EIB PAPERS

Volume 12 No 2 2007

An efficient, sustainable and secure supply of energy for Europe

Meeting the challenge

European Investment Bank

Editorial Policy

The EIB Papers are published each year by the Economic and Financial Studies Division of the European Investment Bank. The journal is divided into two issues and is aimed at encouraging high-quality economic research and debate on matters of European interest. As such the *Papers* are intended to be accessible to non-specialist readers and emphasise policy dimensions rather than technical issues. They present the results of research carried out by Bank staff together with contributions from external scholars and specialists.

Articles will only be accepted for publication on the condition that they have not already been published elsewhere. All articles in the EIB Papers may be freely reproduced and quoted; however, the Editor would appreciate acknowledgment and a copy of the publication in question. They can also be freely accessed and downloaded from our website: http://www.eib.org/efs/

The views expressed in the articles are those of the individual authors and do not necessarily reflect the position of the EIB.

EIB PAPERS

An efficient, sustainable and secure supply of energy for Europe

Meeting the challenge

European Investment Bank

Contents

Preface

Torsten Gersfelt
Vice-President

Ensuring an efficient, sustainable and secure supply of energy is perhaps one of the most important policy tasks of our time – for Europe and, indeed, the global community. European policy makers have recently reinforced their commitment towards achieving this end. Specifically, the European Council of March 2007 endorsed an Energy Policy for Europe that aims, by 2020, at reducing EU Member States' greenhouse gas emissions by at least 20 percent compared to 1990, increasing energy efficiency by 20 percent compared to baseline projections, and raising the share of renewable energy resources in the EU energy mix to 20 percent.

Achieving these targets as planned by 2020 seems ambitious when considering progress on this front so far. However, as the contributions to the companion edition (Volume 12, Number 1) to this edition of the *EIB Papers* (Volume 12, Number 2) suggest, achieving these targets is feasible and does not call for miracles. But it requires a credible and predictable long-term policy framework that induces investment in an efficient, sustainable and secure supply of energy. Yet, even with such a framework in place, difficult choices will have to be made and, moreover, not all means to achieve these targets will be economically meaningful. This is the background against which the contributions to this edition of the *EIB Papers* are set.

These contributions address a variety of issues, including the role of renewables and nuclear energy in an energy mix that strikes an acceptable balance between energy costs and energy price risks and that considerably reduces greenhouse gas emissions; the scope for economically efficient measures to make the supply of energy more secure; the forces driving the extension of the transport system for gas exports to Europe and the implications of that extension for European gas consumers; the strengths and weaknesses of alternative policy instruments in support of renewable energy; the interaction between market failures that adversely affect the environment and market failures that stifle technological progress; the economic rationale for promoting new energy technologies; and barriers that hinder seemingly profitable investments in energy efficiency.

Economic analyses are particularly interesting and fruitful when they question pre-conceived ideas. All contributions to this volume of the *EIB Papers* do this in one way or another. I am thus confident that they enhance our understanding of how we can successfully meet the energy and climate-change challenges in the decades to come.

T. Gersfelt

An efficient, sustainable and secure supply of energy for Europe

Meeting the challenge

The *2007 EIB Conference on Economics and Finance* – held at EIB headquarters in Luxembourg on January 25 – examined challenges towards an efficient, sustainable and secure supply of energy for Europe. Presentations addressed broad policy issues – the credibility and predictability of policy frameworks, for instance – and specific policy questions, such as the rationale for promoting renewable sources of energy, energy efficiency, and new energy technologies.

Speakers included:

Juan ALARIO,
of the EIB

Shimon AWERBUCH,
of the University of Sussex, UK

Bassam FATTOUH,
of the Oxford Institute for Energy Studies, Oxford, UK

Dominique FINON,
of the Centre National de la Recherche Scientifique, Paris

Dieter HELM,
of New College, Oxford, UK

Franz HUBERT,
of the Humbold University, Berlin, Germany

Mark JACCARD,
of the Simon Fraser University, Vancouver, Canada

Machiel MULDER,
of CE Delft, The Netherlands

Armin RIESS,
of the EIB

Dominique RISTORI,
of the European Commission

Joachim SCHLEICH,
of the Fraunhofer Institute for Systems and Innovation Research, Karlsruhe, Germany

Coby VAN DER LINDE,
of the Clingendael International Energy Programme, The Hague, The Netherlands

ABSTRACT

This paper applies portfolio-theory optimisation concepts from the field of finance to produce an expository evaluation of the 2020 projected EU-BAU (business-as-usual) electricity generating mix. *We locate optimal generating portfolios that reduce cost and market risk as well as CO_2 emissions relative to the BAU mix. Optimal generating portfolios generally include greater shares of wind, nuclear, and other non-fossil technologies that often cost more on a stand-alone engineering basis, but overall costs and risks are reduced because of the portfolio diversification effect. They also enhance energy security. The benefit streams created by these optimal mixes warrant current investments of about €250 – €500 billion. The analysis further suggests that the optimal 2020 generating mix is constrained by shortages of wind, especially offshore, and possibly nuclear power, so that even small incremental additions of these two technologies will provide sizeable cost and risk reductions.*

Shimon Awerbuch (http://www.awerbuch.com) was Senior Fellow, Sussex Energy Group, SPRU, University of Sussex until his sudden death in a plane crash on February 10, 2007 – only a few days after presenting this paper at the 2007 EIB Conference on Economics and Finance. He was a financial economist specialising in electric utilities, energy and technology and had previously served as Senior Advisor for Energy Economics, Finance and Technology with the International Energy Agency in Paris. EIB staff who worked with Shimon will remember him as a charming, friendly, and inspiring person and as a dedicated economist enthusiastically presenting his ideas and convictions. We are very grateful to Dr. Spencer Yang – Dr. Awerbuch's co-author – for finishing their joint work.

Spencer Yang (spencer.yang@bateswhite.com) is Visiting Fellow at Sussex Energy Group – SPRU – University of Sussex – Brighton, UK. Spencer Yang is also Manager, Energy Practice, with Bates White, LLC, Washington DC, USA.

Efficient electricity generating portfolios for Europe: maximising energy security and climate change mitigation

Shimon Awerbuch

Spencer Yang

1. Least-cost *vs.* portfolio-based approaches in generation planning

Traditional energy planning in Europe and the United States focuses on finding the least-cost generating alternative. This approach worked sufficiently well in a technological era, marked by relative cost certainty, low rates of technological progress, and technologically homogenous generating alternatives and stable energy prices (Awerbuch 1993, 1995a). However, today's electricity planner faces a broadly diverse range of resource options and a dynamic, complex, and uncertain future. Attempting to identify least-cost alternatives in this uncertain environment is virtually impossible (Awerbuch 1996). As a result, more appropriate techniques are required to find strategies that remain economical under a variety of uncertain future outcomes.

Given the uncertain environment, it makes sense to shift electricity planning from its current emphasis on evaluating alternative technologies to evaluating alternative electricity generating portfolios and strategies. The techniques for doing this are rooted in modern finance theory – in particular mean-variance portfolio theory.[1] Portfolio analysis is widely used by financial investors to create low risk, high return portfolios under various economic conditions. In essence, investors have learned that an efficient portfolio takes no unnecessary risk to its expected return. In short, these investors define efficient portfolios as those that maximise the expected return for any given level of risk, while minimising risk for every level of expected return.

Portfolio theory is highly suited to the problem of planning and evaluating electricity portfolios and strategies because energy planning is not unlike investing in financial securities where financial portfolios are widely used by investors to manage risk and to maximise performance under a variety of unpredictable outcomes. Similarly, it is important to conceive of electricity generation not in terms of the cost of a particular technology today, but in terms of its portfolio cost. At any given time, some alternatives in the portfolio may have high costs while others have lower costs, yet over time, an astute combination of alternatives can serve to minimise overall generation cost relative to the risk. In sum, when portfolio theory is applied to electricity generation planning, conventional and renewable alternatives are not evaluated on the basis of their stand-alone cost, but on the basis of their portfolio cost – that is: their contribution to overall portfolio generating cost relative to their contribution to overall portfolio risk. Portfolio-based electricity planning techniques – pioneered by Awerbuch and Berger (2003), Berger (2003), Awerbuch (2000a), Humphreys and McLain (1998), Awerbuch (1995), and Bar-Lev and Katz (1976) – thus suggest ways to develop diversified generating portfolios with known risk levels that are commensurate with their overall electricity generating costs. Simply put, these techniques help identify generating portfolios that can minimise a society's energy cost and the energy price risk it faces.

This also has important security of energy supply implications. Although energy security considerations are generally focused on the threat of abrupt supply disruptions (see for instance European Commission 2001), a case can also be made for the inclusion of a second aspect: the risk of unexpected electricity cost increases. This is a more subtle, but equally crucial, aspect of energy security. Energy security is reduced when countries (and individual firms) hold inefficient portfolios that are needlessly exposed to the volatile fossil fuel cost risk.

1 Mean-variance portfolio theory (MVP), an established part of modern finance theory, is based on the pioneering work of Nobel Laureate Harry Markowitz 50 years ago. For a recent contribution see Fabozzi *et al.* (2002).

Optimal portfolio mixes are designed to minimise expected generating cost and risk, while simultaneously enhancing energy security.

