A literature review of papers on MTRs with relevance to B&K

A REPORT PREPARED FOR VODAFONE

August 2009

A literature review of papers on MTRs with relevance to B&K

A literature review of papers on MTRs with relevance to B&K

Tables & figures

Executive summary

There is a considerable debate in the economic literature and amongst practitioners on the appropriate level of mobile termination rates (MTRs). In this debate, the relative merits of setting zero MTRs (i.e. Bill and Keep - 'B&K) are often discussed. In this context, Vodafone has commissioned Frontier Economics Europe to undertake a review of the academic literature relating to the efficient setting of MTRs, with specific reference to the efficiency of B&K.

Economic theory on MTRs reflects that cost-based pricing is generally efficient (i.e. welfare maximizing). Departures from this benchmark are usually justified by the existence of call or network externalities. Thus the presence of network externalities supports the imposition of above-cost MTR whereas call externalities require below-cost MTRs.

In this context B&K is efficient only under very specific conditions. It requires the absence of network externalities, the presence of call externalities and, that the cost of origination equals the cost of termination and that the value of calls is shared evenly among senders and receivers. Therefore, the existence of call externalities is a necessary but not a sufficient condition for B&K to be optimal.

It is also important to consider that call externalities are partially internalised by customers through their repeated interaction with other subscribers. In contrast to what is traditionally assumed in the literature, outgoing and incoming traffic are not independent. Outgoing calls usually generate a number of incoming calls. This pattern of reciprocity helps to (partially) internalise call externalities and should be considered in the efficiency of MTRs. That is, only un-internalised call externalities should be relevant for the analysis of efficient MTRs.

While the focus of the discussion of MTRs is often around costs and the existence or otherwise of call and network externalities, it is important to understand that the literature also shows that the level of MTRs has an important impact on inter-network competition, which provides another route by which the level of MTRs may impact on consumer welfare. In particular, excessively low MTRs may be an indication of collusion among operators. If MTRs are set at too low a level then competition between networks may be softened. Specifically, if mobile operators compete in two-part tariffs and price discriminate based on the terminating network, a reduction in MTRs has the effect of increasing the equilibrium subscription fees (alternatively, reducing mobile handset subsidies). This effect is known as the "waterbed effect" and has been empirically tested for the mobile sector. In such circumstances consumers may be left worse off by the imposition of below-cost MTRs.

Given these caveats and that the existing evidence points towards low call externalities, it is quite difficult to support the desirability of B&K on efficiency grounds.

It is frequently argued that B&K could help reduce the gap between large and small operators by reducing existing off-net/on-net price differentials. Such pricing policies generate tariff-mediated network externalities that it is alleged are damaging small operators and even deterring potential entry. The economic literature indicates that the strategic use of the off-net/on-net price differential in a context with call externalities cannot be avoided with B&K. This is confirmed by the experience in "B&K countries" such as the US, where off-net/on-net price discrimination is common.

There are also models analysing the potential for entry deterrence of high MTRs. However, they do not imply that B&K is efficient. The efficient MTR will mainly depend on the existence of un-internalised network and call externalities.

In conclusion, unless there is empirical evidence on the specific conditions under which it is efficient, B&K is likely to result in a loss in market efficiency.

1 Introduction

This report presents an overview of the literature on mobile termination rates (MTRs) and then assesses the implications of this literature for the debate on B&K.

The initial contributions focused on efficiency and operators' incentives to set MTRs in a symmetric context of competition. **Section 2 reviews the basic model considered in this early literature and some recent extensions.**

The basic model analyses competition between two horizontally differentiated operators in the absence of (network or call) externalities. Welfare is maximised when MTRs are cost oriented whereas the MTR that maximises operators' profits may be above or below cost depending on the price structure observed at retail level. A high MTR is profitable only under linear prices and in the absence of on-net/off-net price discrimination. The introduction of non-linear pricing (two-part tariffs) and, in particular, discrimination based on the terminating network reverses operators' incentives. In the latter scenario, a reduction in MTRs is likely to result in higher subscription prices.

The introduction of **call and/or network externalities** exert an impact on both efficiency and operators' behaviour. This is **covered in section 3 of the report**.

Call externalities are present when called parties obtain benefits from receiving calls. Whereas *network externalities* exist when subscribers obtain benefits from a larger network, because they have more people to call/communicate with.

The efficient MTR is above cost when network externalities are present. The opposite happens under call externalities. The combination of both may lead to above or below cost efficient MTRs. The lack of contributions analysing the interaction of these two effects does not permit the drawing of a clear conclusion on how efficiency is affected by the level of MTRs. This is important because in most real situations call and network externalities will co-exist, although the call externality will tend to be internalised.

An interesting property present in the models with call externalities is that operators have a strategic incentive to raise their off-net prices, reducing the positive externality exerted to the customers of rival networks. In this context, there is a number of contributions focused on the **potential effect of the on-net/off-net price differential on competition. Section 4** reviews the main papers in this area.

In this regard, it is usually argued that B&K will reduce the gap between large and small operators by lessening the extent of off-net/on-net price differentials. However, this is not supported by the economic literature analysing the strategic use of off-net/on-net price differentials. For example, under the model of Hoernig (2007), which is covered in more detail in section 4, a large operator may strategically set retail prices to damage the smaller operators even with MTRs below cost.

Moreover, B&K is also not justified by the models that analyse the strategic use of MTRs to deter entry. In these models the efficient level of MTRs will depend, as in the basic context, on costs and the presence of un-internalised call and network externalities. Further, a common feature of these theories is that the use of a high MTR is only profitable in the successful scenario in which entry has been deterred. If entry took place operators would have incentives to lower MTRs.

Section 5 analyses the implications for B&K. Here we show that B&K is likely to result in a efficiency loss and may imply a less intense competition among networks.

Finally, section **6 concludes**. All the references used in the report are contained in section 7.

2 The basic Hotelling model

2.1 MAIN ASSUMPTIONS

The basic model, developed in Armstrong (1998) and Laffont, Rey and Tirole (1998a, 1998b) considers competition for subscribers between two differentiated networks (A and B) where there are no network or call externalities. The degree of differentiation between the two networks is inversely related to the intensity of competition.

Horizontal differentiation: is usually modelled *à la* Hotelling such that the two operators, located at the extremes of a segment representing the market, compete for a share of the consumer base which is taken to be distributed uniformly on a line between the locations of the two suppliers[1]. The transportation cost that consumers have to incur to reach each of the operators is a proxy for the degree of differentiation between the operators.[2] In the absence of transportation costs both operators would be perfect substitutes as far as consumers are concerned.

