Discussion Paper

September 2002, issue 02/03

dot·econ

Economics for a networked world

Caller-called party interaction: Implications for call termination

Dan Maldoom

17 Welbeck Street
London W1G 9XJ
www.dotecon.com

Caller-called party interaction: Implications for call termination

Dan Maldoom*

September 2002

Abstract

Models of telecommunications demand generally assume that call externalities between the caller and called party are internalised. Often, this situation arises as a result of regular, repeated communication. In such repeated bilateral calling relationships, calls in either direction may serve the purpose of communicating news. Therefore, the optimal calling pattern in one direction may depend on the calls received in the opposite direction and vice-versa.

As a result of these interactions, traffic flows in one direction in a bilateral calling relationship may be affected not just by call prices in that direction, but also by the price of calling in the opposite direction. Where the parties in a bilateral relationship can internalise externalities, they may be more willing to plan regular calls to reduce the total costs of calling regardless of direction, making traffic balance more sensitive to the relative costs of calling in each direction.

This situation gives rise to a novel form of substitution, in which ingoing and outgoing calls within regular calling relationships can behave like substitutes, even though they are purchased by distinct consumers.

Further, where this call direction substitution occurs, subscribers have an indirect interest in the cost of being called. Raising call termination rates can lead to increased outbound calling in order to replace reduced inbound calls; this may make some subscribers worse off. If there is competition 'in utility space', operators need to take this effect into account when determining the optimal balance in charging for different services.

Keywords: telephony demand; substitution; call termination.
JEL classification: D11, D49, C72, L19.

*DotEcon Ltd, 105-106 New Bond Street, London W1S 1DN, phone +44-20-7870 3800, fax +44-20-7870 3811, email dan.maldoom@dotecon.com.

1 Introduction

There is currently significant regulatory scrutiny of call termination charges, especially mobile call termination charges, in a number of jurisdictions. The concern is that when the calling party pays (CPP), subscribers take little account of the interests of those that call them. This can create an incentive for operators to raise call termination charges in order to compete for subscribers. Perversely, tough retail competition might lead to call termination rates being too high as a result of the CPP externality between subscribers and those that call them. This effect has been described by a number of authors including Armstrong (1998) and is the key argument advanced in Oftel (2001) justifying regulation in the UK.

Where subscribers and those that call them belong to a common economic unit, such as a household or a company, the CPP externality is internalised. In such cases the chooser of the network should take account of the interests of those calling in, removing the incentive of the operator to raise call termination rates above socially optimal levels. However, whilst such closed used groups provide a competitive constraint on call termination rates, there is a question about its materiality. First, not all subscribers belong to such closed user groups. Second, there is at least a theoretical concern that operators might use on-net tariffs as a means of price discriminating between subscribers within and outside closed user groups.

However, it is not necessarily the case that subscribers and those that call them need to have an explicit common economic interest for call termination rates to affect subscription decisions. A large proportion of voice traffic occurs within the context of repeated bilateral calling relationships, where two parties are in regular communication. In such cases a call in one direction may reduce the need for a call in the other direction. We call this phenomenon *call direction substitution*. Where it occurs, subscribers have an indirect interest in call termination rates, as being called reduces their need to make outbound calls (and incur call charges).

This phenomenon is distinct from the more familiar concept of call externalities, where the recipient of a call enjoys a benefit that the caller does not take account of. It is usually assumed that call externalities are largely internalised as a result of repeated call interaction, for example as described in Taylor (1994), and so can be ignored. However, such internalisation behaviour also gives rise to demand linkages between calls in and out of a network. Therefore, even if we adopt the conventional demand modelling approach of ignoring call externalities, we cannot necessarily ignore call direction substitution.

2 Communication within a calling relationship

Communication commonly occurs within the context of a business or personal relationship, even if this might be a short-term or temporary one. In this case, the parties involved may communicate repeatedly over time, even if this is infrequent or sporadic. When there is repeated communication, it is reasonable to expect one party's calling behaviour to be affected by the frequency with which the other party calls. Whenever this is the case, a connection between incoming and outgoing calls is established. In particular, the pricing of one can be expected to affect demand for the other.

Established calling relationships may be particularly important for mobile phone users. Casual observation suggests that mobile phones tend to be used for more 'informal' types of calls. For example, when calling a person for the first time, it is more common to dial their fixed line, not least as directory services do not list mobile telephone numbers. Thus many calls to mobile phones may be from parties who have an established regular calling relationship with the mobile subscriber.