The purpose of this paper is to describe a portfolio optimisation analysis that develops and evaluates optimal and efficient EU electricity generating mixes for 2020, in an environment of uncertain CO_2 prices. These optimal portfolio mixes are designed to minimise expected generating cost and risk – while simultaneously enhancing energy security – and they can be used as a benchmark for evaluating electricity generating strategies aimed at minimising CO_2 emissions. A key finding of the analysis is that compared to the projected 2020 EU business-as-usual (BAU) electricity generating portfolio, there exist optimal and efficient portfolios that are less risky, less expensive, and that substantially reduce CO_2 emissions and energy import dependency.

In developing these results, we proceed as follows. Section 2 sets out the main principles of the portfolio-based approach to electricity-sector planning. Section 3 describes the data needed for applying such an approach and how we have compiled and estimated them. Using these data, Section 4 identifies optimal EU electricity generating portfolios for 2020 and it presents key features of these portfolios. Section 5 probes deeper into some of the findings, highlighting the role of nuclear energy, the scope for minimising CO_2 emissions, the economic consequences of real-world technology constraints, and the effects of carbon pricing. Section 6 summarises, concludes, and stresses the potential and limitations of our analysis.

Box 1. Electricity generating costs, risks, and correlations

Electricity generating cost and returns

Portfolio theory was initially conceived in the context of financial portfolios, where it relates expected portfolio return to expected portfolio risk, defined as the year-to-year variation of portfolio returns. This box illustrates portfolio theory as it applies to a two-asset generating portfolio, where the generating cost is the relevant measure. Generating cost (€/kWh) is the inverse of a return (kWh/€), that is, a return in terms of physical output per unit of monetary input.

Expected portfolio cost

Expected portfolio cost is the weighted average of the individual expected generating costs for the two technologies:

(1) Expected portfolio cost = $X_1 E(C_1) + X_2 E(C_2)$,

where X_1 and X_2 are the fractional shares of the two technologies in the mix, and $E(C_1)$ and $E(C_2)$ are their expected levelised generating costs per kWh.

Expected portfolio risk

Expected portfolio risk, $E(\sigma_p)$, is the expected year-to-year variation in generating cost. It is also a weighted average of the individual technology cost variances, as tempered by their covariances:

(2) Expected portfolio risk = $E(\sigma_p) = \sqrt{X_1^2\sigma_1^2 + X_2^2\sigma_2^2 + 2X_1X_2\rho_{12}\sigma_1\sigma_2}$,

where: X_1 and X_2 are the fractional shares of the two technologies in the mix; σ_1 and σ_2 are the

2. Portfolio-based approach to electricity sector planning

2.1 Portfolio optimisation basics applied to electricity sector planning

Portfolio theory was developed for financial analysis to locate portfolios with maximum expected return at every level of expected portfolio risk. Box 1 reviews the basics of this theory and explains how this paper applies it to electricity generation mixes. An important point to note here is that in the case of electricity generating portfolios, it is more convenient to optimise portfolio generating costs as opposed to portfolio returns (see Awerbuch and Yang 2007 and Awerbuch and Berger 2003). This choice does not affect the results and conclusions presented in this paper.

In the case of electricity generating portfolios, it is more convenient to optimise portfolio generating costs as opposed to portfolio returns.

Expected portfolio generating cost is the weighted average of the individual technology costs. The expected risk of an electricity portfolio – that is, the expected year-to-year fluctuation in portfolio generating cost – is a weighted average of the risks of the individual technology costs, tempered by their correlations or covariances. Each technology itself is characterised by a portfolio of cost streams, comprising capital outlays, fuel expenditures, operating and maintenance (O&M) expenditure, and CO_2 costs. It follows that for each technology, risk is the standard deviation of the year-to-year changes of these cost inputs.

standard deviations of the holding period returns of the annual costs of technologies *1* and *2* as further discussed below; and ρ_{12} is their correlation coefficient.

Portfolio risk is always estimated as the standard deviation of the holding period returns (HPRs) of future generating cost streams. The HPR is defined as: *HPR=(EV–BV)/BV*, where *EV* is the ending value and *BV* the beginning value (see Brealey and Myers 2004 for a discussion on HPRs). For fuel and other cost streams with annual reported values, *EV* can be taken as the cost in year *t+1* and *BV* as the cost in year *t*. HPRs measure the rate of change in the cost stream from one year to the next. A detailed discussion of its relevance to portfolios is given in Berger (2003).

Each individual technology actually consists of a portfolio of cost streams (capital, operating and maintenance, fuel, CO_2 costs, and so on). Total risk for an individual technology – that is, the portfolio risk for those cost streams – is σ_T. In this case, the weights, X_1, X_2, and so on, are the fractional share of total levelised cost represented by each individual cost stream. For example, total levelised generating costs for a coal plant might consist of ¼ capital, ¼ fuel, ¼ operating costs, and ¼ CO_2 costs, in which case each weight $X_j = 0.25$.

Correlation, diversity, and risk

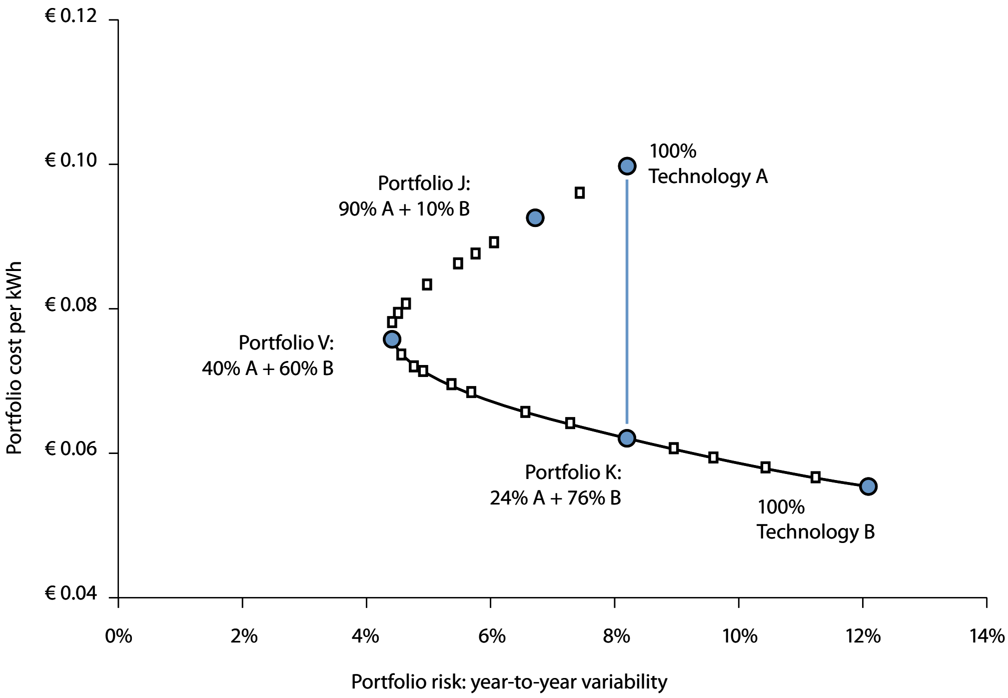
The correlation coefficient, ρ, is a measure of diversity. Lower ρ among portfolio components creates greater diversity, which reduces portfolio risk σ_p (with the notable exception discussed by Roques 2006). More generally, portfolio risk falls with increasing diversity, as measured by an absence of correlation between portfolio components. Adding a fuel-less (that is fixed-cost, riskless) technology to a risky generating mix lowers expected portfolio cost at any level of risk, even if this technology costs more (Awerbuch 2005). A pure fuel-less, fixed-cost technology, has $\sigma_i = 0$ or nearly so. This lowers, σ_p, since two of the three terms in equation (2) reduce to zero. This, in turn, allows higher-risk/lower-cost technologies into the optimal mix. Finally, it is easy to see that σ_p declines as $\rho_{i,j}$ falls below 1. In the case of fuel-less renewable technologies, fuel risk is zero and its correlation with fossil fuel costs is zero too.

For electricity planning, portfolio optimisation exploits the interrelationships among the various technology generating cost components.

Portfolio theory improves decision making in the following way. First, since the investor only needs to consider the portfolios on the so-called efficient frontier, rather than the entire universe of possible portfolios, it simplifies the portfolio selection problem. Second, it quantifies the notion that diversification reduces risk. For electricity planning, portfolio optimisation exploits the interrelationships (i.e., correlations) among the various technology generating cost components. Take for example fossil fuel prices. Because they are correlated with each other, a fossil-dominated portfolio is undiversified and exposed to fuel price risk. Conversely, renewables, nuclear, and other non-fossil options diversify the mix and reduce its expected risk because their costs are not correlated with fossil prices.

The portfolio diversification effect is illustrated in Figure 1, which shows the costs and risks for various possible two-technology portfolios. Technology *A* is representative of a generating alternative with higher cost and lower risk – such as photovoltaics (PV). It has an expected (illustrative) cost of around €0.10 per kWh with an expected year-to-year risk of 8 percent. Technology *B* is a lower-cost/higher-risk alternative – such as gas-fired generation. Its expected cost and risk are about €0.055 per kWh and 12 percent, respectively. The correlation factor between the total cost streams of the two technologies is assumed to be zero. This is a simplification since in reality the capital and variable cost of PV will exhibit some non-zero correlation with the capital and variable cost of gas generation.

Figure 1. Portfolio effect for illustrative two-technology portfolio



As a consequence of the portfolio effect, total portfolio risk decreases when the riskier technology *B* is added to a portfolio consisting of 100 percent *A*. For example, portfolio *J*, which comprises 90 percent of technology *A* plus 10 percent *B*, exhibits a lower expected risk than a portfolio comprising 100 percent *A*. This is counter-intuitive since technology *B* is riskier than *A*. Portfolio *V*, the minimum variance portfolio, has a risk of around 4 percent, which is half of the risk of *A* and one-third of the risk of *B*. This, however, illustrates the point of diversification.

