim is that in a zero profit equilibrium, competition will cause each operator to set the price of usage for each customer equal to the o2uo servat customer's.
[6] See also DeGraba (2002) and (2000) and Hermalin and Katz (2002).

This paper shows these conflicting results are the consequence of different assumptions regarding two different features of customer behavior. The first is that LMRT assume that per minute usage rates do not affect customers' usage, while DeGraba assumes that per minute usage rates do affect customers' usage. The second is that LMRT look primarily at a set of customers in which some customers only originate traffic while others only terminate traffic. DeGraba on the other hand looks at a set of customers in which all customers have a (sufficiently high) positive probability of originating each call.

These two different sets of assumptions suggest that there are two competing effects that determine how access rates affect retail usage rates in equilibrium. One is a rent seeking effect on the part of network operators. The other is an efficient utilization effect.

The rent seeking effect says that when originating networks pay access rates to terminating networks in excess of those necessary to efficiently allocate the cost of a call between interconnected networks,[9] customers that receive more traffic than they originate generate rents, because they generate (in expectation[10]) positive net access rates for the network to which they subscribe. Competition to serve such customers results in rates that enable the customer to capture these rents.

The efficient utilization effect says that networks compete for customers by setting on-net usage rates that offer customers the highest surplus. DeGraba (2003) shows this requires setting usage rates so that parties to a call bear its cost in proportion to the marginal benefit they receive

[9] DeGraba (2003) shows the efficient access rate is the one that allocates the cost of a call between two interconnected networks so that the percentage of the cost borne by each network is equal to the percentage of the marginal benefit received by its customer.

[10] A customer generates positive net access rates for its network if he receives more off-net minutes than he originates. In symmetric models in which all customers speak to all other customers, a customer that terminates more minutes than he originates will typically generate such rents in equilibrium.

from the call. Network operators have an incentive to set on-net rates to efficiently allocate the cost of a call between the two parties regardless of the access rate.

The intuition behind why ONCPP rates aren't always sustainable in equilibrium can be seen in the following simple example of competition between two interconnected telephone networks.[11] Suppose there are two identical networks that have a zero marginal costs of originating and terminating calls, that two customers always share equally in the benefit of each minute of calling in which they are engaged, that the benefit of a minute of conversation between two customers is a decreasing function of the number of minutes consumed between them, and that each customer is equally likely to initiate a minute of conversation as he is to be called. Suppose also the government imposes an access charge of a, which the originating network must pay the terminating network for any inter-network call.[12]

The ONCPP implies that on-net and off-net origination rates will both equal a, and that on-net and off-net termination rates will both equal $-a$.[13] The reason these retail rates will not always be supported in an equilibrium is the following:

In the proposed ONCPP equilibrium, each network earns a zero profit. Because an originating customer is charged a usage fee of a, he consumes minutes up to the point at which his marginal benefit of the last minute equals a. Thus, a usage rate of a that exceeds the efficient rate creates for its customers, resulting in positive prof

Efficient utilization would occur if all origination and all termination rates were equal to $c/2$. Yet a network has the incentive to offer off-net termination rates below $c/2$, on-net origination rates equal to infinity, serve only the customers that terminate calls and earn a positive profit. This would exceed the zero profit that would occur under Bertrand competition.

This paper shows that the structure of on-net usage rates depends on whether customer demand implies that rent seek or assumptions regarding the probabilities with which customers originate traffic. In section 3 I present an example in which some customers only receive calls, some only originate calls and others are equally likely to receive as they are to initiate calls, and determine what rates would prevail in equilibrium.

2. Equilibrium rates and customer demand characteristics

In this section I present a simple example of the benchmark model presented by LMRT. I then consider four extensions to illustrate the interaction between the rent seeking and the efficient utilization effects. The two key assumptions are whether or not usage rates affect customers' usage level, and whether each customer both originates and receives calls or whether a customer either only originates or only receives calls. Each of the extensions corresponds to one of the combinations of these assumptions.