Consumers: choose a supplier on the basis of which provides them with the highest level of utility. This is measured by the value that the consumer gets from the product less any charges made by the supplier. In the absence of call externalities, consumers' utility depends only on the number of calls originated, not on incoming calls.

Operators: provide subscription and call services to consumers. The marginal cost of a call, denoted by C, is made up of the cost of origination C_O and the cost of termination C_T. The sum of these two costs is the marginal cost of an on-net call. The cost of an off-net call is a function of the termination charge to be paid to the receiver's network, denoted by a.

$$C_{on\text{-}net} = C_O + C_T$$

$$C_{off\text{-}net} = C_O + a$$

In addition, mobile operators face a fixed cost per subscriber, denoted by k.[3]

2.2 RETAIL PRICES

Retail pricing: there are several alternatives to consider. If operators can charge separate prices for subscription and usage, operators will set call charges at marginal cost and then compete with each other over the level of the subscription charge. The profit maximizing subscription charge will be inversely related to the intensity of competition/degree of substitution. In the extreme

---

1 The basic model assumes that all consumers in the base will chose either network. There are no consumers which chose not to consume.

2 The models usually assume linear transportation costs, such that the cost of walking distance *x* is *t*x*, where *t* is the unit cost of transport.

3 These costs may represent the costs of including the subscriber in the data base, etc.

case where operators are perfect substitutes (no transportation costs) the fixed fee would equal the cost of subscription, k.

If networks can discriminate (i.e. charge different prices) between on-net and off-net call charges, then:

$$P_{on\text{-}net} = C_O + C_T$$

$$P_{off\text{-}net} = C_O + a$$

If networks cannot price discriminate between on-net and off-net call, the usage price will be a weighted average of the cost of the call, which takes the following form if customers are evenly distributed between both operators:

$$Price_{Blended} = C + (1/2)(a - C_T) = (1/2)C + (1/2)(C_O + a)$$

If networks cannot charge two-part (or non-linear) tariffs as shown above, but rather charge "linear tariffs" then call charges will be set above marginal cost and will increase as the intensity of competition declines.

Efficiency

In this simple framework in which call and network externalities are absent and penetration is fixed and complete, optimal retail prices follow the CPP principle, termination fees are cost oriented. Thus, efficiency requires:

$$a_{efficient} = C_T$$

2.3 OPERATORS' INCENTIVES

There is a wide body of literature that addresses the incentives of operators when deciding the level of MTRs. Armstrong (2002) provides an extensive review of the main results considering the setting of FTM and MTM MTRs when these are decided independently from each other. As we will see below, the incentives of operators are radically different in these two scenarios. Armstrong and Wright (2008) contribute to the literature by analysing the joint determination of FTM and MTM MTRs.

2.3.1 FTM termination rates

If each mobile customer holds only one cellular phone, and assuming CPP at retail level, then FTM termination involves a competitive bottleneck as emphasised by Armstrong (2002), Armstrong and Wright (2008) and Houpis and Valletti (2005) among others. The key is that each mobile network is a monopolist over delivering calls to its subscribers. Hence, the profit maximizing FTM will be set at the monopoly level even if competition is strong at retail level.

To illustrate, consider the simple model in Armstrong (2002) with the following assumptions:

- Assumption 1. All calls made from mobile networks are terminated on the fixed sector;
- Assumption 2: Mobile subscribers gain no utility from receiving calls (no call externalities);

- Assumption 3: Mobile subscribers do not care about the welfare of the people who call them;
- Assumption 4: Mobile subscribers do not pay anything for receiving calls made to them (Consumer Party Pays or CPP);[4]
- Assumption 5: The mobile sector is perfectly competitive.

Let Q(a) be the demand for FTM call for a given MTR of a.

The cost structure for mobile operators is as described in section 2.1. Further, suppose that mobile operators charge two-part tariffs with a usage price of *p*. Assumption 5 implies that in equilibrium mobile operators will make no profits[5] and will set their retail prices to maximize subscriber utility subject to a break-even constraint. Because of assumptions 2, 3 and 4 usage prices will be set at marginal costs and the fixed fee will be set to drive profits to zero.

In this context, even though competition leads mobile operators' economic profits to zero, each mobile network has incentives to maximize its termination profits, given by:

$$(a - C_T)\, Q(a)$$

Therefore, in equilibrium a is set at its monopoly level (a^{mon}). By maximizing access revenue from received fixed calls, mobile operators can compete harder for subscribers subsidizing subscription.

2.3.2 MTM termination rates

In the case of MTM MTRs the literature has focused on the joint determination of MTRs as a way to affect competition at retail level.[6] A classical result is that of Gans and King (2001), which shows that operators can relax the intensity of competition by agreeing on a very low MTR. This result is obtained under the basic framework described in section 2.1 when operators compete in two-part tariffs and price discriminate based on the terminating network.

The intuition for this result is that given the equilibrium prices under two-part tariffs,[7] there is a direct relationship between MTRs and off-net call prices and the difference between on-net and off-net prices is given by the mark-up implied by the MTR. In the particular case when the MTR is below cost (negative access mark-up) customers prefer to join smaller operators over larger ones, since the price for off-net calls is below the price for on-net traffic. This creates an incentive for operators to raise their subscription fees since being small is valued by customers. In this way, a reduction in MTRs allows operators to relax the intensity of competition for subscribers. The effect that MTRs exert on other

4 The alternative scenario where the receiving party also pays for the call is usually known as RPP.

5 That is, not profits above the economically efficient level.

6 See Armstrong and Wright (2008).

7 Usage prices based on perceived costs and fixed tariffs inversely related to the level of competition.

prices in the operator's bundle, fixed fees in this case, is known as the 'waterbed effect.'[8]

The literature has, however, shown that this result may be sensitive to the prevailing retail price structure. Under linear pricing, high access prices result in the 'raising each other cost effect' which reduces consumer welfare.[9]

However, "*this effect partially or totally disappears when providers can operate a price discrimination between on-net and off-net calls (Laffont et al.(1998b)) or when they compete in non-linear prices (Laffont et al. (1998a)). (…) The collusive power of access charges totally disappears in two part tariffs*"[10].