Investors would not hold any mix above portfolio *V* because mixes exhibiting the equivalent risk can be obtained at lower cost on the solid portion of the line. Portfolio *K* is therefore superior to 100 percent *A*. It has the same risk, but lower expected cost. Investors would not hold a portfolio consisting only of technology *A,* but rather would hold the mix represented by *K*. Taken on a stand-alone basis, technology *A* is more costly, yet properly combined with *B,* as in portfolio *K*, it has attractive cost and risk properties. Not only is the mix *K* superior to 100 percent *A*, most investors would also consider it superior to 100 percent technology *B*. Compared to *B*, mix *K* reduces risk by one-third while increasing cost by just 10 percent (€0.005 per kWh), which gives it a higher Sharpe ratio than other mixes.[2] Mix *K* illustrates that astute portfolio combinations of diversified alternatives produce efficient results, which cannot be measured using stand-alone cost concepts. To summarise, portfolio optimisation locates minimum-cost generating portfolios at every level of portfolio risk, represented by the solid part of the line in Figure 1, that is, the stretch between *V* and *B*.

2.2 Portfolio-risk perspective *vs.* engineering-risk perspective

Having sketched the gist of the portfolio approach to electricity generation planning, it is useful to comment on the distinction between unsystematic (or firm-specific) risk, systematic (or market) risk, and risks usually considered in engineering approaches to analysing the pros and cons of alternative generation technologies.

Finance theory divides total risk into two components: unsystematic risk that affects primarily the prices of an asset (these risks can be reduced through diversification) and systematic that affect the prices of all assets. Systematic risk refers to the risk common to all securities and cannot be diversified away (within one market). Within the market portfolio, unsystematic risk will be diversified away to the extent possible. Systematic risk is therefore equated with the risk (standard deviation) of the market portfolio.

In the case of generating technologies and other real assets, diversification and portfolio risk are frequently misunderstood. With some analysts adopting an engineering approach that strives to enumerate all conceivable risks, including those that do not affect overall portfolio risk by virtue of diversification.[3] Ignoring diversification effects in this manner, however, yields a portfolio risk estimate that is systematically biased upwards.

By ignoring diversification effects, engineering risk studies yield a portfolio risk estimate that is systematically biased upwards.

For example, year-to-year fluctuations in electricity output from a wind farm is an unsystematic risk and is probably not relevant for portfolio purposes since it is uncorrelated to the risk of other portfolio cost streams – though this unsystematic risk presents a potential risk to the owner of the wind farm. Certainly in the case of a large, geographically dispersed mix such as the EU generating portfolio, year-to-year wind resource variability can be considered random and uncorrelated to fossil fuel prices or other generating cost components. While it is possible to measure the standard deviation of the yearly wind resource at a given location, its correlation to the output of other wind farms across the continent (see Figure 2), or to many if not most other generating cost components, is arguably zero (that is, $\rho_{12} = 0$ in equation (2) of Box 1). Thus, wind variability at a particular location does not contribute significantly to portfolio risk.

2 Developed by Nobel Laureate William F. Sharpe, this ratio is a risk adjusted performance of an asset and is used to characterise how well the return of an asset compensates the investor for the risk taken.

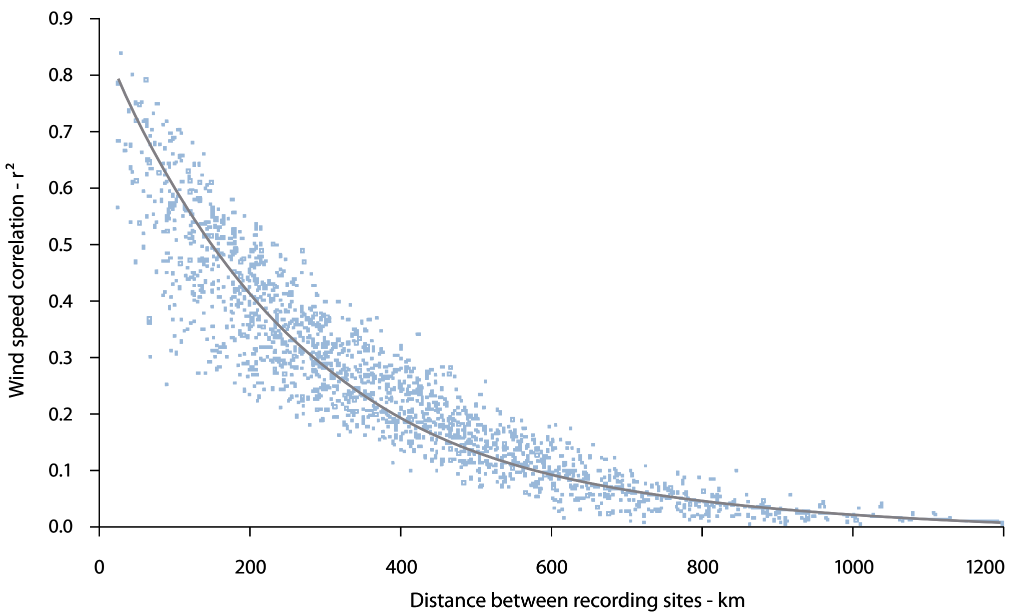
3 For example, Jansen *et al.* (2006, p. 56) develop complex *ad hoc* procedures intended to produce a 'transparent [and] comprehensive' portfolio risk measure by attempting to enumerate and combine various random risks that might affect individual generators, but which cannot be expected to affect overall portfolio risk except possibly in the case of very small generating systems.

Operating costs for wind, solar, and other passive, capital-intensive renewables are essentially fixed, or riskless, over time.

From a portfolio perspective, there is another important point to consider. Operating costs for wind, solar, and other passive, capital-intensive renewables are essentially fixed, or riskless, over time. The finance-theory aspects of these fixed-cost, riskless technologies are developed in Awerbuch (2000b).[4] Perhaps more important is that these costs are uncorrelated to fossil fuel prices. This enables these technologies to diversify the generating mix and enhance its cost-risk performance. Given sufficient geographic dispersion in the wind resources, as would be expected in an EU-wide portfolio, the operating cost of a generating system with 30 percent wind will fluctuate less from year to year than a system with no wind.[5]

The idea that enumerating all conceivable unsystematic risks is misleading in the context of a generating portfolio study holds for other engineering variances – such as annual variations in attained fuel conversion efficiency for a particular gas plant. Some analysts (Jansen *et al*. 2006, for instance) choose to include this risk. Although such yearly efficiency fluctuations might change the accountant's estimate of kWh generating costs at a given site[6], it is reasonable to assume that risk is uncorrelated, making only small contributions to overall portfolio risk.

Figure 2. Onshore wind speed correlation by distance – United Kingdom



Source: Sinden (2005)
Note: Showing 1,770 pairs of wind speed recording sites (surface wind speed), typically based on 30 years of data per pair.

4 Strictly speaking, in the case of capital costs, this statement holds only *ex post*, although, given the short lead times of renewables projects and the large proportion of manufactured components, construction-period risks for these technologies are low even *ex ante*. O&M costs for renewables arguably have the same portfolio risks as O&M costs for conventional technologies. However, because they represent a small share of total cost of renewable generation, their risk contribution is also small. This is further discussed in Awerbuch (2000).

5 Sinden (2005) and Grubb *et al*. (2007) illustrate how geographic dispersion diversifies wind variability.

6 On an accounting basis, kWh generating cost is calculated by dividing annual capital charges plus operating costs by the year's kWh output. Given a fixed capital charge and relatively fixed maintenance costs, therefore, annual wind output variability would cause year-to-year kWh costs to vary. Sunk capital costs are irrelevant in an economic sense, but fluctuations in periodic wind output might change the economic kWh cost estimate on the basis of avoided costs.

3. Data needed for computing optimal electricity generating portfolios

Applying portfolio optimisation to the EU generating mix requires the following inputs: (i) capital, fuel, operating, and CO_2 costs per unit of output (kWh) for each generating technology; (ii) the risk or standard deviation of each cost component; (iii) the correlation factors between all cost components. The following sub-sections will address each input and the way they are used to identify optimal portfolios. A more detailed presentation of the data and estimation can be found in Awerbuch and Yang (2007).

3.1 Technology generating cost

Figure 3 shows levelised 2020 generating cost for various technologies based on TECHPOLE performance and cost data.[7] Fossil fuel costs reflect the most recent projections of the European Commission (European Commission 2006) and the International Energy Agency (IEA 2006).

As for the cost of CO_2, a value of €35/t CO_2 has been used. This can be interpreted as an expected market price of CO_2, assuming that economic policies aimed at internalising the economic cost of CO_2 emissions yield a market price of CO_2 – for example, under the European Union Emissions Trading Scheme. Alternatively, in the absence of such policies, the cost of CO_2 can be interpreted as the shadow price of CO_2, estimated on the basis of the economic cost of CO_2 emissions and of CO_2 abatement cost.[8] As for capital cost, this study assumes full capital cost recovery for new and already installed generating capacity. Although capital costs are sunk from an economic perspective, we assume that electricity producers set prices to recover their sunk costs. This assumption may not hold in day-to-day decision-making, but over time, producers cannot remain viable unless they recover their capital costs. Thus, a full-cost recovery approach is implemented for both existing and new plants.

