ISCR

COMPETITION & REGULATION TIMES

MARCH 2007 ISSUE 22

IN THIS ISSUE

Academic salaries: where angels fear to tread

If you pay peanuts, do you get monkeys? Using data made available by the Performance-Based Research Fund (PBRF), Glenn Boyle finds that for New Zealand universities the answer appears to be 'yes'.[1]

To economists, it is axiomatic that offering employees low pay relative to what is available elsewhere will have adverse consequences for worker quality. But not everybody agrees – witness the commonly heard claim that the job of CEO at some organisation or other could be done just as well by the claimant at a fraction of the cost. Rather more rigorously, some sociologists and psychologists argue that workers are primarily motivated by non-pecuniary features of their employment.

Universities are often cited as providing direct evidence for the latter view. Academics, it is argued, cannot be motivated by money, since if they were they would not choose to be academics. However, the extent to which this notion is reflected in university remuneration systems varies considerably. New Zealand academic pay is, with the limited exception of medicine and dentistry, independent of discipline – that is, a professor of economics is paid from the same scale as a professor of history despite the former having more valuable labour-market opportunities than the latter. In general, academics in business, law and some science areas have more valuable outside opportunities than those in humanities and most sciences, but this is not reflected in New Zealand universities' compensation systems.[2]

By contrast, academic salaries in the US vary from discipline to discipline, reflecting differences in the opportunity cost of working in academia. The justification for this view is succinctly stated by Daniel Hamermesh: 'If a university went ahead and paid equally, lowering economists' pay and raising French professors' pay, it would have a great French staff and a dreadful bunch of economists'.[3]

By and large, however, policy-makers and university administrators in New Zealand have rejected this view: some disciplines are, it is acknowledged, more difficult to hire in than others, but hard work and innovative recruiting strategies ensure there is no significant variation in the quality of those ultimately employed. To the extent there is such variation, it has nothing to do with remuneration. And, in any event, paying academics on a differential basis would be inequitable since they all do the same job.

Which view is correct? Is discipline-specific pay unfair and unnecessary, or is it actually a required response to discipline-specific variation in the economic cost of academic human capital? Until recently, the absence of any consistent performance measure by discipline meant that answers to this question were largely speculative. But the introduction of the PBRF has provided the necessary data.

PBRF

During the 1990s, government funding of an institution's teaching and research activities was bundled into a single bulk grant, the size of which depended on the institution's student enrolments. However, concerns arose that such a system did not allocate research funding to its most productive uses, and so in 2002 the government announced the establishment of the PBRF. Under this arrangement, the funding for research was to be separated from that for teaching and made dependent on performance. Assessment of research performance required all eligible academics to first nominate a 'subject area' (roughly equivalent to a discipline)[4] within which their research would be evaluated, and then to submit to one of 12 peer review panels a portfolio summarising their research activities over the previous six years. The panel assessed each portfolio, with the assessment being made in the context of international (not just local New Zealand) research quality in that subject area; and it then assigned to each portfolio a 'quality category' that corresponded to a numerical score. On completion of this process in early 2004, individual scores were aggregated to obtain overall 'quality scores' (performance measures) for subject areas and academic units.

The PBRF's quality scores thus provide a 'consensus' measure of a subject area's (or discipline's) research performance. A reasonable measure of the opportunities available to researchers in each discipline is the salary paid in the US to a representative academic in that discipline: these not only provide a direct measure of the discipline-specific opportunities available within academia, but also yield an approximate ranking of a discipline's non-academic opportunities.

The difference between US and New Zealand academic salaries is not only high on average, but also highly variable across disciplines: using the end-2003 exchange rate, the mean difference is $20,710 but discipline

to page 2

from page 1

differences range from $90,520 to -$340. If peanuts beget monkeys, then PBRF quality scores should be high in areas where these differences are small (the least 'underpaid'), and low in areas where they are large (the most 'underpaid').

And this is exactly what happens.

Pass the bananas

After controlling for other factors that might affect research performance, the relationship between quality scores and the US-NZ salary differential appears strong. On average, a $25,000 increase in this difference lowers a discipline's quality score by 0.45 points (corresponding to a fall of 15% for the median discipline) and raises its number of the portfolios in the 'R' – the very lowest – quality category by 8.4 percentage points (26% for the median discipline). Put another way, moving from the most underpaid decile to the least underpaid predicts a rise in average quality score of about 0.73 points and a decrease of 14 percentage points in the number of portfolios in the 'R' quality category (approximately 27% and 40% of their respective sample means).

Table 1 shows the extremes. Only one of the top five research performers is in the 20 most underpaid disciplines – and none of the five most 'underpaid' disciplines lies in the top half of performers.

Monkey economics?

In a profession where job motivation is frequently claimed to be untainted by base financial considerations, these results are unlikely to be either popular or universally accepted. Are other stories also consistent with the data? Several possibilities present themselves, but none seems very compelling.

One argument could be that research performance in the most underpaid disciplines is artificially lowered by the need to spend lots of time dealing with large numbers of students. But disciplines with high student numbers typically adopt less-intensive teaching methods, so it is by no means clear that they suffer from less available research time. And a high teaching load seems more likely to be a *reflection* of low research quality than a cause – the most underpaid disciplines are primarily able to recruit only weak or unmotivated researchers who are willing to accept unattractive teaching conditions in exchange for lower research expectations.

Another argument might be that the data simply reflect 'monkey-mimicking behaviour': all disciplines are able to hire researchers of similar quality, but those in the most underpaid areas take advantage of their greater market opportunities by spending more time on outside consulting work, and consequently less time on research. But researchers in other countries also have these consulting opportunities, and so (given that New Zealand research is assessed in an international context) this cannot explain why the most underpaid disciplines in New Zealand also have the weakest research performance.

Perhaps the results simply reflect the greater use of part-time academics – who might be expected to have a lesser research focus – in the most underpaid disciplines. Controlling for part-time participation, however, makes no difference to the results described above.

Finally, could the results be attributable to a 'new-researcher' bias? To be assigned a quality category above 'R', a researcher had also to clear a *quantity* hurdle. As a result, many new researchers received an 'R' despite being heavily involved in research. If such researchers were disproportionately represented in the most underpaid disciplines, then the research performance of these disciplines would be biased downwards. But controlling for average researcher age, or for the proportion of portfolios that were not submitted to a peer review panel because they obviously didn't meet the quantity criterion, has no effect on the results.

Scratching our heads ... or our tails

New Zealand universities apparently get what they pay for. Disciplines in which compensation is high relative to other opportunities are best able to recruit high-quality researchers and/or motivate their researchers to be productive; paying (relative) peanuts attracts mainly monkeys.

Does this imply that New Zealand universities should allocate funding based on market forces? Such a move would be controversial – and painful, in the short term. Nevertheless, the costs of not doing so are clear.[5] Previously, there was no concrete evidence that discipline-independent pay had implications for research quality. Now there is.