Consider the following simple example. Suppose that there are two parties, call them *A* and *B*. Suppose that at random intervals some need arises for one party to communicate with the other. For example, suppose that *A* has an item of news or information to communicate to *B*. Both *A* and *B* will receive some benefit from the communication of this news. The more delayed the news, the less valuable its communication will be to both the sender and the recipient. The costs of time delay will depend on the exact nature of the news item and will vary from case to case. Given this situation, each party must decide on its calling strategy given the costs of calling the other party and *given the calling strategy of the other party*.

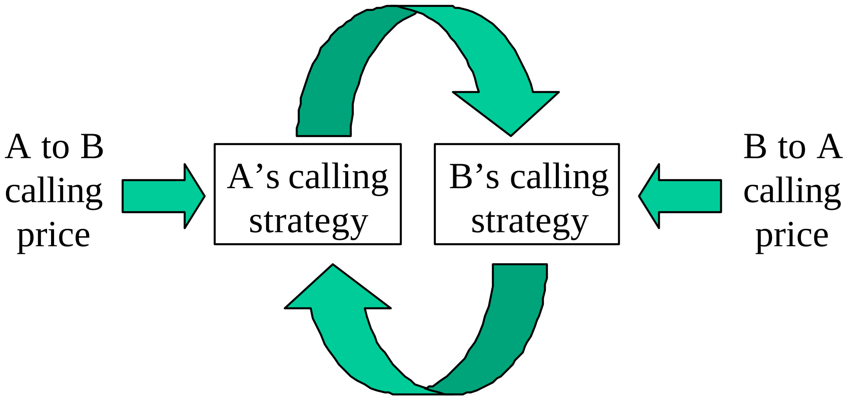
Whilst it might seem that *A* should simply ring *B* immediately whenever *A* has an item of news, this is unlikely to be an optimal calling strategy from *A*'s perspective. In particular, there are alternatives, especially for news items whose communication is of low value to *A* or not time critical. First, *A* could wait until *B* rings, which is lower cost as it avoids paying call charges all together. Clearly the desirably of this strategy depends on the frequency with which *B* is expected to call *A*. Second, *A* could stockpile news until such time that the benefits of communicating (or alternatively, that the costs of further delay) are sufficiently large relative to the cost of calling. At this point, all the stockpiled news could be communicated to *B*.[1]

Therefore, if *A* acts in a self-interested manner and maximises the benefits of communication to itself less its call charges, its calling strategy will be

[1] Here we are implicitly assuming that there are economies of scale in communication, so that multiple "news items" can be communicated as a bundle in less time than if they were each communicated separately. This is a reasonable assumption as there are likely to be fixed time overheads in setting up each communication.

affected not only by the costs of calling *B*, but potentially also by *B*'s calling strategy, as shown in the figure below. In particular, *A's* need to communicate with *B* may be partially met by *B* calling *A*. Therefore, *B* making more frequent and/or longer calls to *A* is likely to be associated with *A* making fewer and/or shorter calls to *B*. In effect, there is a demand for communication that might potentially be met by *A* calling *B* or *B* calling *A*.

Once we consider calling patterns in this way, as the expression of an underlying demand for communication (various and differentiated as this may be), it is immediate that calls from *A* to *B* are substitutes in the following sense. If the price of *B* calling *A* is lowered, *B* will change its calling strategy and make more (and/or longer) calls to *A*. However, this meets part of *A*'s need to communicate with *B* and so reduces *A*'s demand for calls to *B*.



The intrinsically bilateral natural of communication makes this substitution mechanism somewhat unusual. In particular, in more usual settings, we would think of two products being substitutes if there are a sufficient number of consumers who would switch between the two products as their relative prices changed. Our situation is unusual in that *A* chooses to make calls to *B*, whereas *B* chooses to make calls to *A*. We appear to lack a *single* consumer weighing up options on the basis of value for money. However, although two *different* consumers are involved, it is still the case that calls from *A* to *B* and calls from *B* to *A* are substitutes in the sense that demand for one is positively related to the price of the other.

