10

Open Access Markets for Capacity and the Inseparability of Spectrum and Infrastructure

Linda Doyle, Peter Cramton, and Tim Forde

1 INTRODUCTION

The chapter sits within this book in a number of contexts. It is about opening up competition for current and future mobile virtual network operators (MVNOs) and can therefore potentially be seen as a response to a number of network consolidation issues raised by Klass in Chapter 4. It is also about providing existing mobile network operators (MNOs) with alternative ways of managing supply and demand. It provides an example of a wholesale approach that is compatible with the Red Compartida requirements described in Chapter 5 by Mariscal. In addition, this chapter touches on the theme of alternative governances that are present in the chapters by Joyce (Chapter 1), Marcus (Chapter 8), and Weiss and Gomez (Chapter 9). Here the dedicated wholesale network is seen as an opportunity to reformulate the relationships between MVNOs and mobile network operators.

The chapter is organized as follows. Section 2 provides the main motivation for the dedicated wholesale network and open access market approach, from the MVNO perspective. While this book is not just focused on spectrum, but rather on how we use technology to further the public good, it is nonetheless important to include the spectrum perspective. Hence Section 3 briefly looks at the concept of fluid spectrum trading, which, as will become apparent, acts as a further means of motivating the work presented here. Section 4 moves on to describing the open access capacity market that is at the heart of the chapter, and Section 5 furnishes some further details of the market structure. Section 6 returns to the spectrum perspective and looks at future spectrum demands to show how the open access market can play a role in meeting those, and Section 7 concludes.

2 MOTIVATION

A key motivation for the work presented comes from observing the MVNO markets. An MVNO typically does not own network infrastructure or spectrum, and accesses these resources through some kind of wholesale network. However, in most examples around the world the wholesale network is not a dedicated wholesale network.[1] More often than not, the owner of the wholesale network, the MNO, shares that network between its own retail customers, any MVNO entities it owns, and third-party MVNOs.

The problem that arises with this scenario is that it is not always within the interest of the operator of the network, the MNO, to allow third-party MVNOs to expand and grow beyond certain limits. Figure 10.1 attempts to capture this case. So, in Figure 10.1, though MVNO 1 has the potential to expand its business substantially, this expansion is capped by the MNO controlling the wholesale network. The access of the MVNO to the MNO network is completely controlled by the MNO.

There are additional problems with the situation. Simple mechanisms are often used to estimate spare capacity on the MNO network, usually based on average demands. The number of MVNOs that can be supported by the spare capacity is then in turn based on the predicted average demands of the MVNOs. There is no sense that the optimal number of MVNOs, whatever that may be, has been accommodated.

In addition, the MNO–MVNO relationship is not designed for access to capacity for new and various different types of emerging service providers such as those which might result from different sectors such as health, transport, agriculture, and so on becoming increasingly digitized and calling for new types of operator. The relationship is not designed to support access to capacity in given locations only; to support access to capacity on an as needed basis; to dynamically target underused capacity; or any mix of the above.

In an attempt to illustrate these issues, we present three different types of virtual network operators (VNOs), all requiring different capacity for different durations: The “Plain Old Coverage Operator”

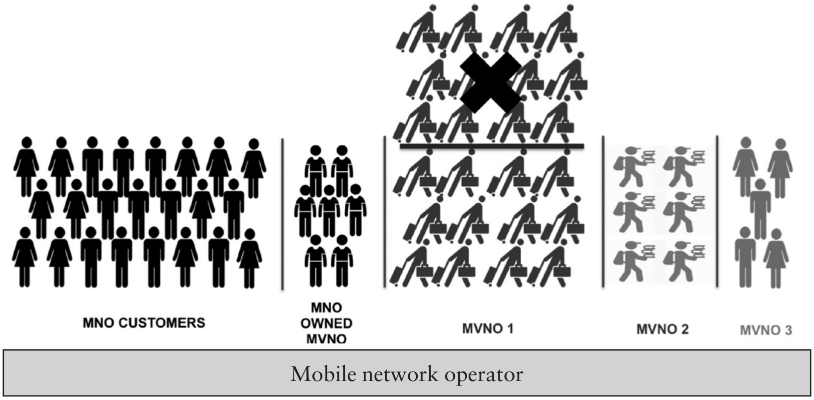

Figure 10.1 MVNO 1 desires to expand but any expansion is blocked by the MNO

requires capacity right across the network all of the time, like very many MVNOs in operation today. However, the "Connected Cow Operator" requires rural area capacity only, and then only once a day. The "Remote Surgery Operator" requires extensive capacity in discrete locations for a given duration. Each of these three examples is of course simplified, but the different flavours of operator nonetheless illustrate the point that as Internet of Things (IOT) and machine-to-machine (M2M) services grow, the virtual network operator of the future may be very different from that which we understand today (see Doyle et al. 2014) and will have very different spatial and temporal demands for capacity than is the case for MVNOs today.

3 THE SPECTRUM PERSPECTIVE

The purist approach to the challenge outlined in Section 2 would focus on the problem from a spectrum sharing perspective, that is, give the different virtual operators the spectrum they need, where and when it is needed, to provide the service. Indeed, a number of chapters in this book focus on spectrum-sharing. Marcus, in Chapter 8, provides an historic perspective on spectrum-sharing, detailing many different approaches, and Weiss and Gomez, in Chapter 9, look at polycentric forms of governance and in doing so showcase an approach to spectrum-sharing that allows service providers to be matched with the spectrum they need for the service they offer, through local negotiations.

