Annex 5 to Three's response to PSSR auction consultation Jan 2017

The Need to Limit Disparities in Spectrum Holdings

Jon M. Peha[1]

Executive Summary

Now that traffic volumes are increasing rapidly, cellular carriers must increase capacity to remain competitive, and the cost of expanding capacity has become a large portion of expenditures. There are strong economies of scale when expanding capacity, because a carrier with more spectrum benefits more from every new cell tower, and a carrier with more towers benefits more from every new MHz of spectrum. While it is technically possible to expand capacity by increasing either towers or spectrum holdings, the cost-effective approach is to increase both types of assets at a similar rate. In the absence of countervailing policies, the big carriers will get bigger, in terms of spectrum holdings, towers, and ultimately market share.

For policymakers, this economy of scale creates a trade-off between two important objectives: reducing the cost of cellular capacity, and increasing competition. This paper shows that any division of spectrum that is Pareto optimal with respect to these two competing objectives will split the spectrum fairly evenly among competing carriers, regardless of how many or few competitors there are. A large disparity in spectrum holdings may yield poor results with respect to both objectives, i.e. the lower cost-effectiveness of a larger number of carriers, and the lower competitive pressure of a smaller number of carriers. Spectrum holdings in the U.K. today exhibit this kind of disparity.

One effective way to achieve a division of spectrum that is close to Pareto optimal is a spectrum cap. This cap should be consistent with policymakers' objectives for competition. Ofcom currently proposes a spectrum cap of 42%, and proposes to exclude the 3.4 GHz band from the cap. This proposed policy is *not* consistent with Ofcom's stated objective of having four credible competitors.

[1] Professor Jon M. Peha, Carnegie Mellon University, peha@cmu.edu, www.ece.cmu.edu/~peha

Author Biography

Jon Peha is a Professor at Carnegie Mellon University, with experience in industry, government, and academia. In government, he served at the U.S. Federal Communications Commission (FCC) as Chief Technologist, in the White House as Assistant Director of the Office of Science & Technology Policy (OSTP), on legislative staff in the Energy & Commerce Committee of the U.S. House of Representatives, and at the U.S. Agency for International Development (USAID) to launch the Telecommunications Leadership program. In industry, he has been Chief Technical Officer for three high-tech companies, and member of technical staff at SRI International, AT&T Bell Labs, and Microsoft. At Carnegie Mellon University, he is a Professor in the Dept. of Electrical & Computer Engineering and the Dept. of Engineering & Public Policy, and former Associate Director of the Center for Wireless & Broadband Networking. Dr. Peha holds a PhD in electrical engineering from Stanford. He is Fellow of the Institute for Electrical and Electronics Engineers (IEEE), a Fellow of the American Association for the Advancement of Science (AAAS), and was selected by AAAS as one of 40 Featured Science and Technology Policy Fellows of the last 40 years ("40@40"). Dr. Peha has received the FCC's "Excellence in Engineering Award," the IEEE Communications Society TCCN Publication Award for career contributions, and the Brown Engineering Medal. He consults on a wide range of technical and policy issues related to information and communications technology.

1 Introduction

Ofcom has recently justified spectrum policies regarding competition [OFCO16b] using an earlier conclusion that cellular carriers need to have 10-15% of the spectrum that is available to cellular carriers to be "credible." This conclusion was based on the observation that as of January 2012, the smallest nationwide carrier in many countries had about 10% of cellular spectrum [OFCO12a, OFCO12b]. 10% would certainly be sufficient in a world where customers choose their provider primarily based on coverage rather than capacity. For someone who uses her cell phone primarily to make telephone calls, coverage is extremely important. Driving through even a small hole in coverage can terminate an important conversation. However, in recent years, customers have become more concerned about capacity, thanks to the rise of smartphones and the high-data-rate applications they enable, and this effect becomes more important with every passing year. The capacity required to meet the demands of cell phone users is doubling worldwide every 20 months [CISC16] and in the U.K. every 16 months [OFCO16b]. As a result, expanding macrocellular capacity has become a major driver of annual expenditures for cellular carriers. This will make it hard for a carrier with 10% of available spectrum to compete with carriers that have far more than 10%. Policy should keep up with this changing reality.

As this paper will show, this need to rapidly increase capacity has yielded new economies of scale. The largest carrier can expand its capacity at a lower cost than its smaller competitors, which means that the large are likely to get larger. The fact that scale decreases costs has advantages for users of cellular services, as long as the large carriers pass these cost savings along to their customers. However, as the large carriers increase market share, there will be less competition, giving carriers less incentive to lower prices and improve quality of service. Moreover, the carrier with the lowest cost per GB of data carried maximizes profits be setting prices based on the higher costs of its competitors rather than its own costs.

This paper explores how best to divide spectrum resources among carriers given this economy of scale. There are two potentially competing objectives for policymakers: increasing competition and lowering the cost of capacity. This paper will show that in any result that is Pareto optimal with respect to these two objectives, spectrum is divided fairly evenly among cellular carriers, regardless of the number of competing carriers. Large disparities in spectrum holdings are therefore not in the public interest. One simple way to divide spectrum in a market economy in a way that is close to Pareto optimal is through some form of spectrum cap.

The analysis in this paper assumes that there is no market failure in spectrum, so that spectrum is available at the price where supply meets demand. This means we assume that carriers are not engaging in a strategy of "foreclosure," even though a common argument for spectrum caps is as a means of preventing foreclosure [ERGA98, BAKE07, CRAM11, DOJ13] rather than addressing an economy of scale. Because there is a limited amount of spectrum available to cellular carriers and few opportunities to obtain spectrum from sources other than a direct competitor, a carrier that is determining the price at which it will buy or sell spectrum "will include in its private value not only its use value of the spectrum but also the value of keeping the spectrum from a competitor" [CRAM11]. As defined by the U.S. Department of Justice, "the latter might be called 'foreclosure value' as distinct from 'use value.' The total private value of spectrum to any given provider is the sum of these two types of value" [DOJ13]. If some carriers do consider foreclosure value in their transactions, this would only increase the need for policies that limit large disparities in spectrum holdings.

This paper is organized as follows. Section 2 will describe the most important assumptions of our analytic model with respect to both the cost and capacity of cellular infrastructure. Section 3 explores the resource decisions that would maximize profit for cellular carriers. This section shows how carriers reduce cost by balancing spectrum acquisitions and tower acquisitions, and the economies of scale for both. Section 4 explores the resource decisions that would serve the public interest, which might involve lowering the cost of capacity, increasing competition, or both. This section derives the Pareto optimal strategies with respect to these two competing objectives. While Sections 3 and 4 assume that one carrier's cost per cell tower does not depend on the choices of that carrier's competitors, which is sometimes but not always the case, Section 5 relaxes that assumption by considering the effects of explicit tower-sharing arrangements between carriers. We discuss these issues in the context of the United Kingdom cellular market today in Section 6. Finally, we summarize conclusions and discuss their policy implications in Section 7.