As Figure 3 shows, a system integration charge is added to wind generation to compensate for 'intermittency costs'. This adjustment is necessary because wind is a variable-output technology. System integration is a complex issue. Many think of wind as intermittent, although there are very few times when wind output is actually zero (Sinden 2005 and Grubb *et al*. 2007). The existing electricity network organisation and protocols do require wind integration to have some extra level of backup capacity to balance the system when wind electricity output is reduced.[9] The costs have been quantified in multiple studies with similar results (Dale *et al.* 2004, DENA Grid Study 2005, and UKERC 2006, for instance). Our analysis follows the results of the UKERC (2006) survey, which estimates the aggregate intermittency costs in the range of €7.5–€12 per MWh (£5–£8 per MWh) for 20 percent wind penetrations. Because intermittency cost estimates in Europe are somewhat lower (DENA Grid Study 2005, for instance, estimated cost at or under €10/MWh), we apply a system integration charge of €10/MWh. This analysis, however, does not include possible associated systematic risks that may become more significant for wind penetrations in excess of 20-30 percent.

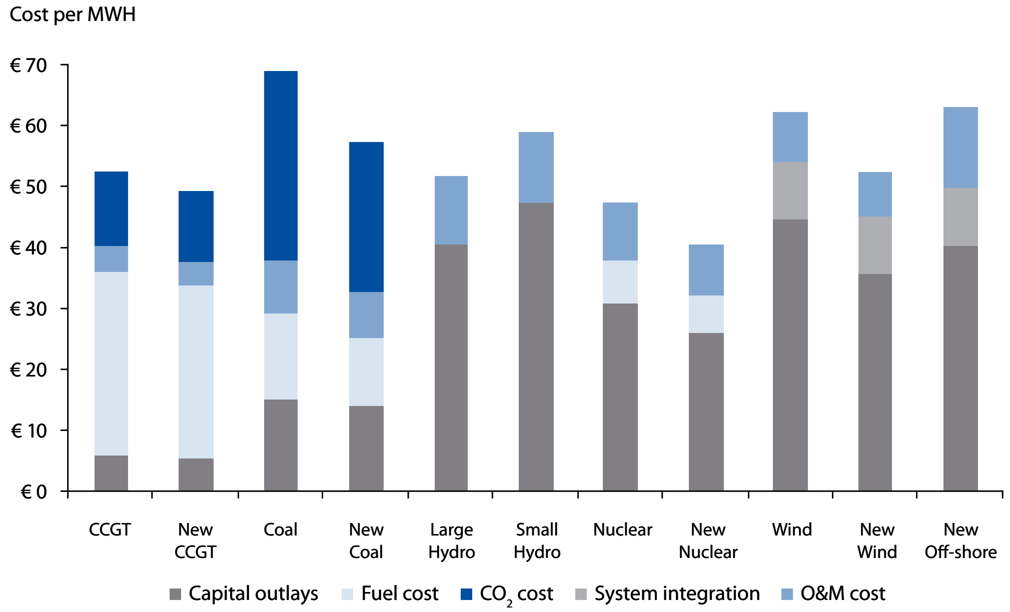
Many think of wind as intermittent, although there are very few times when wind output is actually zero.

7 TECHPOLE database, LEPII, University of Grenoble, CNRS. Fuel input cost, reflecting most recent projections of the European Commission and IEA are shown in Annex Table A1.

8 For example, in its cost-benefit analyses of energy sector projects, the European Investment Bank currently uses a baseline shadow price that rises from €25/t CO_2 in 2007 to €45/t CO_2 in 2030.

9 This being said, new electricity network protocols and information systems have even been proposed in order to exploit wind variability and obviate the need for standby reserve capacity (see – for instance – Awerbuch 2004 and Fox and Flynn 2005). These proposals generally involve matching variable output wind to interruptible load applications to prevent both system balancing and/or backup generation.

Figure 3. 2020 generating costs (€/MWh) for various technologies



Sources: Based on TECHPOLE database, LEPII, University of Grenoble; European Commission (2006), and IEA (2006).
Note: Economic costs of CO_2 assumed to be €35/t CO_2.

3.2 Technology risk estimates

Table 1 summarises our technology risk estimates, expressed as the standard deviations of the holding-period-returns (see Box 1) based on historical data for each cost component.

Construction cost risk varies by technology type and is generally related to the complexity and length of the construction period.

Let us start with capital, or construction, cost risk. This varies by technology type and is generally related to the complexity and length of the construction period. A World Bank analysis covering a large number of projects estimates the standard deviation of construction period outlays for thermal plants (for instance, coal-fired power stations) at 23 percent and 38 percent for large hydro plants (Bacon *et al.* 1996). For the purpose of our analysis, we apply the thermal plant value to the construction-period risk of nuclear plant (but we will consider alternative values in Section 5.1). To some extent, this is an arbitrary simplification. Many believe these risks are significantly higher. Others, however, believe such risks will resolve themselves with experience. The estimates for wind, gas, geothermal, and solar risk were determined from developer interviews as reported in Awerbuch *et al.* (2005). Construction cost risk of existing capacity was estimated at around zero percent. This suggests that 'new' vintage assets are riskier than old ones – for example, risks for a new, not yet constructed coal plant are greater than those for an existing coal plant.

Fuel cost risks have been estimated on the basis of historical (1980-2005) European fossil fuel import prices taken from an IEA database. Annual price observations were used because they eliminate seasonal variations that could potentially bias the results. In practice, electricity producers buy fuel through spot and contract purchases so that the cost of fuel in any calendar period is best measured as the total fuel outlays divided by total fuel delivered. The HPR standard deviations of fuel cost range from 0.14 for coal to 0.24 for oil. Obviously, renewable technologies and geothermal require no fuel outlays and there is thus no fuel cost risk.

Table 1. **HPR standard deviations for generating technology cost streams (in %)**

	Construction	Fuel	O&M	CO_2
Coal	23.0	14.0	5.4	26.0
Oil	23.0	25.0	24.2	26.0
Gas-CC turbine	15.0	19.0	10.5	26.0
Nuclear	23.0	24.0	5.5	–
Hydro-large	38.0	0.0	15.3	–
Hydro-small	10.0	0.0	15.3	–
Wind	5.0	0.0	8.0	–
Wind-offshore	10.0	0.0	8.0	–
Biomass	20.0	18.0	10.8	–
PV	5.0	0.0	3.4	–
Geothermal	15.0	0.0	15.3	–

Source: Own calculation.

Notes: HPR ≡ holding-period-returns; for definition of HPR see Box 1; they measure the year-to-year fluctuation of the underlying cost stream; as a result, the standard deviation is expressed in % while the cost stream itself is measured in €/kWh; construction cost HPRs for existing capacities are not shown as they are estimated at about zero.

The risks of operating and maintenance outlays are difficult to estimate.

The risks of operating and maintenance outlays are difficult to estimate. Typically, estimates can be found in corporate records. But often, these records are not publicly available. Even if they were, maintenance policies may not keep the records in a format suitable for the analysis carried out here. In addition, companies design these records to promote overall corporate objectives, which can result in biased numbers. For example, during periods of poor financial performance, corporate managers may choose to defer maintenance in order to meet specified corporate objectives – such as reducing O&M expense. Thus, maintenance outlays might be arbitrarily recorded as capital improvements and be depreciated over time. In the case of rate-regulated utilities, there is a significant incentive to charging these outlays to capital improvements because they earn a regulated rate of return.

The US Energy Information Agency and the Federal Energy Regulatory Commission databases maintain records covering every generator operated by a regulated utility. This data was used to estimate the HPR standard deviations for O&M costs (along with the correlations between these costs discussed in the next subsection). By using this data, we implicitly assume that the maintenance volatility for a large portfolio of generating assets in the United States will not differ materially from those that would be found for a similar European portfolio. As Table 1 shows, different technologies show different year-to-year fluctuations in maintenance outlays – ranging from 3.4 percent for photovoltaics to 24.2 percent for oil.[10]

This takes us to the risk associated with the last cost category, that is, the cost of CO_2 emissions, which is relevant for fossil fuel technologies. As Table 1 indicates, the HPR standard deviation for CO_2 has been estimated at 26 percent. The approach underlying this estimate will be presented next in the context of discussing the correlation between the cost of different fuels, the correlation between O&M costs of different technologies, and the correlation between the cost of fossil fuels, on the one hand, and CO_2 cost on the other. A more comprehensive presentation of the technology cost and risk estimation can be found in Awerbuch and Yang (2007).

10 In principle, the O&M cost category should include outlays for property taxes, insurance, and other non-maintenance categories. These would most likely exhibit lower risk and potentially dampen the results of Table 1. Because the focus in this paper is on CO_2 risk, we did not pursue this O&M issue further.

3.3 Correlation coefficients

We start here with a brief description of our approach to estimating the HPR standard deviation for CO_2 and the correlation between CO_2 cost and fuel prices. Our estimates are derived using both analytic techniques and Monte Carlo simulation. The analytic approach to estimating CO_2 risk and correlation follows the spirit of Green (2006), who expresses CO_2 price in terms of gas and coal prices. This relationship is used to derive the HPR standard deviation of CO_2 as well as its correlation with fossil fuels. The Monte Carlo approach uses a series of simulations that provide a second set of CO_2 risk and fossil fuel correlation estimates. In the Monte Carlo analyses, we used the volatility and other trends from 18 months of actual data to simulate 20 years of trading. This and its correlation to coal, gas, and oil provides an estimate of annual risk factors for $CO_{2.}$

Both methods provide a range of estimates of CO_2 risk and correlations. We compared the analytical and Monte Carlo results and performed various sensitivity analyses to test the reasonableness and robustness of these estimates. The HPR standard deviation for CO_2 that we use in the portfolio optimisation model (26 percent) is shown in the last column of Table 1.[11] The CO_2 cost/fuel cost correlation coefficient used in the portfolio optimisation is shown in the second-last column (or row) of Table 2 below.