The spectrum sharing perspective is one that applies a highly dynamic mindset to resource usage. To illustrate this we draw on a 2007 paper, titled "Towards a Fluid Spectrum Market for Exclusive Usage Rights," that focused on a very fluid and dynamic approach to spectrum trading (Doyle and Ford 2007). We reproduce some the material from that paper as a means of further setting the scene for this chapter.

In the 2007 paper we built on work that conceptualized what we call the radio spectrum rights continuum as a three-dimensional model: space, time, and frequency. Space, a broader term than "area," captures the three-dimensional notion of electromagnetic radiation. We proposed that such a continuum should be quantized into discrete blocks. Each block represents a unique assignment of spectrum rights at a particular place, for a particular frequency and at a fixed time. The purpose of quantizing the spectrum continuum on the time-axis is to allow each block assignment to be recycled and reassigned at each time interval through some kind of trading process. While the paper did not specify the dimensions of a block, the idea was that the block would be at some level of granularity, so that smaller and bigger players could trade in the market. The paper visualized the process involved through a set of Rubik's Cube–like representations. Each block in each Rubik's Cube–like figure represents a unique spectrum assignment, and each colour represents a unique spectrum consumer. We chose the term spectrum consumer to convey the point that we wanted to represent a more general case rather than tie the process to any traditional form of mobile operator.

In one of these cubes, we showed two licences granted to mobile network operators in Ireland. The allocations consisted of spatially and temporally contiguous blocks of spectrum for a single frequency range that aggregated to form conventional licences, thus reflecting traditional business plans. It is worth noting that the allocations for the Republic of Ireland extend to the entire jurisdiction of the state, and that while these licences are given for a specific time frame, they tend to go on in perpetuity. Hence the RF spectrum rights continuum model emphasizes this general *long-lived-ness* of the licences. Also of note is that spectrum has to be bought well in advance of the market for the services emerging.

Another of our cubes showed another option: a very fluid and flexible market in which spectrum consumers can freely buy and sell exclusive rights to the cubes of spectrum. The services that owners of the usage rights deliver and the technologies used to deliver those services are not proscribed – that is, there is total liberalization of the spectrum. The spectrum goes to those who need (value) it, when they need it. There are no limits or rules as to what blocks can be neighbours and what services can be delivered by neighbouring blocks. The blocks are of dimensions that make it possible for smaller players to participate. In such a very fluid system, the individual blocks of course may be aggregated to form larger assignments. Aggregations in this situation occur because of market drivers and not because of the straightjacket of a specific licence framework. We acknowledge the possibility of market failure, of course, and that interference is a challenge in this very open, liberalized system.

The paper "Towards a Fluid Spectrum Market for Exclusive Usage Rights" discusses interference in detail, describing issues around co-channel interference and adjacent channel interference, and issues around the fact that different topological deployments of dense and sparse networks side-by-side can cause issues, as well as discussing the general meaning of interference. Potential approaches to managing interference are presented. In the main they all focus on giving choice to the spectrum consumer. Choices on how the spectrum consumer might meet boundary conditions range from self-imposing guard bands (from within its own block), to using sophisticated cognitive techniques to sense its footprint and adjust to sculpt its profile in accordance with neighbours. In essence the paper presented an idealist vision that would draw on future technological advances to deliver.

The vision of totally liberalized spectrum trading, which sees spectrum going to those who need it when they need it, big players or small, calls for much complexity in the use of technology and for consumers to dynamically adapt to changing contexts (new and different neighbours); it also raises issues around investment. And while technology has moved on in the ten years since the paper was written, the challenges remain similar. The paper embodies a sense of freedom and openness that is lacking in spectrum management today, as well as a sense of freedom and openness that could well serve the current and future MVNO world.

4 MAKING THINGS REAL – THE OPEN ACCESS CAPACITY MARKET

What we see in the world of the MVNO is *capacity* trading rather than the trading of *naked spectrum* as has just been described. The term "naked spectrum" was coined by Hazlett (2011) and refers to frequency bands only. The spectrum auctions run by many regulators in different jurisdictions around the world trade in spectrum only, though of course these auctions are examples of trading that is highly static and slow in nature (i.e., licences are awarded for long periods of time on the basis of the auctions). The examples discussed in the previous section depicted more dynamic (naked) spectrum trading: the traded commodity is again spectrum, though there are no practical examples of such dynamic systems in operation. Capacity, on the other hand, is something that exists on a network. In other words, it is the result of *coupling* the spectrum with infrastructure.

Capacity trading lacks the purist attraction of spectrum trading because the network technology plays a role in the solution, whereas spectrum trading is technology neutral in principle. The key challenges with spectrum trading, however, are those related to interference. This inability to easily and effectively manage interference is the main reason why highly dynamic spectrum trading has not been implemented. While capacity trading, on the face of it, lacks the full dynamism and technology neutrality of the vision that emerges from the spectrum-trading world, it does offer the attraction of being tractable and implementable, and there are possibilities for implementing capacity trading in a much more dynamic manner. Hence the capacity trading solution presented in this chapter is an attempt to embody that dynamism, as well as the freedom of choice and the opportunity for big or small players envisaged in the fluid spectrum trading scenario depicted in the previous section.