2 Model Assumptions

This section presents some of the most fundamental assumptions underlying our analysis, and the implications of any simplifications made. Section 2.1 describes our assumptions about how carriers maximize profit. Section 2.2 explains our focus on capacity-limited macrocells in both cost and capacity calculations. Sections 2.3 and 2.4 discuss key assumptions in the cost and capacity analyses, respectively.

2.1 Profit Maximization

This paper assumes that carriers are in an equilibrium state where profit has been maximized, e.g. where cost is minimized for a given capacity or capacity is maximized for a given cost, and where costly resources have not been spent to increase capacity that is not (yet) needed to carry customer traffic. This simplification ignores the unique history that produced each carrier's infrastructure. For example, immediately after a merger, a carrier's infrastructure and spectrum holdings are unlikely to be as cost-effective as possible, but the carrier will strive to move towards a maximally cost-effective state over time. This simplification also ignores dynamic elements. For example, where our assumptions might cause a carrier to gradually obtain spectrum and gradually build towers to meet the gradually-increasing customer demand for capacity, a real carrier might occasionally obtain and use a large block of spectrum that is expected to meet both current needs and anticipated needs for the next few years, thereby creating capacity that is temporarily unneeded. Thus, at any instant in time, one carrier may have more excess capacity than another. Over the long term, however, these temporary effects should have little impact.

We also assume that all carriers are deriving similar revenues per unit of capacity. Thus, if two carriers have the same capacity but one has much higher cost, the high-cost carrier will be less profitable. Moreover, this means that a carrier's share of overall cellular capacity should be similar to its share of the overall cellular market, at least in regions where there is no excess capacity that is not in use.

2.2 Capacity-Limited Macrocells

In both our cost analysis and our capacity analysis, we consider a region where all cellular carriers are capacity-limited, i.e. where a carrier that merely deploys the minimum number of cell towers to bring adequate coverage to the region will not have enough capacity. Of course, there are also many sparsely-populated regions that are coverage-limited rather than capacity-limited, i.e. where simply employing the minimum number of towers to bring adequate coverage will also provide more than enough capacity. We assume that a region that is coverage-limited for one carrier is coverage-limited for all carriers, and that the value of spectrum in these sparely-populated regions is negligible compared to the value of spectrum in the capacity-limited regions. Thus, we do not count the cell towers deployed in coverage-limited regions in our analysis. Although these assumptions might not be valid for a new company wishing to enter the cellular market by building an entirely new cellular infrastructure, and therefore in need of spectrum in coverage-limited as well as capacity-limited regions, these assumptions are appropriate when the entities seeking spectrum are nationwide cellular carriers that are trying to expand their existing capacity.

We also consider only macrocells in our cost and capacity analysis. The load on macrocells can be reduced by offloading traffic to inexpensive short-range devices, such as Wi-Fi hotspots, residential femtocells, and even roadside units that provide Internet access to users in moving cars via DSRC [LIGO15]. However, because these devices are short-range, they can only cover a small fraction of the capacity-constrained regions. Even as the volume of traffic offloaded to such devices is increasing, the volume of traffic that cannot be offloaded is also increasing. For the foreseeable future, cellular carriers must continue to expand the capacity of their macrocells to meet user needs and expectations, and this capacity expansion will be a large part of the carriers' expenditures. Thus, the assumptions underlying our analysis are valid even with offloading.

Finally, we assume that the macrocellular networks operated by all of the carriers are comparable in the sense that they use a similar mix of technologies and frequencies, and therefore have a similar spectral efficiency, and a similar number of sectors per cell. Of course, carriers upgrade their technology at different times, so at any given time this may not be the case. For example, in the U.K. today, cellular carrier Three has no 2G technologies in use while its competitors do, so Three has an advantage with respect to spectral efficiency. However, advantages of this kind are temporary, and tend to benefit different carriers in different years. They should not matter when one is considering the long-term sustainability of competition.

2.3 Capacity

We calculate capacity in a way that is traditional for a capacity-limited cellular network, where a carrier's capacity increases linearly with the number of cell towers that the carrier uses, and linearly with the amount of spectrum it holds. Technology is also a factor, including whether it uses second, third, or fourth generation cellular technology.

While it is possible to increase macrocellular capacity by increasing either the number of cell towers or the amount of spectrum, the two do not actually have identical impact as we assume here. The advantages of increasing spectrum are actually somewhat greater than our traditional model would imply. As a result, our analysis may underestimate the disadvantages of carriers with especially low spectrum holdings.

One reason is that resources can be allocated to individual mobile devices in a way that takes advantage of multiuser diversity [CAPO13]. If all devices in a cell were at a single location, then as Shannon's theorem dictates, achievable downstream capacity would increase linearly with downstream spectrum bandwidth, and would depend on SINR at that specific location. However, devices are not all in the same location. Path loss and interference vary from location to location and frequency to frequency, and consequently so does SINR. Thus, device 1 may achieve greater throughput using frequency f_1 than using frequency f_2, while device 2 achieves greater throughput with frequency f_2 rather than f_1. Algorithms that take advantage of this will increase throughput. When such algorithms are used, as is the case with LTE technology, capacity increases with spectrum bandwidth at a rate that is more than linear. The same cannot be said when increasing the number of cell towers.

2.4 Cost

There are two principal costs associated with cellular infrastructure: spectrum and cell towers. Carriers can expand capacity by increasing expenditures on either resource. To make them comparable, we consider the net present value of all such costs over time. Thus, we do not distinguish one-time costs from annual costs, as both are accounted for.

We assume that all carriers can obtain spectrum, and at the same price per MHz, as would be appropriate in the absence of market failure. This may not always be the case in today's spectrum market, especially if some players pursue a foreclosure strategy. Thus, this assumption somewhat understates the benefits of holding spectrum in reserve.