Whenever subscribers in a repeat calling relationship have an underlying demand for communication that can be *partially* met by calls in either direction, this substitution relationship will emerge despite there being different consumers choosing calls in the two directions. Of course, this is not to say that calls in either direction are perfect substitutes. If *A* has important and time-critical news to convey to *B*, there may be little substitute but to call *B* immediately. However, there will be other types of communication that are less important and less time critical. In this case, there may be alternatives to ringing *B* immediately. The ability of telecoms operators to price discriminate between these different types of call is very limited; there is no way to identify which calls are time-critical. Therefore, even if the large majority of calls have few alternatives, this is not to say that there are not also a significant minority of calls where the underlying demand for communication could be met in other ways, including delayed and reversed calling. A significant price interaction between incoming and outgoing calls is entirely consistent with the *majority* of calls having few substitutes provided that there is no systematic means to price discriminate between calls with alternatives and those without.

This substitution between incoming and outgoing calls caused by calling pattern adjustment in response to the charges incurred by the party making the call does *not* in any way rely on their being a common economic interest between the two parties. Calls in one direction may depend on the price in the other direction without there being any sense in which the caller is taking account of recipient's interests. It is simply that regular communication in one direction may partially meet an underlying demand for the exchange of information and reduce the need to call in the other direction.

3 Call direction substitution and tariff balance

We consider a simple stylised model of bi-directional calling and examine how call direction substitution affects the incentives for operators to raise call termination rates to subsidise outbound calls.

3.1 Framework

Consider an interconnected network. Let q_I and q_O be the total volumes of inbound and outbound calls respectively and let p_I and p_O be their respective prices. Suppose that demand for inbound and outbound calls to and from the network are given by linear demand functions

$$q_I = 1 - b\,q_O - g\,p_I$$
$$q_O = 1 - b\,q_I - g\,p_O$$

where for simplicity we suppose that demand functions in both directions are the same. We take the total market size to be 1 without any loss of generality.[2] Here $b, g \geq 0$ where b measures the extent of call direction substitution. Suppose further that $b < 1$, so that any substitution of calls by those in the reverse direction is less than one-for-one.

[2] If alternatively we assumed that each call market was of size a, the results would depend only on the ratio g/a. Therefore, we can take $a = 1$ without loss of generality.

Corresponding to this demand function, outbound callers have an indirect utility function $v(p_O, q_I)$ where q_I represents an expectation about inbound calls. Calls received constitute an externality experienced by outbound callers. For any given expectation about inbound call volumes, this indirect utility function satisfies Roy's identity $\partial v/\partial p_O = -q_O$. This relationship determines the indirect utility function up to the addition of arbitrary functions of q_I. In particular, this implies that

$$v(p_O, q_I) = f(q_I) - (1 - b\,q_I)\,p_O + \frac{g}{2}\,p_O^2 .$$

The function f can be interpreted as the value given to incoming calls when outgoing calls are free. f must be a non-decreasing function, which is sufficient to imply that the marginal external benefit of receiving calls is non-negative as

$$\frac{\partial v}{\partial q_I} = f'(q_I) + b\,p_O \geq 0 .$$

Note that the marginal external value of inbound calls increases as outbound calls become more expensive. A reasonable case to consider is that of f being constant, which can be interpreted as inbound calls having no marginal external value to recipients when outbound calls are free. However, even in this case, the marginal external value of inbound calls will become strictly positive as the price of outbound calls increases above zero provided that $b > 0$.

Consumers decide how many outbound calls to make given expectations about the number of inbound calls they will receive and vice-versa. We suppose that in equilibrium these expectations are correct. Therefore, equilibrium demands for calls are given by

$$q_I^* = \frac{1}{1+b} - g\,\frac{p_I - b\,p_O}{1-b^2}$$

$$q_O^* = \frac{1}{1+b} - g\,\frac{p_O - b\,p_I}{1-b^2}$$

We can think of this as the Nash equilibrium of a game in which callers in each direction simultaneously choose the amount they wish to call.

3.2 Revenue neutral move from symmetric prices

In order to investigate the impact of call direction substitution on call termination rate setting, we conduct the following thought experiment. Suppose that we start with balanced prices $p_O = p_I = p$. As demands in the two directions are symmetric, inbound and outbound traffic flows are initially balanced. We then consider a small perturbation of this situation, moving inbound prices up by $d\,p_I > 0$ and outbound prices down by $d\,p_O < 0$ where these two price changes are such that overall revenues from calls remains unchanged.

If consumer welfare increases, this is a profitable strategy for the mobile operator, as it would be possible to capture some of this increased surplus through higher subscription charges and other retail side tariffs. Conversely, if consumer welfare falls, this is an unprofitable strategy. We do not explicitly model the full competitive process in this paper. However, it would be possible to embed the model here into the "competition in utility space" model of Armstrong and Vickers (2001) or other possible models of mobile competition.