The solution described here is drawn from two major sources, namely the Cramton and Doyle 2016 white paper and the associated *Telecommunications Policy* journal publication. Details from both papers are reproduced in this chapter. In addition, the open access market for wholesale capacity is one that is under development by a company called Rivada Networks, which further illustrates the practical and feasible nature of the system.

The fundamental idea of an open access market is that anyone with a need for capacity can gain access to mobile communications at competitive rates. An open access market for mobile network capacity is a market that is open to all. Indeed, the cornerstone to open access is that use of the network cannot be withheld. Just like the Internet, anyone can use it on a non-discriminatory basis. Of course, the capacity of the network is scarce, so prices are required to assign network resources to users. Hence an open access network adopts efficient pricing. Supply is not withheld. Price is set at the value of the marginal demand.[2]

Figure 10.2 is a very high-level graphic of the open access market in play. There are a number of different components to the market; there are also core ideas on which it is based. The remaining paragraphs in this section of the chapter detail these in turn.

There are two distinct markets on which capacity can be bought and sold. The first is a wholesale market, in which resources are offered to service providers, and the second is a retail market, in which service providers offer services to retail users. Here we concentrate on a wholesale open access market. We therefore use the term *wholesale network* to describe the network on which the wireless capacity is available. The wholesale network will cover a specific geographical area. The wholesale network can be owned by single or multiple entities. One network does not need to equate with one owner. However, what is key is that this is a dedicated wholesale market. For the purposes of this chapter we consider an LTE network. LTE is deployed extensively around the world and therefore cannot be seen as a limiting factor. The solution presented here can, of course, also be applied to any other cellular-like network, and therefore remains valid as new 5G cellular technologies emerge. The purpose of focusing on LTE here is to ground the description of the system in real terms and to discuss limitations.

The *service providers* buy and sell capacity on the wholesale network to service their own users. The service provider can be any entity wishing to provide some kind of mobile/wireless service and does not need to do so universally across the geographical jurisdiction of the wholesale network. Two service providers are shown for illustrative purposes in Figure 10.2: service provider A and service provider B. Service provider A has bought capacity over the entire network while service provider B has bought capacity in a localized area. Service provider A might therefore equate with the “Plain Old Coverage Operator” as discussed in Section 2 (or indeed to a spectrum consumer wanting spectrum across the entire region as discussed in Section 3). Whereas service provider B might better equate from a spatial perspective with the “Connected Cow Operator” (or indeed with a spectrum consumer wanting one block of spectrum).

Note that the service providers illustrated in Figure 10.2 need not be pure virtual network operators. For example, the service provider could be a traditional mobile network operator and use the wholesale network to supplement its capacity. Most network operators today

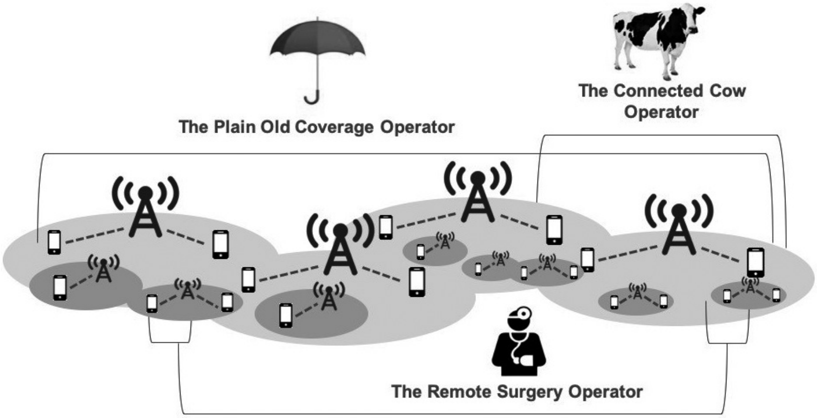

Figure 10.2 A high level view of the open access capacity market

overprovision their networks to accommodate periods of high activity.[3] It should instead be possible to resort to the wholesale network for this "top-up" capacity, thereby reducing the need for every individual operator to overprovision. Top-ups can be bought in specific areas or across the network.

While time dimension cannot be drawn in Figure 10.2, both service provider A and service provider B will have differing capacity demands over time, and hence the scenario in Figure 10.2 should be seen as illustrating one moment in time. Depending on the type of service provider, the demand might vary significantly in space and time. There are of course, service providers for which the demand profile can be engineered to fit specific spatial or temporal profiles. The latter tend to be demand profiles associated with machines as distinct from those associated with humans and may increasingly arise as M2M and IOT services grow.

To respond to different service providers we therefore define *capacity* on a spatiotemporal basis and think in terms of both the demand side and the supply side. The demand profile of each service provider will vary over time and space. We therefore need to think of capacity demand as being some measure of required throughput (e.g., GB/s) for a specific duration at a location. The wholesale network provides the supply. The supply can also have a spatial and temporal variation. The spatial variation of supply will be due to the varying capacity of the physical network as most networks are dimensioned differently in different geographical areas. For example, in most network designs, the capacity in urban areas tends to be higher than in rural areas. In most cases the temporal variation will be slow and reflect, for example, upgrades to the network.