We consider three categories of tower costs. Some tower costs are constant, regardless of spectrum holdings or data rates. This category includes the cost of building a new tower, rental charges for placing a multiband antenna on an existing tower, and the capital expenses for a backup generator. Many tower costs fall in this category. A second category is costs that depend on spectrum holdings, but not data rates. For example, adding a new spectrum band requires additional equipment, with the associated capital and operating expenses. Also, increasing total spectrum bandwidth means increasing the number of power amplifiers and associated operating costs. We assume that costs in the second category increasing linearly with spectrum bandwidth. The third category of costs depend on data rates. The obvious example is backhaul. While a large fraction of backhaul costs fall in the first category of fixed costs, there are also backhaul costs that increase with data rate. We assume that the latter increase linearly, i.e. backhaul costs for a tower are *a*+ *b**(data rate), for some constants *a* and *b*. (Note that if the cost increase that depends on data rate is sub-linear due to even steeper volume discounts, this means that there is even more benefit to having more spectrum and therefore greater data rate per tower as compared to more towers.)

In cases where towers are shared among cellular carriers, it also matters how the costs are shared. In this paper, we employ two different models. In the first model, which we explore in Sections 3 and 4, we assume that one carrier's tower costs depend only on the choices of that carrier, and not the choices of its competitors. This is obviously the case if each carrier owns and operates its own towers. It can also be the case if carriers primarily rent space on towers that are owned and operated by third-party providers. In the second model, which we explore in Section 5, we assume that each carrier establishes a cooperative relationship with one of its competitors for the sharing of towers. In this model, a carrier wanting more towers can work with its partner to add shared towers, or can work alone to deploy its own towers, depending on which strategy best serves its own interests.

3 Strategies to Meet Carrier Objectives: Economies of Scale

In this section, we present the strategies that best serve carriers. We assume that carriers seek to maximize their profit, which means that they will accumulate the cell tower and spectrum assets that minimize cost for whatever capacity that infrastructure provides.

More specifically, each Carrier i wishes to provide capacity C_i Mb/s per square km using S_i MHz of spectrum, where capacity and spectrum holdings may differ from carrier to carrier. Carrier i deploys N_i cell towers per square km in the capacity-limited regions. (Towers in the coverage-limited regions are excluded.) We assume all carriers have a spectral efficiency averaged throughout their cells of e bps/Hz, as would be appropriate if they all employ a similar mix of technologies (e.g. 2G, 3G, and 4G cellular) and a similar mix of frequencies. (The value of spectrum can also depend on its frequency, but we ignore frequency-dependent differences in this analysis.) We assume that all carriers have r sectors per cell, and a frequency reuse of f.

$$C_i = N_i \, r \, e \, S_i / f$$

As discussed in Section 2.4, most infrastructure costs are simply proportional to the number of towers, but we consider three categories of tower costs. Let T_0 be the net present value (NPV) of the fixed costs of deploying and operating a tower, so it includes both CAPEX and OPEX. Let T_{bw} be the NPV of costs that are proportional to bandwidth, and T_{bps} be the NPV of costs that are proportional to data rate. Let M be the cost per MHz divided by the area of the capacity-limited regions. We can then derive carrier i's cost per square km K_i as a function of its spectrum holdings S_i.

$$K_i = M\,S_i + N_i\left(T_0 + T_{bw}\,S_i + T_{bps}\,\frac{r\,e\,S_i}{f}\right)$$

If Carrier i is rational, it will choose the amount of spectrum and number of towers to minimize cost K_i for a given capacity C_i. By combining these equations, we get the following.

$$K_i = M\,S_i + \frac{C_i\,f\,T_0}{r\,e}\,\frac{1}{S_i} + \frac{C_i\,f\,T_{bw}}{r\,e} + C_i\,T_{bps}$$

$$0 = \frac{dK_i}{dS_i} = M - \frac{C_i\,f\,T}{r\,e}\,\frac{1}{{S_i}^2}$$

$$S_i^{\,2} = \frac{f\,T_0}{M\,r\,e}\,C_i$$

This shows that capacity does not increase linearly with spectrum holdings as many people believe when a carrier minimizes its costs. In this case, capacity increases with the square of spectrum holdings. This is because a rational carrier will increase spectrum holdings and the number of towers in the capacity-constrained region together. Indeed, when S_i and N_i are selected to minimize Carrier i's costs, the number of towers in the capacity-limited regions is proportional to spectrum holdings, as shown by the equation below. (Note that this contrasts with the coverage-limited regions, where the number of towers depends highly on whether a carrier holds spectrum at low frequencies, but depends little on how much spectrum each carrier has.)

$$N_i = \frac{M}{T_0}\,S_i$$

The fact that capacity increases with the amount of resources squared shows that there is a large economy of scale. The carrier with the most towers will benefit the most from another 10 MHz of spectrum, and the carrier with the most spectrum will benefit the most from another tower. As a result, a carrier with more spectrum has a smaller cost per capacity, as shown by the following equation. Thus, if a large portion of overall cost comes from providing adequate capacity, as might be expected when utilization per user is increasing rapidly, it will be hard for the carriers with less spectrum to compete when there is a large disparity in spectrum holdings. For example, the market price might make the cost of increasing capacity less than the revenue derived from that additional capacity for the large carriers, making expansion profitable. That same market price might make the cost of expansion less than the revenue derived from that capacity for small carriers, making expansion too expensive to contemplate. In this case, the small carriers will stop growing. With data rate per user doubling every 16 months, a carrier that does not expand will rapidly lose market share. Eventually, it may be acquired by a rival, go bankrupt, or simply become a small niche player that is no longer a competitive factor.

$$\frac{K_i}{C_i} = \frac{M\,S_i}{C_i} + \frac{f\,T_0}{r\,e}\,\frac{1}{S_i} + \frac{f\,T_{bw}}{r\,e} + T_{bps} = \frac{f\,T_0}{r\,e}\,\frac{1}{S_i} + \frac{f\,T_0}{r\,e}\,\frac{1}{S_i} + \frac{f\,T_{bw}}{r\,e} + T_{bps}$$
$$= \frac{2\,f\,T_0}{r\,e}\,\frac{1}{S_i} + \frac{f\,T_{bw}}{r\,e} + T_{bps}$$

Moreover, once a large disparity exists, this disparity is likely to grow, because the carrier with more spectrum is willing to pay more for additional spectrum. If revenues are roughly proportional to capacity, then the value to Carrier *i* of additional spectrum is $\frac{dC_i}{dS_i}$, which increases linearly with Carrier *i*'s spectrum S_i as follows.

$$\frac{dC_i}{dS_i} = \frac{2\,M\,r\,e}{f\,T_0}\,S_i$$

4 Strategies to Meet Policy Objectives: Competition vs. Capacity

In this section, we present the best strategies for spectrum policymakers who are tasked with serving the public interest. This is reflected in two different objectives. First, for a given set of resources dedicated to cellular carriers, the public interest is served when these resources are used to make a great deal of capacity available to cellular users. Second, the public interest is served when competition among cellular carriers is high, as this gives profit-seeking carriers incentive to offer better services at lower prices.

