

George Yarrow, Regulatory Policy Institute

Security of supply: general issues and implications for the water sector

The issues for consideration

Efforts to promote upstream competition in the water sector can be expected to give rise to a number of questions concerning the implications of such reforms for security of supply. It is entirely reasonable that such questions be raised, and be taken seriously, in that the development of upstream competition could be expected to involve some movement away from the current business and regulatory arrangements for addressing security of supply issues. Even when the end states of any process of transition to more market-based arrangements are clear and well-defined, the very fact of transition can give rise to significant problems on account of the simultaneous existence, for a period, of characteristic features of both the old and the new regimes. Such combinations of 'old' and 'new' can potentially cause uncertainty, confusion about responsibilities, and unintended effects when the relevant policy issues have not been adequately thought through.

The discussion below considers some of the relevant issues in the specific context of the water services sector, but begins with a number of general points concerning the meaning of the term 'security of supply' and the expected relationships between security of supply and market competition. These clarificatory points are important because, in public discourse on security of supply issues (in sectors such as energy, as well as water), there can be initial confusion about the meanings of terms, which serves not only to inhibit the development of better understandings of the issues at stake, but also to create 'imaginary' problems that, in reality, simply don't exist.

The meaning of security of supply

Discussions of security of supply typically encompass a range of different issues and concerns, and the distinctions among these issues and concerns, which can be of considerable importance in the development and application of regulatory policy, are sometimes not always made clear.

At its most general, security of the supply refers to the ability of a resource allocation mechanism to ensure that the requirements of consumers or customers are met, period by period. This immediately raises a number of questions:

- Given that demand is influenced by a number of factors, including price, how are the relevant consumer/customer requirements to be defined?
- Are all consumer groups to be treated in a roughly similar way?

- If consumer/customer requirements are not met for a period, how long does this period have to be before it is considered that there is a significant security of supply problem.

This list of questions is not exhaustive, but it does at least provide an initial disaggregation that allows for more precise analysis of the issues.

Supply security in competitive markets

In a typical market context, the answer to the first question – relating to definition of consumer requirements – is usually couched in terms of willingness of consumers to pay for the relevant product or service at levels that will cover the costs of supply. On this basis, security of supply can be analysed in terms of familiar supply and demand factors. In competitive markets, prices will clear the market in such a way that demand-side requirements, at the equilibrium price, are met. If supply conditions tighten, price will rise, but it would not be normal to refer to this as a *security* of supply problem or failure. Flows of the product/service directed toward meeting demand would be reduced, but would typically be secure in the sense of being temporally stable.

Rather, security of supply issues would more usually be considered to occur where there is some sharp, usually temporary and often unexpected, disruption of supplies such that:

- The market fails to clear for a period, implying the existence of excess, unsatisfied demand until clearing is restored.
- The market clears but at price levels or in ways (e.g. black markets) that, for one reason or another, are considered unacceptable in terms of the public policy objectives of the day.¹

For reasons that will be explained below, the distinction between these two cases is important, because they tend to be associated with different economic effects/impacts that have different costs for different groups of consumers. There is, however, an important common consideration in both cases: *security of supply is not necessarily to be equated with continuity of supply over time.*

To explain, particular groups may prefer to contract for the possibility some interruption or reduction in supplies for a period, in return for a lower average price over a more extended

¹ Although the reason most usually cited is affordability of the product/service for low income households, it can be noted that pressures to suppress price increases that might be required to clear a market facing disrupted supplies can come from other sources too. For example, a less elevated public policy objective might be to protect/subsidise a downstream industry that uses the relevant product as an input.

period. Such an arrangement is simply a form of demand management, whereby the consumer responds to temporarily high prices. Not only are such (*ex ante*) commercially negotiated periods of supply interruption not an indication of security of supply problems, the existence of such arrangements actually contributes to greater security of supply in the market in two ways:

- The demand responsiveness tends to make it more likely that the market will clear in the presence of adverse supply shocks: interruptible demands shoulder part of the burden of the adjustment,
- The demand responsiveness tends to reduce the magnitude of the price hike that is necessary to secure market clearing.

