

Demand and capacity management practices in austerity: A study across UK police forces

*Paul Walley (paul.walley@open.ac.uk)
The Open University, Milton Keynes, UK*

*Matthew Adams
The Open University, UK*

Abstract

All public sector organisations within the UK have enduring a period of restricted finance. The UK's police forces are no exception, despite changing patterns of demand. This paper provides fifteen case studies of how police forces have adapted demand and capacity management strategies as a consequence. The findings show there is no standard approach to managing demand, with wide diversity in system design. A distinctive feature of police forces, unlike most other organisations, is how they adjust thresholds for dealing with demand fluctuation. There are some promising ways in which demand can be reduced without compromising quality.

Keywords: Demand, Capacity, Police

Introduction

All public services sectors are under significant financial pressure, originating after the financial crisis from 2007. In a UK context, the Police Service has experienced similar financial pressures to other public services, with a steady real-terms cut in funding from 2015 and other changes to funding from 2009 (see Elliot-Davies et al., 2016). In the period 2009-2016 the number of full-time equivalent officers fell by 14% according to the Institute for Fiscal Studies (Disney and Simpson, 2017). At the same time the patterns of demand have been under considerable change, moving away from car theft, robbery and burglary towards “white-collar crime”, internet offences, sex crime and trafficking (Keene, 2012; Vinod Kumar, 2014; Boulton et al., 2017). This potentially changes the quantity and mix of skills required by forces. These problems have become more obvious in the last two year or so, leading to a series of comments in the 2017 “PEEL” review of UK policing by the Inspectorate of Constabulary (HMIC, 2018):

“[There are] major concerns that policing is under significant stress. On occasions, that stress stretches some forces to such an extent that they risk being unable to keep people safe in some very important areas of policing... About a quarter of forces are all too often overwhelmed by the demand they face, resulting in worrying backlogs of

emergency jobs, with officers not attending incidents promptly, including those involving vulnerable people.”

Source (HMIC, 2018)

In May 2015 the national Police Chief’s Council established the second phase of a project into demand management that reported back in November 2017. Amongst wide-ranging terms of reference there were the following objectives:

- To examine demand beyond recorded crime
- To examine the sources of demand data and how these could be recorded
- To identify ongoing professional good practice and set up a central repository
- To link demand to public value in terms of how the services add value
- To utilise demand forecasting models

The report contained a number of recommendations including the adoption of tools and techniques for assessing risk and prioritisation, emphasis on collaborative working, understanding internal processes to reduce waste and improve productivity and the better use of analytics.

Within the report attention was paid to defining demand, which was split into three types:

1. Public demand is equated with incidents reported by the public (but there is a need to factor in the actual resource consumption needed to meet this demand)
2. Protective demand comes from the need to provide policing cover for events, acting on intelligence or general protective patrols.
3. Internal demand is the demand for resources within policing organisations, including administrative tasks, processes and protocols.

The model that is produced to link this together is replicated in the figure below:

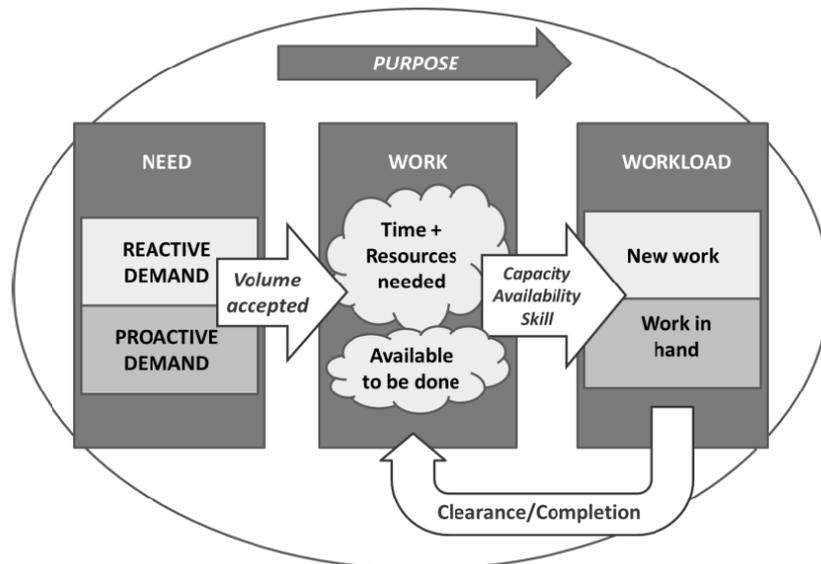


Figure 1 The NPCC representation of demand

Source: NPCC (2017)