More specifically, we assume that the total amount of spectrum S_{tot} and the total number of towers N_{tot} across all carriers is fixed. S_{tot} describes the resources allocated directly by the spectrum policymaker. N_{tot} describes the investment from cellular carriers. Whereas in Section 3 we assumed that carriers seek to minimize the cost for achieving a certain capacity, in this section we assume that they seek to maximize the capacity for a given amount of spectrum and number of towers, which is roughly the same as maximizing capacity for a given cost. As discussed in Section 2.4, most tower costs are fixed, and therefore are directly proportional to N_{tot}. Even more importantly, carriers would be happy to reduce the costs that are fixed (as reflected in T_0), but would not want to reduce the costs that are proportional to the capacity that they are struggling to increase (as reflected in T_{bw} and T_{bps}).[2] Thus, the number of towers is an appropriate measure of non-spectrum costs.

We seek the allocation of these spectrum and tower resources that optimizes the spectrum policymaker's objectives: total capacity C_{tot}, which is the sum of capacities of all cellular carriers, and cellular competition as measured using the Herfindahl–Hirschman Index (HHI), which is the standard measure of market concentration. For the HHI calculation, we assume that market share is proportional to capacity in the capacity-constrained region, for reasons discussed in Section 2.1. The allocation of spectrum is defined by the spectrum shares s_i and tower shares n_i of each carrier, where $s_i = S_i / S_{tot}$ and $n_i = N_i / N_{tot}$.

$$C_i = (n_i \; N_{tot}) \, r \, e \, (s_i \, S_{tot}) / f = \alpha \, n_i \, s_i \qquad \text{where } \alpha = r \, e \, N_{tot} \; S_{tot} / f$$

$$C_{tot} = \sum_{i=1}^{\infty} C_i = \alpha \sum_{i=1}^{\infty} n_i \, s_i$$

$$HHI = \frac{\sum_{i=1}^{\infty} C_i^2}{C_{tot}^2} = \frac{\alpha^2}{C_{tot}^2} \sum_{i=1}^{\infty} n_i^2 \; s_i^2$$

Ideally, C_{tot} should be as large as possible, and HHI should be as small as possible. Of course, these objectives are inherently in conflict, because of the economies of scale that were demonstrated in Section 3. Capacity C_{tot} is optimized by giving one carrier all of the spectrum and all of the towers, which yields the worst possible HHI of 1, i.e. a monopoly.

---

[2] This is why a carrier would consider T_0 when deciding how many towers to employ, but not T_{bw} or T_{bps}, as shown in Section 3.

The most cost-effective allocation of spectrum and tower resources is for each carrier to choose the same spectrum share and tower share, i.e. to let $n_i = s_i$. Indeed, Section 3 showed that a profit-seeking carrier would choose to let $n_i = s_i$ by showing that N_i / S_i is the same constant for all *i*. This means that n_i / s_i is also the same for all *i*, which is only possible when $\sum_{i=1}^{\infty} n_i = \sum_{i=1}^{\infty} s_i = 1$ if $n_i = s_i$. We can show that letting $n_i = s_i$ is also the most cost-effective approach within the policymaker framework of this section as follows. Consider a resource allocation in which spectrum share and tower share are not the same for all carriers. In this case, there must be at least one carrier *j* for which $n_j > s_j$ and at least one carrier *k* for which $n_k < s_k$. If we shift the same amount *x* of tower share from Carrier *j* to Carrier *k* and spectrum share from Carrier *k* to Carrier *j*, where *x* is less than both $n_j - s_j$ and $s_k - n_k$, then we reduce the differences between tower share and spectrum share for both carriers. The equations below show that we also improve the capacity of both carriers. Thus, everyone benefits from reducing the difference between spectrum share and tower share. Let C_j^* and C_k^* be capacities after the shift.

$$C_j^* = \alpha\left(n_j - x\right)\left(s_j + x\right) = \alpha\left(n_j\, s_j + n_j x - s_j x - x^2\right) = C_j + \alpha\, x\left(n_j - s_j - x\right) \quad > C_j$$
$$C_k^* = \alpha\left(n_k + x\right)\left(s_k - x\right) = \alpha\left(n_k\, s_k - n_k x + s_k x - x^2\right) = C_k + \alpha\, x\left(s_k - n_k - x\right) \quad > C_k$$

Spectrum policymakers should therefore prefer the division of spectrum resources that is best for the competing objectives of maximizing capacity and maximizing competition when $n_i = s_i$. While reasonable minds may disagree on how to balance the two competing objectives, we should all agree that a good division of spectrum would be Pareto optimal, i.e. if spectrum holdings s_i for all *i* are good, then it should not be possible to change spectrum holdings in a way that makes results better with respect to one objective without making results worse with respect to the other objective. We will now show that any Pareto optimal solution requires spectrum to be divided fairly evenly among the carriers. More specifically, if there are *d* carriers with spectrum (i.e. for which $s_i > 0$), then *d*-1 of these carriers should have the same amount of spectrum. The last carrier gets whatever is left in S_{tot}, which should be the same as or less than what the other *d*-1 carriers have.

Another way to describe this Pareto optimal result, which will be proven below, is to let the number of competitors *b* be any real number ≥ 1. If *b* is not an integer, then the fraction represents the extent to which the last carrier gets a smaller share of spectrum than the others. For example, if *b*=4.7, then the Pareto optimal market would have 4 carriers with equal amounts of spectrum, equal number of towers in the capacity-limited regions and equal market share, while the fifth carrier had .7 times as much spectrum as the first four. Figure 1 shows how spectrum would be divided among carriers as a function of *b.*