In practice, mobile operators do not usually ask for a reduction of MTRs, which raises the question as to whether the linear pricing model; or the two-part pricing model is the most relevant benchmark. Below we consider several papers that provide alternative rationales for mobile operators' behaviour; resulting in higher MTRs which do not necessarily rely on the imposition of linear and non-discriminatory prices at retail level.

2.3.3 MTM and FTM termination rates

Armstrong and Wright (2008) extend the basic model of Hotelling competition among mobile operators in order to consider the joint determination of MTM and FTM termination rates. In particular, the paper analyses what would be the profit maximizing MTR if mobile operators set a uniform termination charge for both services (FTM and MTM).[11] When this is the case, operators' incentives to lower MTRs are much reduced because of the effect that this reduction has on the revenues of calls from fixed networks.

The model considers a similar framework to that of Gans and King (2001) for MTM calls.[12] On the other hand, FTM termination enjoys a bottleneck as described in section 2.3.1 As we have already seen before, under these modelling assumptions, if set separately operators will set the FTM MTR at its monopoly level (a_{mon}), whereas the MTM MTR will be below cost (a_{below}).

If operators decide on a uniform price for MTM and FTM MTRs the resulting MTR will be below cost as in the case of equilibrium MTM MTRs when set separately. The intuition for this result is that profits are neutral with respect to the FTM MTR[13], but not with respect to MTM: "*Therefore, firms' incentives are*

8 See Schiff (2007) for an theoretical analysis of the waterbed effect.

9 See the analysis under *linear non-discriminatory pricing* contained in section 4.2.3 of Armstrong (2002).

10 Baranes and Flochel (2004), page 2.

11 Armstrong and Wright (2008) explains this constraint by the possibilities of wholesale arbitrage, meaning that "*a mobile network cannot maintain a high FTM termination charge together with a low MTM termination charge, since the fixed network could then "transit" its calls via another mobile network and so end up paying the lower MTM rate (plus a small transit charge)*". Armstrong and Wright (2008), page 3.

12 Model described in section 1.1 under two-part tariffs and network discrimination.

13 That is, these are not affected by the level of this charge. This implies a 100% waterbed effect.

exactly as if there is only MTM traffic."[14] However, when there is potential for market expansion (see section 3.1.1, below), so that the base of mobile customers is endogenously determined in the model instead of being constant, then operators will choose a MTR that lies between a_{below} and a_{mon} . The key for this result is that in the context with market expansion profits are no longer neutral with respect to FTM MTR, presenting a direct relationship with this price. This creates incentives for operators to ask for a high FTM MTR. In this case, the equilibrium MTR may be above *or* below the efficient level depending on a range of factors. Namely, the potential for market expansion and the importance of FTM calls in relation to MTM traffic.

An alternative scenario where operators will have incentives to set too high MTRs is that in which MTM and FTM MTRs are uniform and set unilaterally by each operator.[15] Hence, despite the fact that a high MTR intensifies network competition, thereby reducing operators' profits, operators' incentives to raise the MTR in order to maximize their revenues from call termination and gain market share - due to the impact that MTRs have on rivals' off-net prices, lead to a MTR above the efficient level but below the monopoly threshold.

14 Armstrong and Wright (2008), page 21.

15 The result that mobile operators have incentives to set above cost MTM MTRs if these are decided unilaterally is also observed in Gans and King (2001).

3 Model's extensions

3.1 NETWORK EXTERNALITIES

Network externalities arise when existing subscribers of a network benefit from a new subscriber joining the network. In mobile markets the presence of additional subscribers generates a positive externality to existing subscribers, because it creates the possibility of calling additional people and of being called by these new subscribers.

The literature shows that, in the presence of network externalities, the efficient termination rate should be set above cost. A higher termination rate induces operators to lower their subscription prices, thus promoting network participation at a level consistent with the social interest.

The above result is observed in a number of contributions. For example, Armstrong (2002)[16] and Valletti and Houpis (2005) in the context of FTM call termination. The intuition provided by the former being that: "*a higher termination charge raises the equilibrium mobile subscriber utility via handset subsidies and the like, this in turn increases mobile subscription, which in turn raises the utility of fixed network subscribers because of the network externality effect.*"[17]

3.1.1 Market expansion of the mobile market

In an extension of their model Armstrong and Wright (2008) consider the possibility that the mobile market can be expanded by relaxing the assumption that the number of mobile customers is constant.[18]

Given the way the FTM market is modelled[19] again, in equilibrium, operators will set the FTM MTR at its monopoly level (a_{mon}). Now, the welfare maximizing MTR is above cost since this induces extra mobile subscription, benefiting all users. Nevertheless, the efficient level is still below a_{mon}.[20] An important difference with respect to the case without market expansion is that now the profit neutrality result with respect to FTM MTRs no longer holds. This implies that operators have incentives to cooperatively choose a high FTM MTR.

With regards to MTM MTRs, as in the case without market expansion, operators have incentives to set this price below cost, in order to relax network competition. However, now the efficient MTM MTR is above cost also due to its positive effect on mobile subscription.

---

16 Section 3.1.3.

17 Armstrong (2002), page 343.

18 In particular, they consider the so-called 'Hotelling model with hinterlands.'

19 The base of fixed line customers is given and the traffic flows only from fixed to mobile customers which do not face call externalities.

20 This result is also obtained in Armstrong and Wright (2002) and Valletti and Houpis (2005).

In the numerical analysis contained in section 2.4 of the paper Armstrong and Wright (2008) show that with market expansion the efficient MTR for MTM and FTM services will differ and will be above the cost of call termination.[21] As already commented on above, if operators are constrained to set a uniform MTR for both MTM and FTM calls then the profit maximizing MTR will lie between the profit maximizing MTM MTR and the MTR that maximizes FTM profits (i.e., between a_{below} and a_{mon}).

In a related paper, Schiff (2001) considers the effect of partial consumer participation on operators' incentives to set MTM MTRs and efficiency. Schiff (2001) presents a variant of the basic model with two-part tariffs and uniform prices for on-net and off-net calls. He considers an endogenous customer base and a variant with network externalities.[22]

Endogenous participation is modelled by assuming that potential subscribers have an option value associated with joining the market, which is randomly distributed. Once the decision to subscribe is made, based on expected benefits from joining, the subscriber chooses network in the same way as in the basic model with Hotelling competition. In the absence of network externalities all subscribers still make the same volume of calls. Schiff (2001) models network externalities assuming that the calls made by each subscriber are a linear function of the number of subscribers.

