Population and access to local services

Staff paper - authors: Kyle Thomson, Dr Karen Malam and Lucy Williams

At a glance:

- This paper sets out the basic economic reasons why population has an impact on the services available in a local area and their cost.
- A population threshold varies depending on factors which influence the demand and supply of the service in question.
- On the demand side, the more potential users, or in other words, the larger the market population, the greater the potential demand for the service.
- On the supply side, the minimum price at which a service can be supplied is related to the cost structure of the service. At least at low levels of production, the cost structure of a service is characterised by economies of scale. As a consequence there are some low levels of production which are inefficient (because they cost more than they benefit consumers).
- Services which have to cover their own costs (self-sustaining services) face the additional constraint that the revenue they collect from consumers must be greater than their costs.

Introduction

The objective of this paper is to show why population has an impact on the services available in a local area and their cost. The approach taken is to show that markets tend to have a minimum population before a service provider will enter the market, and then explain why this occurs and how this relates to competition. The paper uses a standard neo-classical micro-economic framework, but makes explicit some often overlooked implications of the model for markets that have similar cost structures but differing levels of demand.

The paper is divided into three sections. The first provides experimental estimates of a minimum population required, or population threshold, of a market for a range of public and private services. The second briefly outlines demand side issues and relates demand to population. The final section focuses on the supply side and shows how economies of scale can cause population thresholds. It also shows how the cost structure interacts with demand to shape the services that are provided and the level of competition in markets.

The concepts discussed in this paper follow on from previous BITRE Information Sheets which have investigated regional access and service provision. These papers have included a discussion of the distribution of the Australian population (BITRE 2019a) and the origins and effects of economies of scale and scope on services (BITRE 2019b and BITRE 2019c respectively). Detailed information on the service location data used in this research can be found in Appendix B of BITRE 2019a.

In investigating regional access to services, it is important to be clear about what we mean by access. Rather than using a single definition, this paper uses a broad and multi-dimensional concept of access that reflects people’s ability to derive benefit from services, as explained in more detail in

Box 1. Access costs are a similarly broad category of costs that span everything that a consumer has to pay, forgo, or would pay to forgo to access a service. These are related to every dimension of access. However, the focus of this paper are costs that have a relationship with space. This is because the particular focus of this investigation is on the relationship between service access and social geography.

Costs related to space are described here as *transport costs* and are a subset of access costs, although they are much broader than simply the direct cost of moving from one location to another. They include any cost that localises markets to a specific geography and so cover every type of cost to the consumer that is related to moving from one location to another to access a service. Example components of transport costs include the monetary costs of transport (bus tickets, fuel etc.), the opportunity cost of time (e.g. missed work or paying a babysitter), the effort of travelling and any uncomfortableness (disutility) of travelling, or the reduced effectiveness of a service due to the time taken to access it.

The most important feature of transport costs is that they localise demand in space. For example, due to transport costs, customers are unlikely to travel from other countries to a local pie shop, even though they invariably sell 'the world's best pies'. This creates the difference between a local service, which is close by and easily accessible, and a non-local service, which is further away and is more costly to access.

Transport costs are particularly important for services, as services are consumed as they are produced. This is related to the defining characteristic of a service: the consumption of a service is inseparable from its production (Thomson et al. 2019). Modern technology does allow for some separation between where a service is produced and where it is consumed, for example phone and internet services. However, if the service needs to be accessed in person, then the implications of inseparability become important as inseparability implies that consumption needs to take place at the same time as production. In other words, services cannot be stored or transported. If a service needs to be accessed in person, either the consumer will have to go to where the service is produced or the producer will have to produce the service at the location of the consumer. In either case, there will be transport costs that shape producers' and consumers' choices.

A market can be defined as the extent of the interaction between buyers and sellers of a good or service. This includes the medium of the transaction, for example a physical market like a grocery shop or an electronic market like a stock exchange. In a market, buyers stop buying when the benefit of an extra unit is less than the cost to them of buying that unit. Similarly, suppliers stop selling when the price received for an extra unit is less than the cost to them of supplying that unit. This leads to the market equilibrium, where the cost of the last unit sold by a supplier (the marginal cost), equals the benefit of the last unit sold to a consumer (the marginal benefit).

Box 1: What is access?

The term ‘access’ may seem intuitive. However, it can be quite difficult to precisely articulate what we mean when we speak about how ‘accessible’ a service is, at least for the purposes of comparing and analysing differing levels of access. In the context of this report, a useful starting point is to think about access as ‘the ability to derive benefits from things’ (Ribot and Peluso 2003). As such, the definition of access relates to the constraints people encounter when they attempt to use services and the extent to which they are able to derive benefit from services. This includes the perspectives of both consumers and producers, because services exist in the context of interaction and exchange that is related to both demand and supply.

Previous research has defined access to services as a number of interrelated dimensions (Penchansky and Thomas 1981). Each dimension of access is best thought of as a characteristic of how the service is accessed; these can include things such as the quality, price or physical location of a service. Importantly, these dimensions cannot be reduced to a binary relationship of ‘having access’ or ‘lacking access’. Rather, they can be thought of as existing on a continuum. Consumers make decisions about which services they consume and will often trade access on one dimension for access on another. Likewise, producers make trade-offs between the dimensions of access they are willing or able to provide to consumers. The dimensions we have found most useful in understanding access to services are set out below.