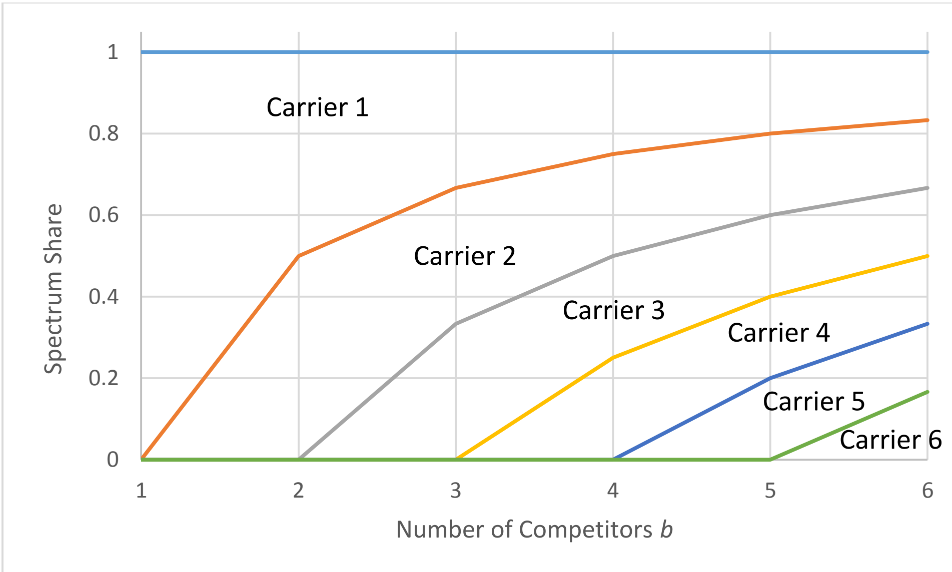

Figure 1: Division of spectrum versus number of competitors *b*

To prove this assertion, we consider a case where the condition above does not hold, and show that this division of spectrum cannot be Pareto optimal. If this condition does not hold, then there must be a carrier that has more spectrum than two other carriers. Without loss of generality, we number the carriers such that $s_1 > s_2 \geq s_3 > 0$. To prove that this is not Pareto optimal, we show that it is possible to shift spectrum from Carriers 1 and 3 to Carrier 2 in a way that improves (decreases) *HHI* without reducing total capacity C_{tot}. In particular, we increase s_2 by a small amount ds_2 while decreasing s_1 and s_3 as follows. Because $\frac{ds_1}{ds_2} + \frac{ds_2}{ds_2} + \frac{ds_3}{ds_2} = 0$, $\sum s_i$ remains constant at 1 even as spectrum holdings change.

$$\frac{ds_1}{ds_2} = -\frac{(s_2 - s_3)}{(s_1 - s_3)}$$

$$\frac{ds_3}{ds_2} = -\frac{(s_1 - s_2)}{(s_1 - s_3)} = -\frac{ds_2}{ds_2} - \frac{ds_1}{ds_2}$$

Let $C_{tot}(s_2)$ and $HHI(s_2)$ be total capacity and *HHI* as a function of s_2, respectively. If spectrum is initially divided such that increasing s_2 from its initial value will decrease $HHI(s_2)$ while $C_{tot}(s_2)$ remains constant, i.e. if $\frac{dHHI((s_2)}{ds_2} < 0$ and $\frac{dC_{tot}(s_2)}{ds_2} = 0$, then that initial division of spectrum cannot be Pareto optimal. We assume here that $s_2 > s_3$, as $s_2 = s_3$ is a minor special case.[3]

[3]Where $s_2 = s_3$, the slope of both $HHI(s_2)$ and $C_{tot}(s_2)$ with respect to s_2 are 0. A slight transfer of spectrum from carrier 3 to carrier 2 therefore has a negligible impact on both objectives, and then the assumption that $s_1 - s_2 > s_3 > 0$ applies.

$$C_{tot}(s_2 + ds_2) - C_{tot}(s_2) = \alpha \left(s_1 + \frac{ds_1}{ds_2}\, ds_2\right)^2 + \alpha\,(s_2 + ds_2)^2 + \alpha \left(s_3 + \frac{ds_3}{ds_2}\, ds_2\right)^2 - \alpha \sum_{i=1}^{3} s_i^2$$

$$= \alpha \left(s_1^2 - 2\, s_1 \frac{(s_2 - s_3)}{(s_1 - s_3)}\, ds_2 + \frac{(s_2 - s_3)^2}{(s_1 - s_3)^2}\, ds_2^2\right) + \alpha\,(s_2^2 + 2\, s_2\, ds_2 + ds_2^2)$$

$$+ \alpha \left(s_3^2 - 2\, s_3 \frac{(s_1 - s_2)}{(s_1 - s_3)}\, ds_2 + \frac{(s_1 - s_2)^2}{(s_1 - s_3)^2}\, ds_2^2\right) - \alpha \sum_{i=1}^{3} s_i^2$$

$$= 2\alpha \left(-s_1 \frac{(s_2 - s_3)}{(s_1 - s_3)} + s_2\, ds_2 - s_3 \frac{(s_1 - s_2)}{(s_1 - s_3)}\right) ds_2 + \alpha\, \frac{(s_2 - s_3)^2 + (s_1 - s_3)^2 + (s_1 - s_2)^2}{(s_1 - s_3)^2}\, ds_2^2$$

$$= 0\, ds_2 + \alpha\, \frac{(s_2 - s_3)^2 + (s_1 - s_3)^2 + (s_1 - s_2)^2}{(s_1 - s_3)^2}\, ds_2^2$$

$$\frac{dC_{tot}(s_2)}{ds_2} = 0 + \alpha\, \frac{(s_2 - s_3)^2 + (s_1 - s_3)^2 + (s_1 - s_2)^2}{(s_1 - s_3)^2}\, ds_2 = 0$$

Thus, total capacity remains constant as spectrum holdings shift using this formula.
HHI changes as follows.

$$HHI(s_2 + ds_2) - HHI(s_2)$$

$$= \frac{\alpha^2}{C_{tot}^2}\left(s_1 + \frac{ds_1}{ds_2} ds_2\right)^4 + \frac{\alpha^2}{C_{tot}^2}(s_2 + ds_2)^4 + \frac{\alpha^2}{C_{tot}^2}\left(s_3 + \frac{ds_3}{ds_2} ds_2\right)^4 - \frac{\alpha^2}{C_{tot}^2} \sum_{i=1}^{3} s_i^4$$

$$= \frac{4\,\alpha^2}{C_{tot}^2}\left(-\frac{(s_2 - s_3)}{(s_1 - s_3)}\, s_1^3 + s_2^3 - \frac{(s_1 - s_2)}{(s_1 - s_3)}\, s_3^3\right) ds_2 + O(ds_2^2) + O(ds_2^3)$$

$$= \frac{4\,\alpha^2}{C_{tot}^2}\; \frac{-(s_2 - s_3)\, s_1^3 + (s_2 - s_3)\, s_2^3 + (s_1 - s_2)\, s_2^3 - (s_1 - s_2)\, s_3^3}{(s_1 - s_3)}\, ds_2 + O(ds_2^2) + O(ds_2^3)$$

$$= \frac{4\,\alpha^2}{C_{tot}^2}\; \frac{(s_1 - s_2)\,(s_2 - s_3)\,[(s_3^2 - s_1^2) + s_2(s_3 - s_1)]}{(s_1 - s_3)}\, ds_2 + O(ds_2^2) + O(ds_2^3)$$

$$\frac{dHHI(s_2)}{ds_2} = \frac{4\,\alpha^2}{C_{tot}^2}\; \frac{(s_1 - s_2)\,(s_2 - s_3)\,[(s_3^2 - s_1^2) + s_2(s_3 - s_1)]}{(s_1 - s_3)} < 0$$

Thus, the derivative of *HHI* is always negative at a division of spectrum where $s_1 > s_2 > s_3 > 0$, so moving spectrum from carriers 1 and 3 to carrier 2 according the formula above improves *HHI* without changing total capacity. Indeed, this can continue until either s_3 falls to 0 or s_2 rises to equal s_1. This means that the initial division of spectrum cannot be Pareto optimal.