One of the key themes of the report is to identify opportunities to reduce demand placed upon the service through a combination of selecting out demand that police do not need to attend and identifying those incidents that can be addressed through less resource

intensive solutions. An emphasis is also placed on attending incidents where those creating demand have some level of vulnerability (e.g. victims of domestic or sexual abuse).

Existing literature

In a previous paper (Walley & Jennison-Phillips, 2017) we drew attention to the perceived differences in the challenges of demand and capacity management between the public and private sectors. Literature from the private sector has long been established, with emphasis on the nature of demand in the service sector. Service demand is often instantaneous or unplanned, with seasonal and random variation (Lovelock 1992) making high utilisation a challenge. Many services also have degrees of variability that make responsiveness and flexibility necessary characteristics of service operations (Frei, 2006). The literature initially focuses on the capacity management aspects of the problem, with capacity adjustments to match demand a desirable practice (Sasser, 1976; Lovelock, 1992). Where capacity is not adjusted, demand management practices become more important. However, within the private sector there remains a desire to meet all profitable demand, so mechanisms such as price adjustments influence the timing of demand to quieter times, instead of always refusing demand when capacity limits are reached. Fitzsimmons and Fitzsimmons (2006) identify five types of demand management:

1. Reservation systems and overbooking
2. Offering complementary services that deflect demand from the core service
3. Segmenting demand so that level capacity can be better utilised by using it to deliver counter-cyclical services
4. Price incentives and price manipulation to influence the size and timing of peaks and troughs in demand
5. Promotion for off-peak times, to deflect demand away from the peak.

Within the public sector, increased demand is not usually met with an increase in revenues or resources unless there is some direct payment for service. This has significant consequences for the ways in which capacity is planned. In many cases capacity levels in the medium term are fixed and the system simply has to cope with whatever demand enters the system (Walley, 2012). Hence, public services adopt cost-centric rather than revenue-centric approaches to capacity management. Figure 2 shows how these approaches are manifested (Walley & Jennison-Phillips, 2017).

The literature on demand and capacity management within policing prior to the NPCC report (NPCC, 2017) is especially sparse. First, there are very few studies of demand for police services at all. One study (Boulton et al, 2017) highlighted the sheer diversity of the demand, with the single biggest identifiable category being concern for welfare (19% of incidents), with public nuisance (18%) and acquisitive crime (17%) also being significant. However, the biggest single category was “other” (28%), showing how the police have to deal with a wide range of rarely occurring situations. The latest crime figures from the Office for National Statistics (ONS) show that there were increases in homicides (14% increase), public order (24%) and robbery (17%). An 8% increase in knife crime has more recently become more of a topic for discussion, with a debate about the impact of the availability of police resources as a possible reason why this type of offence has become more prevalent. There is much discussion of emerging threats in policing (see Ransley and Mazerolle, 2009). The NPCC report cited previously identified 85 separate types of demand coming into police systems.