Dimension	Consumers	Producers
Time	The availability and time taken to access a service.	The timing of a service and the operating hours.
Space	The travel time and travel costs of consuming the service.	The location of the service.
Price	The consumer’s expectation of prices and ability to pay for the service.	The price set by the service provider.
Quantity	The amount of the service available to consume.	The quantity of the service produced by the service provider.
Quality	The extent to which the service directly satisfies consumers’ needs.	The degree to which the service directly satisfies consumers’ needs and meets government and industry standards.
Acceptability	The degree to which a service is adapted to allow a consumer to benefit from a service.	The degree to which the service provider responds to the varied consumer needs to allow them to benefit from a service.
Information	The consumer’s knowledge of the nature and availability of services.	The dissemination of information about the services available and their features.
Awareness	The consumer’s understanding of their own needs and the knowledge of how to satisfy them.	The providers’ understanding of consumers’ needs and how to satisfy them.

Source: BITRE review of access literature.

A detailed discussion can be found in the BITRE staff paper *What is Access?* (Reoch & Thomson 2019).

A market does not necessarily have a spatial component as the medium: the buyers and the sellers are potentially spread across the entire world. However, for services that are accessed in person, transport costs mean that the market is localised in space and has spatial boundaries. These boundaries, or the market area, mark the extent of the furthest buyers from the furthest sellers and are dictated by the benefits and costs of consuming a service. Where there is a transport cost, that cost is included in the cost to buyers and sellers in the transaction, causing the market to be geographically localised.

Many government services are not strictly characterised by markets, as the consumers do not buy the service and the producer does not sell the service. However, there are still consumers and suppliers who interact through some medium, be it physical such as a Centrelink office, or electronic such as *myTax*. Although they do not strictly form a market, government services that are accessed in person have a similar spatial component and boundaries to market services that are accessed in person. These boundaries essentially mark the extent of the furthest consumers from the furthest suppliers and are again dictated by the benefits and costs of consuming the government service. While not strictly a market, rather than invent a new term that covers both the spatial extent of private and public services, we use market area to describe the spatial extent of transactions and participants, whether public or private.

Experimential estimated population thresholds

A population threshold is the minimum population of a market before a good or service could be supplied. Empirically we can see that there is a population threshold which depends on the type of service in question. Note that there has been no adjustment of the population to take into account how different demographic groups are users of different services. In practice, we would expect this could be refined considering the population within groups that the service is limited to (for example, school aged children or the elderly etc) or the regional variation of these groups.

A detailed explanation of the method used here to estimate the population of a market and population thresholds can be found in Appendix A. However, in summary: the method is to define the geographical market of services of a given type, calculate the population in the market and divide this between the number of services in that market. The distribution of market populations per service point can then be examined to find the minimum population at which we observe a given service type. This distribution is summarised in Table 1.

Table 1 summarises all services of a given type across Australia. The bottom 10th percentile of the market population provides an approximation of the population threshold without being influenced by extreme outliers. This is used to represent the population threshold of the minimum population required for a service to be present.[1] The median market population represents a usual market population for a service of a given type, while the 90th percentile is an indication of the maximum population per service provider. The Market Area Limit represents the geographic size of a market for a given service type, and is defined as the distance at which the resident population is no longer a statistically significant predictor of the number of services in a city, town or village.

These descriptive statistics are based on the observed distribution of services and population as it exists now, not an analysis of the fairness or efficiency of each service. As such, the statistics represent what is and should not be assumed to represent an outcome that is either necessarily

[1] While this represents a simple descriptive statistic, the results correspond with other entry thresholds for the first firm derived through ordered probit regression in the manner of Dranove et al. 2003. For the sake of providing results for services with too few service points to estimate a catchment (and given the very similar results where this can be calculated) descriptive statistics have been used rather than modelled estimates.

equitable or efficient. For some services this outcome is the result of market forces and firm strategy, while for many others it represents the complex decision making and trade-offs of government and not-for-profit providers.

Table 1: Estimated distribution of market populations by service

	Estimated market population per service point			Estimated Market Area Limit (km)
	10th percentile	Median	90th percentile	
Centrelink - Access Point	300	3,800	103,400	63
Australia Post	300	3,300	13,900	52
Centrelink - Agent	400	8,300	121,900	49
Schools - Government Primary	400	3,600	10,100	26
Employment Services	1,000	7,400	42,300	82
Schools - Government Secondary	1,100	13,100	30,400	65
Public Hospitals - Very small	1,300	6,500	154,800	79
Aged Care - Home Care (Low)	1,400	13,200	47,900	57
Aged Care - Residential Care (Low)	2,300	8,800	24,200	32
Public Hospitals - Small	2,400	12,200	186,400	72
Aged Care - Residential Care (High)	2,700	9,700	25,100	67
Schools - Catholic Primary	3,600	15,400	31,300	42
Major Grocery Retailers	3,700	10,800	21,400	86
Schools - Independent Primary	4,200	16,800	43,800	87
Schools - Independent Secondary	5,000	21,900	53,700	86
Private Hospitals	5,300	31,600	92,500	61
Aged Care - Home Care (High)	6,200	27,000	82,000	31
Centrelink - Customer Service Centre	6,300	57,500	148,300	51
Public Hospitals - Small (with Surgery/Obstetrics)	8,300	36,700	241,000	58
Schools - Catholic Secondary	8,500	42,000	98,000	20
Schools - Government Special	17,000	52,800	128,700	35
Aged Care - Transition Care	23,100	158,300	534,600	44
Medicare	26,800	78,500	198,300	88
Schools - Independent Special	27,200	103,500	365,100	74
Schools - Catholic Special	29,400	291,100	951,500	69
Public Hospitals - Medium	47,600	157,900	306,100	72
Public Hospitals - Large	81,300	225,800	407,400	70
Public Hospitals - Principal Referral	197,000	589,200	1,332,500	61

Source: BITRE analysis of the estimated resident population 2016 (ABS 2017) and service locations. For details on the service location data, see Appendix B of BITRE 2019a.