That is, those customers who are less able to reduce demand and are therefore less price sensitive are, to at least a certain extent, protected in relation to security of supply risks by the existence of customers who are more price sensitive.

Another situation that is sometimes referred to as raising security of supply issues occurs when there is an anticipated, future fall away in established sources of supply to a market, or anticipated inability of established sources to meet a growing demand. A good example is the discussion of potential sources of energy 'when UK continental shelf gas runs out'.

There is some potential sense in this classification (of long term developments as 'security of supply issues') if there is cause to expect that a major change in the pattern of supply will have significant implications for the likelihood of major, but temporary, disruptions of supply.² More often than not, however, there is little or no basis for such an expectation, and it is simply the result of applying a 'common sense' but false intuition to the effect that, since we don't know precisely how future supplies will be delivered, we must assume that those supplies will be less reliable/secure than current sources of supply.

Failure of market clearing

Failure of a market to clear does not necessarily have major costs. For the period of the imbalance, alternatives to price as a rationing mechanism will inevitably develop. For durable goods, stock draw-downs and order backlogs provide initial mechanisms for adjusting to imbalances. For non-durable goods, first-come first-served queuing systems may be observed.

² Increased reliance on imported fuels tends to be *assumed* to be associated with less reliability in flows, presumably on the basis that foreigners are less to be trusted than natives. For UK energy, however, the evidence indicates that, at least over recent decades, the major shocks have tended to have domestic origins, with industrial action at the top of the list of causes (i.e. natives having been less dependable than foreigners).

In terms of general economic efficiency effects, the 'welfare costs' of imbalances can be estimated by familiar welfare triangles, and the losses will depend upon factors such as the magnitude and duration of the imbalances and the elasticities of supply and demand curves. There can also be equity issues in the event that the rationing system is such as to lead to some groups being particularly disadvantaged in their ability to secure supplies. However, non-price rationing is not invariably bad for the more vulnerable consumer groups who may be of particular concern to policy makers (as illustrated by the extreme case of war-time food rationing, where alternatives to the price mechanism are deliberately introduced for equity reasons). It is therefore a matter for empirical determination as to whether there are particular groups who are likely to be seriously harmed by a failure of a market to clear.

Security of supply issues in the energy sector

In the energy sector, on the other hand, there are specific types of phenomena that can give rise to very high costs of imbalance; and high costs can occur even if the duration of the imbalance is very short. These matters are frequently discussed in debates on energy policy, but it is important to recognise their special character, which means that mechanistic translations of energy sector experience into other context, such as water supply, are best avoided.

In electricity, the problems are very basic, and stem from the fact that, unlike for most markets, if there is a significant imbalance in supply and demand, even for a very short time period, *the whole of the supply is lost* across potentially large geographic areas. This follows from the physics of electricity supply systems: minor imbalances can be absorbed by adjustments to voltage and frequency – manifest in a phenomenon such as the Sunday joint taking a bit longer to cook than normal – but the tolerances are very limited, and, if they are violated, the supply system tends to crash.

It is because of this 'balancing externality' that the costs of imbalance in electricity are very high, and that considerable effort and resources are devoted to maintaining short term security of supply, for example by maintaining sufficient reserve capacity on the system to respond to equipment failures that could disrupt the energy flows.

The gas system is less sensitive to imbalances, but it faces a related problem. If supply falls relative to demand, pressures in the pipeline system can fall and can cause risks of explosions. Again, therefore, the costs of supply disruptions, even for relatively limited periods, can potentially be high.

In both gas and electricity markets, therefore, the function of 'systems operations' has been developed, which includes, among other things, taking responsibility for addressing the 'balancing externalities'. In GB, the system operator for both electricity and gas is the National Grid. This institutional development/innovation has had significant success, and,

as confidence in market arrangements has grown, there has been some tendency for the scope of systems operations to narrow somewhat, at least in relation to balancing activities.

Some commentators have also argued that, in addition to the various balancing activities undertaken by the system operator, it is necessary for additional policy measures to be undertaken in order to ensure security of supply in the face of temporary imbalances/disturbances/shocks in/to the supply-demand balance. Examples of such proposed, additional measures include capacity payments in electricity (which involve payments, at regulated prices, to generating capacity made available to the system, and which were used in under the early post-privatisation pooling arrangements) and 'strategic' storage in gas.