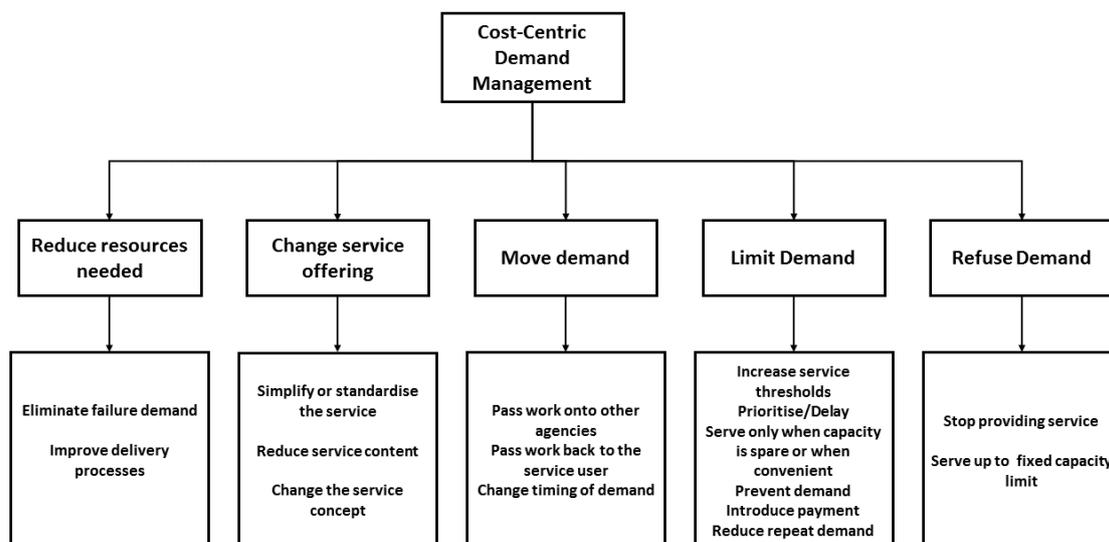


Figure 2 Cost-centric demand management practices (Walley and Jennison-Phillips, 2017)

One study has previously looked at demand and capacity management practices inside one police force, focusing mainly on custody suites (Ritchie and Walley, 2015).

Table 1 Capacity Management Strategies Observed in policing

Method	Comments
Chase demand	Assets (i.e. cells) are fixed, so in medium-term planning only adjustment of staff offers capacity change. There is limited focus on low utilization of cells, and to an extent staff. Therefore the attention to chasing demand is limited.
Increasing customer participation	In context of custody the arresting and investigating officers are the customer. The arresting officers have a specific influence on demand and throughput time. There has been awareness and engagement with arresting officers to complete all processing where possible in advance of entering the custody system.
Scheduling work shifts	A precedent has been set that although there are three shifts, these are consistent in their staffing. This is reflective of limited understanding of short term demand variability.
Creating adjustable capacity	Staff are paid both a shift and a rota allowance. However both through precedent and local affiliation the flexibility in staff assignment this is meant to provide the force is rarely used. Most of the flexibility is done through good-will agreements.
Sharing capacity	The rollout of super-custodies is meant to deliver shared capacity in terms of physical assets and staff. Further the use of ‘clusters’ (geographically close facilities) also provides some short-term capacity sharing. Hence capacity is shared across similar facilities more than switching resource from one type of service to another.
Using part-time staff	Part time staff have been in use for many years, and are inseparable from the full-time staff in their roles.

Source: Ritchie and Walley (2016)

Research Method

Between October 2018 and February 2019 all forces that are members of the Centre for Policing Learning and Research at the Open University were invited to participate in a study of police demand and capacity management practices. Out of twenty forces and agencies, fifteen were able to participate on the study within the timescales set. In most cases forces have been anonymised in the research, with the exception of specific cases where permission to identify them has been given. Forces have been labelled by letter (A – O) where single examples are provided.

Each force was visited for at least one day by the research team to gather the basic information. The contact centre for each force was observed and demand entering the system tracked to establish how work entered the system from 999 and 101 calls and then processed through to dispatch. Where more detailed cases were generated a considerable amount of follow-up information was obtained after visits, including samples of demand data, performance reports from contact centres and other reports of improvement or demand reduction work.

Information was also obtained through direct discussion in structured interviews where the opinions of force officers and staff were obtained to build up a view of the perceptions of staff responsible for aspects of demand and capacity management about the situation their force is in. In most cases, officers and staff from both operational and planning roles were seen, and this allowed contrasting views to be observed. Where permission was obtained, interviews were recorded with the condition that responses were anonymised both in terms of the force and interviewee.