The existence of population thresholds has direct consequences in terms of access because it means that people who live in markets with small populations do not have as many local services. This shapes the spatial access to services and begs the questions of why population thresholds exist, and why they vary between services. The following sections approach these questions through the lens of economic forces that act on supply and demand.

Demand for local services

While a population is readily observable, a population threshold is more a reflection of the underlying demand for services. The relationship is fairly straightforward: where there are more people, there is more demand. On the one hand, for services which each person needs to consume at some point in their life, such as obstetrics, education services or postal services, demand directly relates to the number of people. On the other, for specialist services which only a proportion of people will use, for example surgery, oncology or employment services, the greater the number of people, the more likely there will be people needing the service.

There are of course exceptions as some groups in society are more likely to require services and these groups may be spatially concentrated. Aged care is a good example: populations that have more senior citizens (e.g. seaside towns) will require more aged care services than a population with a young demographic (e.g. university campuses). Similarly, other individual traits like income and preferences have an impact on demand, especially for private services.

Demand itself is related to the benefits the service provides. When consumers pay for the service themselves, the maximum they would be willing to pay is the benefit of that service to them. Ordering all consumption from highest to lowest willingness to pay per unit (i.e. private benefit per unit) creates the demand curve. When a government or not-for-profit organisation pays for a service on behalf of consumers, again, the maximum these providers would be willing to pay is the benefit they perceive from the consumption of the service. This can be quite different from private willingness to pay, either because of imperfect information, or because of positive and negative externalities.

There are additional benefits to consumers from a *local* service, that is, a service that is provided in the local area and is easily accessible. These can be broken down into two categories of benefit. First, the transport cost savings, as people do not have to go to an alternative location. Second, the benefits to people who use a local service who would not pay the transport costs to access it at an alternative location (derived demand). High transport costs make a local service more attractive as the cost of an alternative is more expensive. Similarly, the frequency, predictability, urgency and severity of a consumer’s needs all influence the demand for a specifically local service. These characteristics are briefly described below.

- The frequency with which the service is accessed directly affects the transports costs, as the transport costs need to be paid each time the service is consumed. Services that are regularly accessed, like schools and retail services, have a high frequency, and so the total transport costs of non-local services are much higher than they would be if the services were accessed once a year, instead of daily or weekly.
- Predictability is also an important factor in the total transport costs of a non-local service. More predictable needs and wants have lower access costs as consuming them can be planned. This reduces the opportunity cost of the consumer’s time and there is a greater choice in how to access the service, for example in the mode of transport and in the particular service provider used. For example, imagine the difference between the non-monetary costs (e.g. discomfort, wait time) of attending a hospital emergency department as compared with a scheduled medical appointment.
- The urgency of a need increases the demand for a local service. Where a service is addressing a need which cannot be put off or delayed, the time savings and increased opportunities to access a local service increase the benefit it brings. Trivial as it might sound, public toilets are a good illustration of a service which is useful when the need arises but have no value if they are not close enough to be used in time.

- There is also a higher demand for a local service when the severity of the need, or in other words, the cost of not using a service, is high. At one extreme medical services where non-use would otherwise lead to death have large benefits and so a high demand. In contrast, for some discretionary services, for example a pedicure, the cost of not using the service is low and so demand for a local service is correspondingly lower.

While we observe the population threshold, what we are in fact seeing is a demand threshold. This in turn relates to the benefits to individuals of consuming the service. For a specifically local service, the benefit is related to the frequency, predictability, urgency and severity of the need that the service addresses.

The supply of local services

The fact that population thresholds reflect demand, which is in turn influenced by the benefits of having a service provided locally, is only one side of the story. The other is the cost of providing the service. At its most basic, for a service to exist someone must pay for it. This means that the total cost of the service must be covered, either by consumers or by a government or not-for-profit organisation on their behalf. This is true regardless of whether a producer is public or private, or whether the consumer pays the costs of the service or another decision maker in government or a not-for-profit agency does. In addition, the proportion of the population demanding services does vary between service types, resulting in differences in a minimum market size required to cover costs.

The impact of economies of scale

The reason that there is some minimum level of demand (i.e. a population threshold) is related to the pervasiveness of economies of scale in service production. These are situations where it is cheaper per unit to produce more of a good or service than to produce less, or more formally, where the average cost of production falls with increased production. This has an important impact on whether or not it is viable to provide a local service, and explains why services with higher costs require larger populations to be viable.