There is no economic consensus on these issues, however. It is generally agreed on the theoretical side that wholesale electricity market designs with and without capacity payment mechanisms can work in principle, and that the detail of the relevant factual context is what matters in determining how things work out in practice. In England and Wales, the capacity payments system under the old pooling arrangements did not, in practice, function as intended.³ On the gas side, the chief concern with strategic storage options is that they can 'crowd out' private provision, and therefore either fail to enhance security of supply or else do so only at excessive cost (see the 'high prices and policy failures' section below).

³ The approach required the regulator to set a price that was, in effect, an estimate of the "marginal value of capacity to the system". This is a highly uncertain exercise, since it effectively requires the regulator to determine an option value (the option provided by the existence of an extra unit of available capacity to generate an extra unit of electricity, if required to do so), leading to a high likelihood of mistakes (cost-based price control is difficult enough for regulators, option-value based price control is an order of magnitude more difficult to get even approximately right). More important, the form of the payment arrangements was such as to create incentives to *withhold* capacity from the system, so as to achieve higher revenues. Not only did this add to consumer bills, but the incentive structure actually served to tighten the demand/capacity balance, contrary to policy intentions. One lesson to be learned here is the sensitivity of outcomes to the specific details of the factual context. Although the option pricing was always bound to be a bit of a guess, the capacity payments system had a clear logic in the context of a *perfectly competitive* market structure (see Vickers and Yarrow, *The British Electricity Experiment*, Economic Policy 1991, for an account of the theory). Like nearly every other market, however, wholesale electricity is not, and probably never can be, perfectly competitive (which is no bad thing, since the concept of perfect competition is entirely static in nature, and the great bulk of the benefits of competitive markets arise from their dynamic properties). A second lesson is that, since factual contexts are constantly changing, it is always important to check the "robustness" of policy options across a full range of contexts that might eventuate. Measures such as capacity payments, designed to affect market conditions at times of system peaks/stress, when circumstances are often (and almost definitionally) 'abnormal', are notoriously prone to non-robustness flaws.

Networks and markets

An aspect of the security of supply issues that arise in the energy sector which is more relevant to water supply stems from the distinction between network activities and supply activities. Each category of activity here can be a source of supply-demand imbalances. For example, insufficient power in a particular area of the country may be the result of a transmission equipment failure, and may occur even though, on a national basis, there is more than enough generating capacity available to the system as a whole to meet national demand.

It can also be noted that the two activities are not fully separable influences on supply-demand balances. One or more major power plants located in a particular region may fail, but if there is sufficient transmission capacity it may be possible for the system operator to accommodate the shock without any loss of load. However, if the import capacity of the relevant region is constrained, the generating plant failure may lead to loss of load because of inadequacy of reserve generating capacity in the region.

Put at its most simple, supply of the commodity to the end consumer depends upon both availability of the commodity itself (electricity, gas, treated water) and of the means to transport it to the customer. Failure in *either* supply *or* in transportation will lead to a failure to deliver to the consumer. Thus, if it is decided as a matter of policy to rely upon competitive markets to secure supply-demand balance in provision of the commodity, there is still a role for regulation in relation to the provision of network capacity where the latter is developed and operated in monopolistic conditions.⁴ For example, the electricity transmission network is developed and operated in the context of prescribed security standards, which require the system to be robust (in terms of being able to make required bulk, longer distance transfers of power) against a specified number of transmission line failures. The imposition of such security standards on regulated, monopolistic network activities is not at all inconsistent with the development of competition in wholesale and retail supply of the underlying commodity (whether electricity, gas or water).