Table 1: Research performance rankings and the value of researchers' market opportunities (n = 41)

Subject area	Quality score ranking	'Underpayment' ranking
Philosophy	1st	36th
Anthropology and archaeology	2nd	35th
Earth sciences	3rd	23rd
Ecology, evolution and behaviour	4th	21st
Biomedical	5th	14th
Computer science, information technology, information sciences	26th	5th
Marketing and tourism	30th	4th
Law	20th	3rd
Management, human resources, industrial relations, international business and other business	31st	2nd
Accounting and finance	34th	1st

1 This article is based on: G Boyle. 2006. 'Pay Peanuts and get Monkeys? Evidence from Academia' (a draft version is available at www.iscr.org.nz/navigation/research.html).

2 Some academic staff (not always in the disciplines with the greatest outside opportunities) receive salary premia, but these are small and infrequent.

3 D Hamermesh. 2004. *Economics Is Everywhere* (2nd ed.) McGraw-Hill. New York.

4 There are 41 subject areas; some (such as 'archaeology and anthropology' and 'history, history of art, classics and curatorial studies') combine two or more disciplines.

5 After expressing positive thoughts about the quality of the research on which this article is based, a US blog ('Marginal Revolution') surmised: 'I expect the author will soon leave New Zealand'.

Glenn Boyle is the Executive Director of ISCR. He is also a professor of finance – the most 'underpaid' discipline in New Zealand universities – so there is a signif-icant possibility that he is himself somewhat bananas.

ISCR Competition & Regulation Times is the newsletter of the New Zealand Institute for the Study of Competition and Regulation Inc.
PO Box 600, Wellington, New Zealand. Ph: +64 4 463 5562, fax: +64 4 463 5566, e-mail: iscr@vuw.ac.nz, website: www.iscr.org.nz
The ISCR editorial team for this issue was Glenn Boyle and April-Mae Marshall.
The views expressed in **ISCR Competition & Regulation Times** are the views and responsibility of the contributing authors.
ISSN 1175-2912

Is overconfidence the X FACTOR in teamwork?

You find overconfident colleagues tiring and obnoxious? You prefer to avoid working with them? Such negative reactions, while understandable, may nevertheless be wrong-headed – as Mike Webb points out.[1]

The problems associated with working in teams are well known and long standing. In a team environment, the incentive for any individual to exert optimal effort is diminished by the so-called free-rider problem: the benefits of each person's work are shared amongst team members while the costs are borne individually. Consequently, each team member exerts too little effort because the costs of working harder exceed the benefits to the individual – despite the fact that the team as a whole (and therefore the individuals within it) would be better off if everyone worked harder.

Possible responses to the free-riding problem in teams have focused on the obvious roles of altruism, peer pressure, and completely specified contracts. Now it appears that help may also come from an unexpected quarter: the well-documented propensity of individuals to be overconfident and to overestimate their own abilities.[2] Biases of this kind are often detrimental when working alone – overestimating the benefits, or underestimating the costs or risks, of a particular course of action results in too much effort being expended relative to the potential rewards, thereby making the individual worse off. However, the presence of annoyingly optimistic individuals has advantages for a team.

A glass more than half full

The mechanism by which it does so is simple. Excessively confident workers over-estimate the gains to the team of their efforts. And because they receive (as a team member) a share of these perceived greater gains, they are able to justify expending greater effort than would otherwise-identical individuals who are realistic about their abilities.

In short, their optimistic beliefs about the benefits of their work help overcome the free-rider problem by inducing them to exert effort at a level closer to that which is optimal from the perspective of the team as a whole.

The benefits of individual overconfidence are even greater if worker synergies are present. In that case, the greater effort put in by over-confident workers not only directly reduces the extent of the free-rider problem. It also does so indirectly, by increasing the productivity of other workers – thereby (rationally) inducing those workers to work harder as well. The cumulative effect is to make the team unambiguously better off, even if the overconfident members have no greater ability than others.

Can overconfident workers themselves also be made better off by their own folly? Because the anticipated direct gains from extra effort are over-optimistic, the worker's share of these gains normally falls short of the costs incurred – and so the worker ends up worse off even though the team as a whole is better off. But the additional indirect gains from synergies add a further dimension that results in enhanced individual worker welfare, so long as the perception bias is not too great. Although overconfident workers must endure the cost of exerting more effort than is justified by their individual abilities, they can nevertheless end up better off as they share the benefits of their teammates' increased effort.

Leading, learning, motivating

Do overconfident workers make the best leaders? Perhaps surprisingly in light of the above, this is only the case if the bias is small. Otherwise, a leader whose self-confidence is too extreme exerts too much effort, thereby making the team worse off. In most cases, an unbiased leader is better placed to anticipate and exploit the benefits of overconfident followers.

Interestingly, overconfidence can be self-perpetuating, even when team members have the opportunity to learn their true abilities through the realised performance of their team. When a team contains overconfident members, the team performs well because of the extra effort exerted by all members – but overconfident members attribute this success to their perceived greater ability. This allows the effects of overconfidence to continue longer than they otherwise might.

The advantages of overconfidence in team situations can also provide interesting incentives. Leaders have an incentive to build up and maintain the self-esteem of team members because this can lead to the latter considering it in their best interest to work harder – which benefits the team. Similarly teams can benefit from 'talking up' the chances of a risky venture succeeding, because overestimating the chance of success can lead to higher effort being exerted.

1 This article is primarily based on: S Gervais and I Goldstein. 2005. 'The Effects of Biased Self-Perceptions in Teams' (available at http://ssrn.com/abstract=843508).

2 For example, the 'above average' effect – in which the vast majority of participants in studies consider themselves to be above average in abilities as diverse as driving skills, intelligence, and management talent. See, among others: O Svenson. 1980. 'Are we all less risky and more skillful than our fellow drivers?' *Acta Psychologica*, 47 pp143-148; and also L Larwood and W Whittaker. 1977. 'Managerial Myopia: Self-Serving Biases in Organizational Planning' *Journal of Applied Psychology* 62 pp194-198.

Mike Webb is a research assistant at ISCR.

Ebbs and Flows in New Zealand Irrigation

From 1990, the assessment and administration of New Zealand rural irrigation schemes were subject to some fundamental changes. René Le Prou describes then, now, and some implications.[1]

Before 1990, irrigation-scheme development in New Zealand took place within a centralised structure (see Figure 1).

Farmers would propose a scheme to the District Committee of Officials for Irrigation and Rural Water Supplies, which undertook initial analysis of whether irrigation was required to supplement rainfall. The National Water and Soil Conservation Authority then investigated the proposed scheme by, for example, polling all farmers in the community to assess the level of support for the scheme. If at least 60% provisionally agreed to it, then development would proceed to the economic evaluation stage, undertaken by the Ministry of Agriculture and Fisheries. The requirement for approval was that the expected return on investment be at least 10%.

Once approval was granted, the scheme was administered by the Ministry of Works and Development – which took responsibility for design, organised construction, and owned and controlled the scheme once it was operational.[2]

The Ministry also provided significant funding. In 1975, for example, scheme headworks and both off-farm and on-farm distribution works received subsidies ranging from 33% to 100% of total cost. Moreover, any remaining costs were eligible for long-term subsidised loans from the Rural Banking and Finance Corporation (RBFC).