In all these models, Schiff (2001) finds that it is still efficient to price calls at marginal cost and compete over the level of the rental charge. Schiff finds that an endogenous market size without externalities intensifies competition relative to the basic model (because networks compete for new subscribers as well as for market share) but profits, consumer surplus and hence total welfare are maximised by cost-based access charges.

With regards to the profit maximizing MTR, Schiff (2001) shows that an endogenous market size increases the incentive for networks to price reciprocal access at cost.

In the presence of network externalities Schiff (2001) shows that the networks will profit maximise by pricing access below marginal cost (even though they are charging uniform on-net and off-net prices), while consumer surplus is maximised by pricing access above marginal cost. The intuition of this result is that externalities make competition even fiercer in a non-linear way. Adding a customer, when access is priced above cost, creates profits directly and increases the volume of calls by existing customers, which multiplies the profit. The networks would choose to mitigate competition by setting the price of access

---

21 The efficient FTM MTR will be above the efficient MTM MTR when market expansion possibilities are large (see table 7 of Armstrong and Wright (2008)).

22 The former is modelled by assuming that potential subscribers have an option value associated with joining the market, which is randomly distributed. Once the decision to subscribe is made, based on expected benefits from joining, the subscriber chooses network in the same way as in the other models discussed here. All subscribers still make the same volume of calls. Schiff models network externalities assuming that the calls made by each subscriber are a linear function of the number of subscribers.

(and calls) below cost to offset the effect of the network externality. Total welfare on the other hand is maximised with access priced above marginal cost, because this leads to a lower rental charge, which in turn drives up the penetration rate.[23]

3.2 CALL EXTERNALITY EFFECT

In the basic models the value of a call accrues entirely to the caller, i.e., the receiver does not benefit from receiving calls. Recently, the economic literature has extended the basic model to include call externalities, i.e. by considering the more realistic scenario in which the recipient of the call also benefits. Therefore, under call externalities calls generate value to both callers and recipients.

In the basic setup with call externalities[24] a call is assumed to generate a value u to the sender and βu to the receiver, with $\beta > 0$[25]. Thus β is the ratio between the recipient and the caller's valuation of a call.

Under this model, efficient network utilisation implies that the total costs of the call should be recovered from both parties in proportion to the benefits each receive. This means that with call externalities operators should charge both callers and receivers, and RPP becomes efficient.[26]

The efficient retail prices in this case will be a function of the total cost of the call and the call externality ratio. In particular,

$P_{Caller} = C/(1+\beta)$

$P_{Receiver} = \beta C/(1+\beta)$

It is important to note that the parties share the total costs of the call in proportion to the benefits. This could imply, for instance that the receiver pays a retail price above the costs of terminating the call but overall, the retail price would just recover the total costs of the call. Thus it should be noted that for the purpose of determining optimal retail charges and hence optimal MTRs the actual cost of termination may be of limited relevance as the more important factors are the total cost of calls and the size of the call externality.

3.2.1 Optimal MTRs

If operators set retail call prices at costs, either because of regulation or as a consequence of competition, then the efficient termination fee would be:

$$a_{efficient} = C_T - \beta C/(1+\beta)$$

23 Dessein, W. (2001) shows that the welfare result is not completely general, but is true provided that two duopolists offer a larger net surplus to customers than a monopolist.

24 See Jeon, Laffont and Tirole (2004).

25 Note that in the absence of call externalities ß = 0.

26 Note that the caller also pays a part of the call costs according to the benefits he/she gets, so the optimal charging scheme is not pure Receiver Party Pays.

In this case the efficient termination fee is below cost and *decreases* as the call externality becomes larger. The intuition for this result is that as the call externality increases the receiver should pay a larger fraction of the cost of the call and this is achieved by setting a lower termination fee. In other words, an increase in the call externality has a positive impact on the willingness to pay of the receiver and thus the terminating network needs to charge a lower termination fee in order fully to recover its costs.

Notice that in the analysis above we have disregarded the ***effects of competition between operators on the efficient MTR.*** The literature shows that in the presence of call externalities the characterisation of the welfare maximizing MTR may be extremely complicated and the results may be sensitive to the assumptions of the model. This is, for example, the case with linear tariffs as shown by the analysis of Berger (2004). In this case, the welfare maximizing MTR is found only through a graphical analysis as it is not possible to solve the problem analytically. This analysis reveals that the welfare maximising termination charge is lower than the profit maximising charge and may be below zero.

Berger (2005) considers competition with two-part tariffs. In the symmetric case, i.e. when both operators share the market evenly, the welfare maximizing MTR decreases with the size of the call externality. As in the case of linear tariffs, the efficient MTR may be below cost. In the context of this model, operators have incentives to set below cost termination rates -even below the welfare maximizing MTRs.

Baranes and Flochel (2004) consider a slightly different model from the ones commented on above, by assuming that networks are vertically differentiated, such that they differ in their quality. Further, operators compete in non linear pricing and can discriminate between on-net and off-net prices. In this model, consumers face call externalities, such that they have the same valuation for the calls they send and receive. The aim of the paper is to analyse the incentives of operators *when deciding on their MTRs unilaterally.*[27]

Baranes and Flochel find that, although in equilibrium the access charge chosen by operators is above the terminating cost,[28] the incentives to set high MTRs are reduced in order to internalise off-net call externalities for its customers. That is, by setting a lower MTR consumers receive more off-net calls from the alternative network. In the presence of call externalities, this raises the value of the network allowing it to charge higher subscription fees. This contrasts with previous results found in the literature when considering incentives of operators when setting MTRs unilaterally, like in the context of proposition 1 of Gans and King (2001). The intuition is that, in the presence of call externalities, operators' incentives to raise MTRs to maximize their access revenues are countered by the negative

27 In contrast with the papers considered above, Baranes and Flochel (2004) do not impose reciprocity on the access charge. Instead, MTRs are unilaterally decided by each operator.

28 Because of the double marginalisation effect faced by providers, which is the main effect in proposition 1 of Gans and King (2001).

impact that MTRs exert over the off-net price charged by other operators which reduces the call externality enjoyed by their own subscribers.