Longer term issues in energy

In relation to the longer term issues of supply-demand balancing in energy, the main current of policy thinking since privatisation has been that this is a matter that can be left to market forces. However, since networks are regulated monopolies, and since network development and access is an important factor influencing the timing and location of new

⁴ Not all network capacity provision is necessarily regulated. For example, there is 'merchant interconnection' in the energy sector (although it is more common in gas than in electricity), LNG shipping networks can be (and are) competitive, and there is competition in gas storage.

investment in electricity generation, gas fields and terminals, storage facilities, etc, there has been a recognised role for the provision of information about system development. The National Grid seven year (electricity) and ten year (gas) statements are examples of this kind of information provision, which is aimed at facilitating better co-ordination among potential investment projects that are, in economic terms, *complementary* to one another (rather than *substitutable* or *competitive*, where coordination would generally be regarded as potentially anti-competitive).

More recently there has developed a much greater propensity on the part of central government to seek to influence the detail of investment in electricity generation (e.g. preferences against or in favour of coal, nuclear, wind, etc.), driven in large part by environmental policy considerations. There is a strong argument that these policy influences have, in fact, created potential security of supply problems in the energy sector where none had existed previously, since the 'problems' appear to be significantly correlated with the prevalence of new, 'command and control' interventions. In gas, for example, where such policy interventions have been much less in evidence, there is very little indication of long term security of supply problems, notwithstanding the substantial (and faster than expected) decline of UKCS output: capital markets have been willing to invest in the development of alternative sources of supply, whether by increased interconnection to the continent or in new LNG terminals at which gas transported by sea from a variety of different sources across the globe can be landed in the UK.

These developments point to the importance of *policy/regulatory failure* as a source of security of supply problems in the energy sector, an issue whose importance is difficult to over-state.

Temporarily high prices and policy failures

Ironically perhaps, at least from an economics perspective, the major problems associated with security of supply in sectors such as energy have oft times tended to arise from intolerances within the public policy system for high prices, even on a temporary basis. Thus, in a competitive market system, a disruption to supply can be expected to lead to an increase in market prices, with the size of the price hike being greater the lower is the elasticity of demand for the commodity in question.⁵ Some prices, however, are more

⁵ This illustrates again the importance of developments/innovations in demand management, which can be stimulated by effectively competitive retail markets: a well functioning market can promote higher demand elasticities, leading to lower demand spikes (in the face of supply shocks) and to greater security of supply. The tradition of utility regulation has, for most of its history, been focused on supply side responses only; and measures nominally aimed at increasing security of supply have tended to crowd out developments on the demand side which might have been much cheaper and much more effective.

politically sensitive than others (and not necessarily because of inherent characteristics of the products and services themselves – it may just be a matter of a history of intervention in some economic areas and not in others).

In the relevant circumstances, references to “security of supply problems” tends to be code for “a period of unacceptably high prices”. The lessons of economic history indicate that political intervention to hold down prices below competitive equilibrium levels does not have a track record of success: it is nowadays better understood that the likely effect is to dampen, if not destroy, supply side incentives, to the great disadvantage of consumers in the longer term.

Intervention motivated by a desire to address “security of supply problems” sounds much more consumer-friendly than “price control”, however; and the notion that the intervention is restricted to those limited periods when there are supply disruptions appears, at first glance, to be an example of well-targeted intervention. There is, however, a large element of illusion in all this, or at least a temptation to be misled by use of language: measures directed at holding down ‘peak’ prices *are* forms of price control, and they can therefore be expected to have the traditional effects associated with all attempts to hold down prices to below market clearing levels.

Competitive markets and security of supply

It is a matter of historical record that competitive markets have an excellent record in ensuring sustained flows of products and services to meet the requirements of consumers/customers, and it is a major achievement of economic thought over the centuries that it is possible to offer a coherent account of why this is so. Common sense intuition suggests to many people that supplies will be more secure if there is someone in charge of, or responsible for, ensuring security of supply; but such intuitions are not only wrong – being both contrary to the historical experience and explicably wrong in terms of theory – but potentially harmful if translated into economic policy making.

A major reason competitive markets do so well in ensuring sustained supplies is attributable to the ‘feedback’ properties of the price mechanism. Increased scarcity of a product leads to a higher price, and higher prices provide signals and incentives for (a) increasing supply and (b) reducing demand. If there is a disruption to supply, the resulting price hikes therefore provide an instant intensification of incentives to respond in appropriate ways – i.e. of finding ways to increase supply and to reduce demand – to correct the imbalance. Moreover, the strength of the incentives is broadly proportionate to the significance of the disruption.