However, it is possible that additional spectral resources could be added to the network. This can happen as the wholesale network gets additional spectrum over its lifetime through traditional auctions that are held by national regulatory authorities. But it can also happen by gaining access to shared spectrum, through new spectrum-sharing schemes. Consider the Licensed Shared Access (LSA) framework defined in Electronic Communications Committee Report 205 (2014) as one potential example. LSA was conceived as a means of providing access to licensed bands that otherwise would not be possible in some European countries because of existing incumbents. An LSA licensee is granted exclusive use of the spectrum at a given time and location while protection is provided to the incumbent. It is similar to the licensing of RRBSs in Canada, as described by Taylor in this collection. With LSA, the aim is for the licensee to get long-term access to the spectrum. Much of the work on LSA has focused on the 2.3 GHz band, where LSA can be used to create sufficient scale for the deployment of mobile services on the band across Europe. However, the approach is a general one and can be applied to many bands. Testing of LSA has taken place in Spain, France, Italy, Finland, and the Netherlands. If LSA were to become widely adopted, the wholesale network operator would be able to add spectrum from this sharing scheme to its network (provided of course the network was suitably dimensioned for the frequency bands in question). This spectrum could be used to offer additional capacity on the market over the lifetime of the LSA licence. Hence LSA or any other spectrum sharing scheme could provide a means of adding to the capacity on the network.

If demand profiles of different service providers are correlated in space and time, then congestion may occur, and this is where the open access market comes in to play. As shown in Figure 10.1, the *independent system operator* (ISO) manages the auctions (step 1). This term is borrowed from the electricity markets. Independent means that the ISO has no ownership interest in the market participants and does not take a position in the market. The ISO qualifies market participants and establishes any limits on each participant's bidding activities, reveals supply curves for the open access network, conducts the forward and real-time auctions, operates the open access market, settles all transactions in a timely manner consistent with market rules and supply and demand realizations, provides information on market performance to market participants and the market monitor (defined below), and improves the market as problems are identified.

There are two market mechanisms depicted in Figure 10.2, namely a forward market and a real-time market. A *real-time market* is a market that is conducted "on the fly," based on up-to-date events, and is about allowing the service providers to access the capacity they want "now." Real-time evokes the idea of a "near instantaneous response" (RealWireless 2019). In the open access capacity market, this instantaneous way of doing things means that up-to-the-instant knowledge about demand profiles is to hand. If the area over which the demands are being made are also "small," the services providers can, in principle, make very accurate bids. For the moment let us define the space dimension to be the smallest location that can be independently controlled by the network. In Figure 10.2 we have depicted this to be at the cell level. There is of course the question around what real-time means. In reality, real-time is as fast as it takes to trade and for capacity to be directed/redirected to the users of the service providers who successfully bought the capacity and away from those who sold it. In an LTE network, information must be sent to the edge of the network (i.e., to the base station, or enodeB, as it is termed) to effect the change in resource distribution. This is shown as steps 2 and 3, following the auction in Figure 10.2. A highly conservative estimate of time to trade and effect change in the LTE network is one hour. In the current network management regimes, changing spectrum ownership on a one-hour basis would be considered exceptionally dynamic.

The purpose of using a real-time market is, in principle, to ensure that the supplier can exactly match the type of spatiotemporal demand profile of the service provider. This benefits both the service provider seeking the capacity and the wholesale network operator supplying the capacity. The service provider is better able to bid for the exact resources needed, for the resolution (in space and time) at which the real-time market operates makes it easier for the service providers to estimate the demands of their users. The service provider thus pays only for the resources needed and does not have to buy extra resources just in case. The wholesale network operator is able to support many more service providers, because each service provider will not need to build overprovisioning into its demand bids and in principle, this means accommodating more demands.

One characteristic of real-time markets is that they may involve increased risk. Prices become more volatile closer to real-time (shorter time intervals) as market participants have fewer and fewer options; this makes supply and demand curves steeper and prices thus more sensitive to even modest quantity changes. Therefore, to allow the service providers to make plans and better manage risk, the real-time market can be complemented with a forward market. A *forward market* is a market dealing in commodities for future (forward) delivery at prices agreed upon today (date of making the contract). These markets can be used to hedge against sharp fluctuations in prices. In our case this means that service providers can hedge their bets by acquiring capacity in advance of the real-time market.

It does not make sense for the forward market to operate at the granularity of the real-time market because when looking ahead at expected demand, a service provider will not be able to predict the demand at the fine-grained level of the real-time market. The service provider will however be able to estimate expected demand over a larger area and for a longer interval. In addition, the auction process would be very onerous if this were the case. Hence the forward market, which deals in GB/timeslot/location, operates at much longer time scales and over locations covering a greater area. It is also important to note that there can be more than one forward market. The position taken by the service provider can be successively refined over each forward market. Each successive forward market will cover a smaller geographical area and shorter time interval than the previous to facilitate this fine-tuning. In general, current mobile service providers think about longer-term usage trends on a yearly basis. It therefore makes sense to offer forward products on the open access market that allow the purchase of capacity for year-long periods. A second level of refinement can happen through medium-term products. There are seasonal variations in demand that manifest themselves on a monthly basis, for example. Growth/decline in customer bases may also be evident at the monthly level. Therefore, the ability to buy or sell for the month ahead makes for a reasonable medium-term product. It may be the case that a forward market for a week ahead or a day ahead will make more sense as service providers evolve in a future world of the Internet of Things, for example, and as demand profiles change, but for the moment the yearly and monthly options seem sensible, and hence we select two forward markets. Figure 10.2 captures the larger geographical areas associated with the forward markets – in this case the forward market is conducted over the area much greater than the cell-level area of the real-time market.[4]

In summary, step 1 sees the auctions conducted by the ISO. Once the results are known, they are communicated to the LTE network in step 2, which then manages the access each service provider gets, step 3, on the basis of market share. This last step is carried out using techniques that are possible with an LTE network and that do not break LTE standards. These techniques, however, are beyond the scope of this chapter.