For simplicity, we take marginal costs for the network operator to be zero as a convenient normalisation. The network operator retains the entirety of the retail price of outbound calls, but only call termination receipts in respect of inbound calls. Suppose that at the initial symmetric prices, the network operator retains a proportion r of the overall retail price of inbound calls through a call termination charge, with the originating operator retaining $1-r$.

The change in profit is given by

$$d\pi = d\,p_I \cdot q_I + r\,p_I \cdot d\,q_I^* + d\,p_O \cdot q_O + p_O \cdot d\,q_O^*$$

which we require to be identically zero. It is easy to show that this requires that inbound and outbound price changes are in the ratio

$$\Delta = \frac{-d\,p_O}{d\,p_I} = \frac{1-b+\left[2b-(1+r)\right]g\,p}{1-b+\left[(1+r)\,b-2\right]g\,p}.$$

If the originating operator's retention on inbound calls is zero, then $r=1$ and so $\Delta=1$; the increase in the price of inbound calls and the reduction in the price of outbound calls are the equal as demand relationships in the two markets are the same. However, in general, as $r \leq 1$

$$(1+r)\,b-2 \leq 2b-(1+r)$$

and so $\Delta \geq 1$. When the originating operator obtains a retention, the revenue-neutral reduction in the price of outbound calls is greater than the increase in the price of inbound calls. A simple interpretation of this is to consider the model as one in which the network operator sets the retail price of inbound calls, but incurs an input 'cost' due to the originating operator's retention. An increase in the inbound price reduces inbound volumes and so reduces retention 'costs'. This cost saving translates into outbound prices falling by more than inbound prices increase.

3.3 Welfare impact

We now consider the welfare impact of a revenue neutral move from initially balanced prices p. The change in welfare is given by

$$dv = -q \cdot dp_O + dq_I^* \cdot \left(e + \frac{1+2b^2}{1+b} p \right)$$

where $e = f'(q) \geq 0$ is the marginal external value of the initial volume of inbound calls at zero outbound price. Here

$$q = \frac{1-g\,p}{1+b}$$

is the initial volume of traffic in each direction when prices are balanced. The positive term $-q \cdot dp_O$ is the welfare gain enjoyed by consumers of outbound calls due to lower prices. The question is whether this is outweighed by the adverse impact of reduced inbound calling due to higher inbound prices used to fund reduced outbound prices.

The reduction in the equilibrium volume of inbound calls, dq_I^*, is given by

$$\frac{dq_I^*}{dp_I} = -g\,\frac{1+b\Delta}{1-b^2}.$$

The greater the extent of call direction substitution (i.e. the closer b is to 1), the greater will be the reduction in inbound call volumes, as not only will there be an own-price effect from increasing inbound call prices, but also some traffic will reverse direction. Putting this all together,

$$\frac{dv}{dp_I} = \Delta\frac{1-g\,p}{1+b} - g\,\frac{1+b\Delta}{1-b^2}\left(e + \frac{1+2b^2}{1+b} p \right).$$

The sign of this expression tells us whether welfare is increased (when positive) or decreased (when negative) from moving in a revenue neutral manner towards unbalanced prices. Putting this expression over a common denominator and regrouping terms gives that this sign is the same as the sign of the expression

$$S = \underbrace{\frac{1}{g}}_{\text{outbound consumer surplus}} - \underbrace{e\,\frac{b+1/\Delta}{1-b}}_{\text{call externalities}} - \underbrace{p\left[1 + \frac{(b+1/\Delta)(1+2b^2)}{1-b^2}\right]}_{\text{call direction substitution}}.$$

This gives some immediate comparative statics results and an interpretation of the three individual terms:

- In the absence of any call direction substitution ($b = 0$) or call externalities ($e = 0$), $S \geq 0$. Increasing inbound charges whilst lowering outbound charges is welfare-increasing for subscribers. Consumers enjoy lower outbound charges without any countervailing disbenefit. Given that this increase in consumer welfare can be partially captured by operators through higher subscription charges (or other multipart tariff structures), there will be an incentive for the network operator to unbalance tariffs and raise call termination rates;