As gas becomes more expensive, electricity generation shifts to coal, putting upward pressure on CO_2 prices – be they market prices or shadow prices.

As can be seen from these correlation coefficients, there is a negative correlation between CO_2 and coal prices and a positive correlation between CO_2 and gas. This is the expected result. Intuitively, as gas becomes more expensive, electricity generation shifts to coal, putting upward pressure on CO_2 prices – be they market prices or shadow prices. Conversely, rising coal prices shift generation to gas, which emits about half as much CO_2. As a result, the price of CO_2 falls with rising coal prices.

Table 2 also shows the correlation coefficients for the various fuels, indicating a positive correlation between fuels – with the notable exception of biomass. Although the data used for this analysis do not obtain a negative fuel correlation for nuclear, a number of researchers (Awerbuch and Berger 2003 and Roques 2006) find a negative correlation between nuclear and fossil fuels, suggesting a greater diversification potential than that resulting from our analysis.

The estimated O&M correlation coefficients are shown in Table A2 in the Annex.

Table 2. Fuel and CO_2 HPR correlation coefficients

	Coal	Oil	Gas	Uranium	CO_2	Biomass
Coal	1.00	0.27	0.47	0.12	-0.49	-0.38
Oil	0.27	1.00	0.49	0.08	0.19	-0.17
Gas	0.47	0.49	1.00	0.06	0.68	-0.44
Uranium	0.12	0.08	0.06	1.00	0.00	-0.22
CO_2	-0.49	0.19	0.68	0.00	1.00	0.00
Biomass	-0.38	-0.17	-0.44	-0.22	0.00	1.00

Source: Own calculation.

11 While our CO_2 risk estimates are statistically robust, it is important to note that they are based on just 18 months of CO_2 trading. Because the results of the CO_2 risk and correlation estimates were relatively consistent over various unrelated estimation procedures, we are relatively confident in applying them to the analysis.

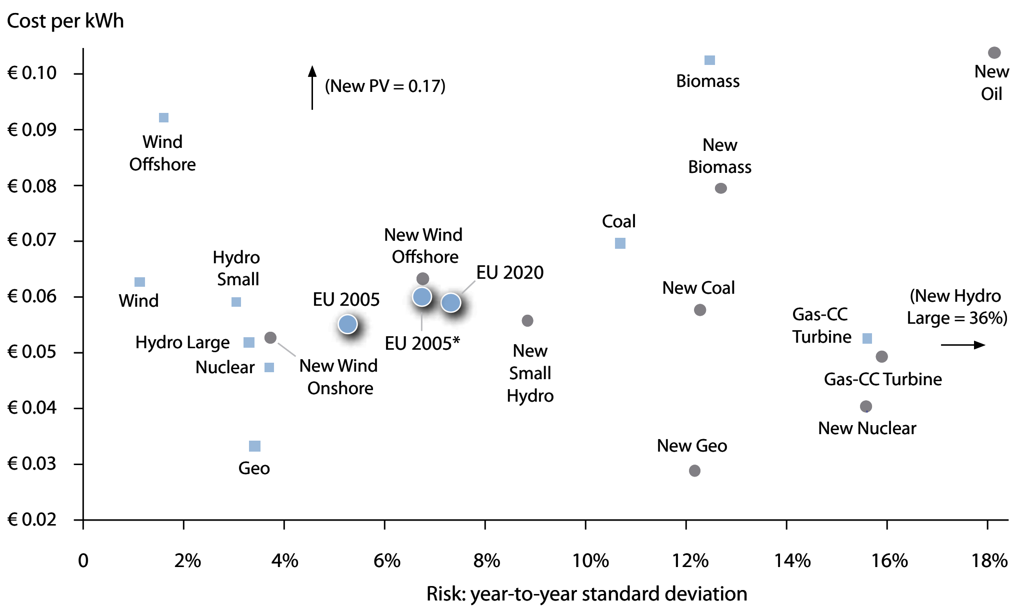
3.4 Total technology cost and risk

The previous sub-sections described the cost and risk inputs for the various generating technologies. These are combined using equation (2) in Box 1 to produce a total HPR standard deviation for each technology, where the weights (X_1, X_2, … etc.) are given by the proportional values of the levelised cost components, that is, capital, fuel, O&M, and CO_2 outlays.

Figure 4 shows the costs per kWh for each of the generating technologies in 2020 along with its risk, with the added assumption that CO_2 costs €35 per tonne. For comparison, Figure 4 also shows the cost-risk combination of the projected EU 2020 BAU mix; in addition, it pictures two variants of the EU 2005 mix: one assuming CO_2 cost of €15 per tonne and the other €35 per tonne. The former reflects the approximate price of CO_2 in 2005 and the latter enables a direct comparison between the 2005 mix and 2020 BAU mix. This comparison shows that relative to the 2005 mix, the 2020 BAU mix slightly reduces electricity generating cost from 5.98 €-cents to 5.87 €-cents per kWh. This cost reduction is attained by increasing expected risk from 6.8 percent to 7.3 percent. Compared to the 2005 EU mix, the 2020 BAU mix represents a cost-risk trade-off that few investors would make: a cost reduction of less than 2 percent would come with an increase in risk of almost 9 percent.

The business-as-usual EU electricity mix for 2020 represents a cost-risk trade-off that few investors would make.

Figure 4. Cost and risk of existing and new EU generating alternatives in 2020



Source: Own calculation.

Notes: Estimates for individual technologies and the EU 2020 BAU mix are based on a CO_2 emission cost of €35/t CO_2. For comparison, the Figure shows the actual EU 2005 generation mix for €35/t CO_2 (EU 2005*) and for €15/t CO_2 (EU 2005). See text for details.

The results also show that compared to existing vintages, new vintages exhibit lower cost and larger risk (in Figure 4, new vintages lie to the southeast of existing vintages). The cost decline is because new-vintage technologies increase energy efficiency and, thus, lower cost. For example, electricity produced by new coal plants cost 5.8 €-cents per kWh, which is 1.2 €-cents less than for existing coal plants. Risk for new vintages increases because the construction-period risk of existing vintages are sunk or zero while new generating assets yet to be constructed are exposed to construction-period risk. The largest differences between the new and existing vintages show up in capital-intensive technologies such as nuclear, wind (especially offshore), and geothermal.

CO_2 prices also increase the risk of the fossil alternatives to the extent that the holding-period-return risk of the CO_2 exceeds that of fossil fuels.

Not unexpectedly, the inclusion of CO_2 charges increases the generating cost of fossil alternatives relative to non-fossil technologies. CO_2 prices also increase the risk of the fossil alternatives to the extent that the HPR risk of the CO_2 exceeds the HPR risk of the fossil fuel. As shown in Table 1, the HPR standard deviation for CO_2 is 26 percent as compared to 14 percent for coal fuel and 19 percent for natural gas fuel. Observe that with €35/t CO_2, the standard deviation of existing coal technology rises from 5.6 percent to 10.7 percent, while the risk of existing gas generation increases much less from 14.3 percent to 15.7 percent (see Table 3). The increase for new coal is also smaller than for existing coal because the risk of new coal includes the construction-period risks, reducing the fractional share of CO_2 outlays (that is, the weight of CO_2 outlays in equation (2) of Box 1).[12]

Table 3. The effect of CO_2 costs on coal and gas generating cost-risk

	CO_2 cost per tonne					
	€0.00		€15.00		€35.00	
	Cost (€/MWh)	Risk (%)	Cost (€/MWh)	Risk (%)	Cost (€/MWh)	Risk (%)
Coal	3.8	5.6	5.1	6.2	6.9	10.7
Coal – New	3.3	11.7	4.3	10.3	5.8	12.3
Gas-CC	4.0	14.3	4.6	14.9	5.3	15.7
Gas-CC – New	3.8	14.7	4.3	15.2	4.9	16.0
Oil	8.2	20.2	9.2	18.8	10.6	17.8
Oil – New	8.0	20.8	8.7	19.3	10.2	18.3

Source: Own calculation.

In the case of oil-fired electricity generation, the HPR fuel price risk is 25 percent (slightly lower than CO_2). Because of the low correlation between CO_2 and oil (0.19 as shown in Table 2), the inclusion of CO_2 charges reduces overall risk of this technology as the proportional weight of CO_2 outlays rises as a share of total costs.

The general outcome is that our 26 percent estimate for the CO_2 HPR risk and our estimated CO_2-fossil fuel correlations, along with the addition of CO_2 charges, do not significantly raise total HPR risks of new fossil generating assets and in some cases lowers them. This is contrary to widely held beliefs. Of course, higher CO_2 risk estimates (or higher correlation with fossil fuels) will affect even new assets to a greater extent.