I analyze the model in the context of voice telephony, instead of as internet usage as LMRT did, because some extensions require all customers to initiate calls. The institutional details are not especially important. The important elements of the analysis are that i) Two customers jointly consume a good (in this example called phone conversation) and share the benefits of this good; ii) To consume this good each customer must purchase an input from one firm and the cost of the input is the same regardless of whether both customers purchase from the same firm or they purchase from different firms; and iii) For each unit of consumption one customer can be identified as the "originator" and the other can be identified as the "terminator" so that origination may be priced differently from termination and access charges can be appropriately assessed.

customer is the expected benefit he receives from the minutes of calling consumed less expected payments made to or received from his network. The payoff for each network is the expected profit it earns, which includes the expected revenue received from customers less expected net access payments/receipts. I consider only subgame perfect Nash equilibria.

2.a.1. Example 1. $v(1) = v(2) = 4$, $a = 3$, *A* always originates the call.

In this example each customer receives a benefit of 4 from each of the 2 minutes consumed. Customer *A* always initiates the call, and *B* receives it. This structure is equivalent to the benchmark case in LMRT, where *A* is the website and *B* is the customer.[18] The access rate is set at 3.

Under these assumptions the ONCPP rates can be supported in equilibrium. Each network sets both origination rates at 3 and both termination rates at -3. At these rates, customers subscribes to each

$-(1.5 + \varepsilon)$ for $0 > \varepsilon > 1.5$. The magnitude of on-net termination is irrelevant. At these rates the only subgame perfect continuation is for *B* to subscribe to network 1 and for *A* demand curve for each customer. Observation shows how this one change causes ONCPP rates to be unsupportable in equilibrium.

Observation 1. The ONCPP rates cannot be sustained in equilibrium under the assumptions, $v(1) = 4$, $v(2) = 2$, $a = 3$, and *A* initiates the call with probability 1.

Proof: The ONCPP says all origination rates equal 3 and all termination rates are -3. Customers subscribe to each network with probability .5. At these rates only one minute is consumed, so the surplus for *A* is 4-3 = 1, and the surplus for *B* is 4 + 3 = 7.

Suppose WLOG network 1 deviates by setting its on-net origination rate equal to $.5 + 2\varepsilon$, its on-net termination rate to $-(.5 + \varepsilon)$, its off-net origination to $3 - \varepsilon$ and its off-net termination rate at -3, for $0 < \varepsilon < 1$. The only subgame perfect response is for both customers to subscribe to network 1. It is a dominant strategy for *A* to subscribe to network 1 under these rates, and *B*'s unique best response to *A* subscribing to network 1 under these rates is to also subscribe to 1.

When both customers subscribe to network 1, *A* earns a payoff of $6 - 2(.5 + 2\varepsilon) = 5 - 4\varepsilon$. *B* earns a payoff of $6 + 2(.5 + \varepsilon) = 7 + 2\varepsilon$, and network 1 earns a payoff of 2ε, which is greater than 0. Thus, there is a deviation which invalidates the proposed ONCPP equilibrium. *QED*

Observation 2. There is an equilibrium in which both networks set on-net usage rates equal to 0 and both off-net usage rates greater than 4 when $v(1) = 4$, $v(2) = 2$, $a = 3$, and A always initiates the call.

Proof: Suppose customers adopt the following deci

Lemma 2. There is no profitable deviation by network 1 in which exactly one customer subscribes to each network.

Proof. If customers are on different networks there is no consumption, because both off-net usage rates on network 2 are above 4. Therefore both customers would earn a zero payoff. □

Lemma 3. There is no profitable deviation for network 2.

Proof. Any deviation by network 2 that would allow it to earn a positive profit would require the two customers to jointly earn less than 12. Decision rule i) implies they would both subscribe to network 1 in that case. □ *QED.*

While this equilibrium captures the notion that efficient network usage plays a key role in the level of equilibrium rates, it is somewhat unsatisfying. It involves each network setting both off-net origination and off-net termination rates that eliminate all inter-network calling if customers were to subscribe to different networks. In essence the rent seeking behavior in this example is defeated by the networks effectively disconnecting from each other, rather than setting on-net usage rates equal to their off-net counterparts.