5 Impact of Tower Sharing on Spectrum Strategies

One of the ways for carriers to reduce the costs of cell towers is to enter into agreements with other carriers to share towers, or even share towers with government systems intended for public safety agencies [HALL11, PEHA13], and thereby share some of the fixed costs. In Sections 5.1 and 5.2, we consider tower sharing and how it might affect the results of Sections 3 and 4, respectively.

We assume in this section that each carrier has an arrangement to share towers with exactly one other carrier, so carriers share in pairs. When two carriers share a tower, each pays a fixed cost per tower T_{0sh} that is significantly lower than the fixed cost T_0 without sharing. The cost per bandwidth and cost per data rate are the same. If both carriers in a pair consider it cost-effective to add a tower when paying their portion of the cost, they do so. If one carrier does not consider it worth paying its portion of the cost, but the other considers it cost-effective to add a tower at full cost, the latter adds a tower that is not shared. This model is somewhat similar to the current market in the UK, where there are four large cellular carriers; British Telecom-EE and Three have a tower sharing agreement, and Telefonica O2 and Vodafone have a tower sharing agreement.

5.1 The Carrier Perspective

If two cooperating carriers have similar capacity requirements, then they can both benefit from relying on shared towers, and acquiring the same amount of spectrum. In this case, the analysis is similar to what was presented in Section 2, except with lower fixed cost per tower of T_{0sh} instead of T_0. Thus, the carriers would choose to have more towers and less spectrum when towers are shared, i.e. $N_{sh}/S_i = M/T_{0sh}$ instead of the M/T_0 derived in Section 2, where N_{sh} is the number of shared towers per sq km in the capacity-limited region.

If Carrier *i* needs more capacity than Carrier *j*, then Carrier *i* may be not be able to rely entirely on towers shared with Carrier *j*. The most cost-effective approach for Carrier *i* is to have more spectrum and more towers than Carrier *j* is willing to share. Consider the case where Carrier *j* is willing to pay for up to N_{sh} shared towers. Because Carrier *i* will choose to have N_{sh} towers or more, i.e. $N_i \geq N_{sh}$. its total cost is simply the cost derived in Section 3 minus the savings from sharing which is a constant.

$$K_i = M\,S_i + N_{sh}\left(T_{0sh} + T_{bw}\,S_i + T_{bps}\,\frac{r\,e\,S_i}{f}\right) + (N_i - N_{sh})\left(T_0 + T_{bw}\,S_i + T_{bps}\,\frac{r\,e\,S_i}{f}\right)$$
$$= M\,S_i + N_i\left(T_0 + T_{bw}\,S_i + T_{bps}\,\frac{r\,e\,S_i}{f}\right) - N_{sh}\,(T_0 - T_{0sh})$$

This only differs from the K_i derived in Section 3 by a constant $N_{sh}(T_0 - T_{0sh})$, so $\frac{dK_i}{dS_i}$ is the same as in Section 3, which means that Carrier *i* will choose to acquire the same amount of spectrum as derived in Section 3, and the same total number of towers as well.

$$S_i^{\,2} = \frac{f\,T_0}{M\,r\,e}\,C_i$$

$$N_i = \frac{M}{T_0}\,S_i$$

Thus, tower sharing does not change the fundamental observations from Section 3. With or without sharing, (i) there are economies of scale in the provision of adequate capacity in capacity-limited regions, (ii) carriers who have more capacity can add capacity at lower cost, and (iii) carriers with more spectrum should willing to pay more per MHz for additional spectrum than carriers with less spectrum.

5.2 The Policymaker Perspective

This section explores the Pareto optimal allocation of resources when carriers share all their towers, and the objectives are maximizing both capacity and competition (as quantified through the *HHI*), as in Section 4.

In any Pareto optimal solution, two carriers sharing all of their towers should have the same amount of spectrum. Since they have the same number of towers N_{sh}, their combined capacity does not depend on how spectrum is divided between them, and *HHI* is best if they have equal spectrum and therefore equal capacity. The question is then how to divide spectrum among each pair of carriers that share towers. Let s_i is the spectrum share for the *i*th pair, with each of the carriers in that pair getting half that spectrum. Let n_i be the number of towers that these two carriers are sharing.

$$C_{tot} = \sum_{i=1}^{\infty} \left(\alpha\, n_i\, \frac{s_i}{2} + \alpha\, n_i\, \frac{s_i}{2}\right) = \alpha \sum_{i=1}^{\infty} n_i\, s_i$$

$$HHI = \frac{1}{C_{tot}^2} \sum_{i=1}^{\infty} ((\alpha\, n_i\, s_i\, /2)^2 + (\alpha\, n_i\, s_i\, /2)^2) = \frac{\alpha^2}{2\, C_{tot}^2} \sum_{i=1}^{\infty} n_i^2\, s_i^2$$

The equation for total capacity is identical to that of Section 4, and the equation for *HHI* is simply that of Section 4 divided by 2. Thus, spectrum divisions among carriers that were Pareto optimal in Section 4 are identical to the spectrum divisions among pairs of carriers that are optimal in the scenario considered here. In any Pareto optimal resource allocation with *d* carriers, each of which share their towers with one other carrier, *d*-2 of these carriers will have the same amount of spectrum, and the last pair of carriers will split what is left of the spectrum, and their share will be the same as or less than the first *d*-2.

6 The Current UK Cellular Market

There are currently four prominent nationwide cellular carriers in the UK. This is not just an accident of history; it is the result of deliberate policy decisions, even as market forces were pushing towards consolidation. In particular, the European Commission blocked the proposed merger of O2 and Three to maintain this number of competitors at four, with Ofcom support. As Ofcom stated in 2016 [OFCO16b], "we believe that the existence of at least four credible MNOs is important for the UK mobile market. This is consistent with our views in previous documents and is a position maintained by the EC's recent decision to block the proposed merger of H3G and O2, which would have reduced the number of MNOs to three. We agree with the EC's conclusions."