4. Portfolio optimisation of EU electricity generating mix

4.1 Efficient multi-technology electricity portfolios – an illustration

As previously stated, the aim of this study is to evaluate whether there exists feasible 2020 generating mixes that are 'superior' to the 2020 EU-BAU mix by virtue of reducing risk or the cost of

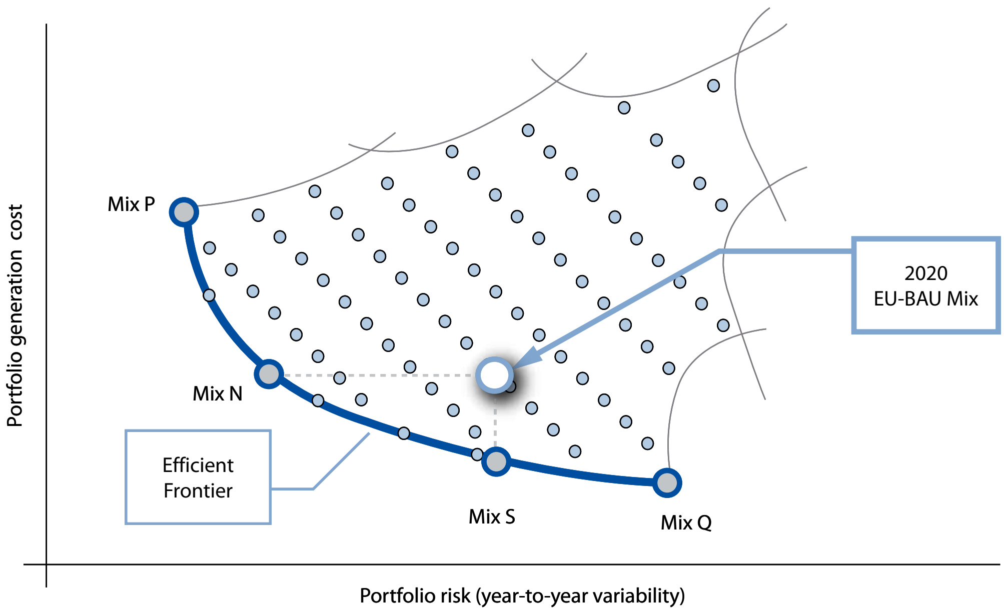
12 Note that the risk for new coal decreases slightly as CO_2 costs move from €0 to €15 per tonne; this is undoubtedly caused by the negative correlation between CO_2 cost and coal prices. As CO_2 cost rise to €35, however, the magnitude of the price overwhelms the negative correlation, and overall risk rises again.

producing electricity. To prepare for the interpretation of the results of our portfolio optimisation model, it is useful to offer a general illustration of possible results.

Figure 5 shows an infinite number of different generating mixes that could meet the 2020 electricity needs with a unique mix of the various technology options. The different portfolios all have different cost-risk as represented by the blue dots. Interestingly, technology shares do not change monotonically in any direction in Figure 5 so that two mixes with virtually identical cost-risk characteristics (that is, two mixes located close to each other in cost-risk space) can have radically different technology generating shares. Indeed, Awerbuch and Berger (2003) show that costs and risks of the EU generating mix projected for 2010 are virtually identical to a mix consisting of 100 percent coal. Likewise, radically different mixes can have nearly identical cost-risk characteristics, that is, they could be virtually co-located in the risk-cost space. The intuition for this is straightforward: there are many ways to combine ingredients in order to produce a given quantity of salad at a given price.

Radically different mixes can have nearly identical cost-risk characteristics, that is, they could be virtually co-located in the risk-cost space.

Figure 5. Feasible region and efficient frontier for multi-technology electricity portfolios



The blue curve (*PNSQ*) is the so-called efficient frontier (EF), the locus of all optimal mixes. There are no feasible mixes below the efficient frontier, and along it, only accepting greater risk can reduce cost. The blue-dot mixes in Figure 5 are sub-optimal or inefficient because it is still possible to reduce both cost and risk by finding mixes on the efficient frontier by moving below or to the left. As we will show below, the 2020 EU-BAU mix lies above the efficient frontier.

Although an infinite number of possible generating portfolios lie on the efficient frontier we focus on four typical optimal mixes *P, N, S, Q*. Taking the 2020 EU-BAU mix as the benchmark, they are defined as follows:

- Mix *P* is a high-cost/low-risk portfolio. It is usually the most diverse (see, for example, Stirling 1996 and Awerbuch *et al.* 2006).

- Mix *N* is an equal-cost/low-risk portfolio, that is, it is the mix with the lowest risk for costs equal to that of the 2020 EU-BAU mix.
- Mix *S* is an equal-risk/low-cost portfolio, that is, it is the mix with the lowest costs for a risk equal to that of the 2020 EU-BAU mix.
- Mix *Q* is a low-cost/high-risk portfolio. It is usually the least diverse portfolio.

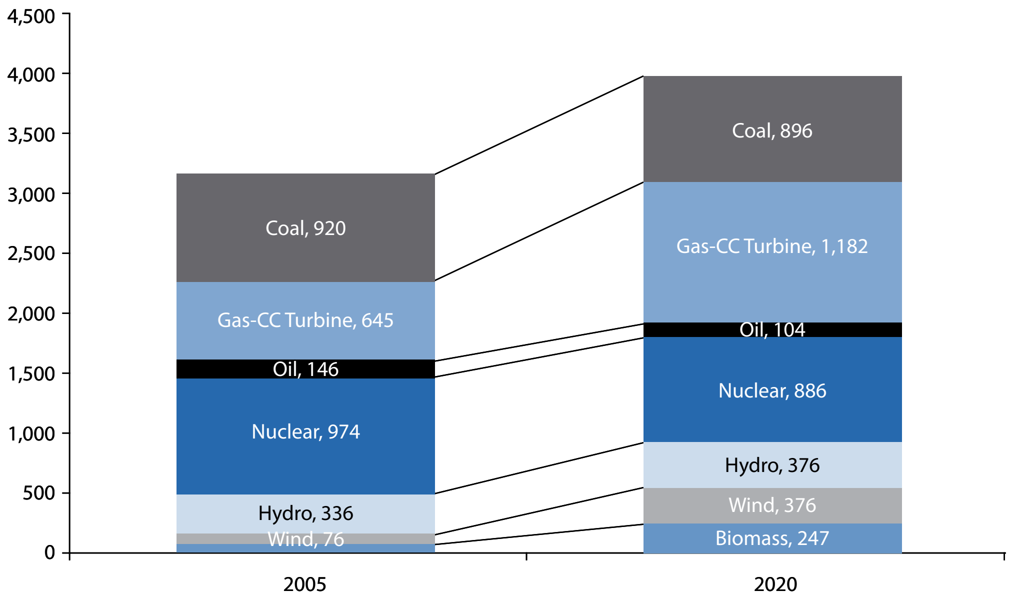
The portfolio analysis does not advocate any particular generating mixes, but rather displays the risk-cost trade-offs across many mixes. Although it may turn out that solutions in the region of the 2020 EU-BAU mix – for example, solutions between portfolios *N* and *S* – may be the most practical, we do not claim that our optimisation results help set technology targets for 2020. Rather, the idea is to highlight and quantify the trade-offs between generating mixes.

4.2 Efficient multi-technology electricity portfolios for 2020 – results

An aggressive technology deployment would likely be difficult to attain in practice …

The portfolio optimisation evaluates the 2020 EU-BAU mix shown in Figure 6 against two cases: 'Baseline' and 'Realisable' case. These cases differ in the extent to which future technology choices are constrained because of upper (and lower) bounds, representing either maximum attainable deployment levels for each technology or maximum resource limits, as in the case of renewables such as wind or hydro (see Awerbuch and Yang 2007 for a more detailed discussion). The Baseline represents aggressive technology deployment levels that would likely be difficult to attain in practice. Its purpose is to help explore practical policy limits and identify policies that may be worth pursuing. The Realisable case, however, represents a set of upper technology limits that could be attained in practice given sufficiently focused policies and accelerated resource deployments. Table 4 shows Baseline case and Realisable case lower and upper limits for the share of alternative technologies in the overall generation mix. For each set of constraints, we compute efficient electricity generation mixes and analyse the level of CO_2 emissions associated with them.

Figure 6. 2005 and 2020 EU-BAU generation mix (in TWh)



Source: European Commission (2005).

Table 4. Lower and upper technology limits (in % of electricity mix)

	Baseline case		Realisable case	
	Lower limit (%)	Upper limit (%)	Lower limit (%)	Upper limit (%)
Coal	3	52	5	35
Gas-CC Old	5	16	10	16
Gas-CC New	0	50	0	20
Oil	2	8	2	5
Nuclear	15	52	15	33
Hydro	8	13	8	11
Biomass	2	22	2	13
PV	0	5	0	1
Geo	0	½	0	0
Wind-onshore	2	32	2	7
Wind-offshore	0	40	0	7

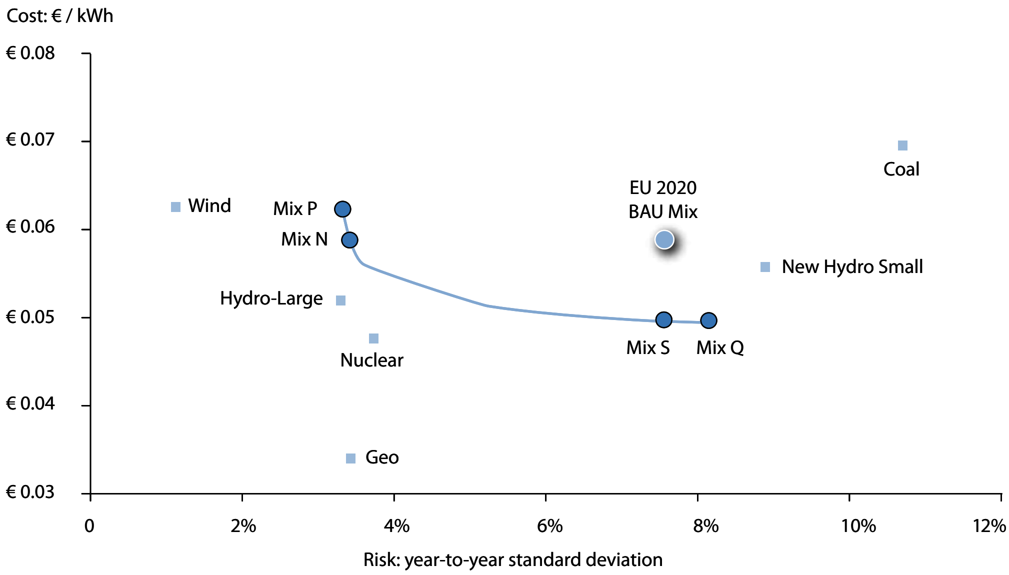
4.2.1 Efficient portfolios – Baseline case

… but if attained, it would minimise cost and risk and reduce CO_2 emissions.