DeGraba (2003) considers a model in which operators compete for two customers that face call externalities, in particular, the value of the call is evenly shared among both parties. Operators compete offering usage prices for on-net/off-net outgoing and incoming calls and face the same cost for originating and terminating the call. In this simplified framework, DeGraba (2003) finds that B&K is efficient. It should be noted however that this result is mainly driven by the assumptions considered in the model on the size of the call externality and terminating costs.[29] Further, in one of the extensions of the model, considering a random allocation of the value of the call, DeGraba (2003) finds that if one of the parties receives most of the benefit from the call then "*imposing all of the cost on the calling party will be relatively more efficient*"[30] than B&K. This is an essential observation, because, as we show below, existing evidence shows an uneven distribution of the call value in favour of the calling party.[31]

3.2.2 Connectivity breakdown

An interesting result of the literature on call externalities is that their presence may lead to a **"connectivity breakdown"** given their effect on operators' incentives to set retail tariffs. That is to say, operators may have incentives to set the prices in such a way that calls to rival networks are prohibitively costly.

The issue of connectivity breakdown has been analysed under CPP and RPP retail tariff regimes. Here we will focus on CPP.

Connectivity breakdown in a CPP context: Jeon, Laffont and Tirole (2004) provide the basic model to understand why call externalities may generate a connectivity breakdown.[32] In the Jeon, Laffont and Tirole (2004)'s model with CPP the equilibrium prices for on-net and off-net calls take the following form:[33]

$P_{on\text{-}net} = (Co + C_T)/(1 + \beta)$, the same for both operators; and,

$P^i_{off\text{-}net} = (1 - s_i)(C_0 + a)/(1 - s_i(1 + \beta))$, for operator *i*, with market share s_i

In contrast to the case without call externalities, the price of off-net calls depends on the size of the network. The larger network sets higher off-net prices.[34]

29 The value of the call is evenly distributed among senders and receivers; and the cost of terminating a call equals the cost of origination.

30 DeGraba (2003), page 18.

31 See section5.1.

32 The models considered by Berger (2005) and Armstrong and Wright (2007) are also illustrative.

33 These are the equilibrium prices under a two-part tariff structure and off-net/on-net price discrimination. Berger (2004) analyzes linear prices.

34 The derivate of the off-net price with respect to the market share of the operator can be written as $\frac{\partial p^i_{off-net}}{\partial s_i} = \frac{\beta(c_0 + a)}{(1 - s_i(1+\beta))^2} > 0$ if $\beta > 0$ and $(C_0 + a) > 0$.

Furthermore, if the network externality is large enough, $\beta > ((1/s_i) - 1)$, connectivity breakdowns since operators would set an infinite price for off-net prices. The reason is that the existence of call externalities generates incentives to increase off-net prices in order to reduce the numbers of off-net calls and thus damage the rival network, if their customers value highly incoming calls.

This connectivity breakdown's result is common to most models including call externalities.

3.2.3 Call propagation

Cambini and Valletti (2008) consider a more realistic approach to modelling consumers' behaviour and include the possibility of "reciprocal" communication patterns, such that each outgoing off-net call results in a fraction x of incoming calls. Comparing their results with Jeon et al. (2004), the authors show that networks will have reduced incentives to use off-net/on-net price discrimination to induce a connectivity breakdown when outgoing and incoming calls are complements.

In a symmetric equilibrium without reception charges, the off-net price takes the following form:

$$P_{off\text{-}net} = (C_0 + a - x\,(a - C_T))(1 - \beta(1\text{-}x))$$

Note that if propogration (x) is close enough to 1, then the effect of call externalities may become rather insignificant. So when discussing the importance of call externalities, one should consider this call propagation effect which does tend to diminish the significance of call externalities.

The paper also analyses the incentives of operators when setting MTRs and shows that, under some circumstances operators may "*achieve first-best allocations via negotiated access charges that internalize externalities.*"[35]

3.3 NETWORK AND CALL EXTERNALITIES

In reality, both types of externalities will be present to some extent and the regulator will have to weigh the importance of each. However, no paper combines access and network externalities in a model. So *a priori* it is unsure what result/effect would be when both are present An interesting pointed raised by Armstrong and Wright (2007) is that:

"*the presence of call externalities will amplify the impact of network externalities, since users will receive more calls when there are more mobile subscribers.*"

This suggests that the combination of both network and call externalities could result in above-cost MTRs. In other words, despite the fact that call externalities, when considered alone, lead to below cost MTRs, call externalities widen the importance of network externalities, which require higher MTR.

35 Cambini and Valletti (2008), page 17.

Notwithstanding this, in the absence of a more rigorous analysis it is difficult to assess the potential impact on MTRs when call and network externalities co-exist.

3.4 HETEROGENEITY

3.4.1 Consumers

In the models described above, under two part-tariffs it is efficient for networks to price calls at (perceived) marginal cost and for them to compete over the level of the rental charge. However, this may change as we relax the assumption that all subscribers are homogenous, with the same demand to make calls once they have joined a network characteristics. If subscribers differ, either in terms of the volume of calls they would make at a given call price, or in terms of the volume of calls they receive, then it no longer is the case that it will be efficient for networks to price calls at marginal cost.

This is an aspect of pricing dealt with by Dessein (2001) and by Valletti and Houpis (2005). The specific insight that these papers bring is that they show that when the marginal subscriber makes fewer calls than the average caller then it will be efficient to price calls above marginal cost and reduce rentals.

Both Laffont, Rey and Tirole (1998a, 1998b) and Armstrong (1998) argue (without formal proof) that once customers are heterogeneous in their consumption and access prices differ from marginal cost then the market outcome may resemble the collusive outcome created by linear pricing, even if two-part tariffs are used in practice. Dessein (2001), however, demonstrates that this is not always the case. He shows that the profit neutrality result holds even in the presence of customer heterogeneity. Moreover, he extends Schiff's (2001) result by showing that in the presence of customer heterogeneity and externalities, networks would choose to price access below marginal cost while welfare is maximised by pricing access above marginal cost. Valletti and Houpis (2005) note specifically that results are sensitive to the way in which heterogeneity is modelled. If the differences between subscribers are additive then marginal cost pricing will remain efficient, while other formulations tend to result in pricing calls above marginal cost.

Schiff's paper is a good example of this. An endogenous participation rate is explained by customers having an "option value" from subscription which is randomly distributed, but unrelated to the calls they make if they become subscribers because of the additive structure that Schiff has chosen. Hence in Schiff's model, even in the presence of externalities, the marginal customer makes the same number of calls as the average customer, so the conditions for marginal cost pricing still hold. By contrast, in Dessein (2001)'s model, customers are split into low and high calling (and receiving) behaviour. Inevitably marginal customers are drawn from the low-use group. In these circumstances it becomes efficient to raise call charges above marginal cost and lower rental charges.