Of relevance to this paper, questions were asked on the following themes:

1. How well do forces and agencies understand their levels of demand?
2. Have forces changed practices involving prioritisation and response as a consequence of demand/capacity imbalances?
3. Are forces trying to reduce demand?
4. What capacity and demand management practices have been adopted across forces?

Findings

In this section we summarise the key findings across the case studies.

Demand Measurement and Forecasting

Forces generally measured the volume of calls coming into their control centres and used this to measure demand. In all but one of the forces this data collection was carried out by a variety of human interventions and automated processes. One force that had fully automated the process through a bespoke piece of software but this was used primarily as a costing tool. . It can work out the amount of demand placed on the force by an individual or business and highlight where costs are being incurred.

Forecasting of call volumes was used in all 15 forces to determine staffing levels for the control centre. These predictions had a variety of successes with one force (F) being an example of good practice in accurately predicting between 97-98% of future calls. Few forces translated this data into hard resource requirements outside of the control centre environment, especially officers needed to meet the demand and any other policing resource, such as investigative requirements. Force M was an exception as they have measured the demand across 30 different teams and created resource models to meet the demand. This has allowed force M to predict their future resource requirements. Generally forces lacked detailed understanding of demand once the call goes beyond the control room.

Call Journeys

Although there was some variation in the nature of demand across forces because some were rurally based and others were metropolitan, the basic task of meeting demand is largely the same in all forces. The call journey requires three basic steps:

1. Call handling, where calls are picked up by an operator and assessed for urgency and risk. The outcome is to reject or filter out those calls that do not need to be attended to and pass on all others graded by level of priority.
2. Dispatch, where work is allocated to available officers.
3. Attendance, where officers attend to an incident. More incidents are attended remotely now, e.g. by phone or email.

There was an unexpected amount of variation in the call journeys in each force. The variation mainly occurred through differences in the following aspects of the system design:

- The level of division or resource sharing between 999 (urgent), 101 calls (non-urgent) and dispatch.
- The number of filtering steps before dispatch, including whether or not a switchboard is used.
- The means used to assess the risk of each call.
- Thresholds for response decisions for common incident types, such as shoplifting.
- The levels of one-touch or call handler resolution.
- The levels of integration between call handling and dispatch.
- The points where demand is re-graded, e.g. by dispatch
- The skill sets used within the contact centres (specialist staff, police officers etc.)
- The grading systems and response targets used
- The levels of additional back office support, such as assessment units and mental health support
- The types of response resources used, such as diary cars or booked appointments

Figures 3 and 4 show the basic flow of calls at force A and Force B:

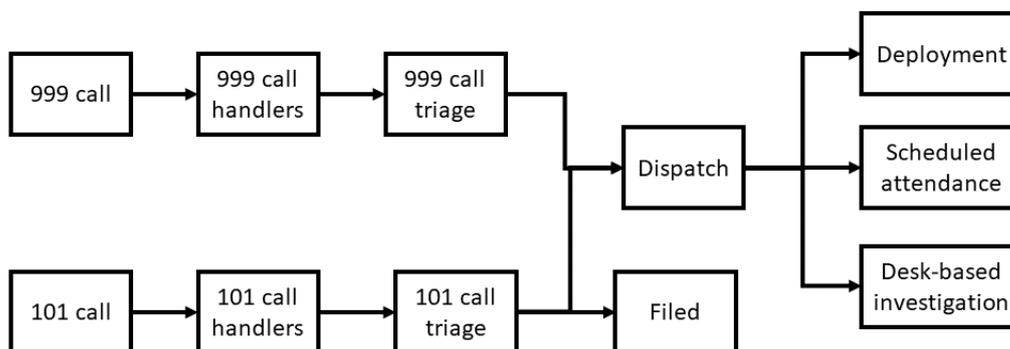


Figure 3 Call Journey at Force A

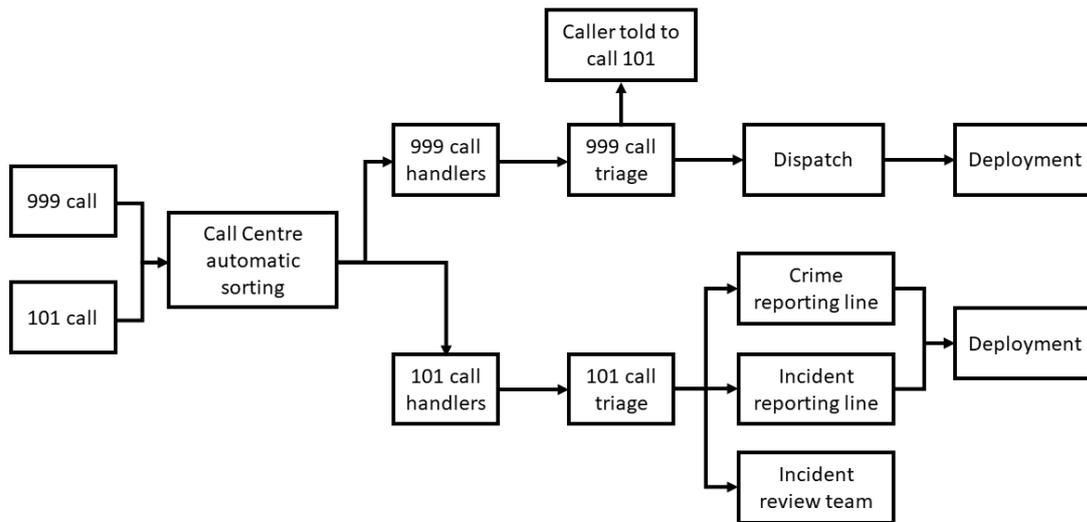


Figure 4 Call Journey at Force B

Prioritisation and Response

In all 15 forces the national decision making model was fundamental in making and justifying decisions across all ranks and departments. This changes slightly in the call centres where a method called THRIVE (threat, harm, risk, investigative opportunities, vulnerability and engagement) is predominately implemented. THRIVE has been adopted across 13 out of the 15 forces. There was evidence that all forces were simplifying their prioritisation systems, such as the number of levels of job grading, partly to improve the ways in which low-priority demand was dealt with quickly. There were clear trends towards remote resolution, where incidents would not be attended in person and would be classed as advice only. As such most forces now have just three main types of demand: urgent to be attended in person, attended soon and some form of bookable demand, such as diary car. All forces dispatched officers in a timely manner for the respective first priority category. Issues started to arise in the second category often called priority or prompt. The variety of different incidents in this category left it up to the individual dispatchers to reassess the vulnerability and risk before choosing which incidents should be dispatched to first. This problem was identified by force J, who decided to split the priority category into priority high and priority low. Hence the priority list in force J looks like this:

1. Immediate (15mins)
2. Prompt (1 hour)
3. Scheduled
4. Diary
5. Not Dealt with or closed

One of the issues that needed to be addressed in many forces was that of re-grading work where dispatch were unable to send officers within the target time. In many instances, across most forces, a call would be graded as a priority but there would not be the resource to immediately allocate to the work. There were many comments about the problem of the level of unresolved calls still being handled at any one time. In practice, once a response was going to be missed the dispatch team would re-grade the call, usually to a lower grade, including “not deal”. The actual frequency of this occurrence is difficult to objectively measure, partly as there was little desire to highlight this within the control systems.

Forces reported much demand that was unnecessary or was demand that should not be the responsibility of the police. There were common problems associated with demand associated with non-police matters such as noisy neighbours, inconsiderate parking, fly tipping and other civil matters. All forces had a concept of avoidable demand even if they didn't use that specific term. The definition provided by forces was very similar. The general consensus was that avoidable demand was demand that the police shouldn't be dealing with. This definition does differ from the formal definition of avoidable demand (demand arising from behaviours that can be changed). The lack of a consistent design archetype for control centres did mean that some failure demand was generated by some of the systems in place. For example, many forces had all 101 calls arrive at a switchboard before the work was passed to call handlers in the contact centre. However the role of the switchboard varied where some merely filtered out calls that were routine contact with office staff but others deflected demand that was deemed inappropriate or unnecessary. There was also significant variation in how work passed through from call handling to dispatch.