Sitting at the top of this structure was the government of the day, an entity capable of producing perverse outcomes: more than one commentator suggested that farmers in marginal electorates were more likely than others to receive approval for requested schemes. Government control also meant that abrupt change could occur in the sector, for reasons divorced from irrigation. For example, the criterion for scheme approval changed significantly in 1980, apparently because of concerns about the budget deficit rather than any fundamental change in the cost of irrigation-scheme capital.

By 1990, 53 irrigation schemes were administered centrally within a hierarchical structure – a 'bureaucratic maze' in which there was a significant imbalance between the number of farmers using irrigation schemes and the number of bureaucrats administering the schemes.[3]

Inefficiencies became apparent, particularly in design. Experience from early community irrigation in the flat and pastoral Canterbury Plains showed that border-dyking (water distributed through ditches) was effective there – and a 'one-size-fits-all' view encouraged government administrators to impose this large and expensive design on other, less favourable, areas.

This tendency was exacerbated by the funding system. For example, RBFC grants were not available for pumps and motors, which were major costs in a spray or trickle irrigation system. But they were available for costs associated solely with border-dyking (such as land grading).

… and then there was light

From 1990 onwards the government began to privatise community irrigation schemes, with the vast majority eventually being sold to farmer groups. As a result, the process for developing a new community irrigation scheme is now very different from that which operated previously (see Figure 2).

Today the decision on whether or not an irrigation scheme is economically viable is made by farmers – a scheme is developed if a sufficient number of farmers believe that they will benefit from it. The design and construction, and most of the funding, are undertaken privately. Tenders are typically called for the design of a scheme, with farmers choosing the tender they find most attractive. And, once construction is completed, the scheme is controlled by the owner-farmers.

The only direct government involvement in this process is through the Ministry of Agriculture and Forestry's Sustainable Farming Fund. The fund exists to 'support projects that will contribute to the financial and environmental performance of the land-based productive sectors', and provides grants of up to $200,000 per project per year from an annual budget of around $9.5 million. Compared with the pre-1990 government funding on offer, this involvement is minor: private-sector finance now bears most of the costs of a community irrigation scheme.

Water rights and the RMA

Privatisation of irrigation schemes was accompanied by the passing of the Resource Management Act 1991 (RMA), which had significant implications for water rights associated with the schemes. The RMA gives councils the responsibility to grant water for a period of up to 35 years (which provides some certainty for those wishing to invest in an irrigation scheme). But it also allows them to change consent conditions at any stage (which has the opposite effect). Clearly, a successful irrigation scheme requires property rights to water that are well defined and enforceable. Such rights also facilitate the trading of water rights, which helps ensure that scarce water is used in the most productive and valuable manner.

Before 1990, trading in water rights was not possible, because the value of these rights was capitalised into land values – that is, water rights were tied to land ownership and hence were not separately tradable. By contrast, section 136 of the RMA explicitly allows for the trading of water rights, although only in certain circumstances. First, transfers must take place within the same water catchment. Second, resource consents are granted for specific water uses, thereby precluding the trading of rights across uses. Third, trading must be 'expressly allowed' by a council's regional plan or must be otherwise approved by the council.

Despite these restrictions, the trading of water rights within community irrigation schemes has increased significantly since its inception. Irrigation companies will usually facilitate trading by registering supply offers and demand requests, with individual farmers then arriving at an agreed trading price – elaborate mechanisms are apparently not necessary for successful water trading.[4]

Figure 1: Centralised administration schemes in New Zealand

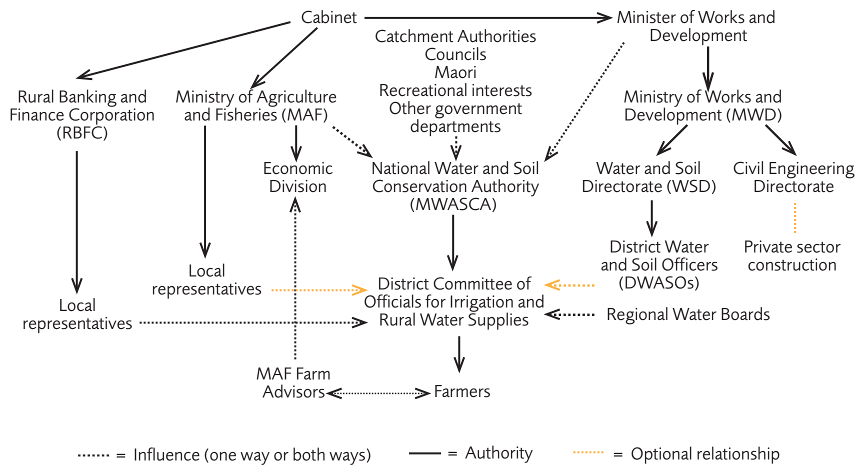
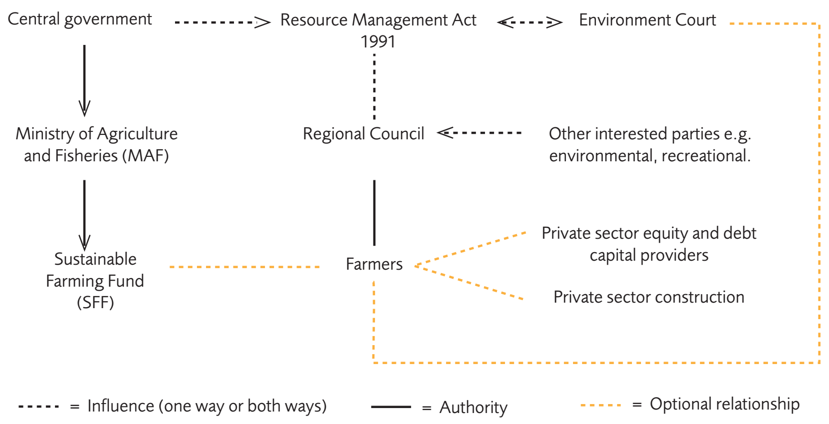


Figure 2: Decentralised administration of irrigation schemes in New Zealand



From there, to here, to …?

The pre- and post-1990 situations in New Zealand irrigation are classic examples of, respectively, centralised and decentralised organisational forms.

Before 1990, the sector was characterised by extensive government involvement, a large bureaucracy, and little role for farmers or other investors. Few scheme-design alternatives were considered and the adoption of more efficient irrigation methods was discouraged.

Since 1990, by contrast, the sector has been characterised by a much-reduced role for central government, a flat hierarchy that is more local in nature, and a large degree of independence – particularly in terms of ownership and control. Irrigation schemes are now adopted if (and only if) accountable parties such as farmers and outside investors consider them sufficiently profitable. A scheme may grow iteratively in stages, and there is competition in design and construction. In addition, irrigation water rights are traded and priced, with prices now reflecting water's relative scarcity. While the benefits and costs of these changes are difficult to quantify, qualitative performance indicators support the conclusions of scholars who emphasise the general advantages of decentralised systems.[5]

Are there ways by which New Zealand irrigation arrangements could be further improved? Two possible reforms are apparent. First, water rights could be better defined (thereby removing impediments to investment) by invoking a 'presumption of renewal' beyond the initial 35-year period.[6] Second, water usage could be directed to more valuable ends by permitting trading across uses and catchments. Although this would require a central registry that records the ownership details of water rights and the transfers of these rights, the benefits seem likely to outweigh the costs.