3.4.2 Operators

The basic model and its extension including call and network externalities assume that operators are horizontally differentiated. There are only few models that consider further asymmetries among operators. Therefore, it is still unclear how the introduction of asymmetries among operators may affect market outcomes.

Baranes and Flochel (2004), which was described above in the context of call externalities, depart from the standard assumption of horizontal differentiation and consumers' homogeneity, by considering vertical differentiation a la Mussa and Rosen (1978) and heterogeneous consumers. In this context the high quality operator has higher incentives to distort (upwards) both the off-net price and the access price. The reason is that in the (separating) equilibrium characterized by Baranes and Flochel (2004) high quality consumers adopt the high quality network whereas low quality consumers subscribe to the low quality network. This implies that in equilibrium both operators face slightly different conditions in order to keep their targeted consumers. In particular, the high quality network has to set prices such that high type consumers are better off subscribing to the network than (1) staying out of the market (individual rationality constraint) and (2) subscribing to the low quality network. The relevant constraint for the low quality network is only the first (individual rationality).

As a consequence, the low quality network does not have incentives to distort the off-net price upwards in order to reduce the call externality faced by high quality customers. Its only incentive to raise the MTR comes from the maximisation of access revenues. By contrast, the high quality network has an additional effect when deciding its off-net prices and access charge. By increasing off-net prices it reduces the positive externality received by clients of the lower quality network.

Carter and Wright (2003) combine vertical and horizontal differentiation. In their model callers do not take into account the different on-net and off-net prices and base their decisions on a weighted-average price. In such scenario, a reciprocal termination charge above cost benefits the smaller operator relative to the larger one.

The key to understanding this result is that since smaller firms face a bigger proportion of off-net calls, above cost MTRs, make their customers face higher per-minute prices. This means that their callers will tend to call less than callers on the bigger network. Hence, there is an outflow of calls from the bigger network which generates an access deficit.

Section 4 describes some more papers on asymmetry in the context of the potential anti-competitive effect raised by off/on-net price differentials.

4 On-net/off-net price differential as an anti-competitive tool

There are a number of contributions that focus on the potential use of on-net/off-net price differentials to distort competition. Below we comment two well-known papers are those of Hoernig (2007) and Calzada and Valletti (2008). We will also cover the recent contribution by Lopez and Rey (2008).

In Hoernig (2007) a large and a small operator compete for mobile customers in the presence of call externalities. In equilibrium, either with linear or two-part tariffs, the off/on-net price differential increases with the termination charge. The off-net equilibrium price under two-part tariffs is as in Jeon, Laffont and Tirole (2004) shown above:

$P^{i}_{off\text{-}net} = (1 - s_i)(C_0 + a)/(1 - s_i(1 + \beta))$, for operator i, with market share si

The expression above indicates that the off/on-net price differential is mainly driven by the presence of call externalities, represented by parameter β. Even if the MTR is zero (B&K) the on-net/off-net differential would be present.

It is also important to emphasize that the above strategy is not anti-competitive, as it is not designed to damage rivals regardless of one's own profitability. Rather the on-net/off-net price differential arises as the optimal profit-maximising strategy for each operator, small or large, when individually setting its own prices.

The paper also considers the possibility that the large firm engages in predatory behaviour by increasing (decreasing) the off/on-net price differential above (below) the equilibrium level if the termination fee is above (below) cost. Hence, the predatory outcome is not conditional upon the MTR being above cost or even positive, as **the crucial element is the existence of call externalities** rather than a positive MTR. Hoernig (2007) does not analyze the welfare maximizing MTR.

Gabrielsen and Vagstad (2008) focus on the role of access charges to create tariff-mediated network externalities and induce a collusive outcome. The paper challenges the contributions of Gans and King (2001), and Laffont et. al. (1998b) questioning the incentives of operators to set low termination charges. Gabrielsen and Vagstad show that, in a setting where operators can create tariff-mediated network externalities, there are exogenous switching costs and calling clubs , then a high access charge increases the perceived switching costs which allows operators to impose a higher subscription price. In their model efficiency dictates cost-based MTRs since they do not consider neither call nor network externalities.

Calzada and Valletti (2008) analyze the question of whether incumbent operators may coordinate on a high reciprocal access charge in order to deter entry. The key for this strategy is that in their model a high access charge reduces profits for all operators (not only for the smaller ones), which makes entry less attractive. **If entry took place, incumbent operators would have incentives to reduce**

access charges, even below cost, resulting in higher profits for the whole industry. Thus, the strategy to set high termination charges in order to deter entry is only rational from an ex-ante view point. If allowed to re-negotiate, incumbent operators would reduce MTRs once entry has taken place.

Lopez and Rey (2008) analyze a related question, whether an incumbent operator may strategically set the MTR in order to prevent entry. In this model all consumers are initially attached to the incumbent operator who faces potential competition from an entrant. The incumbent and the entrant are horizontally differentiated *à la* Hotelling and there are switching costs, which creates a certain preference for the incumbent operator. In this paper consumers do not experience either network or call externalities. However, tariff-mediated network externalities may be created by the on-net/off-net price differential.

Lopez and Rey (2008) find that under certain conditions the incumbent operator may impose an access mark-up in order to make entry unprofitable. However, the paper also finds the opposite, i.e., market foreclosure by subsidizing termination.[36] Further, the model presents a number of **temporal inconsistencies**:

- It is strange that a monopolistic mobile operator uses the interconnection charge in order to deter entry. In the pre-entry stage, the access charge is redundant since there is only one active network. Thus, it is difficult to understand how the entrant observes the access charge.
- One can suppose that the access charge may be announced by the incumbent once the entrant has communicated its entry into the market. Given the entry barriers existing in mobile markets it is not credible that an operator that has invested in spectrum is going to exit the market just because the existing monopolist threats with a high MTR.
- Moreover, once the new operator has entered the market the incumbent would have an incentive to reduce the access charge, since a high access charge is only profitable provided it successfully keeps the entrant out of the market.[37]

36 Sse proposition 4 of Lopez and Rey (2008).

37 This point is emphasized in the abstract of the paper.

5 Implications for B&K

In the above sections we have provided a descriptive overview of the literature on MTRs. This section considers the implications of these models for B&K.