It is important to note that, in speaking of ‘feedback’, the economic mechanism under discussion is not at all similar to a physical mechanism in which the adjustments only occur

after, and in response to, an external shock. Economic mechanisms are *intelligent* in the sense that responses are made in anticipation of future shocks, or, more accurately, are made to the perceived likelihoods of such shocks. What is required for this to work is that the relevant supplier in a market is contractually bound⁶ to customers in ways that make it costly for the supplier to fail to deliver (i.e. ‘firm’ customer rights, rather than an arrangement under which a supplier can simply say: “sorry chaps but there is a drought that’s not our fault”, and then reduce supplies without financial cost to itself). In the words of the SkyBet advert: “it matters more when there is money on it.”

This simple and powerful feedback and incentive mechanism is absent from command and control systems, and constitutes a central defect of the latter in addressing security of supply issues. It follows that any shift from reliance on command and control to reliance on competitive market mechanisms can be expected to tend promote greater supply security.

Why then the worries that almost invariably accompany market liberalisation in the regulated sectors of the economy?

Causes of policy failures

One source of regulatory failure is simply the false intuition that, if no-one is in charge or responsible for achieving certain, desirable outcomes, then those outcomes are less likely to eventuate. This is a view/intuition that can be expected to be well represented among groups with whom regulators necessarily have to deal. Unfortunately, it all too frequently makes an appearance within regulatory agencies themselves. Whatever the source, there is nothing to be done other than to respond with reasoned arguments based on evidence and experience.

Another source of policy failure is political resistance to market clearing prices when those prices are particularly high, as they will tend to be in periods of major supply constraints. This is a very serious problem that has bedevilled electricity market liberalisation in the USA. For example, Joskow⁷ has recently identified it as one of the three major factors that have contributed to deficient investment in electric systems.

⁶ In circumstances where markets are insufficiently developed to rely on contractual negotiation between customers and suppliers, and where the relevant continuity/security standards are imposed by regulation, the requirement is for the regulator to set clear standards and rewards/penalties around those standards on an *ante* basis, and then credibly to commit to their enforcement. This provides the incentives required for effective risk assessment and management, which are often lacking in long-term planning approaches that, in practice, tend to be associated with ‘soft constraints’ in relation to performance standards.

⁷ P. Joskow, “Capacity payment in imperfectly competitive electricity markets”, *Utilities Policy*, September 2008.

It can be noted at this point that interventions motivated by resistance to market clearing prices can take different forms. For example:

- The simplest approach is to impose a price cap that directly limits how high prices can rise in conditions of scarcity.
- However, public policy might instead seek to provide for the holding of ‘reserves’ with the intention of releasing those reserves in times of resource scarcity.⁸

Whatever the approach the general effect on incentives is the same: the intervention dampens incentives for suppliers to make provisions for the relevant periods of resource scarcity, since the returns from doing so will be reduced (potentially by very substantial amounts). The regulatory intervention therefore tends to crowd out supplier effort.

There is nothing at all mysterious about this: it is a first-course-in-economics point, to the effect that, if prices are held below market clearing levels, the predictable consequence is a strangled supply side. Twentieth century European economic history illustrates.

How damaging are high prices?

In competitive, wholesale electricity markets, spot prices can spike to levels that are orders of magnitude higher than average levels measured over longer durations. For the most part, however, retail electricity consumers do not observe such spikes, because suppliers provide ‘smoothed’ pricing paths, backed by longer-contracts in wholesale markets. Evidence on the relationships between spot and forward contracts does not indicate the existence of substantial hedging premiums, and therefore one of the fears of critics of market-based approaches to electricity supply – that the cost of hedging volatility would have significant upward effects on consumer prices – has proved unfounded: in comparison with the various other factors that might influence consumer prices, hedging costs can safely be put in the ‘trivial effects’ category.