5 SOME MARKET DETAILS

Given that the open access market is at the heart of this chapter, it is important to furnish some additional details about the market. The forward and real-time markets can be structured as *two-sided auctions* in which capacity is traded. The wholesaler provides a supply curve indicating the quantity it would like to sell, at each geographical area, at various prices (more at higher prices); the service providers buying network resources, at each given location, provide a demand curve indicating the quantity they would like to buy at various prices (more at lower prices). The supply and demand curves are aggregated across all market participants for each location, and the winning bids are determined. The intersection of the aggregate supply and demand curves determines the clearing price (P*) and the quantity traded (Q*), at the given location, as shown in Figure 10.3.

There is a clearing price per geographical area. Just as in the case of locational marginal pricing for electricity markets, in locations where there is congestion, clearing prices will be high. In locations in which there is no congestion, the clearing price will be zero (or some low price floor). The price floor is meant to ensure that some usage-based revenue accrues in areas with surplus capacity. This strengthens incentives for network investment in low-demand areas. Consistent with the open access principle, the price floor ideally is a nominal amount at or near zero.

The forward markets are financial markets (cash settled) and allow participants to take positions well in advance of real-time. Financial markets allow the goods to be traded back and forth without ever being accessed/consumed/depleted. The real-time market is physical; that is, it involves the physical delivery of wireless capacity. It is at this stage that the capacity is actually consumed by the end users of the

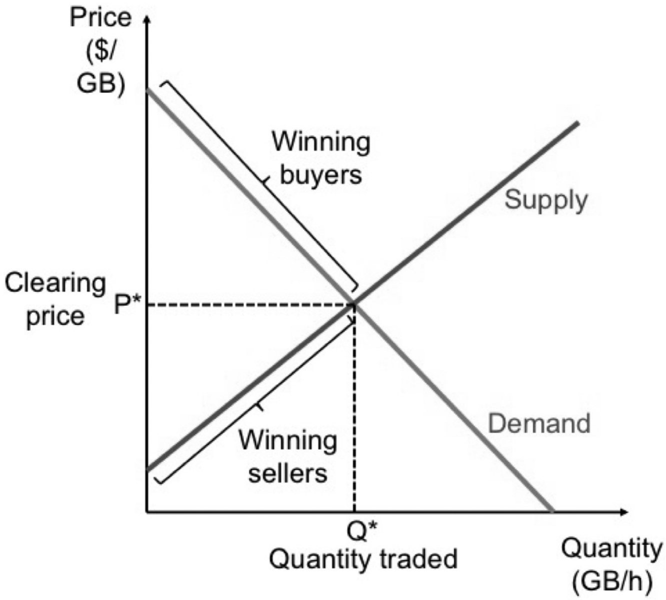

Figure 10.3 Market clearing at a particular time and location

service providers as they run their devices on the network. Deviations from forward positions are settled at real-time prices.

Yearly and monthly auctions involve greater volume. In electricity markets, 80 to 90 percent of volume transacts in forward markets. We anticipate a similar split in mobile communications markets. When auctioning many related products in infrequently conducted forward auctions, bidders find it helpful to learn about market prices and likely winnings during the auction, while they can still adjust their bids. This learning about prices and winnings is called *outcome discovery*. To allow outcome discovery in these auctions, the ISO collects bids with a *simultaneous ascending clock auction*. Details of this type of auction can be found in Milgrom (2004). In each round the auctioneer asks bidders to express a piecewise-linear demand curve for a range of prices. The real-time auction is conducted in a single round (sealed bid). This is because outcome discovery is less important in real-time and the auctions must occur quickly.

The number of products that are placed on the market needs to also be selected. This again is an economic choice that involves trading off the *simplicity* of fewer products with the *flexibility* of more products.

Currently there is clear difference in demand between peak and non-peak hours. It therefore makes sense that service providers be able to take different forward positions on peak and non-peak hours, resulting in two different hourly products. There are 1 + 12 = 13 forward auctions each year and 365 × 24 hourly auctions.

Of course the service provider may still not get the demand right. One important point to note here is that the real-time market is settled on the basis of actual usage during the real-time interval. In the types of systems we are looking at, the usage of the network is scheduled. If a service provider buys capacity in a specific location for a specific time period, it is because the service provider expects its users to require capacity in that location over that time period. That service provider's users will be scheduled on the network as they make demands in that location. If the service provider has got its predictions wrong, and no users materialize in that location, then users from other service providers can be scheduled on the network.

The real-time market should however provide incentives that motivate service providers to estimate as best they can their real-time demand at the location and then bid that quantity as a function of price in the real-time market. A system of penalties for deviations from real-time plans is a common method for inducing bidders to balance supply and demand in real time. There are many ways to do this. We propose a tolerance model and provided the hourly usage of the service provider is no greater than this, no penalty applies; the service provider pays the clearing prices times the number of GB/hr used. A second cost is added however if the deviation is greater than the tolerance. In our case this is equal to the price times the penalty factor times the squared deviation. This can be fine-tuned to get desired behaviour.