Decision rules i) and ii) are essentially correlated strategies. The equilibrium behavior for the customers is to subscribe to the same network even though they are indifferent between the networks. Note that (as in the well known "battle of the sexes" game) they do not want to subscribe to each network with probability .5, because that creates a positive probability they will *ex post* subscribe to different networks and earn 0 surplus.

To see why off-net origination rates must be greater than 4, suppose each network sets off-net origination equal to $4 - \varepsilon$. Such a pricing strategy could be broken by network 1 simply setting off-net termination to – (3 – ε), and raising its on-net origination to 4. In this case *B*

Observation 4. If off-net termination rates are constraine originates 1 minute, then expected joint surplus is only .5[(4-3) + (4+3)] = 4 , so again no mutually beneficial deviation is possible. Last, if *A* originates 2 minutes, then the joint surplus is .5[(6-6) + .5(3+4) = 3.5. □

Lemma 6. There is no profitable deviation by network 1 in which a) if both customers subscribe to network 2 then one customer would be better off by deviating to network 1 and b) given that one customer subscribes to network 1, the other is better off by subscribing to network 1 and c) network 1 earns a positive payoff.

Proof. If *A* subscribes to network 1, then he can receive at most a surplus of 7 when *B* originates

The importance of Observation 4 is to show that when customers originate and terminate traffic with the same probability, there are fewer rents to seek than when one customer always receives calls. In this example, because the rents are sufficiently small, (in fact given equilibrium rates the expected rents are 0) the deadweight loss created by using usage rates to give customers the rents he generates exceeds the value of the rents, and thus do not defeat an equilibrium in which on-net usage is priced efficiently.[19]

The reason for restricting off-net termination rates to be no less than -3 is illustrated in the following example. Suppose given the proposed equilibrium, network 1 deviated by setting an off-net termination rate equal to -10, an off-net origination rate greater than 4 an on-net origination rates equal to ε and an on-net termination rate of 0. If both customers subscribe to network 2 they would each receive a payoff of 6. WLOG *A* could subscribe to network 1, and with probability .5 terminate one minute of calling from *B*, thereby receiving an expected payoff of .5(4 + 10) = 7, which is greater than the 6 he would receive by staying on 2. *B*'s best response to this would be to subscribe to network 1 as well, and earn a payoff of $6 - .5\varepsilon$, as opposed to the .5(4 – 3) = .5 he could earn by remaining on network 2.

The problem with this example is that the proposed equilibrium is broken because network 1 can offer an off-net termination rate that would never be profitable to honor, but because it induces a tipping equilibrium, it will never have to be honored. The restriction prevents network 1 from offering such non-profitable rates.[20] I maintain this assumption in the remainder of section 2.

---

[19] So for example a network could deviate by setting off-net termination equal to $-(3 - \varepsilon)$, and all other usage rates greater than 4, turning any customer into a net terminator of traffic. However, any customer that subscribed under these conditions would earn an expected surplus of only .5(4+3) = 3.5.

[20] A more detailed justification might be that while we are modeling static equilibria that occur instantaneously, a richer model would account for the fact that a network would tend to attract customers over time. So a network that attempted to use a strategy in which it paid customers off-net termination rates in excess of the access charges would have to absorb the losses implied by such payments for a period of time, and such losses (when discounted for time) could exceed the eventual profits of the eventual tipping equilibrium

Lemma 6 points out an interesting complication that is introduced when customers have downward sloping demand curves. Since each customer chooses the quantity of minutes to consume, and that quantity depends on usage rates, each customer's surplus depends not only on the usage rates he faces, but on the rates the other customer's face as well. Thus, each customer cares about the network to which the other customer subscribes. So each customer's subscription choice must not only be optimal with respect to the rates charged by networks, but optimal with respect to the subscription choices of the other customer as well. Hence the need for decision rules i) and ii).

Assume that customers adopt decision rules i) and ii) from the proof of observation 4. Under the proposed rates each customer earns a surplus of 8, each network earns a profit of 0, and the allocation is Pareto efficient.