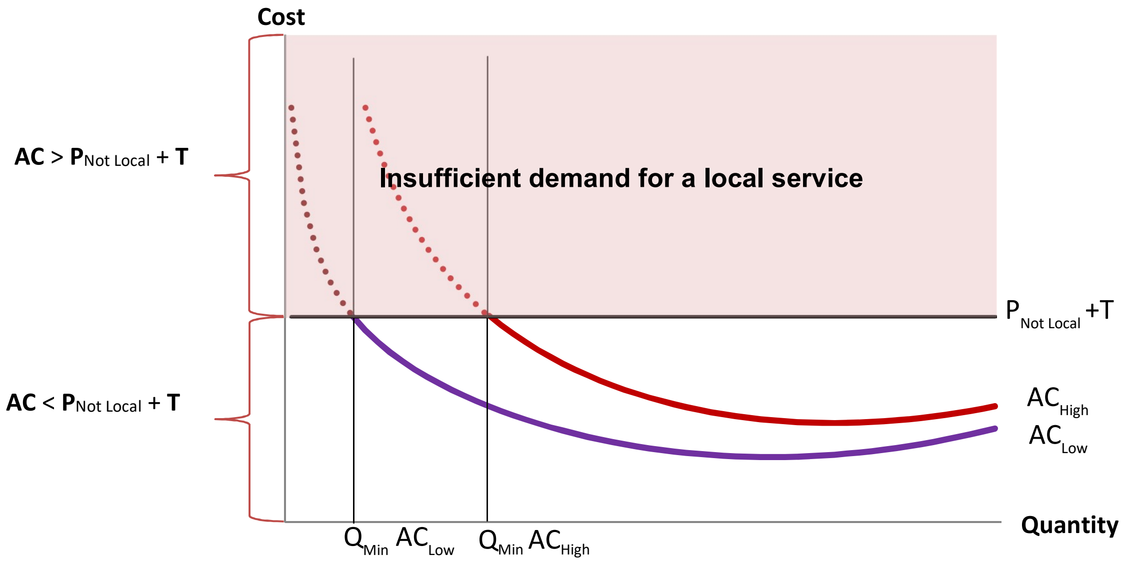
Although they can be caused by other factors, economies of scale most often exist because there is some indivisible fixed cost of production, such as the cost of a lease, basic equipment or minimum staffing. In short, an indivisible input has to be bought or consumed in a fixed quantity which is larger than is needed to produce just one unit of a good or service. As more is produced this fixed cost is split between additional units, causing the average cost per unit to fall. This means that it is on average cheaper per unit to produce more than to produce less. A detailed discussion can be found in the BITRE Information Sheet *Economies of scale and regional services* (BITRE 2019b).

There are logical limits to the amount any decision maker would be willing to pay for a local service. The most a decision maker will be willing to pay is up to the amount that they would have to pay elsewhere, including the extra transport costs involved. Again, the transport costs include all of the costs related to using the service at an alternative location. Where the average cost of producing a local service is above below the cost of the alternative, on average the costs of providing a local service will outweigh the benefit. This means the *average cost must always be below the cost of the cheapest alternative for a service to be viable*.

Economies of scale are defined by the average cost decreasing as more is produced. Because the average cost decreases as more is produced there may be some lower levels of production where a local service costs more than the alternative (including the transport costs). However, there may be larger levels of production where the local service is cheaper than the alternative. This is shown below in Figure 1, which plots average cost curves that exhibit economies of scale with the constraint that the average cost must be lower than the next best alternative, $P_{Not\ Local} + T$. The region where there will be no demand for a local service is shaded pink.

Figure 1 shows two average cost curves to illustrate the effect of higher costs on the population threshold. Both cost curves exhibit economies of scale by decreasing as quantity increases. However, the first average cost curve, AC_{Low} is lower at every point in production than the second cost curve, AC_{High}. The minimum quantity for each average cost curve is shown along the horizontal axis and corresponds with the point at which average cost is the cost of the next best alternative, formally: $\frac{C(Q)}{Q} = P_{Not\ Local} + T$. Demand must be high enough that a quantity greater than the minimum can be supplied, otherwise a local service is more expensive than the cost of the alternative. The population threshold is the level of population that corresponds to the minimum level of production required for the cost of the local service to be equal to the cost of an alternative service.

Figure 1: Economies of scale and population thresholds



Comparing the minimum quantity of the low average cost curve, $Q_{Min}\ AC_{Low}$, to the minimum quantity of the high average cost, $Q_{Min}AC_{High}$, it can be seen that holding the alternative $P_{Not\ Local} + T$ constant the minimum quantity increases as the average cost curve shifts outwards.[2]

This result is driven by the fact that average cost is falling as output increases. Without these economies of scale the minimum quantity produced could possibly be one unit. With economies of scale, and if the best alternative is at some point cheaper than a local service, then the minimum quantity must be greater than one unit. For higher cost structures the minimum quantity increases because of economies of scale. Given a relationship between the size of a local population and demand for services, this explains why services with higher cost structures require larger populations for them to be provided locally.

[2] For simplicity the constraint shown in Figure 1 is constant, however in practice it is likely to increase with quantity, as more distant consumers pay greater transport costs (i.e. T is not constant, but increasing with quantity).

Some further implications that flow from this framework are:

- Holding transport costs and the cost of an alternative service constant, as the total cost of running the service increases, the minimum quantity increases.
- All else being equal, if transport costs increase then the minimum quantity decreases. This is because the greater the access costs the more is saved by having a local service.
- As a town becomes more isolated (i.e. transport costs increase) the quantity (population) required to make a local service worthwhile decreases.

Self-sustaining providers and markets

The following section considers the additional constraint imposed on self-sustaining service providers, or in other words, providers which only derive revenue from their consumers. The discussion to this point has not drawn a distinction between who pays for the service. In practice there are of course differences between service providers that do not need to recover the costs of providing a service (e.g. they are government funded or have a cross subsidisation of costs), and those that do. There are organisations in each of the for-profit, not-for-profit and government sectors that receive subsidies and are not required to be self-sufficient. However, in each sector, and especially in the for-profit sector, there are organisations that need to cover the costs of providing the service through revenue generated by charging customers.