More details of the open access market approach, including finer details of the auction processes, can be found in Cramton and Doyle (2016).

6 BACK TO SPECTRUM

This book is very much about spectrum and how we use spectrum for the public good. Hence it is important to return again to spectrum and attempt to understand the value of the open access capacity market in this respect.

It is widely accepted that spectrum is a highly valuable resource and that the demand for spectrum will continue to rise, as shown by predictions such as that of Cisco (2016). It is useful to look at this in some detail. To do this we draw on a 2016 study (Tech4i2, Real Wireless, Trinity College Dublin CONNECT, and InterDigital 2016), commissioned by the EU, titled "Identification and Quantification of Key Socio-Economic Data to Support Strategic Planning for the Introduction of 5G in Europe," in which a chapter was devoted to the study of spectrum requirements needed to deliver future 5G services. The study examined different use cases for the year 2025 to develop an understanding of the expected spectrum demand.

Of the uses studied, the one we will discuss here is "the motorway use case." This is relevant here because it draws heavily on mobile communications. The motorway use case involved the study of a typical motorway junction to understand at a high level what spectrum would be required in 2025. The expected total number of devices per square kilometre at the junction, the operating data rate/usage rate of the devices, and spectral efficiency were taken into account. Based on the services envisaged in various EU 5G research projects rather than on speculation by the authors of the report – with added input from open workshops – vehicle-based smart hubs, augmented reality glasses, tablets, and on-board video systems were among the types of devices considered for the motorway use case, as well as the myriad communication systems that are (or will be) part of the typical car. The devices were densely deployed geographically and proportionally assigned to three frequency sub-ranges (Sub-1 GHz, 1–6 GHz, and above 6 GHz). The spectrum estimates within each sub-range were calculated by multiplying the number of devices by their respective occupancy of the spectrum in bits per second according to the scenario and divided by the assumed spectral efficiency of the technology used for each device type. The spectrum demand was added across all device types to yield a total spectrum estimate for the use case.

Most importantly for this chapter, the spectrum requirements were estimated based on different network operator scenarios. These scenarios assumed four operators,[5] as well as five different sharing arrangements spanning from the case in which the four operators operate independently, to the case in which there is 100 percent spectrum sharing between operators. The scenarios that involve 20, 50, or 75 percent sharing are ones in which different densities of incumbents exist in the bands, therefore limiting the potential for full sharing to different degrees.

Figure 10.4, reproduced from the socio-economic study, summarizes the spectrum needs for the motorway use case, derived from analysis of what might happen at an actual motorway junction across all network operator arrangements. In an exclusive licensing environment in which the operators function completely independently, the spectrum needed is equal to the total-use-case–driven demand estimate multiplied by the number of operators in the environment. This is of course an extreme scenario as in reality operators tend to serve a percentage of the market rather than each operator having an expectation that it should be capable of supporting 100 percent of the market. In a fully shared environment, the spectrum needed is equal to the total-use-case–driven demand estimate.

Obviously one message from Figure 10.4 is that the spectrum demand for 2025 is extremely large, even when sharing is taken into account. The need for more spectrum is well established of course, but these results become more startling when placed side by side with Figure 10.5, generated by RealWireless (2019), which summarizes the *maximum amount of spectrum* available within in the three spectrum sub-ranges.

For the sub-1 GHz range and the 1–6 GHz range, not enough spectrum physically exists to respond to the demand unless full sharing is possible, and even then, it means that mobile communications will completely dominate spectrum usage in these ranges. In the range above 6 GHz there is enough spectrum. It remains to be seen how much of this will be set aside for international mobile telecommunications (IMT), which is the collective term for 3G, 4G, and 5G. Currently there is no spectrum in this range allocated to IMT, though a number of bands are under consideration.

The socio-economic study does not specify how the sharing scenarios might be implemented, though the study lists different potential sharing approaches. However, if we accept that this level of demand will emerge,[6] we need to return to the focus of the chapter to ask what 100 percent sharing looks like.

One answer is that the open access capacity approach is an example of how 100 percent sharing could be achieved, especially for the sub-1 GHz and 1–6 GHz bands. It is not the only way, of course, but importantly it is a way that can start today based on current technologies. There are many options for customizing LTE network parameters, which can be used to direct capacity based on market share. The technical details are beyond the scope of this chapter, but an

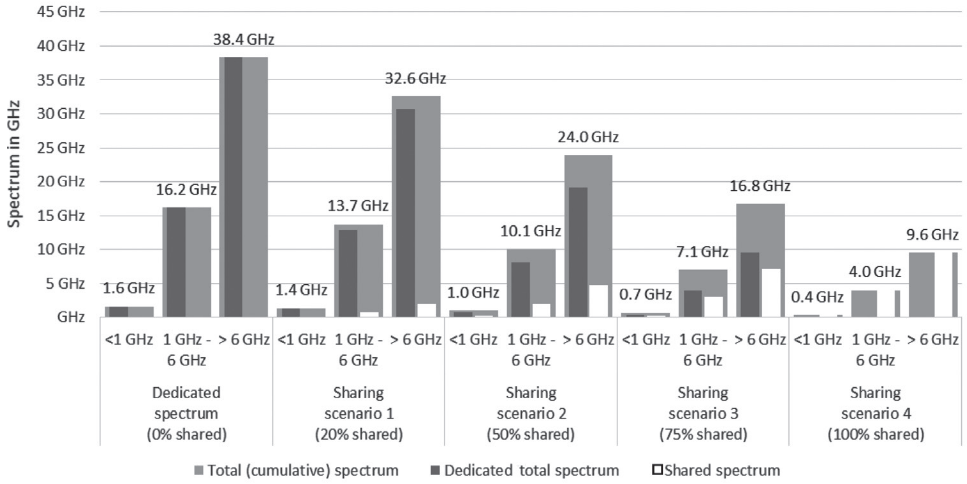

Figure 10.4 Spectrum demand for the 2025 motorway use case, for different sharing conditions