Lemma 7. There is no deviation by network 1 in which both customers subscribe to network 1, jointly earn a surplus of at least 16, and network 1 earns a positive profit.

Proof. If both customers subscribe to network 1, network 2 earns a zero profit. Since the market can generate at most a surplus of 16, if the two customers earn a joint surplus of at least 16, network 1's payoff must be non-positive. □

Lemma 8. There is no deviation by network 1 in which exactly one customer subscribes to network 1,[21] and network 1 earns a positive payoff.

Proof. Suppose *A* subscribes to network 1. If off-net rates on network 1 are such that a call is always completed, then *B* (on network 2) receives an expected payoff of 8 and network 2 earns a profit of 0. *A* (on network 1) has the option of subscribing to network 2 and earning a surplus of 8. Thus, he must earn at least a payoff of 8 by subscribing to network 2. If rates are such that a

Lemma 9

subscribe to network 1. Given that both residential customers are subscribing to network 1, *I*'s best response is to subscribe to network 2, where he earn 2(6+6) = 24.

Under this allocation of customers, network 1 earns a positive payoff of 6ε by deviating. Thus, the ONCPP rates do not constitute an equilibrium. *QED*

This example suggests there may be a broad set of conditions under which ONCPP rates cannot be sustained in equilibrium. The proposed ONCPP rates impose an inefficiency between residential customers that are equally likely to call each other. In this case it seems a network could always deviate from the ONCPP rates by offering on-net rates that eliminated this inefficiency for residential customers while setting off-net rates that left (essentially) unchanged the relationship between residential customers and dial up ISPs.

Observation 7. There exists an equilibrium in which each network sets on-net usage rates to 0, and off-net origination and termin

b) if all three customers subscribe to network 2, no customer could increase his payoff by unilaterally subscribing to network 1.

It is clear that these two rules, if played by all three customers constitute best responses to each other for all rates set by the networks. Under these rules and the proposed equilibrium rates, all customers would subscribe to network 2. Each network would earn a payoff of 0 and each customer would earn a payoff of 2(4 + 2) = 12.

Because network 2 sets all off-net usage rates greater than 4, any customer that unilaterally subscribes to network 1 will complete 0 minutes of calling and therefore earn a payoff of 0.

There is no deviation that network 1 can make that will both yield a positive payoff to itself and a joint payoff to the three customers of at least 36 (since the proposed equilibrium yields a Pareto efficient allocation).

Similarly any deviation by network 2 that would yield a positive payoff would have to result in the three customers earning a joint payoff less than 36 from subscribing to network 2 and therefore (under i')) subscribing to network 1. *QED*

Observation 7 is simply the extension of observation 2, demonstrating again the potential for a tipping equilibrium, and suffering from the same problem as the equilibrium in observation 2 that pricing effectively eliminates inter-network calls. To address this problem I expand on the restriction made in the previous section that a network cannot make a payment to a customer for off-net usage in excess of revenue the network receives from other sources for providing the off-net usage. Thus, a network cannot set its off-net termination rate less than -3 and can not set its off-net origination rates less than 0. Under these restrictions there is a more palatable equilibrium, in that networks do not set prices that effectively preclude internetwork calls. However the equilibrium is still a tipping equilibrium.

Observation 8 There exists an equilibrium in which each network sets its off-net origination rate equal to 3, its off-net origination rate equal to -3, its on-net origination rate equal to .5, and its on-net termination rate equal to -.5.

Proof: Assume customers adopt the decision rules i') and ii') outlined in the proof of observation 7. Under the proposed equilibrium rates all customers subscribe to network 2. *I* receives a payoff of 2(6 + 1) = 14, and *A* and *B* he unilaterally subscribed to 1. But ii’) along with lemma 10 implies *I* would need to earn more than 14 to unilaterally deviate. □

Lemma 12. Any deviation by network 1 in which *A* has a unilateral incentive to subscribe to network 1 (if all three customers are subscribing to network 2) must involve network 1 setting an off-net origination rate less than .5.