- If there are call externalities from inbound calls ($e > 0$) but this does not affect outbound calling ($b = 0$), then $S = 1/g - e/\Delta$, which is negative for sufficiently large e. Large call externalities on inbound calls make a move from balanced tariffs welfare-reducing due to reduced inbound calling. This will reduce subscription revenues, so there is no incentive for an operator to raise call termination rates in this case;
- If there is call direction substitution, then for all b sufficiently close to 1, $S < 0$ and so unbalancing tariffs is welfare-reducing. Let the critical value of b at which S first becomes negative be b^c. This is subject to the following comparative statics results:
  - The higher the initial price of calls, the lower is b^c. Higher prices reduce traffic volumes and hence the size of the welfare benefit received from lower outbound call prices;
  - The smaller g, the price sensitivity of calls, the higher is b^c. Smaller values of g imply that greater consumer surplus is associated with outbound calls and the greater the benefit enjoyed by consumers from lower outbound prices. This increases the incentive to unbalance prices;
  - Reducing the originating operator's retention increases r and so decreases Δ, the ratio of the outbound to inbound price movements required to maintain revenue neutrality. This in turn reduces b^c. The smaller the originating operators' retention, the smaller is the reduction in outbound tariffs achieved (for unchanged revenue) from a given increase in inbound tariffs, as there is a smaller cost saving from reduced volumes of inbound calling.

4 Conclusions

Where there are repeated bilateral calling relationships, it is reasonable to expect interaction between calls in each direction. A call initiated in one direction allows both parties to exchange news and may suppress demand for calls in the reverse direction for some time afterwards. We call this call direction substitution.

Call direction substitution is different from conventional notions of demand substitution as used in market definition, as it does not arise from individual agents choosing between alternatives. Rather it arises as result of the equilibrium behaviour of pairs of agents.

Call direction substitution can reduce the incentive for operators to distort tariffs by raising call termination rates and lowering retail prices. As call direction substitution becomes stronger, it will eventually eliminate any such incentive.

5 References

Armstrong, M, 1998, 'Network interconnection in telecommunications', *Economic Journal*, 108, 545-564.

Armstrong, M, and J Vickers, 2001, 'Competitive price discrimination', *RAND Journal of Economics*, 32(4), 579-605.

Oftel, 2001, 'Review of the Charge Control on Calls to Mobiles', 26 September, available at http://www.oftel.gov.uk/publications/mobile/ctm0901.htm.

Taylor, L, 1994, *Telecommunications Demand in Theory and Practice*, Kluwer Academic Publishers.

DotEcon Discussion Papers

DP No. 02/01: Roger Salsas and Christian Koboldt, ‘Roaming free? Roaming network selection and inter-operator tariffs’, August 2002.

DP No. 02/02: Réka Horváth and Dan Maldoom, ‘Fixed-mobile substitution: A simultaneous equation model with qualitative and limited dependent variables’, August 2002.

DP No. 02/03: Dan Maldoom, ‘Caller-called party interaction: Implications for call termination’, September 2002.

DP No. 03/01: Christian Koboldt, Dan Maldoom and Richard Marsden, ‘The first combinatorial spectrum auction – lessons from the Nigerian auction of fixed wireless access licences’, May 2003.

DP No. 07/01: Dan Maldoom, ‘Winner determination and second pricing algorithms for combinatorial clock auctions’, December 2007.

DP No. 10/01: Richard Marsden, Eimear Sexton and Arisa Siong, ‘Fixed or flexible? A survey of 2.6GHz spectrum auctions’, June 2010.

DP No. 11/01: Christian Koboldt, Dan Maldoom and Roger Salsas, ‘How strong are merchant constraints on interchange fees?’, April 2011.

Further DotEcon publications, including the DotEcon Perspectives series and publicly available reports, presentations and submissions can be downloaded at www.dotecon.com.

About DotEcon

DotEcon is an economic consultancy advising private and public sector clients in:

- Competition cases, regulatory proceedings and commercial litigation
- Public policy design and regulatory impact assessments
- Design and implementation of auctions and trading mechanisms
- Bidder support for high-value transactions
- Econometric analysis and data mining
- Business strategy and decision support

Founded in 1999, the company focuses on providing high-value advice to governments and leading companies using rigorous microeconomic techniques. DotEcon works for clients across the world, and is involved with many high-profile policy debates and business decisions. DotEcon has a particular focus on network industries such as telecommunications, transport, energy and payment systems.

DotEcon is the recipient of a Queen’s Award for Innovation 2011.

dot·econ

Economics for a networked world