Demand Reduction

All 15 forces have taken steps to improve their ability to manage demand or indeed reduce demand. These practices have resulted in both success and failure but demonstrate that forces are striving for improvement.

Forces have set up protocols that identify types of demand that should not be dealt with by the police. The actual topics of these protocols varies considerably. These are often incidents such as fly tipping or noise complaints that can be dealt with by another agency in a more appropriate manner. It is not the case that forces are just refusing calls for service but they are educating the public and advising them to contact one of their partners. However, the way that forces have approached this varies between blanket policies that state that they will refuse to deal with a particular call. For example, force A refuse to deal with noisy parties and lost property. By contrast, most other forces will conduct a risk assessment on the call before deciding if they will refuse deployment or not. Some forces have taken this protocol and applied it to calls for service that included a crime. This was approached in two contrasting ways displayed by force A and force B. Force A have started to only respond to shoplifting if the value of the theft is over a stated value. Force B has taken a different approach and they assess the call based on solvability factors. If there are no solvability factors and no vulnerability or risk they will not investigate the crime. Forces have also moved towards increased use of telephone resolution to close an incident in the control room and therefore prevent an officer being dispatched.

Three forces had adopted comparatively sophisticated methods of reducing failure demand (Seddon 2009; Randle and Kippin, 2014). Gloucestershire Constabulary were one of the forces that conducted an in-depth study of their unnecessary demand. They studied a sample of non-urgent demand and discovered that, for every 100 calls that could have been resolved in one contact, the demand created was 160 actual contacts. The number of contacts per incident varied quite considerably, with up to seven extra unnecessary contacts on a single incident (see Walley and Jennison-Phillips, 2018).

Capacity Strategies

All forces were able to study demand profiles broken into short (say 15 minute) blocks of calls coming into the contact centres. This did allow them to make adjustments to centre staffing levels at any particular time of day or week. A key measure for these centres is the response time for emergency calls and systems were designed to maintain good

performance in this area. (Some forces did not regularly achieve the target response time of 10 seconds for emergency calls). Forces also made attempts to meet locally-set targets for response to non-emergency calls through capacity strategies. However, these strategies were not carried through as effectively in response functions due to the underlying demand-capacity mismatch. There were capacity adjustments to cope with weekly seasonal demand fluctuation, but as one officer put it “we need to choose when we are most short of capacity”. Demand management or demand deflection is still used to move demand, through the use of appointments. However, the most interesting aspect of current practice concerns the flexing of response thresholds to cope with demand at peak times. The threshold of what would be deemed a response call is adjusted, usually at dispatch, to maintain the balance between demand and resource availability. On a quieter day a less serious or urgent call may be attended. On a busy day an identical call might be graded as “not attend”, re-graded as less urgent than the original call handler grade or left in the system until the incident de-escalates.

Discussion

The findings support the idea that police demand and capacity management fits within the developing theory of cost-centric demand management practices. There is work to reduce the level of resource needed to meet demand through the elimination of failure demand and through attempts to streamline processes. However, this work was seen to be the most challenging to implement as it requires system and behavioural change for it to be effective. As a consequence it was only those forces that were developing high-level demand management strategies that were taking this approach.

It is also clear that changes to the service offering are happening, usually as an emergent strategy that finds ways to reduce resource commitment to some types of demand. In particular, remote resolution through telephone advice and recording only of incidents is happening more frequently. The most significant change is where demand will not be met if there is no policing benefit – such as if attendance will not result in catching a criminal.

Another difference to existing theory comes from how systems adapt to random and seasonal demand variation. Whereas their current theory points towards chase capacity strategies that flex capacity to meet demand, in the case of policing the threshold of what they are prepared to do flexes to limit demand to what they can cope with. There is also some flexing of capacity, but this is mainly limited to the most obvious seasonal peaks and troughs. The elements of chase adjustments are part of a formal plan whereas the threshold adjustments are unplanned but happen relatively consistently when demand exceeds capacity.

All forces declared strategies to move demand from themselves to other agencies, and there was probably some naivety in the belief that demand coming to them as the “wrong place” could easily be shifted elsewhere. In practice the problem is more complex, especially where incidents involve those who are vulnerable. In such cases demand can only be met in cooperation with partner agencies, sometimes having to both deal with an incident at the same time. There was no consistent pattern in how this was achieved, with a variety of initiatives that identified specific themes of partnership working.