5.1 B&K IS UNLIKELY TO BE EFFICIENT

As we have seen above, departures from cost-based pricing are only justified by the existence of call or network externalities. If there are call externalities and these are not internalized in other ways,[38] sharing the total costs of the call between the called and the calling party (i.e. RPP) becomes the efficient pricing mechanism at the retail level. In this case, optimal call termination prices would be below cost, with B&K (i.e. MTRs equal to zero) optimal under very specific assumptions. For example, in the simple context considered in section 3.2.1, a zero MTR is efficient when the ratio of the cost of termination to the cost of originating the call equals the ratio between the recipient and the caller's valuation of a call as shown in Figure 1.

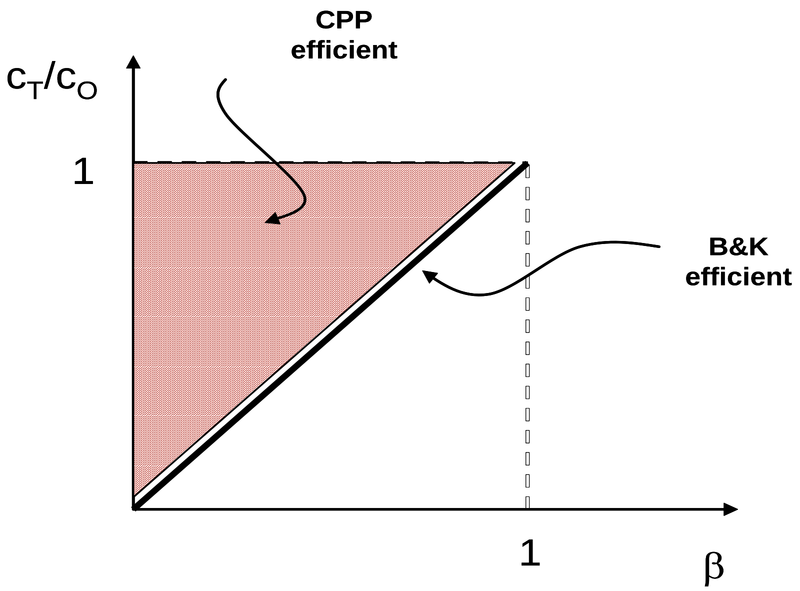

Figure 1: Efficiency of B&K

C_T/C_O is the ratio of the cost of termination to the cost of origination

Source: Frontier Economics

Thus, the optimality of B&K requires information on origination and termination costs and on the average relative valuation of a call for both the calling and called parties. B&K cannot therefore be justified solely on the existence of call externalities.

---

38 For example, through reciprocal calling patterns as considered in Cambini and Valletti (2008).

There is not much public information regarding the importance of call externalities. A study by Ofcom in 2005[39] showed that, in their decision on network subscription, consumers do not assign much value to the possibility of being called. Only 2% of responders considered the price of others to call them in their choice of the network. This evidence suggests a low call externality.

In a recent paper, Sandbach and Van Hooft (2008) test the empirical importance of residual call externalities[40] by matching the predictions of the models[41] with data on retail prices. They find that the estimated size of the call externality is small, not being statistically different from zero.

In addition, while B&K may reduce some transaction costs (for instance the need to bill for interconnection), it also creates other costs. For instance, in order to avoid the "hot potato" problem (i.e. the incentive of the initiating network to deliver the call at the point of interconnection closest to the originating customer) the regulator may need to specify the interconnection points and set regulated charges in the case traffic is delivered at different locations.

B&K may also generate an inefficiently high level of traffic, which could even generate negative call externalities. This is because low termination rates and low off-net call prices help proliferation of certain type of calls which actually harm consumers (for instance marketing calls or SPAM[42]). In this respect, mobile customers in the US have recently filed a lawsuit against 6 mobile-phone carriers and a top mobile virtual operator in Mississippi federal court due to the imposition of charges for unsolicited messages received by subscribers[43].

5.2 B&K MAY DAMPEN COMPETITION

A further feature identified in the economic literature on MTRs is that it has been found that the intensity of competition among existing operators may be affected by the level of the MTR, because of the impact that MTRs may have on the profitability of marginal customers, therefore, on retail prices.

In section 2.3 we have seen that under quite general assumptions, a decrease in MTRs is going to result in higher retail prices for other mobile services (e.g. subscription fees in case of two-part tariffs or on-net prices in case of linear tariffs with network discrimination). Since B&K implies a reduction of MTRs a movement towards this system will probably result in higher subscription/other retail charges and lower consumer welfare.

---

39 Ofcom (2005), Annex F.

40 "The external value which the parties themselves cannot internalise", Sandbach and Luke van Hooft (2008), page 3.

41 They consider an extension of Armstrong and Wright (2007) allowing for the existence of limited calling circles.

42 A study by Ofcom carried out in 2003 found that 36% of mobile subscribers at least occasionally chose not to answer calls from an unrecognized or unidentified source (Ofcom, 2003. Page 10).

43 See RCRWireless News. May 16, 2008.

In this sense, Genakos and Valletti (2008) have empirically analysed the effect of MTRs on the cost of representative usage bundles using a cross-country panel database, finding that "*a regulated percentage reduction in fixed-to-mobile termination rates is associated with an almost equal percentage increase in the expenditure necessary to buy a given usage bundle.*"[44]

---

44 Schiff (2007), page 412.

6 Conclusions

In this report we have provided a broad overview of the literature on MTRs and analysed its implications for B&K. Our conclusions can be summarised in the following points:

- In the *basic model*: the efficient MTR is cost oriented but a high MTR will intensify competition as long as operators price discriminate among on-net and off-net calls. And a below cost MTR may dampen retail competition and damage the consumers' welfare.

- The introduction of *call and or network externalities* make the efficient MTR depart from costs. Network externalities increase the efficient MTR whereas call externalities ask for a reduction in the access charge. In this context, B&K is efficient only under very specific conditions that require detailed information about the size of call externalities.

- It is important to consider only call externalities that are not internalized through *reciprocal communication patters*. Otherwise, estimated call externalities will be biased upwards.

- Recently, a number of papers have emerged analyzing the use of *on-net/off-net price differentials* as a way to distort competition in a asymmetric context. These papers show that such differentials may exist even if there are not interconnection payments (B&K). The evidence of the USA, where off-net/on-net price differentials are observed in a B&K context, supports this result.