1 This article is based on the author's MA thesis: 'Centralised versus decentralised decision making: the case of New Zealand irrigation'. For more details, see *The Administration of New Zealand Irrigation: History and Analysis* (available at www.iscr.org.nz/navigation/research.html).

2 A 'design and construct' model was used for irrigation schemes in New Zealand during this period: schemes would be constructed privately but designed and run publicly.

3 H Morton. 1978. Presentation to the First Annual Conference of the New Zealand Irrigation Association.

4 Trading across water uses would be more difficult to facilitate, possibly requiring a central registry recording the ownership details of water rights and the transfers of these rights.

5 See, for example: J Kornai. 1992. *The Socialist System: The Political Economy of Communism*. Oxford University Press; and also J McMillan. 2002. *Reinventing the Bazaar: A Natural History of Markets*. W W Norton & Company. New York.

6 See also: K Counsell and L Evans. 2005. *Essays on water allocation in New Zealand: the way forward.* ISCR research paper (available at www.iscr.org.nz/navigation/research.html).

René Le Prou is a former Masters student and research assistant at ISCR.

INVEST in haste ... REPENT at leisure

A fundamental insight of real-options analysis is that there can be value in waiting to invest, as doing so provides the opportunity to obtain more information about investment profitability. But some researchers have pointed out that real-world phenomena may significantly reduce the advantages of waiting. William Taylor assesses this argument.

If waiting is valuable, then investment requires an expected-return premium over and above the weighted average cost of capital (WACC) in order to provide compensation for sacrificing the ability to wait any longer. Firms that fail to take account of this premium will tend to make inferior investment decisions and therefore will be less valuable than firms that do. Similarly, regulators that ignore the value of waiting may, as a result, set allowed rates of return too low and thus discourage investment in vital infrastructure.

Simple real-options models suggest that the magnitude of this 'waiting' premium can be substantial – several percentage points or more. But are such models too simple to be useful in the real world? Are there in fact good reasons to think that they significantly overstate the true value of waiting, and that WACC approximates a sufficient return on investment after all?

Not what you'd expect

One common argument against the 'waiting' premium is that the presence of competition eliminates the value of waiting. By waiting, a firm provides competitors with the ability to invest first and obtain any available 'first-mover' profits. But it's easy to over-state this argument. In particular, it assumes that all investments can be characterised as an 'early bird gets the worm' situation. While in some cases there may well be advantages to investing before one's competitors, there are other situations where it is beneficial to let these competitors go first. For example, suppose the proposed investment is in a new and highly uncertain market or technology – by letting a competitor go first and observing its fortunes, the firm can eliminate much of its investment risk. When such 'second-mover' advantages are present, competition actually *reinforces* the value of waiting.

Many investments are likely to contain elements of both first- and second-mover advantages, so the total effect of competition on the value of waiting depends on which advantage dominates. Only when pre-emption is essential will the strategic importance of waiting become negligible. And although the benefits of pre-emption are likely to depend on industry structure, recent research indicates that this relationship is not a one-way street. While firms in competitive industries generally invest faster than those with less competition, firms in the *least* competitive industries actually invest the fastest.[1]

How's your credit?

Investment must be paid for, whether out of internal (retained profits) or external funds (sale of new equity or debt securities). Clearly, the ability to invest depends on the availability of such financing, and on its price and its terms. Firms that must pay a high price to obtain new external financing (perhaps because they are deemed to be high risk) will rely on internal funds, and therefore on the profitability of existing assets.

Most real-options models ignore the financing problem, instead simply assuming that investment will be paid for somehow. But firms that rely on internal financing run the risk that this funding may not be available in the future: for example, an adverse shock to profits may deplete internal funds to the extent that investment becomes impossible. In this situation, waiting is less valuable because of the risk that the investment opportunity may in effect disappear.

Although the possibility of financial constraints weakens the advantages of waiting to invest, these advantages do not disappear entirely. Even a severely constrained firm benefits from acquiring new information – the optimal waiting time is simply shorter than if it were unconstrained.[2]

Time to build

In the standard real-options world, investment occurs instantaneously at the commencement of the project; subsequent investment is not required. In practice, most projects take 'time to build' – that is, they require implementation and construction over a period of time. As a result, typical investment expenditure is ongoing rather than a one-off lump sum.

Investments that begin, and then take time to complete, can of course be abandoned before completion if new information suggests that this would be the optimal strategy. Such projects are, in effect, more 'reversible' than simple lump-sum projects, since the remaining investment cost can be 'recovered' by abandoning the project. But the more reversible the project, the lower the incentive to delay its launching in order to acquire more information about its prospects. A longer implementation period thus decreases the value of waiting to invest.[3]

This phenomenon is exacerbated if the risks surrounding the project's ultimate cost are primarily of the 'technical' variety (uncertainty about the time needed for completion and the quantity of inputs required).[4] These kinds of risk are generally only resolved by having construction commence, so that delaying investment provides no potential for additional information and hence isn't valuable. By contrast, 'input price' risk (uncertainty about the price of inputs) remains, whether or not construction is currently active. So there is value in waiting to gain new information about this even for projects that take 'time to build'.

In the case of major infrastructure projects (the building of transmission investment or a new stadium for a sports event), a certain level of capacity must be in place at a known date in the future.[5] When a project has a long or uncertain construction period, the value of waiting is reduced because waiting may leave insufficient time for completion by the required date.

Reverting to type

Real-options models typically assume that shocks to project value follow a 'random walk' process – that is, value deviates unpredictably around a constant expected growth rate. Thus the value of a project can be subject to repeated adverse shocks, and the incentive to delay investment arises from the desire to minimise the probability of such an outcome.

However, some projects are more accurately thought of as mean-reverting: when their value lies above or below a long-run mean, it tends to revert back towards that mean. As a result, negative shocks tend to be followed by positive shocks, which lowers the potential magnitude of adverse outcomes and hence would seem to lessen the value of waiting.

But this is not the whole story. If mean project value exceeds the investment cost, then – even if project value is currently low – the project is likely to eventually be worth more than it costs. In such a case, the value of waiting can be even greater than in the standard situation.

Costly information

The decision on whether to invest or delay requires calculation of a project's profitability. But unlike the holders of financial derivatives (on which real-options theory is based), investment managers cannot continuously re-evaluate project profitability. The complexity of most real-world projects means that such calculations are time-consuming and costly.

Unlike the holders of financial derivatives (on which real-options theory is based), investment managers cannot continuously re-evaluate project profitability. The complexity of most real-world projects means that such calculations are time-consuming and costly.