Source: Reproduced from Smart 2014/0008

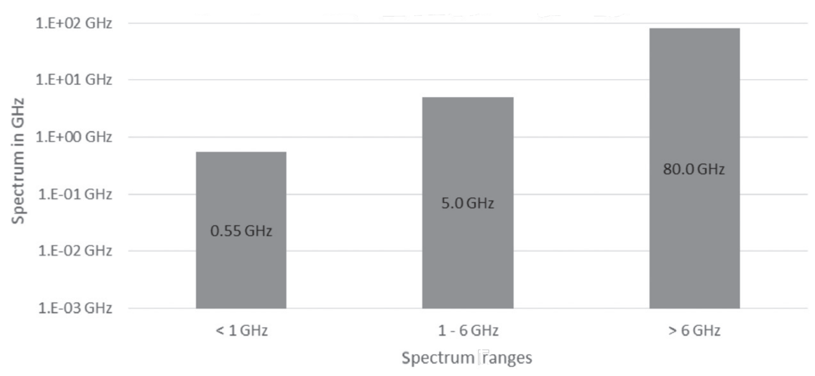

Figure 10.5 The maximum amount of spectrum (total spectrum) available in each sub-range

Source: Real Wireless

examination of sections of the LTE specifications will reveal the many options for this, as can be seen, for example, in the 3rd Generation Partnership Project document TR 22.852 V13.1.0 (3GPP 2014). It is worth noting that LTE is so widely deployed that opportunities for significant benefit from the solution presented here exist today (e.g., companies such as Rivada Networks already have working solutions) and well into the future, and that those opportunities can evolve with the network as the principles and framework remain the same, even as the technology advances.

The motorway scenario is for 2025 and so is some way off. But if we work through the motorway scenario, and possibly consider future trends, we can see how the open access capacity might work. A wholesale network could comprise any infrastructure covering the location. In the future this infrastructure would comprise traditional cells as well as communications infrastructure that comprises road-side furniture and other elements. A set of service providers might exist supporting various motorway services ranging from entertainment, to connected car, to those related to maintenance and safety. It is entirely feasible that different traffic profiles might exist among those service providers and that capacity demands could vary temporarily and spatially. At 100 percent sharing, the capacity would be optimally divided among service providers.

7 CONCLUSION

At the centre of this chapter is an open access market for capacity, which we have presented to show that it is possible to design a very flexible system for accessing capacity that might suit large and small players and that works on large and small scales. It was presented as a means to take aspects of the ideology of highly fluid spectrum trading and make it real and feasible, by focusing on both spectrum and infrastructure and by using current technologies. It was also presented as means of driving sharing to the level that is needed to fulfill future demands for wireless and mobile applications. And it was presented as part of the journey toward a world in which we think much creatively about ownership and control of spectrum and infrastructure.

The wholesale network at the heart of this chapter is a neutral or dedicated wholesale network. It provides access to additional capacity for MNOs (so that they need not fall back on overprovisioning of their own networks) at as spatial and temporal granularity that can match demands. It provides a way of supporting existing MVNOs without conflict of interest. The fact that the wholesale network is neutral and the fact that any interested party can participate in the market removes the hold the MNO wholesaler has on deciding who has access to the wholesale network. It also facilitates the emergence of new types of

MVNOs, especially those that provide IoT or M2M services and therefore have very different demand patterns than are typical today. The granularity of spatial and temporal access that is possible means that a heterogeneous range of existing or future services and service providers can be accommodated, and in a manner that ensures capacity it used in an optimal fashion.

The system described here can be implemented with current technology, and companies such as Rivada Networks are in the process of rolling out such a system. The system can also be implemented within many regulatory frameworks. So, for example, though not adopted by the Mexican government, this approach is highly compatible with the Red Compartida approach as described in Chapter 5. The system can also be implemented at any scale. While there may be economic constraints related to the scale of deployment needed to provide economic returns, it is currently possible to convert any one mobile network into a neutral wholesale network that operates an open access market.

Finally, it is possible to evolve the system as technology changes. There are many different ways in which technology might change. The approach described here can be generalized to any cellular network that allows for spatial and temporal access to capacity. Though of course the idea of an open access market itself is more generally and widely applicable. Advances within cellular networks can easily be accommodated and may in fact contribute to the attractiveness of the solution. One example that bears this out is network slicing. Network slicing allows multiple virtual networks to be created on top of a common shared physical infrastructure. The virtual networks are then customized to meet the specific needs of applications, services, devices, customers, or operators. So for example, a slice that suits low-latency or low-powered applications can be created. While early forms of network slicing are quite static, the open access market structure could potentially be used to bid for different slices on a per cell basis. This means it could be an open access market for more than capacity and support the type of competition discussed by Cave and Webb in Chapter 11.