Proof. With an off-net origination rate equal to .5 the highest surplus *A* could earn by unilaterally subscribing to network 1 is (6 – 1) + .5(4 + 3) + .5(6 – 1) = 11, which is the same as the 11 he could earn by subscribin

Proof. With its off-net termination rate equal to -3 and its off-net origination rate equal to .5, network 1's maximum payoff would occur if it set both on-net usage rates equal to 2. With only *A* and *B* subscribing to network 1 under these rates, each customer would earn a surplus of (6-1) + .5(6-4) + .5(6-4) = 7, which is less than the (6-1) + .5(4-3) + .5(6+6) = 11.5 *B* could earn if he subscribed to network 2 instead of network 1. If only *A* and *B* are subscribing to network 1 at these rates its payoff would be 2(1 – 6) + 2(2+2) = -2. Thus, network 1 would always earn a negative payoff from such a deviation. □

Lemma 15. There is no deviation in which *I* and exactly one residentia originate 0 minutes to *I*. Thus, there is no deviation in which network 1 earns a profit and *I* and exactly one residential customer subscribe to network 1. □

Lemma 16. There is no deviation in which, if all three customers subscribe to network 1, none has a unilateral incentive to subscribe to network 2, and network 1 earns a positive profit.

Proof. With both *A* and *B* on network 1, (and off-net origination rate on network 1 less than .5) if *I* remained on network 2, he would earn a surplus of 2(6 + 6) = 24. With all three customers on network 1, assuming each residential customer originated 2 minutes worth of calling to *I*, network 1 would have to set its on-net termination rate equal to -3 for *I* to receive a surplus of 24. To earn a positive payoff network 1 would have to set on-net origination greater than 3. But, this would cause *A* and *B* to originate at most only 1 minute each to *I*. If *A* and *B* each originated 1 minute to *I*, then *I* would need to receive 16 in on-net termination payments. In this case network 1 would have to set its on-net origination rate equal to 8, which would cause the residential customers to originate 0 minutes. Thus, there are no rates that would prevent I from unilaterally subscribing to network 2 under which network 1 could earn a positive payoff. □

QED

Observation 8 provides several interesting insights. First, *I* receives the same surplus he would receive under ONCPP rates. This observation reflects what I believe to be the robust result of the LMRT analysis, which is that the imposition of an inefficiently high termination charge endows net terminators of traffic with a rent. When there is competition among interconnected homogenous networks, the net terminator of traffic will capture all of the rents from access charge, regardless of the network to which he subscribes. (Note that in this case he

terminate more minutes than they originate are endowed by the interconnection regime with a positive rent, and those that originate more minutes are endowed with a negative rent.

I have shown that when it is assumed that customers' usage levels are not affected by usage rates, then competition among non-differentiated networks results in usage rates that allow net terminators of traffic to capture all of the rents they generate, and force net originators to pay for the negative rents they generate. In this case the ONCPP prediction is born out, so on-net rates reflect off-net access charges.

However, when it is assumed that customers' usage levels are affected by usage rates, setting usage rates to compete for rents creates inefficiently low on-net consumption. When

References

DeGraba, Patrick, 2000, "Central Office Bill and Keep as an Efficient Interconnection Regime," Federal Communications Commission, OPP Working Paper Number 33.

DeGraba, Patrick, 2002, "Central Office Bill and Keep as a Unified Inter-Carrier Compensation Regime," *Yale Journal on Regulation*, 19 (1) pp. 39-84.

DeGraba, Patrick, (2003), "Efficient Interconnection Rates for Interconnected Competing Networks," *Journal of Economics and Management Strategy* Vol 12. no. 2 pp

Hermalin, B. and M. Katz , "Network Interconnection with Two-Sided User Benefits," July 2001 mimeo, University of California at Berkeley.

Laffont, Jean-Jacques, Scott Marcus, Patrick Rey, Jean Tirole, (2001) "Internet Interconnection and the Off-Net-Cost Pricing Principle," *American Economic Review* May. Vol 91 No.2 pp. 287-291.

Laffont, Jean-Jacques, Scott Marcus, Patrick Rey, Jean Tirole, (2002) "Internet Interconnection and the Off-Net-Cost Pricing Principle," Working Paper.