Costly evaluations lower the value of waiting in two ways.[6] First, evaluation costs directly increase the project's cash outflows and thus lower its value. When the acquisition of further information about the project is costly, the opportunity cost of delaying investment to acquire this information is greater – and hence the incentive to wait is lower. Second, and more subtly, higher evaluation costs lead to less frequent evaluations (since doing so continuously would be prohibitively expensive), and therefore to a lower probability of choosing the best time to invest. Since much of the value of waiting stems from having the flexibility to invest on exactly the right date, this effect reduces the incentive to wait in the first place. This is particularly important when project value is mean-reverting, as the 'temporary' nature of value shocks means that getting the timing of investment exactly right is crucial.

Although these real-world factors drive a wedge between the true value of waiting and that predicted by simple theoretical models, the size and sign of this wedge is frequently unclear. Even when the waiting value is unambiguously smaller than its theoretical counterpart, it is extremely unlikely to be zero. Overall, explicit recognition of the option to wait still seems likely to result in more accurate investment decisions than a simple reliance on the WACC.

---

1 E Akdogu and P MacKay. 2007. 'Investment and competition' *Journal of Financial and Quantitative Analysis* (forthcoming).

2 G Boyle and G Guthrie. 2003. 'Investment, uncertainty and liquidity' *Journal of Finance* 58 pp2143-2166; and also G Boyle and G Guthrie. 2006. 'Hedging the value of waiting' *Journal of Banking and Finance* 30 pp1245-1267.

3 A Milne and A E Whalley. 2000. 'Time to build, option value and investment decisions: a comment' *Journal of Financial Economics* 56 pp325-332.

4 R Pindyck. 1993. 'Investments of uncertain cost' *Journal of Financial Economics* 34 pp53-76.

5 G Boyle, G Guthrie and R Meade. 2006. *Real options and transmission investment: the New Zealand grid investment test* (available at www.iscr.org.nz/documents/git-iscr.pdf).

6 G Guthrie. 2007. 'Missed opportunities: optimal investment timing when information is costly' *Journal of Financial and Quantitative Analysis* (forthcoming). A draft version is available at www.iscr.org.nz/navigation/research.html.

William Taylor is a Masters student at ISCR. His thesis examines the effect that managerial share ownership has on the investment decisions made by managers on behalf of their firms.

Regulatory disharmony in 'uncommon' markets: a tale of EU telecommunications

It's 17 years since the European Union's Directive 90/387/EEC expressed the intention to 'create an open and borderless internal market allowing free movement of services' within Europe. But, as Bronwyn Howell finds, it hasn't got very far.

Despite substantial legislative and regulatory effort – and strong political intent – a pattern of national (fixed-line) markets and limited cross-border entry prevails in Europe. Even emergent mobile-telephony markets, most of which developed after the adoption of Directive 90/387/EEC, are predominantly national or regional in nature. Most firms have a presence in only a few countries (Figure 1).[1] Those firms with the largest multi-national presence have tended to develop regional, rather than EU-wide, coverage patterns. Vodafone's activity is concentrated in Western Europe, with a presence in Germany, the United Kingdom, France, Italy, the Netherlands, Spain, and Portugal; and Telenor is present only in the Scandinavian countries. Prices, terms and conditions vary greatly, even in adjacent countries and for regulated services, and one of the most fraught issues currently concerns the very high costs of mobile roaming – that is, sending and receiving calls when in a country other than the one in which the customer has entered into an access agreement with a mobile network operator. This is a far cry from the pan-European market envisaged in 1990, and plainly not an environment conducive to the seamless communication required to support the operations of a common economic market.

What went wrong

An institutional analysis of the development of EU regulatory institutions[2] suggests that the ongoing fragmented nature of European telecommunications markets is a consequence of the design of the telecommunications regulators established or sanctioned by the European Commission, and of the varied histories of the telecommunications markets in each of the 25 member countries.

At the core of the EU regulatory framework is the European Commission in Brussels, which oversees a common framework of legal rules and policies governing the European telecommunications sector. The framework has been designed to underpin the formation of a 'common market' based upon common rules and processes. Yet each of the 25 member states maintains its own quasi-independent national regulatory authority (NRA), which is accountable to its own national parliament and has the power to set and enforce local regulatory terms and conditions (as long as these arrangements are within the broad frameworks established by Brussels).

Centrifugal forces

Although the individual NRAs collaborate via the European Regulators Group (ERG) to 'contribute to the development of a common regulatory culture',[3] the EU framework precludes forcing member states to agree. It also precludes granting the Commission specific powers beyond 'recommending' common approaches. (Some commentators have noted that it was deemed inappropriate to 'freeze ... a policy vision based on a specific market design reflecting a political agreement achieved at a specific moment in time'.[4])

The wide range of discretionary powers granted to the NRAs has allowed the volatile telecommunications sector to respond rapidly, unpredictably and divergently in response to changes. It is therefore unsurprising to find that the evolutionary pattern in the sector has been dominated not by the pursuit of central common objectives but by the localised interactions of actors, institutional arrangements, legal rules, cultures, values, norms, and attitudes of each of the 25 distinct sub-markets.[5]

Each of the individual NRAs has emerged from the interaction of the technological, political and market forces within its own territory. Some were in place prior to 1990. Others have emerged subsequently, as new nations have emerged from the Eastern Bloc and as other countries have deregulated and privatised many parts of their economies. This varied pattern of emergence has led to each NRA having a different range of responsibilities – some regulate postal services, railways, and even content (broadcasting) in addition to telecommunications. They also have different accountabilities and different processes: some are accountable directly to their governments; the decisions of others are subject to appeals in the courts or specially appointed tribunals.

The national ambit of the NRAs ensures that each will necessarily have a primary focus upon serving its own national interests – which may not always be consistent with the EU objectives of forming a common market. If it is in the national interest of one country to implement a policy that is at odds with the common framework, then (short of a European Court action when the policy violates a directive) the Commission is powerless to force common policies upon member NRAs. This is illustrated by the current impasse between the Commission and the German government, which has announced its intention to grant incumbent Deutsche Telecom a 'regulatory holiday' from the obligation to provide access to competitors on its new fibre-optic network, in apparent defiance of EU Directive 95/62/EC harmonising on open access to incumbent firms' infrastructures.

Moreover, each NRA regulates a market environment that is characterised by a unique history. This history has determined the identity and nature of interaction of sector participants, and has contributed to shaping not only the incentives and abilities for actors to participate in the market, but also their incentives and abilities to participate in shaping the rules (formal and informal) that govern the market and the regulatory processes. Each national market has begun under different political environments, with different incumbents in place, and has undergone patterns of entry and evolution that vary according to the different natures of historic infrastructure investment, consumer tastes and preferences, and commercial opportunities. The evolution of the NRAs themselves and the nature of the formal rules they develop and implement will have been shaped principally by local interactions. Longstanding interactions between individuals in a specific sector increase the likelihood that future agreements and formal rulemaking will favour those more familiar with local norms and rules, thereby creating entry barriers for those without such historic involvement.

Under these circumstances, the lack of widespread cross-border entry in wireline services would appear to be a logical consequence – as would the limitation in the spread of wireless services to geographic regions with common institutional arrangements, legal rules, cultures, values, norms, attitudes, and long-established trading histories. Furthermore, in the absence of changes to central rule-making from Brussels, repeated future interactions amongst the same sets of actors will most likely increase the degree of diversity in the markets of the 25 member states rather than encouraging and promoting regulatory or market convergence.