- There are also some contributions focused on the *potential role of MTRs as an instrument for entry deterrence*. Nevertheless, these models do not show that B&K is efficient. They also present an inconsistency problem: high MTRs are not commercially possible unless they guarantee exclusion. If entry took place, incumbent operators would have incentives to re-negotiate the access charge.

7 References

Armstrong, M. (1998) "Network Interconnection in Telecommunications," *Economic Journal*, 108: 545-564.

Armstrong, M. (2002): "The Theory of Access Pricing and Interconnection", *Handbook of Telecommunications Economics*, Volume 1, Edited by M.E. Cave et. al. Elsevier Science B.V.

Armstrong, M. and Wright, J. (2007): "Mobile Call Termination", September 2007.

Armstrong, M. and Wright, J. (2008): "Mobile Call Termination", October 2008.

Baranes, E. and Flochel, L. (2004): "Competition in networks with call externalities", mimeo.

Berger, U. (2004). "Access Charges in the Presence of Call Externalities". Mimeo, Vienna University of Economics.

Berger, U. (2004): "Access Charges in the Presence of Call Externalities". *Contributions to Economic Analysis & Policy* 3.1: Article 21.

Berger, U. (2005): "Bill-and-Keep vs. cost-Based Access Pricing Revisited". *Economics Letters*, 86(1), 107-112.

CRA and Gilbert Tobin (2007): *Economic study on IP interworking*

Calzada, J. and T. Valletti (2008): "Network Competition and Entry Deterrence", *Economic Journal*.

Cambini, C. and Valletti (2008): "Information exchange and competition in communications networks", *Journal of Industrial Economics*. Vol 56: 707 – 728.

Carter, M. and Wright, J. (2003): "Asymmetric Network Interconnection", *Review of Industrial Organization* 22: 27-46.

DeGraba, P. (2000): "Bill and Keep at the Central Office as the Efficient Interconnection Regime", OPP Working Paper Series No.33, December, FCC.

DeGraba (2003): "Efficient Inercarrier Compensation for Competing Networks when Customers Share the Value of a Call", unpublished version.

Dessein, W. (2001): "Network Competition in Non-Linear Pricing".

Elliott, D. (2004): "Two-way access charges and on-net/off-net differentials: What are the incentives on operators and can large MNOs use high access chargers to foreclose the market?", Frontier Economics, October 2004.

Gans, J. and S. King (2001): "Using "bill and keep" interconnect arrangements to soften network competition", *Economics Letters*, vol. 71: 413-421

Genakos and Valletti, T. (2008). "Testing the 'Waterbed' Effect in Mobile Telephony", *CEIS Tor Vergata, Research Paper Series*, vol. 6, Issue 2, No. 110.

Harbord, D. and M. Pagnozzi (2007a): "On-Net/Off-Net Price Discrimination and 'Bill and Keep' vs. 'Cost-Based' Termination Pricing", European Mobile Challengers' Workshop, 8 November 2007.

Harbord, D. and M. Pagnozzi (2007b): "On-Net/Off-Net Price Discrimination and 'Bill-and-Keep' v. 'Cost-Based' Regulation of Mobile Termination Rates", mimeo.

Hoerning, S. (2007): "On-Net and Off-Net Pricing on Asymmetric Telecommunications Networks," *Information Economics & Policy*, 19(2): 171-188.

Jeon, D.S., Laffont, J.-J. and Tirole, J. (2004). "On the 'Receiver-Pays' Principle", *RAND Journal of Economics*, vol. 35: 85-110.

Julien, B. and Rey, P. (2008): "Notes on the Economics of Termination Charges", IDEI Report # 6.

Julien, B., Rey, P. and Sand-Zantman, W. (2008), "Mobile Call Termination revisited: Consumers Heterogeneity and Calling Pattern", mimeo, Toulouse School of Economics.

Laffont, J.-J., P. Rey and J. Tirole (1998a): "Network Competition I: Overview and Non-discriminatory Pricing", *RAND Journal of Economics* 29: 1-37.

Laffont, J.-J., P. Rey and J. Tirole (1998b): "Network Competition II: Price Discrimination", *RAND Journal of Economics* 29: 38-56.

Littlechild, S. (2006): "Mobile Termination Charges: Calling Party Pays versus Receiving Party Pays," *Telecommunications Policy*, 30(5-6): 242-277.

Lopez, A. and Rey, P. (2008), "Foreclosing Competition through Access Charges and Price Discrimination", mimeo, Toulouse School of Economics.

Mussa M., and S. Rosen (1978): "Monopoly and Product Quality", *Journal of Economic Theory* 18, 301-317.

Ofcom (2003): "Consumers' Use of Mobile Telephony". 14 August 2003. 27 October 2003.

Ofcom (2005): "Wholesale mobile voice call termination markets – a proposal to modify the charge control conditions". Explanatory Statement and notification of proposals. June 2005.

Ordover J. (2008): "Recovering fixed and common costs for mobile networks in Europe", Compass Lexecon

Sandbach, J. and Van Hooft, L. (2008): "Using on-net/off-net price differential to measure the size of call externalities and its implications for setting efficient MTRs"

Schiff, A. (2001): "Two-way Interconnection with Partial Consumer Participation", University of Auckland, Working Paper # 223.

Schiff, A. (2007). "The 'Waterbed' Effect and Price Regulation", *Review of Network Economics*, vol.7, Issue 3.

Taylor, W. (2002): "Customer Demand Analysis", *Handbook of Telecommunications Economic*s. Vol. 1, Ed. by M. Cave et al. 2002. Elsevier Science B.V.

Valletti, T. and Cambini, C. (2005): "Investments and Network Competition", *RAND Journal of Economics*, Vol. 36, No.2: 446-467.

Valletti, T. and G. Houpis (2005): "Mobile Termination: What is the 'Right' Charge?", Journal of Regulatory Economics, Vol: 28, Pages: 235 - 258

Wright, J. (2002a): "Access Pricing under Competition: An Application to Cellular Networks," *Journal of Industrial Economics*, 50: 289-315.

Wright, J. (2002b): "Bill and Keep as the Efficient Interconnection Regime?" *Review of Network Economics*, 1(1): 54-60.

FRONTIER ECONOMICS EUROPE
BRUSSELS | COLOGNE | LONDON | MADRID

Frontier Economics Ltd 71 High Holborn London WC1V 6DA
Tel. +44 (0)20 7031 7000 Fax. +44 (0)20 7031 7001 www.frontier-economics.com