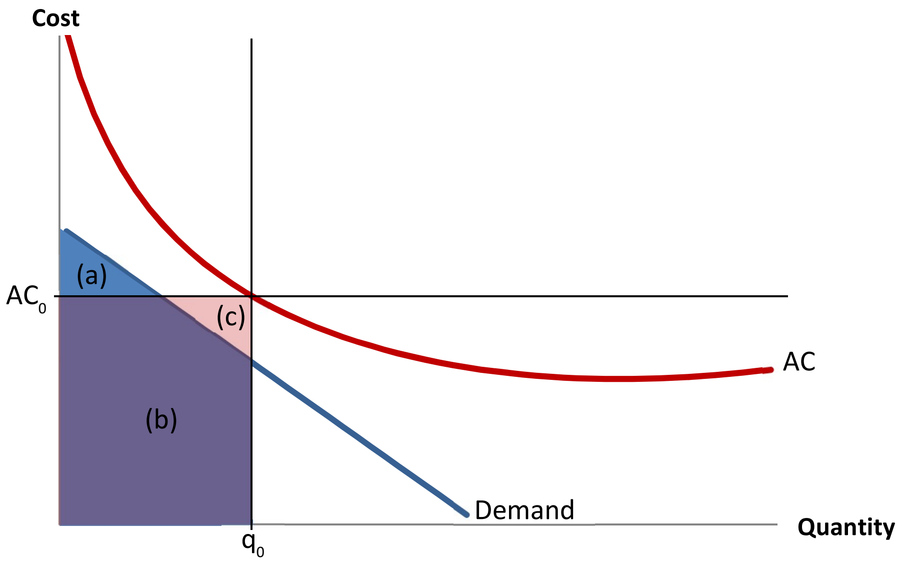
The need to cover the cost of the service based only on revenue derived by charging for the service is an additional constraint. Not only does the demand have to be greater than the total cost, but the total revenue gathered from consumers must be greater than the total cost.

The following section discusses the implications of this additional constraint on whether or not a local service will be provided and the prices that will be charged. Following on from the sections above, we assume that demand is at least in part some increasing function of population and that there are returns to scale at the lower end of service producers' production functions. For simplicity, the market used in the following scenarios is made up of a single service type with a fixed market population.

No producers and no market

The first example to consider is a very small population resulting in such low demand that even a single service provider cannot generate sufficient revenue to cover costs. In this situation, for any price that the marginal consumer would be willing to pay, the average cost of producing that unit of service would be higher, causing the producer to make a loss. This (non) market is illustrated in Figure 2, which shows a demand curve (solid blue line) that is always below the average cost curve (solid red line).

Figure 2: No entry



Source: Adapted from Webb 1976.

In this situation there is no price that consumers would be willing to pay which covers the average cost, so a self-sufficient organisation could not survive and no service would be provided. Depending on the total willingness to pay and total costs, this could be either an efficient outcome, where the total cost of providing the service is greater than the benefit, or an inefficient outcome, where the total benefit is greater than the total cost.

To see this we have taken a quantity, q_0, and shown the average cost at that point, AC_0. The total cost of producing q_0 is the area of the square made up of the areas labelled (b) and (c) on the figure above (i.e. AC_0 x q_0). The total benefit of producing quantity q_0 is the large area under the demand curve (b), as well as the area of the blue triangle in the top left of the chart (a), which represents the consumers' willingness to pay less the cost, or in other words, the consumer surplus. As the area (b) is both part of the total cost and part of the total benefit, the difference between the total benefit and the total cost is simply (a) – (c).

If (c) is greater than (a) then the total costs are greater than the total benefits, and not providing a local service is efficient. However, if the area of (a) is greater than (c) then the total benefits of producing a service are greater than the total costs, and this results in a market failure.

The market failure arises because there are economies of scale and because there is no single price that can capture the high willingness to pay for a service of some consumers and at the same time be low enough to satisfy enough consumers with low willingness to pay. If the benefits outweigh the collection and opportunity cost of public funding this situation may require government intervention.

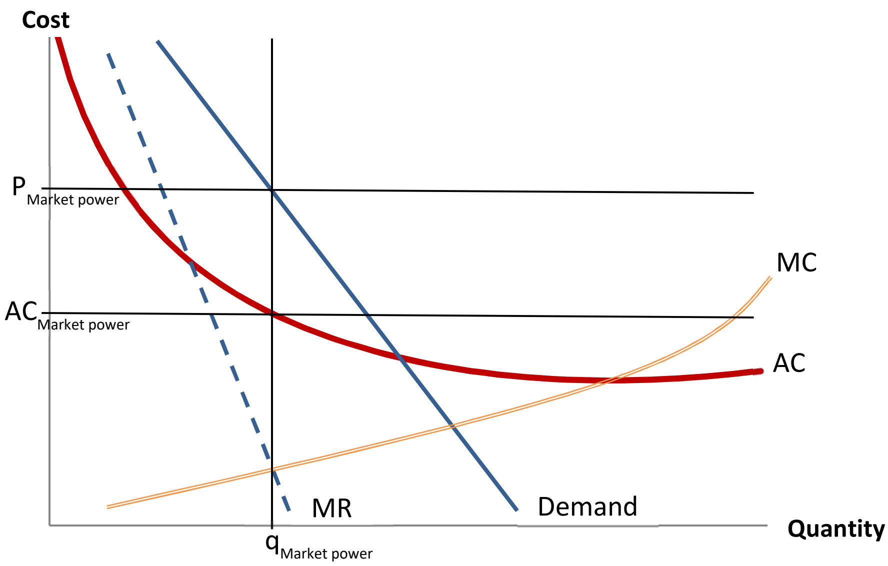
A simple alternative solution is to try to capture the high willingness to pay of some consumers and use this to cross subsidise low willingness to pay consumers, or in other words, to price discriminate, so that there is sufficient demand to support the service. To put this strategy into a concrete example, consider a local club in a small and relatively isolated town. Assuming that the club is in the situation above, there is no single price for meals that would be able to cover the total cost (overheads, staff time, ingredients, etc.) and which would attract enough customers for the service to survive. However, the club is able to price discriminate by giving a discount to seniors card holders. This means that the club can survive by offering seniors cheaper meals (low willingness to pay due to lower incomes) which are cross subsidised by the other customers (higher willingness to pay due to higher incomes).