At the present time, one of the responses to austerity is to limit demand by simply refusing to attend demand where there is a known capacity constraint. Most of the forces had made announcements about the type of demand that they were no longer willing to deal with, including some types of lesser crime.

Conclusion

This research suggests that police forces in the UK are in a state of continuing adaptation and change with regards to the management of demand and capacity. All forces appear to share the common problem that demand outstrips their current effective capacity to attend to all demand. All forces have had to adapt their demand and capacity management practices as a response to austerity, with more sophisticated approaches to both managing demand and making adjustments to capacity. Our main conclusion is that, however, there is no one single approach that has been adopted, consistent with operations management theory. At present there is instead a wide variety of solutions that are being developed across the forces, where the theory would normally predict a more consistent response.

References

- Disney, R. and Simson, P., (2017) Police workforce and funding in England and Wales, *Institute of Fiscal Studies, IFS Briefing Notes BN208*, London.
- Fitzsimmons, J.A. And Fitzsimmons, M.J. (2006) , *Service Management: Operations, Strategy, Information Technology*, New York, McGraw-Hill.
- Frei, F.X., (2006), “Breaking the Trade-Off Between Efficiency and Service”, *Harvard Business Review*, November, 93-101.
- Howgego, J. (2019), “A UK police force is dropping tricky cases on advice of an algorithm”, *New Scientist*, 8 January: <https://www.newscientist.com/article/2189986-a-uk-police-force-is-dropping-tricky-cases-on-advice-of-an-algorithm/>
- Keene, S.D. (2012), Emerging threats: financial crime in the virtual world, *Journal of Money Laundering Control*, Vol. 15 Issue: 1, pp.25-37.
- Klassen, K.J And Rohleder, T.R. (2002), Demand and capacity management decisions in services; how they impact on one another, *International Journal of Operations and Production Management*, 22 (5), p527-548.
- Lovelock, C.H. (1992), “Strategies for Managing Capacity Constrained Service Organisations” in Lovelock (ed), *Managing Services: Marketing, Operations and Human Resources*, Prentice Hall, N.J., p154-68.
- NPCC (2017), *Better Understanding Demand – Policing the future*, NPCC Performance Management Coordination Committee, London.
- ONS (2018), Crime in England and Wales: year ending September 2018, ONS, London.
- Randle, A. and Kippin, H. (2014), *Managing Demand: Building future public services*. RSA, London.
- Ransley, J and Mazerolle, L, (2009), “Policing in an era of uncertainty” *Police Practice and Research, An International Journal*, Volume 10, 2009 - Issue 4, pp365-381
- Ritchie, R. And Walley, P. (2016), “The challenges of public sector demand and capacity management: an exploratory case study of police services”, in Radnor et al. (eds.), *Public service Operations Management, a research handbook*, Routledge, Oxford, pp. 117-13
- Sasser, W.E. Jr (1976), Match Supply and Demand in Service Industries, *Harvard Business Review*, 54 (6), p133-140.
- Seddon, J (2009), “Failure Demand – from the horse’s mouth”, *Customer Strategy*, Issue 1, Vol 2 Winter 2009 pp 33-34.
- Vinod Kumar TK (2014) “Differing services, rising expectations, and greater demands”, *Policing: An International Journal of Police Strategies and Management* 37(1): 170–189
- Walley, P. (2012), “Does the public sector need a more demand-driven approach to capacity management?”, *Production Planning and Control*, Online March 2012, ref: DOI: 10.180/09537287.2012.2012666886, pp1-14
- Walley, P. and Jennison-Phillips, A, (2017) “Demand management in the public sector: developments and issues”, Euroma Conference, July, Edinburgh.
- Walley, P. and Jennison-Phillips, A, (2018) “A Study of Non-Urgent Demand to Identify Opportunities for Demand Reduction”, *Policing: A Journal of Policy and Practice*, pay034, <https://doi.org/10.1093/police/pay034>