So what?

There are lessons to be had from the European experience.

Firstly, whilst regulators and legislators can impose common rules on different markets, the ways in which the rules play out will differ as institutions, inter-relationships, cultures, values, norms, and attitudes differ. It cannot be presumed, for example, that regulations or firm structures adopted in one set of market conditions will accurately predict the outcomes of the same regulations or firm structures in a different market environment. Different market histories lead to different relationships between actors and institutions, and therefore to different outcomes.

Secondly, it is far from clear that federal regulatory arrangements in which individual authorities have autonomy to interpret and enforce centrally determined principles will lead to either common processes or the development of single markets. This point is especially pertinent in a globalising world where federations of nation states form coalitions and trading blocs. As long as individual states retain the ability to determine their own rules, national differences and biases will most likely persist, conspiring against the formation of a single 'common market'.

The EU telecommunications experience would suggest that, despite the laudable objectives of international policy, local and national interests and differing historical development-paths make the pursuit of such goals problematic.

Figure 1: A fragmented mobile market – corporate presence in the EU

	DE	UK	FR	IT	ES	NL	PL	SE	IE	BE	CH	DK	AT	PT	NO
Vodafone	✓	✓	✓	✓	✓	✓			✓	?				✓	
Orange		✓	✓		✓	✓	✓			✓	✓				
T-Mobile	✓	✓				✓	✓						✓		
Telefonica	✓	✓			✓				✓						
Hutchison		✓		✓					✓			✓	✓		
TeliaSonera								✓				✓			
KPN	✓					✓				✓					
Telecom Italia				✓				✓							
Telenor								✓				✓			✓

Source: Ewan Sutherland (op. cit.) p5.

1 Ewan Sutherland. 2006. *European Union 2006 Telecommunications Review*. Lecture at the NordICT PhD Summer School, Skagen, Denmark. 30 August.

2 Bronwyn Howell. 2006. *An Institutional Economics Analysis of Regulatory Institutions in the Telecommunications Sector* (available at www.iscr.org.nz/navigation/research.html).

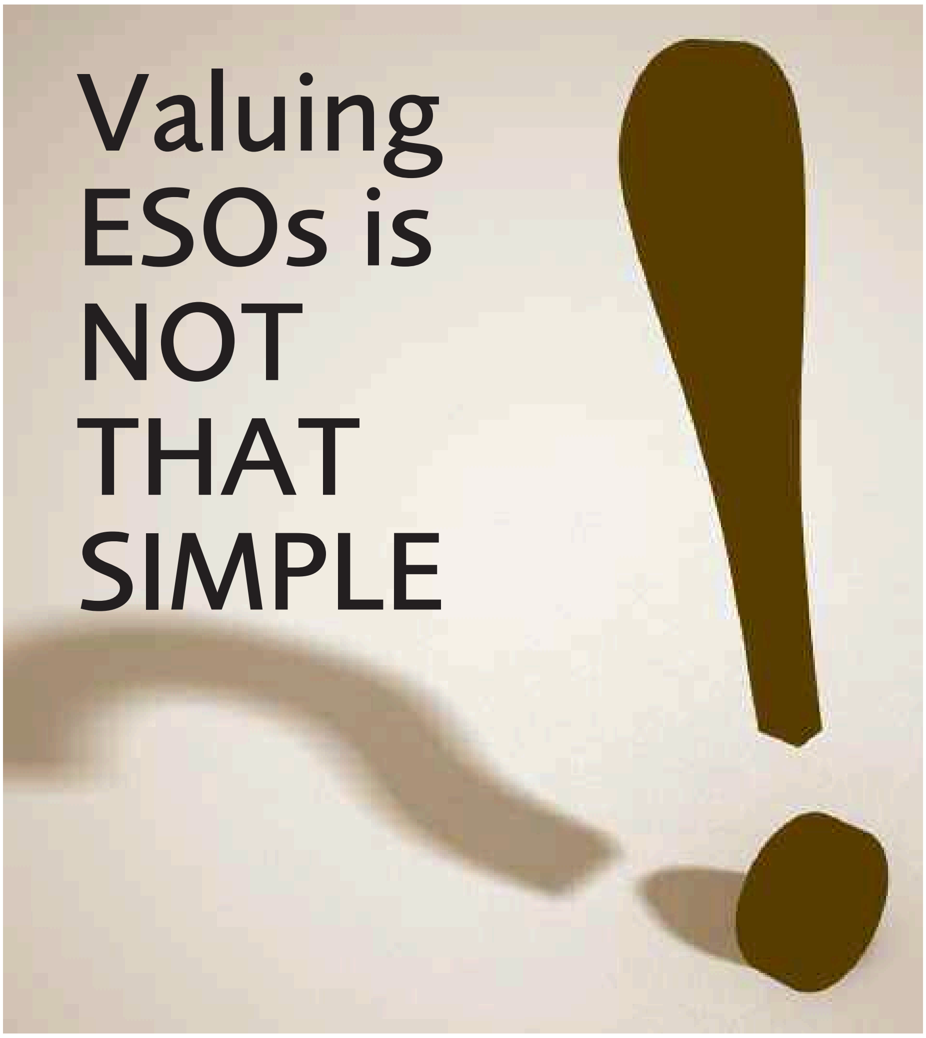
3 Jens Arnbak. 2006. *Tasks and Status of National Regulatory Authorities in the EU*. Lecture at the NordICT PhD Summer School, Skagen, Denmark. 30 August. (Professor Arnbak, of the Technical University of Delft, was formerly the Telecomunications Regulator in the Netherlands.)

4 Christian Hocepied and Alexandre de Streel. 2005. 'The ambiguities of the European electronic communications regulation' in E J Dommering and N A N M Eijk (eds). 2005. *The Roundtable Expert Group on Telecommunications Law*. University of Amsterdam (available at www.fundp.ac.be/pdf/publications/53992.pdf).

5 Framework extrapolated (by Joop Koppenjan and John Groenewegen. 2005. 'Institutional design for complex technological systems' *International Journal on Technology, Policy and Management* Fall pp11-34) from Oliver E Williamson. 1998. 'Transaction cost economics: how it works, where is it headed' *De Economist* 146(1) pp23-58.

Bronwyn Howell is a programme director at Victoria Management School, and a research associate of ISCR.

Valuing ESOs is NOT THAT SIMPLE



From 2007, New Zealand firms must report the cost of employee stock options (ESOs). Accounting regulators envisage the use of existing option-pricing models for this purpose, but these models assume an option that is continuously tradable. Glenn Boyle looks at what is and isn't relevant to ESO valuation, and concludes that the tradability assumption is not trivial.[1]

ESOs provide remuneration to holders (usually executive employees) if their firm's stock price rises above some pre-specified threshold. This seems like an attractive feature in that it provides a direct link between executive pay and shareholder wealth. In practice, however, the reputation of ESOs has been soured by their association with exorbitant compensation packages, repricing in favour of the holders, and general corporate malfeasance. In essence, they have come to be seen as a symptom of managerial excess rather than as a solution to it.