Speaking generally, where there are significant risks, market arrangements can be expected to develop to remove the bulk of risk from those most averse to risk. Since it is the latter groups who tend to be of most political concern – for example, low income households who could be heavily hit by spikes in utility prices – competitive markets themselves are potentially a source of solutions to the ‘political’ problem. Moreover, because incentives are stronger and the information available is greater in competitive market conditions (than

⁸ In cases such as gas and oil, the commodities can be held in storage sites. In the case of electricity, where storage options with current technologies are more limited, the equivalent is to hold reserves of generating capacity that can be brought into use as and when directed.

under monopoly, whether private, public or political), there can be a general expectation that such solutions will tend to be superior to those devised by a central planner, government department or regulatory agency.

Consider, for example, the electricity sector again. A first, important point to recognise is, as indicated earlier, that continuity of supply at some designated level is not a pure public good: different consumers can choose different levels of continuity. The most familiar mechanisms here are interruptible contracts, which afford the supplier the right to cease supply to a particular user for a designated period. These arrangements are particularly associated with large industrial users of electricity, but it is also possible for small domestic consumers to choose levels of interruptions by, for example, trips that automatically cut supply when frequency falls below a given limit. Developments in control technology provide increasingly sophisticated possibilities in this area (e.g. instead of an all-or-nothing choice between flow and interruption, control devices can be programmed to shed parts of the domestic load, for example by adjusting thermostatic controls – whether for heating or refrigeration – by a degree or two).

The overall effect of these factors is, in the current state of markets and technologies, that domestic energy customers tend to purchase at smoothed prices, whereas large industrial customers may prefer interruption (i.e. a lower level of security of supply, in return, of course, for a lower average price). Perhaps surprisingly, however, when given the choice of whether to purchase at smoothed prices or at volatile wholesale prices, many small customers opt for the latter (see, for example, Lynne Kiesling's 2009 Beesley Lecture on promoting innovation in the electricity industry⁹). The underlying issue here seems to be one of control. If the customer is of the view that he/she has sufficient control over short-term usage levels, the attraction of the volatile price profile is that it provides much more scope for user-driven actions to reduce the average price paid. Again, since technological developments are tending toward innovations that enhance the information available to, and the control that can be exerted by, the demand side, the prospects for greater demand management are relatively bright.

At the risk of excessive repetition, it can be further noted that, taking a longer term perspective:

- The technological developments referred to, which tend toward increasing demand responsiveness, are themselves being promoted by market liberalisation, being a response to anticipated energy prices (e.g. the more volatile are energy prices, the

⁹ Forthcoming in *Economic Affairs*, 2010.

greater the payoffs from spreading domestic load, and hence the greater the payoffs to the development of technologies that facilitate load spreading).

- Increased demand responsiveness to price signals will itself serve to mitigate price fluctuations. Again, this is elementary textbook stuff: the price increase caused by, say, an adverse supply shock will be lower the higher is the price elasticity of demand.

The position in water

Water shares with electricity and gas the characteristic that security of supply to end consumers depends upon the operation of a dedicated transportation network, as well as achieving a balance between supply and demand in the relevant commodity market. There are also similarities in that weather factors may play a major role in creating potential imbalances: in energy, extreme cold weather can drive demand to levels that put pressure on supply capabilities, whereas in water long dry spells can create drought conditions which threaten the capability of the supply system to meet demand at 'acceptable' prices.

In contrast, in water there is no direct equivalent of the potentially very large costs of imbalance or disequilibrium that can occur in the energy sector. When they occur, 'water shortages' do not *necessarily* threaten the integrity of the whole of supply (although if the shortage is caused by local equipment failures, there may be an all-or-nothing aspect to the local supply situation). The more likely outcome is reduced availability, along the lines of most markets, with excess demand met by some or other form of quantity rationing (e.g. restrictions on water use). Similarly, if excess demand leads to reduced pressures and flows in pipes, the negative consequences would appear to be of lesser magnitude than in gas, where explosions are possible.

This is not, of course, to say that imbalance/disequilibrium costs are absent or negligible in water; only that the magnitudes appear to be rather lower than in energy, and in electricity in particular.