Producers with market power

In the second example, consider a population that is large enough to support one or more service providers, but small enough that the demand curve crosses the marginal cost curve in a production region where the marginal cost is lower than the average cost. This is shown Figure 3.

Economic theory suggests that in a competitive market the price is equal to the marginal cost. However, if the price were equal to the marginal cost in this scenario, the price would be below the average cost and the producer would make a loss. This means that in the long run, if a market exists it cannot be competitive. Instead, an existing producer can take advantage of there being a fixed cost of entering the market and the transport costs associated with accessing alternative services to set prices that deter the entry of additional producers. This leads to some degree of market power, which allows for a non-competitive outcome where the price is higher than the marginal cost.[3]

Figure 3: Market power



Source: BITRE analysis.

A for-profit producer with market power maximises profit by setting the marginal revenue equal to marginal cost. In this case marginal revenue (the blue dashed line MR), is not the price and instead takes into account the fact that for a producer with market power, producing more lowers the price.

A small population can support a producer (or producers) with market power so long as demand is high enough that the service provider is able to set the price of the service above the average cost. The result is that a service provider or providers enter the market, but that the price is higher than in a perfectly competitive market. Essentially the smaller population means that while a service is available, the market is not perfectly competitive and the price is higher.

This scenario is just one discrete example. However, this case represents a point on a continuum of population and cost structures where market power can be exercised by producers. On this continuum we have, at one extreme, a single producer carrying out business as a monopolist, and at the other, multiple producers which form an oligopoly but fall short of perfect competition.

---

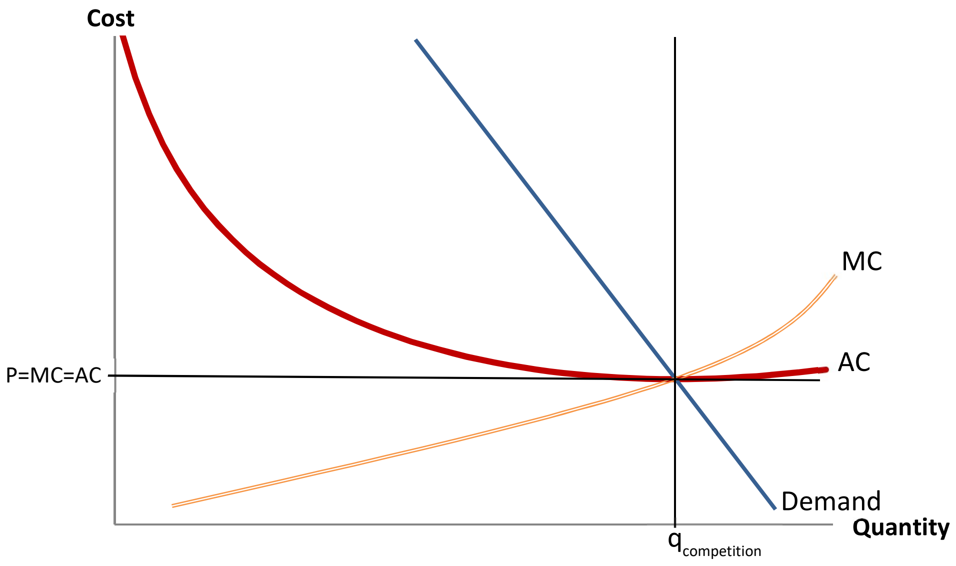
[3] This is a simplified single firm example of the equilibrium in the Dixit & Stiglitz (1977) model.

The potential for long run perfect competition

Finally, consider a large population that generates a high demand for services. In this example producers would be able to make a profit, at least temporarily. This scenario reflects any situation where demand is high enough that it intersects the average cost curve at a point where it is not falling (i.e. where there are no longer economies of scale).

In this example perfect competition is possible. If a producer is making an economic profit, then additional producers have an incentive to enter the market. They have an incentive to do this until competition has reduced the individual demand faced by each producer to the point where it is no longer possible to make a profit. This long term demand curve is illustrated below in Figure 4, which also indicates the region that corresponds to a short term profit. In this situation the service would exist, and the price would be efficient, as price is equal to marginal cost, and relatively low, at price equal to average cost.

Figure 4: Perfect Competition



Source: BITRE analysis.

The observation that in the long run firms will have an incentive to enter a market until the market is no longer profitable is borne out by the stability of the catchment populations of for-profit businesses, relative to non-self-sustaining organisations.

Hence, service providers that are not able to cover their costs if the population does not meet the population threshold, creates a lower boundary. If the population is large enough to make a profit it attracts other firms to enter, reducing the population per firm and creating the upper limit.

The key insight of this section is that the additional constraint faced by self-sufficient organisations means that the population of the market affects the market structure. At one extreme a very small population will not result in a local service being provided at all, while at the other a large population allows many service providers and potentially perfect competition to exist. In between these two extremes lie a range of scenarios where producers in the market will have a greater or lesser market power, which correspond to higher or lower prices.

Conclusion

This paper has addressed the underlying economic reasons that population has an impact on the services available in a local area and their cost. We began by showing the distribution of market populations for various service types and their population thresholds. The second and third sections of this paper have then explained why this threshold exists and how population affects competition from the perspective of demand and supply respectively.