One commonly cited reason for the failure of ESOs to live up to their potential is the absence of an accounting standard that requires firms to treat ESOs as a compensation expense. As far as reported profits are concerned, ESOs have been a free lunch – which has encouraged the granting of too many of them on too generous terms. So accounting authorities have recently taken steps to ensure that the costs of ESOs will in future be recognised in financial statements: all New Zealand firms must this year begin recognising ESOs at their grant-date 'fair value'.[2]

Sorry, what number was that?

Such a requirement begs the question of what constitutes fair value. The methods envisaged by accounting regulators for achieving this involve using models that calculate the equilibrium price of *market-traded* options. In such models, investors diversify their portfolios and thus are essentially risk-neutral with respect to firm-specific risk. Hence, in the language of modern finance, the value of any option equals its expected payoff discounted at a rate that includes a premium for non-diversifiable (market) risk only.

However, ESOs are generally *not* tradable, as allowing their sale would undo the reasons for granting them in the first place. As a result, ESO holders are typically under-diversified and over-exposed to the granting firm – and the value (to the recipient) of ESOs equals the expected option payoff discounted at a rate that includes a premium for both market *and* firm-specific risk. In short, not being able to trade ESOs lowers the value placed on them by their holders relative to that of otherwise-equivalent traded options.

That a lack of liquidity lowers an asset's value is unsurprising. What may be surprising is that this phenomenon is not *directly* relevant to the value that investors and accountants are interested in – the cost incurred by the firm in granting ESOs. To understand the difference, suppose an employee's remuneration contract includes the use of a $50,000 car. If the employee does not drive, or lives only a short distance from the workplace, then the value the employee places on the car is likely to be considerably less than $50,000, and may approach zero. But the cost to the firm is still $50,000. And exactly the same principle applies to ESOs. The opportunity cost to the firm is the value of the ESO in the marketplace (that is, the expected ESO payoff discounted at a rate including a premium for systematic risk only). This is because the funds potentially used to pay out the option could otherwise have been invested elsewhere in the market; the particular circumstances of the ESO holder are irrelevant.

Nevertheless, the inability to trade ESOs *does* have an indirect effect on their cost to the firm. To continue with the company car analogy: the total cost to the firm depends on the employee's usage policy, insofar as a car that has been only lightly used will generally have a greater resale value than one with many kilometres on the clock. With ESOs, exposure to the firm's unsystematic risk leads holders to pursue an exercise policy that differs from the one they would have chosen if the options were able to be traded. This typically results in the ESO being exercised earlier than an otherwise-equivalent traded option, as exercise represents the only way for under-diversified holders to liquidate their position. Early exercise changes the expected option payoff, and so the cost to the firm equals this revised expected payoff discounted at the systematic-risk-adjusted rate used by the market. Thus, the cost to firms of granting ESOs differs from the value of otherwise-equivalent options traded in the market not because the characteristics of ESOs make them less valuable to recipients, but because

Figure 1: Market value and firm cost for a hypothetical ESO

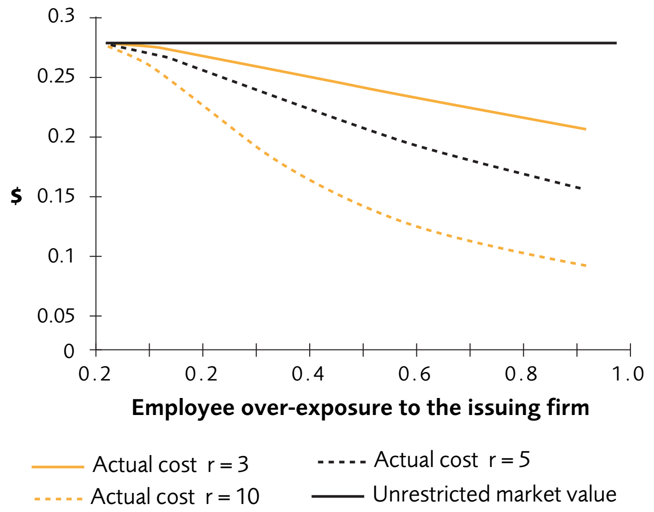
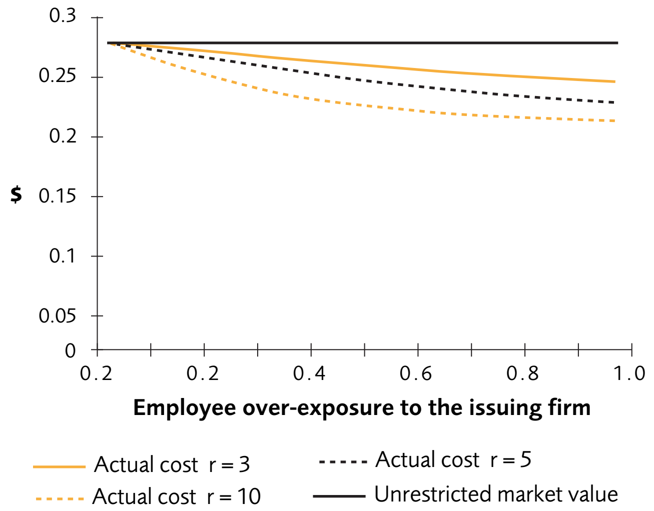


Figure 2: Market value and firm cost for a hypothetical ESO (vests in 18 months)



these characteristics affect the optimal exercise policy of recipients – thereby changing the option payoff distribution.

In summary, neither the unrestricted market value nor the subjective employee value represents the cost of the ESO to the firm: the former assumes an exercise policy, and hence an expected payout, different from that actually followed by the employee; the latter assumes an opportunity cost of funds different from that faced by the firm. Instead the actual cost to the firm is the present *market* value of the expected liability created by the employee's exercise policy.

Cashing up – the cost of tradability

In theory, calculating the actual cost of ESOs is straightforward and consists of two steps: identifying the employee-specific exercise date; and estimating the present market value of the payoff generated by exercising the option at this date.

In practice, however, the first step is problematical precisely because the exercise date is employee-specific – that is, it depends on unobservable characteristics such as the employee's degree of under-diversification and tolerance for risk. If the choice of exercise date is sensitive to these characteristics, then models of market-traded options (in which the optimal exercise-date depends only on market conditions and hence is the same for all investors) may yield estimates of ESO value that significantly diverge from their actual cost to the firm.

Figure 1 illustrates this point for a hypothetical in-the-money, vested ESO with moderate volatility and dividend yield. If this asset were traded in the market then its theoretical value would be $0.28 (the horizontal line), regardless of the specific characteristics of the employee to whom it is granted. However, the remaining three curves demonstrate that matters change once the ESO's non-tradability is incorporated. Now the cost to the firm falls as the employee's under-diversification rises (0.0 corresponds to perfect diversification, and 1.0 to perfect concentration in the issuing firm). Moreover, the extent of this fall is greater for employees who are less tolerant of risk (r is an index of employee risk aversion).[3] Employees who are under-diversified and risk averse are more likely to choose an exercise date that is sub-optimally early from the market's perspective, thereby lowering the firm's expected payoff and hence the cost of granting the ESO.