This section discusses the 2020 Baseline optimisation results and compares their risk-return characteristics and CO_2 emissions to those of the projected 2020 EU-BAU mix. The results indicate that the optimal Baseline portfolios minimise cost and risk and reduce CO_2 emissions. This is shown in Figure 7, which illustrates the risk and return for the projected 2020 EU-BAU and for the typical optimised mixes under Baseline assumptions. The efficient frontier *PNSQ* shows the location of all optimal mixes.

The EU-BAU mix has an overall generating cost of 5.9 €-cents per kWh and a risk of 7.6 percent. By comparison, mix *N*, the equal-cost/low-risk mix, cuts risk nearly by half, to 3.4 percent. Alternatively, mix *S*, has the same risk as the BAU but reduces generating costs by 0.9 €-cents per kWh, which equates to an EU-wide cut in annual electricity outlays of €36 billion.[13]

Figure 7. Efficient frontier for 2020 electricity generation mix – Baseline case



Source: Own calculation.
Notes: For CO_2 cost of €35 per tonne.

13 Based on an annual consumption in 2020 of 4,006 TWh (€0.009/kWh × 4,006 × 10^9kWh = €36bn).

Mix *P*, the minimum-risk mix, reduces risk slightly relative to mix *N*. But this seems to represent an unattractive cost-risk trade-off over mix *N*. Similarly, mix *Q*, the minimum-cost mix, does hardly reduce cost relative to mix *S*, but comes with a noticeable increase in risk. It thus seems that in cost-risk terms, the practical range of policy interest generally runs from mix *N* down to mix *S*.

Policy makers tend to view climate change mitigation as an objective that is detrimental to low-cost electricity. But such beliefs are based on stand-alone cost concepts, not portfolio costs.

Policy makers tend to view climate change mitigation as an objective that competes with cost and, indeed, it is widely believed that low-carbon electricity generation will increase cost. But such beliefs are based on stand-alone cost concepts. The Baseline results, however, show that in addition to reducing cost and/or risk relative to the EU-BAU mix, the optimal mixes also reduce CO_2 emissions, in contradiction to widely held beliefs that climate change mitigation policies inevitably increase cost.[14] This is illustrated in Figure 8, which shows technology shares and portfolio risk on the left vertical axis, CO_2 emissions on the right axis, and portfolio generating cost along the top of the graph. The low-risk mixes, *P* and *N* reduce annual CO_2 to 199 million tonnes, which is 85 percent lower than emissions in the BAU mix (1,273 million tonnes of CO_2). They accomplish this primarily by substituting wind for gas and coal. Indeed, the share of onshore wind is 32 percent, its permissible upper limit (Table 5). Mixes *S* and *Q*, the low-cost mixes, reduce CO_2 emissions to 472 and 549 million tonnes, respectively, by incorporating larger shares of nuclear generation, which reaches its 52 percent upper limit in both mixes. This result – that is, that optimal low-risk mixes increase wind shares relative to the BAU while optimal low-cost mixes increase nuclear – tends to hold for the Realisable case, too, as we will see next.

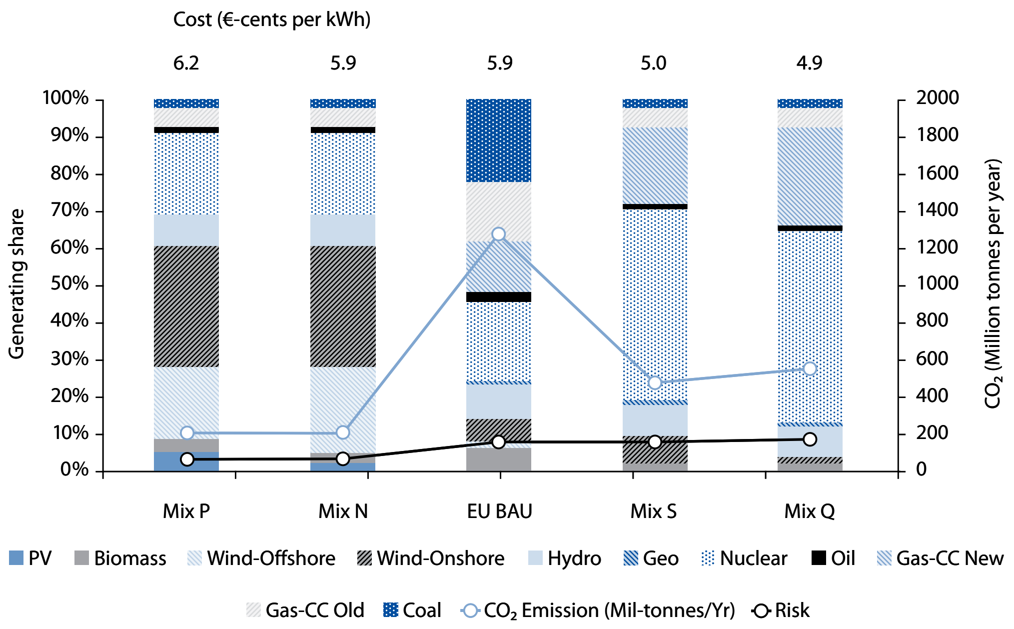
Table 5. Optimal portfolio shares and CO_2 emissions in 2020 – Baseline case

	EU-BAU	Mix P	Mix N	Mix S	Mix Q	Technology bounds	
	Share in electricity generating (%)					Lower (in %)	Upper (in %)
Coal	22	3 ^{L}	3 ^{L}	3 ^{L}	3 ^{L}	3	52
Gas-CC Old	16	5 ^{L}	5 ^{L}	5 ^{L}	5 ^{L}	5	16
Gas-CC New	13	0 ^{L}	0 ^{L}	19	27	0	50
Oil	3	2 ^{L}	2 ^{L}	2 ^{L}	2 ^{L}	2	8
Nuclear	22	22	22	52 ^{U}	52 ^{U}	15	52
Hydro	9	8 ^{L}	8 ^{L}	8 ^{L}	8 ^{L}	8	13
Biomass	6	4	3	2 ^{L}	2 ^{L}	2	22
PV	0	5 ^{U}	2	0 ^{L}	0 ^{L}	0	5
Geo	0	0	0	0	0	0	½
Wind-onshore	6	32 ^{U}	32 ^{U}	9	2 ^{L}	2	32
Wind-offshore	1	19	23	0 ^{L}	0 ^{L}	0	40
CO_2 emissions in million tonnes per year							
	1,273	199	199	472	549		

Source: Own calculation.
Notes: ^{L} and ^{U} indicate that technology share is at Lower or Upper bound; results for €35/t CO_2.

14 This is true only to the extent that the underlying generating costs shown in Figure 7 reflect all economic cost. However, since the costs shown in Figure 7 do not fully incorporate some economic costs such as investment grants that benefited some of these technologies (e.g., wind and nuclear), the resulting climate change mitigation may cost more than what Figure 7 suggests.

Figure 8. Technology shares, portfolio risk and cost, and CO_2 emissions – Baseline case



Source: Own calculation.
Notes: Results for €35/t CO_2.

4.2.2 Efficient portfolios – Realisable case

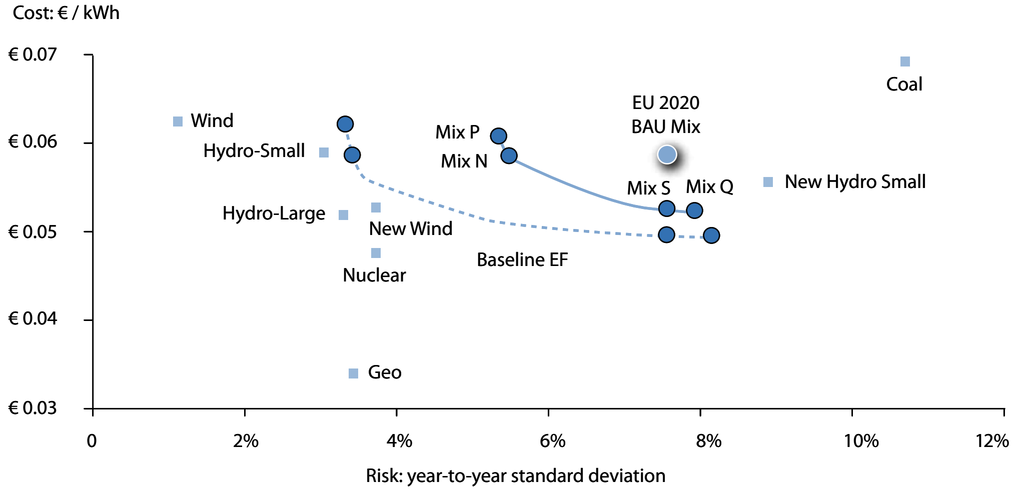
From a portfolio perspective, a less aggressive but more realistic technology deployment is costlier and riskier.