Nevertheless, there are large disparities in the spectrum holdings of these four carriers, with the largest holding roughly 50% more spectrum than the next largest, and roughly three times that of the third and fourth largest [OFCO16b]. The HHI of spectrum shares is 0.303, which is only a little less concentrated than the 0.333 one would obtain from dividing the spectrum evenly among just three cellular carriers. In contrast, there is less difference in the number of cell sites held by each carrier, although the one with by far the most spectrum also has the most cell sites.[4]

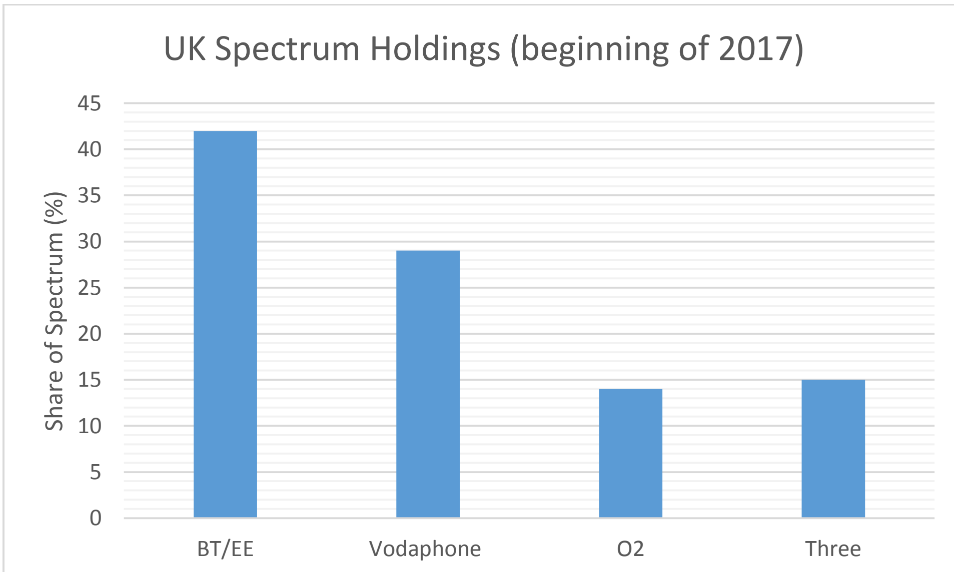

Figure 2: Spectrum Holdings in the UK

✄

In terms of amount of traffic carried, there is a close race between BT/EE and Three, with the other two well behind [OFCO16b]. This may seem surprising given the model presented in this paper. However there are historical reasons why this might make sense, and I can speculate that they form at least part of the answer. As described in Section 2.1, this paper assumed that all carriers use similar technology, and have grown their infrastructure to minimize cost for the given capacity. At the moment, Three has more recent and therefore more spectrally efficient technology, e.g. it expends no spectrum providing inefficient 2G services. This is an advantage in the short term, although when equipment is replaced and upgraded, as it inevitably must be for all carriers, this advantage will disappear. Moreover, BT/EE is the result of a relatively recent merger. Immediately after a merger, one would not expect the synergies of the two infrastructure that grew up separately to be fully exploited. For example, cell towers would initially be using some but not all of the spectrum held by the newly merged company. Thus, it takes some time before capacity reflects spectrum and cell tower assets.

If Ofcom hopes that four strong cellular carriers will serve the UK market in the long term as it stated [OFCO16b], then Ofcom should be concerned about the existing disparity of spectrum holdings, even if Ofcom believes the risk has not yet become apparent in traffic carried. Over time, the larger carriers will

---

[4] BT/EE reportedly has 18,000 sites now [OFCO16b], Vodafone and O2 have announced plans to have 17,500 sites [VODA17], and Three reports having ✄ [THRE17].

benefit from economies of scale. Because one carrier has far more spectrum than its rivals, it gains far more capacity for every new cell tower. Because that same carrier also has more towers, it gains more capacity for every MHz of spectrum it obtains. In general, we should expect the big to get bigger, unless there are explicit policies in place to promote competition, as will be discussed further in the next section.

7 Conclusions and Policy Implications

We find that when increasing capacity to meet customer expectations becomes a significant cost for cellular carriers, there are strong economies of scale. A carrier with more cell towers benefits more from any spectrum it has, and a carrier with more spectrum benefits more from every tower it has. Thus, the most cost-effective strategy for a carrier is to increase both spectrum holdings and number of towers together over time. Consequently, a carrier with large spectrum holdings will generally also be a carrier with many towers, and a carrier with enough capacity to support a large customer base will generally have large holdings of both spectrum and towers. This is radically different from markets where carriers are driven primarily to improve coverage rather than capacity. If it is coverage alone that attracts customers, then a carrier needs some spectrum, especially at lower frequencies, but particularly large spectrum holdings are not important. If capacity is what attracts customers, then the amount of spectrum matters a great deal.

Moreover, we can expect the big carriers to get bigger. Since revenue tends to be proportional to capacity in a capacity-limited region, and large carriers can increase their capacity more than small carriers with every MHz of spectrum they obtain, large carriers will generally bid more in spectrum auctions than their small competitors. Ofcom has suggested that carriers with little spectrum who are unable to outbid their rivals in spectrum auctions could compensate by building more towers [OFCO12a, OFCO12b]. This is certainly possible technically, but our analysis shows that this would be a highly unprofitable strategy. For a carrier with far more towers and far less spectrum than its competitors, obtaining more spectrum is the least expensive way to expand capacity. If such a carrier cannot afford additional spectrum, then the carrier should simply stop expanding capacity rather than adopt an even more costly approach. When data usage per user is increasing rapidly, to stop expanding it so surrender market share to the large carriers. In the absence of countervailing forces from policymakers, this will make competition increasingly difficult to sustain over time. Eventually, a carrier with small market share but valuable spectrum and infrastructure assets may be acquired by a rival.

These economies of scale create a trade-off between two important objectives for policymakers: increasing competition and lowering the cost of capacity. A small number of large carriers can exploit economies of scale to reduce costs, but with little competition, these carriers have little incentive to pass those cost-savings on to consumers. Policymakers must determine the right balance, and this should be reflected in spectrum policy and antitrust policy. Whatever the number of carriers, we find that the public interest is best served when spectrum is split fairly evenly among them. In particular, without tower sharing, any division of spectrum among *d* carriers that is Pareto optimal with respect to our two policy objectives would give the same amount of spectrum to *d*-1 carriers, and spectrum disparities among them are contrary to the public interest. The last carrier would get whatever spectrum is left, and less than the first *d*-1. Our results with tower sharing are only slightly different.