Other differences include the following:

- There are very different levels of interconnection: the water transportation system is much more fragmented than the electricity grid or the gas pipeline system. This implies lower levels of capacity to deal with, say, local supply shortages by import of bulk supplies from neighbouring, less constrained areas. Imbalances can, therefore, be expected to exhibit a more localised pattern than in energy, even when there is a common cause across a number of areas (e.g. dry weather conditions); and, to the extent that there are markets in water, differences among regional price fluctuations might be expected to be higher.

- There is no equivalent in the energy sector to the role played by the EA in water.

Existing provisions in water

The Water Act states the following:

- It shall be the duty of every water undertaker to develop and maintain an efficient and economical system of water supply within its area and to ensure that all such arrangements have been made—
 - for providing supplies of water to premises in that area and for making such supplies available to persons who demand them; and
 - for maintaining, improving and extending the water undertakers water mains and other pipes, as are necessary for securing that the undertaker is and continues to be able to meet its obligations under this Part.

This bundling of responsibilities for securing supplies of water and for maintaining the transportation network at an appropriate level reflects the current, vertically integrated organisation of water supply. It is obvious, however, that the responsibilities in the two areas (commodity and network) are distinct and separable. The same general duties could, therefore, be imposed in circumstances where there is business separation between network and supply operations.

In meeting its duties, each water undertaker:

- Establishes a planned level of service, specified in terms of likelihoods of certain types of restrictions on the use of water – e.g. that a hosepipe ban be imposed (on average) not more than once in twenty years. These standards can vary with local circumstances.
- Prepares and obtains Defra approval of a Water Resources Management Plan, extending over 25 years, setting out how it plans to meet the target level of service. Undertakers again have flexibility in how they predict supply and demand.

That undertakers can set different service level standards is an entirely sensible approach given the points above about lack of interconnection and the resulting greater geographic segmentation of the market than in energy. Security of supply is not different from any other economic ‘good’ in that demand for it is likely to be negatively related to its cost to the consumer. Thus, in a region where increasing security of supply in the longer term (i.e. reducing the scale of possible imbalances, and reducing the probability and the duration of periods of quantity rationing or high prices) is particularly costly, it is sensible to recognise

that it may be preferable for an undertaker to seek, after consultation with its customers, to offer lower service standards than other undertakers.

Twenty five year planning is a different matter. Given the level of uncertainties involved, there is a danger of a large element of pretence taking hold, or of decisions being driven by speculations that turn out not to be closely related to reality. In general, rather than seeking to predict and provide on a long-term basis – which involves long-run forecasting of a type that historically has had a very poor track record – experience suggests that a preferable approach would be for undertakers to be encouraged to increase their flexibility, adaptability and responsiveness, so that they can more quickly adapt to changing economic circumstances.

By way of comparison, one of the effects of liberalisation in energy was substantially to reduce the time it took to construct power stations. Such a development potentially has high option value in that it allows investment decisions to be taken closer to the time at which forecasts of the supply-demand balance are likely to indicate that additional capacity is required. This can mistakenly be interpreted as an increased security of supply problem, since forecasts are likely to show potential demand/capacity imbalances occurring in the nearer future that was the case under predict-and-provide arrangements. In reality, however, the nearer term imbalances are no more than a reflection of increased flexibility on the supply side, which is a factor that contributes to greater security of supply.

Existing arrangements and potential liberalisation

There seems to be no particularly strong reason why, at a general level, existing mechanisms associated with security of supply issues should constitute significant obstacles to the development of upstream markets in water. Clearly, if there are changes in the structure of supply, there will need to be changes in the allocation of the relevant responsibilities, for example so that responsibilities in relation to the supply of the *commodity* water are attached to retail businesses, not to infrastructure businesses (the latter may be assigned responsibilities in relation to the security of the system for *transporting* water). However, the nature of current arrangements does not, as a matter of fact, appear to be such that it would lead automatically to a suppression of price signals in circumstances where water was in relatively short supply.

Of course, the incentives to develop new sources of water supplies will be affected by the service standards that are set. If relatively high probabilities of periods of quantity rationing are allowed (e.g. frequent hose pipe bans), it is a natural consequence that any market value of water in rationing periods will tend to be lower than it would be if higher standards of service were imposed, and hence that the rewards for developing new supplies will be correspondingly muted. That is as it should be: if service standards are low, it implies that

the value attached to security of supply is low, and that less investment in peak capacity is justified.