Demand is highly related to population size and the supply is characterised by economies of scale. Because there are economies of scale there are small levels of production which are on average more expensive than the next best alternative service, but larger levels of production which are on average cheaper. If there is only a small amount of demand due to a small population, the correspondingly small level of production may not be cost effective. If there were higher demand due to a larger population, the service would be on average cheaper than the alternative and cost effective.

The final section also discussed consequences of the need for self-sufficient service providers to collect revenue from their customers that covers their total costs. This leads to the conclusion that very small populations may have no local service provider. At the other extreme, a large population will attract many service providers (and potentially perfect competition). In between these two extremes lie a range of scenarios where producers in the market will have greater or lesser market power, which correspond to higher or lower prices.

Appendix A: Estimating market size

This appendix outlines the method used to derive catchment population estimates for each of the services included in this research.

Like many studies in retail and service catchment areas, this research uses a Voronoi (Thiessen) trade area model to delineate mutually exclusive areas for each service point of each service type. Studies using or testing a Voronoi (Thiessen) trade area model include West 1981; Ghosh & McLaferty 1987; Jones and Mock 1984; Jones and Simmons 1993; Pearce 2000 and Lin et al 2014.

A Voronoi diagram is created by calculating the midpoint between sites on a plane, then connecting the midpoints to delineate the region on the plane closest to each site. Only the location of a set of sites is required to generate a Voronoi diagram, making this method attractive in the delineation of catchment areas where there is very limited information, for example where the actual behaviour of consumers is not known (Boots and South 1997). This is a lower information requirement than other common methods such the convex hull of actual consumers' home locations, the two-step floating catchment area approach or a gravity model, all of which either require information on, or make more complex assumptions about, consumer behaviour.

Although a Voronoi (Thiessen) trade area model requires very limited information to generate a working approximation of a catchment area, the underlying assumption that consumers use the closest service is difficult to defend. In practice, as consumers have different wealth and preferences, market boundaries are not hard and markets for the same product generally overlap, especially near the edges. This means that if consumers have a choice of where to consume, any mutually exclusive geographic representation of markets can only ever be approximate.

Rather than attempt to disaggregate overlapping market areas we have used the combined market area of the City, Town or Village (CTV) in which the services are located. This research uses a specific definition of cities, towns and villages, one which directly relates to service provision. In essence the definition combines two concepts: the idea of a population centre and a service centre in which one or more services are located. Where a location is both a population centre and a service centre it is considered to be a CTV. The advantage of using both population and services to define place is that this geography specifically relates to service provision and we are able to include towns and villages with very small populations that are not usually identified in statistical geography. In this way we have identified 2,450 CTVs across Australia. More information about how these have been defined can be found in Appendix A of the BITRE publication *An introduction to where Australians live* (BITRE 2019a). The services used to define CTVs include postal, retail, education, health and other government services.

Combining the catchment areas of the CTV has been achieved by combining the catchment areas (individual Voronoi polygons) of each site located in a CTV to form a joint catchment area. This means that the market areas described here are not the geographic markets of individual services, but rather the combined and overlapping geographic markets that make up the market area of a service centre. The population per organisation estimates are therefore the population of a service centre’s combined market divided by the number of service providers in that service centre. This reduces the complexity of overlapping markets and the underlying assumption that consumers purchase from the closest CTV is more defensible, although it does not remove this issue entirely as the markets of service centres are still likely to overlap.

A further practical limitation is that in principle the extent of catchment areas delineated using a Voronoi (Thiessen) trade area can extend an infinite distance where a site is not surrounded by other sites. Even where they are bordered by other sites the size of the catchment area is half the distance to the next site. This is problematic in the Australian context because the distance between a consumer and a service can be very large, so large that it is not realistic that consumer is within a service’s (or in this case, a service centre’s) market. In order to consistently estimate who is in the market area of a given service centre for a given service across Australia it is necessary to estimate a limit to the spatial extent of a market, or Market Area Limit.

From the perspective of the consumer the spatial extent of a market is a very difficult and complex question as it depends on their needs and resources. For example, how far would a person go for potentially lifesaving treatment? Some medical services have market areas larger than a single country, with some people coming to Australia for medical treatment and some Australians going overseas. Certainly these are extreme cases rather than the norm, however across every type of service there is very wide variation in the behaviour of individual users. This makes the consumer perspective, which depends on individual circumstances, problematic to measure and compare across services.

The producer perspective is an easier problem, as it is founded on the aggregated behaviour of consumers rather than the behaviour of individuals. Producers have either implicit or explicit expectations about consumer behaviour, at least at the margin when the last producer is considering entering a market. This leads to a testable question:

> At what distance is the population no longer relevant for a producer’s decision to enter the market?

Taking for example Australia Post offices, we can phrase this question with respect to a given distance, say the 6th kilometre from services in a service centre, as follows:

> Given the population within 5 kilometres of a post office, is the population of the 6th kilometre from post offices in service centres a significant predictor of the number of post offices in those service centres?