The divergence of an ESO's actual cost from its market value can be economically significant. For example, the cost of granting this ESO to an employee who holds an extra 25% of his wealth in the firm and is moderately risk averse (r = 5) is $0.25, approximately 11% less than its market value. For a highly risk-averse employee (r = 10) the cost falls to $0.20, some 29% below its market value.

The situation depicted in Figure 1 has obvious implications for ESO valuation. On the one hand, so-called objective methods of valuation that implicitly assume ESOs are traded can yield significant over-estimates of the cost of ESOs to the firm. On the other hand, eliminating this problem requires knowledge of employee characteristics of under-diversification and risk aversion, thereby introducing a subjective element.

Interestingly, this problem is generally less acute for ESOs that are subject to vesting requirements. Figure 2 shows that the combined effects of non-tradability, under-diversification and risk aversion have a considerably smaller effect on the cost of the hypothetical ESO. This result might at first seem surprising, insofar as restricting the dates on which ESOs can be exercised unambiguously reduces their value to employees. But the reason is straightforward: employees who would otherwise have exercised sub-optimally early from the market's perspective (that is, prior to the vesting date) are now forced to wait until a later date that offers a higher expected payoff, thereby increasing the granting cost to the firm.

The potential benefits to investors and markets of requiring firms to value and expense ESOs are difficult to deny (although some have tried). The practicalities of doing this accurately, however, are likely to remain controversial for some time yet. The challenge is to obtain accurate values without creating the scope for manipulation.

1 This article is based on: G Boyle, S Clyne and H Roberts. 2006. 'Valuing employee stock options: implications for the implementation of NZ IFRS 2' *Pacific Accounting Review* 18 pp3-20.

2 Institute of Chartered Accountants of New Zealand. 2004. NZ IFRS 2: *Share-based Payment*.

3 Values of r between three and five are often thought to be reasonable for most investors. See: J Cochrane. 2001. *Asset Pricing*. Princeton University Press.

Glenn Boyle is the Executive Director of ISCR.

guaranteeing the lights stay on
What's it Really Worth?

Power outages are inconvenient and costly. Firms lose production, face material damage, and incur restart costs; households lose leisure time and suffer from stress; and public services may shut down. Reducing the number of outages therefore seems attractive – but doing so creates costs that eventually are borne by electricity users. Resolving the trade-off between security and cost in a socially optimal way requires information on the value of electricity outages. Michiel de Nooij explains how this might be done.[1]

Because there is no trading market for claims on electricity outages, direct price data cannot be used to infer the value of supply security. So other techniques have necessarily developed.

An obvious one is survey data. In the simplest settings, respondents are asked what they would be willing to pay for a specific increase in reliability. Unfortunately, more sophisticated – and more complex – survey techniques are needed to overcome the framing difficulties typically associated with survey questions, and this limits the usefulness of the survey approach.

A second, and more promising, technique is the 'production-function approach'. This calculates the value foregone as the result of an electricity outage. For example, companies lose the value of output not produced, while households lose leisure time (which, following the insights of Gary Becker,[2] can be proxied by the net marginal wage).

The Dutch evidence

Applying the production-function approach to data from the Netherlands shows the value of lost load (VOLL) – the economic cost (in €) of a 1kWh shortfall in electricity availability. (See Figure 1.)

For the Netherlands economy as a whole, VOLL is €8.6/kWh. But this conceals considerable variation across sectors. Perhaps surprisingly, VOLL is highest in the government and construction sectors (€33.5/kWh and €33.1/kWh respectively) and lowest in manufacturing (€1.9/kWh). Inter-sector differences arise primarily because of differences in productivity and electricity usage. For example, manufacturing creates more value-added than construction – but it also uses considerably more electricity, which outweighs this greater productivity. Households' VOLL, €16.4 /kWh, is roughly midway between the extremes.

VOLL also varies substantially across time (see Table 1), being greatest in the early evening and lowest during the post-midnight period. It also varies substantially across regions, with urban areas showing higher VOLL than rural ones.

Topsy turvy

In electricity-supply shortfalls in most western countries, the manufacturing sector receives priority supply, service sectors appear further down the list, and households are placed at the bottom. But the evidence – at least in the Netherlands – shows that, if the objective is to minimise economic costs (which it should be), then this is precisely the wrong way round. Manufacturing should be the first sector to lose electricity supply (since it has the lowest VOLL), and the government sector should be the last.

Transmission grid companies usually face political and social pressures to improve reliability – but they have no direct financial incentive to maintain supply security in a socially optimal way, and so have no reason to consider the variation across customers. However, if they were to be rewarded for preventing damage, they would be more likely to make decisions that reflect an optimal trade-off between their costs and the social benefits of more reliability.

Figure 1: Netherlands' VOLL by region (Monday to Friday)

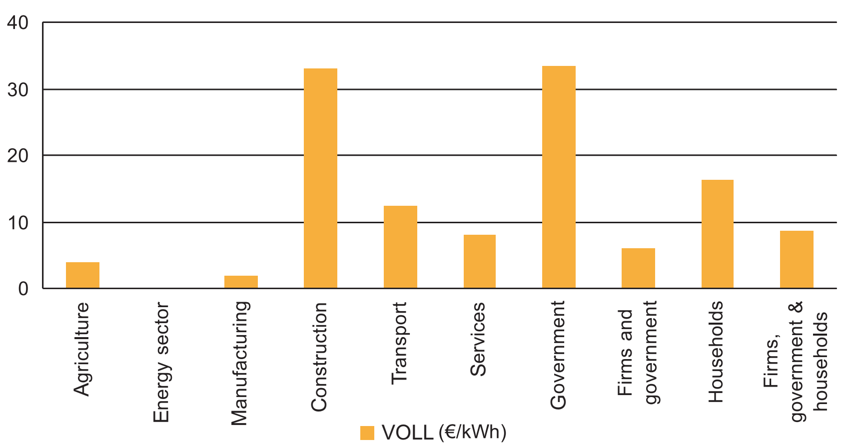


Table 1: Netherlands' VOLL by time of day (€/kWh)

Monday-Friday	day	(8:00-18:00)	8.0
	evening	(18:00-24:00)	8.9
	night	(24:00-8:00)	2.7
Saturday	day	(8:00-18:00)	8.7
	evening	(18:00-24:00)	12.5
	night	(24:00-8:00)	3.9
Sunday	day	(8:00-18:00)	10.3
	evening	(18:00-24:00)	12.5
	night	(24:00-8:00)	3.9

1 This article is based on: Michiel de Nooij, Carl Koopmans and Carlijn Bijvoet. 2007. 'The value of supply security, the costs of power interruptions: economic input for damage reduction and investment in networks' *Energy Economics* 29 pp277-295.

2 G E Becker. 1965. 'A theory of the allocation of time' *The Economic Journal* 75 (September) pp493-517.

Michiel de Nooij is a researcher at SEO Economic Research, University of Amsterdam. During 2006 he was a visiting scholar at ISCR.