ARTICLE IN PRESS

Telecommunications Policy xxx (2017) 1–10

Contents lists available at ScienceDirect

Telecommunications Policy

journal homepage: www.elsevier.com/locate/telpol

Cellular economies of scale and why disparities in spectrum holdings are detrimental

Jon M. Peha

Carnegie Mellon University, Dept. of Engineering & Public Policy, 5000 Forbes Ave., Pittsburgh, PA 15213, USA

ARTICLE INFO

Keywords:
Cellular
Spectrum
Competition
Economy of scale
Market concentration
Spectrum cap

ABSTRACT

Now that traffic volumes are increasing rapidly, the cost of expanding capacity has become a large portion of expenditures for Mobile Network Operators (MNOs). This paper uses an engineering-economic model to show that there are strong economies of scale when expanding capacity, because an MNO with more spectrum benefits more from every new cell tower, and an MNO 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, we find that the cost-effective approach is to increase both types of assets at a similar rate. In the absence of countervailing policies, the big MNOs are well positioned to 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 derives the Pareto optimal division of spectrum with respect to these two competing objectives, and shows that any Pareto optimal assignment will split the spectrum fairly evenly among competing MNOs. This is not simply a method of ensuring that there are many competitors; spectrum should be divided fairly evenly regardless of whether the number of competitors is large or small. 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 MNOs, and the lower competitive pressure of a smaller number of MNOs. One effective way to achieve a division of spectrum that is close to Pareto optimal is a spectrum cap, provided that this cap is set at a level consistent with other policies and policy objectives, including antitrust policy.

1. Introduction

Until recently, cell phones were primarily for phone calls and texts. For someone who uses her cell phone primarily for phone calls, coverage is extremely important. Driving through even a small hole in coverage can terminate an important conversation. To attract such customers, Mobile Network Operators (MNOs) needed spectrum throughout the areas that the MNO intended to serve, and preferably some of that spectrum would be at low frequencies. This probably explains why Ofcom found (Ofcom, 2012, Jan. 12; Ofcom, 2012, July 24) that in the years before 2012 an MNO had been a “credible” competitor if it held 10% or more of the cellular spectrum in that MNO's country. Although Ofcom still justifies spectrum policy decisions on that conclusion (Ofcom, 2016, Nov.), the situation has changed as smartphones became pervasive, and smartphone applications emerged that transferred large amounts of data. Today's cell phone users still want coverage, but now many want to know whether their share of an MNO's capacity will be enough. The capacity

E-mail address: peha@cmu.edu.
URL: http://www.ece.cmu.edu/~peha

http://dx.doi.org/10.1016/j.telpol.2017.06.002
Received 16 February 2017; Received in revised form 29 May 2017; Accepted 3 June 2017
Available online xxxx

Please cite this article in press as: Peha, J. M., Cellular economies of scale and why disparities in spectrum holdings are detrimental, Telecommunications Policy (2017), http://dx.doi.org/10.1016/j.telpol.2017.06.002

required to meet the demands of cell phone users is doubling worldwide every 20 months (Cisco, 2016). As a result, expanding macrocellular capacity has become a major driver of annual expenditures for cellular MNOs.

As this paper will show, this need to rapidly increase capacity has yielded new economies of scale. The largest MNO can expand its capacity at a lower cost than its smaller competitors, which means that the large can easily get larger. This result may help to explain why large MNOs around the world have generally been growing, and many small MNOs have been seeking opportunities to merge. It is also consistent with the profitability often observed in large MNOs, e.g. in the U.S. the two largest MNOs (Verizon and AT&T) have significantly greater margins than the third and fourth largest (T-Mobile and Sprint) (Forbes, 2015). The fact that scale decreases costs has advantages for users of cellular services, as long as the large MNOs choose to pass these cost savings along to their customers. However, if the large MNOs increase market share, there will be less competition, giving MNOs less incentive to lower prices and improve quality of service.

This paper explores how best to divide spectrum resources among MNOs 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 MNOs, regardless of whether the number of competing MNOs is large or small. 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,* where an MNO is free to put together a portfolio of frequencies that meets the company's needs, provided that the total does not exceed some upper bound. Proposed mergers, such as AT&T with T-Mobile in the U.S. and O2 with Three in the U.K., have sparked intense debates over whether the goal of antitrust policy should be to have at least three competing MNOs or at least four. This paper shows that spectrum policy is also important, and that spectrum policy and antitrust policy should be in sync on this issue.

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 MNOs are not engaging in a strategy of "foreclosure," even though a common argument for spectrum caps is as a means of preventing foreclosure (Ergas & Ralph, 1998; Baker, 2007; Cramton, Kwerel, Rosston, & Skrzypacz, 2011; U.S. Dept. of Justice (DOJ), 2013) rather than addressing an economy of scale. Because there is a limited amount of spectrum available to MNOs and few opportunities to obtain spectrum from sources other than a direct competitor, an MNO 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" (Cramton et al., 2011). 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" (DOJ, 2013). If some MNOs 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 describes 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 MNOs. This section shows how MNOs 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 MNO's cost per cell tower does not depend on the choices of that MNO's competitors, which is often but not always the case, Section 5 relaxes that assumption by considering the effects of explicit tower-sharing arrangements between MNOs. Finally, we summarize conclusions and discuss their policy implications in Section 6.

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 MNOs 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. All of our analysis considers only facilities-based MNOs, rather than Mobile Virtual Network Operators (MVNOs) that only resell services. It is facilities-based MNOs that operate infrastructure, and make decisions about spectrum and tower acquisitions to meet capacity needs. Moreover, it is competition among facilities-based MNOs that motivates service providers to pass savings in infrastructure cost on to consumers.

2.1. Profit maximization

This paper assumes that MNOs 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 MNO's infrastructure. For example, immediately after a merger, an MNO's infrastructure and spectrum holdings are unlikely to be as cost-effective as possible, but the MNO 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 an MNO to gradually obtain spectrum and gradually build towers to meet the gradually-increasing customer demand for capacity, a real MNO 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 MNO may have more excess capacity than another. Over the long term, however, these temporary effects should have little impact.

We also assume that all MNOs are deriving similar revenues per unit of capacity. Thus, if two MNOs have the same capacity but one