One policy that will naturally produce a division of spectrum just like this is a spectrum cap. For example, if no carrier can acquire more than 30% of the spectrum, then there will always be at least four carriers with spectrum, and their spectrum holdings will probably be similar to the Pareto optimal division described above. If the cap is 42%, as Ofcom currently proposes [OFCO16b], then we may eventually see just three carriers with significant capacity and market share in the capacity-limited regions. This can also be Pareto optimal, albeit with less competition. However, the European Commission blocked the merger of Three and O2, with the support of Ofcom [OFCO16b], and the U.S. Government similarly blocked the merger of AT&T and T-Mobile. This demonstrates a preference to have at least four large nationwide carriers. If four carriers is indeed the long-term goal, then a 42% cap is dangerously high. Worse yet, if a nation has a spectrum policy allows three but not four carriers to obtain the spectrum they need to expand capacity at costs that are consistent with competitive prices, and an antitrust policy that prevents that fourth carrier from merging with one of the other three, then we may see a poor result with respect to *both* policy objectives: effective competition and efficiency. We may get the effective competition one would expect with just three significant carriers, and the lower efficiency of four. Achieving a poor result with respect to both objectives is precisely what it means to yield results that are not Pareto optimal, as discussed in Section 4.

When imposing a cap in response to economies of scale in macrocells, it is important to include all of those licensed bands that are suitable for macrocells with current technology. Unlicensed bands and very high frequencies would be excluded. Bands like 3.4 GHz in the U.K. that are not usable today because they have not been cleared, but will be usable in a few years, should be included. Such bands could be included at the time of the auction, or they could be included at the time that the bands are expected to be cleared, but if regulators do not make clear before the auction that the bands will be included in caps and when, then regulators create an unnecessary and painful dilemma for themselves and the industry. Once the band is cleared, the regulator may be forced to choose between two bad options: increasing the cap beyond what the regulator has already concluded is the value that best serves the public interest and damaging competition unnecessarily, or adding this band to the cap as it has been defined and forcing one or more carriers to divest of spectrum that they have already obtained and incorporated in their long-term plans. The best argument in favor of excluding this band seems to be that additional spectrum will be cleared and auctioned to cellular carriers in the future, and if that new spectrum comes soon enough and goes mostly to the carriers with less spectrum, there is a *possibility* that the carrier with the most spectrum will not exceed the cap. Thus, excluding 3.4 GHz from the cap may prove to be harmful, or it may prove to be insignificant, but excluding the band will not be beneficial. Why take the chance?

References

[BAKE07] J. B. Baker, "Beyond Schumpeter vs. Arrow: How Antitrust Fosters Innovation," *Antitrust Law Journal,* vol. 74, 2007. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=962261

[CAPO13] F. Capozzi, G. Piro, L. A. Grieco, G. Boggia and P. Camarda. “Downlink packet scheduling in LTE cellular networks: Key design issues and a survey.” *IEEE Communications Surveys & Tutorials*, vol. 15, no. 2, pp. 678-700, 2013. http://ieeexplore.ieee.org/document/6226795

[CISC16] Cisco, *White paper: Cisco VNI Forecast and Methodology, 2015-2020*, June 1, 2016. https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.html

[CRAM11] P. Cramton, E. Kwerel, G. Rosston and A. Skrzypacz, “Using Spectrum Auctions to Enhance Competition in Wireless Services,” *Journal of Law and Economics*, vol. 54 pp. 167-78, Nov. 2011. http://www.stanford.edu/~skrz/spectrum-auctions-and-competition.pdf

[DOJ13] U.S. Department of Justice, “*Ex Parte* Submission of the United States Department of Justice,” in the matter of Policies Regarding Spectrum Holdings, WT Docket No. 12-269, April 11, 2013. http://apps.fcc.gov/ecfs/document/view?id=7022269624

[ERGA98] H. Ergas and E. K. Ralph, “New Models of Foreclosure: Should Antitrust Authorities Be Concerned?,” Feb. 1998. http://ssrn.com/abstract=55775

[HALL11] R. Hallahan and J. M. Peha, "The Business Case of a Nationwide Wireless Network that Serves both Public Safety and Commercial Subscribers," *Telecommunications Policy*, vol. 35, no. 3, April 2011. https://www.ece.cmu.edu/~peha/profitability_of_public_safety_network.pdf

[LIGO15] A. Ligo, J. M. Peha, P. Ferreira, J. Barros, “Comparison between Benefits and Costs of Offload of Mobile Internet Traffic Via Vehicular Networks,” *43rd Telecommunications Policy Research Conference (TPRC)*, 2015. https://www.ece.cmu.edu/~peha/Vehicular_Network_Offloads_TPRC_2015.pdf

[OFCO12a] Ofcom, S*econd Consultation on Assessment of Future Mobile Competition and Proposals for the Award of 800 MHz and 2.6 GHz Spectrum and Related Issues*, Jan. 12, 2012. https://www.ofcom.org.uk/__data/assets/pdf_file/0025/55276/combined-award-2.pdf

[OFCO12b] Ofcom, *Assessment of Future Mobile Competition and Award of 800 MHz and 2.6 GHz*, Annex 3, July 24, 2012. https://www.ofcom.org.uk/__data/assets/pdf_file/0029/47387/annexes1-6.pdf

[OFCO16a] Ofcom, *Smartphone Cities: Measuring 4G Mobile Broadband and Voice Performance*, Dec. 16, 2016. https://www.ofcom.org.uk/research-and-data/broadband-research/smartphone-cities/smartphone-cities-dec16

[OFCO16b] Ofcom, *Award of the 2.3 and 3.4 GHz Spectrum Bands: Competition issues and auction regulations*, Nov. 21, 2016. https://www.ofcom.org.uk/__data/assets/pdf_file/0026/93545/award-of-the-spectrum-bands-consultation.pdf

[PEHA13] J. M. Peha, "A Public Private Approach to Public Safety Communications: The best way to make rapid and affordable progress in implementing the government's FirstNet plan is to start by taking advantage of the existing commercial infrastructure," *Issues in Science and Technology*, National Academy Press, vol. 29, no. 4, pp. 37-42, Summer 2013. http://www.ece.cmu.edu/~peha/Peha_first_steps_for_FirstNet.pdf

[THRE17] Three UK, personal communications, Jan. 2017.

[VODA17] Vodafone UK, *Our Network Coverage.* Accessed Jan. 20, 2017. https://www.vodafone.co.uk/our-responsibilities/our-network