The question can then be asked again of the 7th kilometre, taking into account the population within 6 kilometres, and so on.[4]

This was investigated using a linear regression model by testing the significance of the marginal population living an additional kilometre from services on the number of service providers of a given type in a service centre. The model controlled for the population in the proceeding kilometres, and in addition included two measures of isolation: the shortest distance to a CTV providing the same service (1) and the shortest distance to a CTV of equal or greater size (2).[5] The model was run across a panel of all CTVs for each service type and specified as follows:

$$N_{services} = \beta_0(constant) + \beta_1(Si) + \beta_2(CTVi) + \beta_3\left(\sum\nolimits_{0\ to\ j} pop\right) + \beta_4(pop_k)$$

Where:

$\boldsymbol{N_{services}}$ is the number of service providers in the CTV

$\boldsymbol{Si}$ is the distance to the nearest alternative CTV providing the given service type

$\boldsymbol{CTVi}$ is the distance to the nearest CTV of the same size (within +/-10 per cent) or larger

$\boldsymbol{\sum_{0\ to\ j} pop}$ is the total population located in kilometre_0 to kilometre_j to which services in the CTV are the closest

$\boldsymbol{pop_k}$ is the population in the k^{th} kilometre to which services in the CTV are the closest

Due to the lumpy nature of the population distribution, in practice the significance of each additional kilometre declines gradually and intermittently, so that there is usually a distance at which an additional kilometre will not be significant but the next more distant kilometre is significant. To encompass the full extent of the market we use the last significant kilometre for each service type within 100 kilometres, where significance is tested at a 99 per cent confidence level.

This process generates a Market Area Limit, which can be interpreted as the distance at which populations living further than this distance from a service located in a service centre are not significant predictors of the number of service providers in a service centre. The results for each service type are reported in the final column of Table 1 in the body of the report. As this process resulted in 2,800 individual regressions, only results (and not individual descriptive statistics) are provided.

---

[4] A similar method is used by Guerra et al. (2012) to test the distance around a transit station that best predicts direct demand at that station.

[5] In (1) the distance is calculated from the population weighted centroid of the CTV to the service point weighted centroid of the alternative CTV and in (2) between the population weighted centroid of the CTVs.

References

Australian Bureau of Statistics (ABS) 2017, *Regional population growth, Australia, 2015-16*, Cat. No. 3218.0.

Boots, B and South, R 1997, ‘Modeling retail trade areas using higher-order, multiplicatively weighted Voronoi diagrams’, *Journal of retailing*, vol.73, no.4, pp.519-536.

Bureau of Infrastructure, Transport and Regional Economics (BITRE) 2019a, *An introduction to where Australians live*, Information Sheet 96, BITRE, Canberra.

Bureau of Infrastructure, Transport and Regional Economics (BITRE) 2019b, *Economies of scale and regional services*, Information Sheet 99, BITRE, Canberra.

Bureau of Infrastructure, Transport and Regional Economics (BITRE) 2019c, *Economies of scope and regional services*, Information Sheet 101, BITRE, Canberra.

Dixit, A and Stiglitz, J 1977, ‘Monopolistic competition and optimum product diversity’, *The American Economic Review*, vol.67, no.3, pp.297-308.

Dranove, D, Gron, A and Mazzeo, M 2003, ‘Differentiation and Competition in HMO Markets’. *The Journal of Industrial Economics*. LI(4): 433–454, December 2003.

Ghosh, A and MacLafferty, SL, 1987, *Location strategies for retail and service firms*. Lexington Books.

Guerra, E, Cervero, R and Tischler, D 2012, ‘Half-mile circle: does it best represent transit station catchments?’, *Transportation Research Record: Journal of the Transportation Research Board*, vol.2276, no.1, pp.101-109.

Jones, K and Simmons, JW 1993, *Location, location, location: analysing the retail environment*, Nelson Canada, Scarborough, Ontario.

Jones, KG, and Mock, DR 1984, ‘Evaluating retail trading performances’ in RL Davies and DS Rogers (eds), *Store Location and Assessment Research,* pp. 333-360, John Wiley & Sons, New York.

Lin, T, Palmer, R, Xia, J and McMeekin, D 2014, ‘Automatic generation of station catchment areas: A comparison of Euclidean distance transform algorithm and location-allocation methods’, in *11th International conference on fuzzy systems and knowledge discovery (FSKD)*, Aug 19-21 2014, IEEE, Xiamen, China.

Pearce, J, 2000, ‘Techniques for defining school catchment areas for comparison with census data’, *Computers, Environment and Urban Systems*, vol.24, no.4, pp.283-303.

Penchansky, R and Thomas, JW 1981, ‘The concept of access: definition and relationship to consumer satisfaction’, *Medical Care*, vol.19, no.2, pp.127-140.

Ribot, JC and Peluso, NL 2003, ‘A theory of access’, *Rural Sociology* vol.68 no.2, pp.153–181.

Reoch, A, Thomson, K 2019, *What is access?*, Bureau of Infrastructure, Transport and Regional Economics (BITRE) Staff Paper, BITRE, Canberra.

Thomson, K, Malam, K and Williams, L 2019, *What are services and who provides them?*, Bureau of Infrastructure, Transport and Regional Economics (BITRE) Staff Paper, BITRE, Canberra.

Webb, MG 1976, *Pricing policies for public enterprises*, Macmillan, London.

West, DS 1981, ‘Testing for market preemption using sequential location data’, *The Bell Journal of Economics*, vol.12, no.1, pp.129-143.

ISBN: 978-1-925843-29-3

February 2020

Use of the Coat of Arms

The Department of the Prime Minister and Cabinet sets the terms under which the Coat of Arms is used. Please refer to the Department's Commonwealth Coat of Arms and Government Branding web page http://www.pmc.gov.au and in particular, the Guidelines on the use of the Commonwealth Coat of Arms publication.

Contact us

This publication is available in PDF format. All other rights are reserved, including in relation to any Departmental logos or trademarks, which may exist. For enquiries regarding the licence and any use of this publication, please contact:

Department of Infrastructure, Transport, Regional Development and Communications
Bureau of Infrastructure, Transport and Regional Economics (BITRE)
GPO Box 501, Canberra ACT 2601, Australia

Phone: (international) +61 2 6274 7210
Fax: (international) +61 2 6274 6855
Email: bitre@infrastructure.gov.au
Website: www.bitre.gov.au