Spectrum is the lifeblood of wireless and mobile communication applications. Access to spectrum can be a roadblock to or an enabler of new services and encourage innovation. The world of dynamic spectrum access sought to facilitate access to spectrum for small and large players, for short- to long-time durations, and over areas of different sizes. The full vision of the world of dynamic spectrum access remains unrealized. This chapter has aimed to reawaken that vision in a different guise, namely in the guise of highly dynamic access to capacity, through an open access market.

ACKNOWLEDGMENT

This material is partly based upon works supported by the Science Foundation Ireland under Grant No 13/RC/2077.

NOTES

1 One exception is the Red Compartida network in Mexico, as described by Mariscal (Chapter 5).

2 The open access market model is not new. Open access is the foundation of today's restructured electricity markets. Many modern wholesale electricity markets, such as those in the United States, operate on this open access principle and price energy at every time and location. Pricing energy at every time and location is called *locational marginal pricing* (LMP) in the real-time market. Locational marginal pricing is a way for wholesale electric energy prices to reflect the value of electric energy at different locations, accounting for the patterns of load, generation, and the physical limits of the transmission system. LMP is a mechanism for using market-based prices to manage transmission congestion. Prices are determined by the bids/offers submitted by market participants. The charge for transmission usage is the incremental cost of the redispatch required to accommodate that transmission usage. Locational marginal prices differ by location when transmission congestion occurs – areas that have more congestion will have higher prices. If there is no transmission congestion, the charge for transmission usage is zero (except for other charges to recover portions of the embedded cost of the transmission grid, etc.). Open access markets in the electricity sector work extremely well. The high level of price transparency not only leads to efficient short-run decisions but also provides a wealth of market information for longer-term planning, including future network investments. Open access is the key force that has led to competitive wholesale electricity markets that have provided reliable electricity supply while saving consumers many tens of billions of dollars from an efficient and competitive market for electricity, as described by O'Connor, Philip, and O'Connell-Diaz (2015).

3 Real Wireless has carried out an analysis for the UK, based on data supplied by different operators (4G.co.uk 2017).

4 Note that the wholesaler does not need to release all capacity on the forward market. The wholesaler can decide to release only a certain percentage. However, all capacity will of course be available in the real-time market.

5 Four is the number typically considered as offering competition in EU states.

6 It is worth commenting on any bias that may exist in the spectrum demand example chosen. The study on the "Identification and quantification of key socio-economic data to support strategic planning for the introduction of 5G in Europe" does provide a justification for the results, and some sensitivity analyses are included. It also includes other use cases, in the areas of health and utilities, that will make lower demands on spectrum, though those demands will still be high. The use case most needy of spectrum was described here, because it is the most relevant in a mobile communications world. However, even if a different use case were selected, the total demand would be the sum of *all of the demands* (those considered in the report and others) as 5G is very much about *digitalization of all verticals*. It is also worth noting that historically, spectrum demand has been significantly underestimated.

REFERENCES

3GPP (3rd Generation Partnership Project). 2014. "TR 22.852 – study on Radio Access Network (Ran) Sharing Enhancements." http://www.tech-invite.com/3m22/tinv-3gpp-22-852.html.

4G.co.uk. 2017. "4G frequency bands – Which UK networks will my phone work on?" January 2017. https://www.4g.co.uk/4g-frequencies-uk-need-know.

Cisco. 2016. "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2015–2020." http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.pdf.

Cramton, Peter, and Linda E. Doyle. 2016. "An Open Access Wireless Market Supporting Competition, Public Safety, and Universal Service." White Paper. http://www.cramton.umd.edu/papers2015-2019/cramton-doyle-open-access-wireless-market.pdf.

Doyle, Linda E., and Tim Forde. 2007. "Towards a Fluid Spectrum Market for Exclusive Usage Rights." 2nd IEEE International Symposium on

New Frontiers in Dynamic Spectrum Access Networks (DySPAN '07), Dublin, 620–32.

Doyle, Linda E., Jacek Kibiłda, Timothy K. Forde, and Luiz DaSilva. 2014. "Spectrum without Bounds, Networks without Borders." *Proceedings of the IEEE* 102(3): 351–65.

Electronic Communications Committee. 2014. ECC Report 205 – Licensed Shared Access. https://www.ecodocdb.dk/download/baa4087d-e404/ECCREP205.PDF.

Hazlett, Thomas. 2011. "Creating Efficient Spectrum Property." Spectrum Markets: Challenges Ahead Workshop, Evanston.

Milgrom, Paul Robert. 2004. *Putting Auction Theory to Work*. Cambridge: Cambridge University Press.

O'Connor, Philip R., and Erin M. O'Connell-Diaz. 2015. "Evolution of the Revolution: The Sustained Success of Retail Electricity Competition." COMPETE Discussion Paper. https://sites.hks.harvard.edu/hepg/Papers/2015/Massey_Evolution%20of%20Revolution.pdf.

RealWireless. 2019. "Real Wireless: Independent Wireless Experts." http://www.realwireless.biz.

Tech4i2, Real Wireless, Trinity College Dublin CONNECT, and InterDigital. 2016. "Identification and Quantification of Key Socio-Economic Data to Support Strategic Planning for the Introduction of 5G in Europe." SMART Number: 2014/0008 https://connectcentre.ie/wp-content/uploads/2016/10/EC-Study_5G-in-Europe.pdf.