The risk of things going wrong is, therefore, more likely to derive from policy inconsistency over time. Suppose that high service standards are adopted, but that policy is tolerant of failure to reach those standards. In such circumstances, anyone investing in non-regulated activities associated with developing incremental peak supplies of water will achieve lower prices than might be expected on the basis of the nominal standards specified.

Anticipations that this may turn out to be the case can then be expected to lead to under-investment, in line with classic reasoning concerning policy credibility. In this case any under-investment will be associated with reduced security of supply (i.e. with increased incidence of periods of quantity rationing).

The bottom line of all this is that a more rigorous and consistent enforcement of service standards is likely to be required in order for peak water pricing to develop in ways that provide appropriate incentives to undertakers. This does not necessarily mean that Ofwat would need to mandate service standards and to develop elaborate incentive schemes around the provision of quality of service in sectors of the market open to competition (although that would be an approach likely to be appropriate for supplies that continue to be regulated). The constraints on undertakers could be 'hardened' via the contractual relationships between retail supply businesses and consumers, with the regulatory role being the more limited one of helping to establish the institutional frameworks in which such contracting can be efficiently managed.

Current arrangements in more detail

Ofwat nominally holds undertakers accountable for achieving the standards of service they have established; using a Security of Supply Index (SoSI) to compare the actual security of supply with the planned level of service, and taking regulatory action if the planned service levels are not achieved. This is not the place to comment in detail on the SoSI approach, but it is worth noting that, if such an approach is to be maintained, time consistency requires that Ofwat maintain a degree of stability in the way that the index is calculated.

This might be difficult to achieve in circumstances where the index seeks to capture a wide range of factors and where there is some arbitrariness in the way that different factors are combined. Complexity and arbitrariness are things that tend to lead to tinkering aimed at 'improving' the measurements, and, although the motives here are admirable, the unintended consequences can be highly adverse (for the time consistency reasons given). A simpler and more basic set of measurements may, therefore, be preferable if they are associated with greater stability; and this is therefore one of the issues that will require

some regulatory attention if greater reliance is to be placed on competitive market mechanisms in the water services sector.

Transitional arrangements

There appear to be no strong reasons why the broad features of the current Ofwat approach to security of supply could not be retained during a process of transition to competition. However, in these circumstances:

- The current approach would be better reinterpreted as a monitoring activity, designed to provide information to Ofwat to assist the regulator in carrying out its functions, and
- Instead of companies making undertakings to Ofwat on an occasional basis, when perceived problems arise, they should make commitments to their customers on a continuing basis in regard to the reliability of supplies of water, both now and into the future.

If, when companies fail to meet service standards, compensation is paid to consumers or financial penalties are levied by Ofwat (depending upon the approach taken), a more normal market incentive mechanism will have been developed which can be expected to contribute to achieving efficient levels of supply security.

In relation to service standards themselves, Ofwat can obviously be expected to play a significant role in establishing their initial levels. However, there would be considerable merit in allowing suppliers and consumers freedom to negotiate alternative levels of service, so that customers can, if they so wish, choose to have a more interruptible supply in exchange for a lower price – a point that brings the discussion back to the centrality of demand management.

Security of supply can always be increased by gold plating supply structures. The problem is that such solutions tend to be very expensive, either for consumers or for taxpayers (depending upon how the gold plating is financed); and, precisely because they are expensive, they tend to come under political pressure when times get tough. Gold plating then turns into tin plating, and security of supply may come to be underprovided (for further discussion of this overprovision/underprovision cycle, which is familiar in many public services, see Yarrow, “A theory of privatization, or why bureaucrats are still in business”, *World Development*, 1999). In short, there is a certain non-robustness to the gold plating option – it relies on a fair wind in terms of consumer/taxpayer willingness to pay.

In contrast, improved demand management arrangements tend to increase the *flexibility*, and the longer-term robustness of the system, to shocks: they are good for all seasons, times and purposes. They are also relatively cheap when compared with other options, at least in the early stages of their development, when many, previously unexploited opportunities for benefit might be readily available.

December 2009