To recall, compared to the Baseline case, the Realisable case considers technology deployment levels that can be attained by 2020, assuming focused policy efforts. This case incorporates upper bounds for renewables based on the 'Realisable' scenarios developed by Ragwitz *et al.* (2005), who estimate the realisable market potential for renewable energy technologies as "the maximum achievable potential, assuming that all existing barriers can be overcome and all driving forces are active" (Ragwitz, personal communication 2006). Compared to the Baseline case, the Realisable case has less latitude to search for optimal solutions because it is limited to a smaller feasible region. As a consequence, optimal Realisable mixes are costlier and riskier, and they emit more CO_2 than optimal Baseline mixes.

Figure 9 shows the cost and risk results for the Realisable case (solid line). There are mixes on the efficient frontier that exhibit lower cost-risk than the projected EU-BAU mix. However, as the Realisable case is more constrained, the efficient frontier is shorter, riskier, and more costly relative to the Baseline. The tighter resource limits – particularly the penetration levels for onshore wind and nuclear – increase the cost of mixes *S* and *Q* and the risk of mixes *P* and *N*.

For example, the cost of mix *S* rises by 0.3 €-cents/kWh (6 percent) relative to the Baseline. This increase in cost equates to an increase in total annual outlays by EU electricity consumers of €12 billion. This figure represents about 0.1 percent of the current GDP of the EU. To illustrate the impact of tighter technology deployment limits on risk: with less wind resource available, the optimisation cannot reach the low risk levels of the Baseline. For example, in mix *N*, lower limits for wind (in particular) increase coal and nuclear shares, thereby raising risk by some 60 percent, from 3.3 percent in the Baseline case to 5.4 percent in the Realisable case.

Figure 9. Efficient frontier for 2020 electricity generation mix – Realisable case

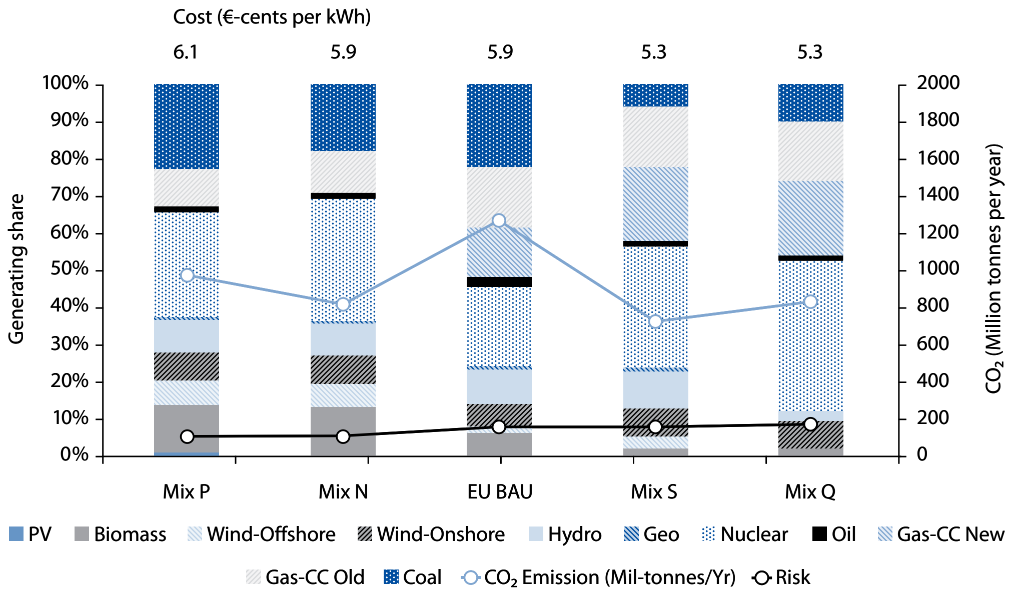


Source: Own calculation.
Notes: Results for €35/t CO_2.

Optimal shares of nuclear hit their upper limit in all optimal mixes except one.

Compared to the Baseline case, the Realisable case is characterised by significantly lower shares of nuclear, wind, and – in some cases – new and existing gas-fired power plants (see Table 6 and Figure 10). This is driven by the lower upper technology bounds for most technologies, as can be seen by comparing the right-hand column of Table 6 to that of Table 5. Further, as Table 6 indicates, wind hits its upper limit in all the optimal Realisable mixes, while offshore wind hits the upper limit in the low-risk mixes *P* and *N* where offshore wind is required to balance and complete the mix. Nuclear is at its upper limit in all except mix *P*. The results of Table 6 suggest that additional deployment of these technologies could lower cost, risk, and CO_2 emissions. As a comparison of the last row in Table 6 with the last row in Table 5 shows, the Realisable case reduces annual CO_2 emissions at best by 548 million tonnes (mix *S*) while they might fall by as much as 1,074 million tonnes under Baseline assumptions (mixes *P* and *N*).

Figure 10. Technology shares, portfolio risk and cost, and CO_2 emissions – Realisable case



Source: Own calculation.
Notes: Results for €35/t CO_2.

Table 6. Optimal portfolio shares and CO_2 emissions in 2020 – Realisable case

	EU-BAU	Mix P	Mix N	Mix S	Mix Q	Technology bounds	
	Share in electricity generating (%)					Lower (in %)	Upper (in %)
Coal	22	22	17	5 L	10	5	35
Gas-CC Old	16	10 L	11	15	16 U	10	16
Gas-CC New	13	0 L	0 L	20 U	20 U	0	20
Oil	3	2 L	2 L	2 L	2 L	2	5
Nuclear	22	29	33 U	33 U	33 U	15	33
Hydro	9	9	9	10	10	8	11
Biomass	6	13 U	13 U	2 L	2 L	2	13
PV	0	1 U	0 L	0 L	0 L	0	1
Geo	0	0 U	0 U	0 U	0 U	0	0
Wind-onshore	6	7 U	7 U	7 U	7 U	2	7
Wind-offshore	1	7 U	7 U	5	0 L	0	7
CO_2 emissions in million tonnes per year							
	1,273	981	825	725	836		

Source: Own calculation.
Notes: L and U indicate that technology share is at Lower or Upper bound; results for €35/t CO_2.

Table 7 summarises the changes in technology generating shares and CO_2 emissions for the typical optimal mixes relative to the 2020 EU-BAU. The low-risk mixes *P* and *N* show large percentage increases for nuclear, biomass, and wind, coupled with significant percentage reductions for gas, oil, and coal (in mix *N* only). The low-cost mixes *S* and *Q* show large percentage rises for gas, nuclear, and wind (in mix *S*), coupled with large reductions of coal, oil, and biomass.

The low-cost electricity mixes show large increases in the share of gas, nuclear, and wind, coupled with large reductions of coal, oil, and biomass.

Table 7. 2020 EU BAU electricity generation mix *vs*. optimal Realisable mixes

	Mix P	Mix N	EU-BAU	Mix S	Mix Q
Portfolio risk	5.3%	5.5%	7.6%	7.6%	7.9%
Portfolio cost in €/MWh	61	59	59	53	53
	% change from EU-BAU			% change from EU-BAU	
Annual CO_2	-22%	-35%	1,273m tonnes	-45%	-34%
Coal	0%	-22%	897 TWh	-78%	-57%
Gas-CC	-66%	-61%	1,182 TWh	+19%	+22%
Oil	-31%	-42%	104 TWh	-42%	-42%
Nuclear	+29%	+50%	886 TWh	+50%	+50%
Hydro	-4%	-4%	376 TWh	+12%	+8%
Biomass	+115%	+115%	247 TWh	-70%	-70%
Wind	+85%	+85%	303 TWh	+67%	+0%
Other	265%	-7%	12 TWh	-7%	-7%
Total			4,006 TWh		

Source: Own calculation.
Notes: Results for €35/t CO_2.

In practice, the move from the 2020 BAU mix to the Realisable mix *S* is probably the most attractive of the realisable possibilities. If new policies were to redirect investment so that mix *S* is achieved, this would have the highly desirable effect of cutting annual electricity costs by €24 billion[15] and CO_2 emissions by 548 million tonnes without changing risk.

However, other moves involving alternative risk choices are possible. For example, to the left of the 2020 BAU mix in Figure 9 lies mix *N*. Compared to the BAU mix, mix *N* cuts the portfolio risk by about one-third while simultaneously reducing annual CO_2 emission by 448 million tonnes, or 35 percent. This move produces no cost reductions and while CO_2 reductions are not as large as when moving from the BAU to mix *S*, risk is significantly reduced. Obviously, comparing the risk-cost and CO_2 combinations of *N* against *S* requires knowledge of societal preference functions.

Over the long run, further technology deployment may make it possible to move closer to Baseline mix *S* from the BAU mix (or the Realisable mix *S*). The decline in CO_2 emissions would be 46 percent higher (801 versus 548 million tonnes a year), accompanied by 33 percent greater cut in the EU's electricity bill (€36 compared to €24 billion).

4.3 A summary of key results

The results in this section highlight the importance of focused technology deployment policies designed to move the EU generating mix away from the BAU mix and closer to electricity generating portfolios such as the Realisable mix *S*. This mix would reduce annual EU electricity cost by around €24 billion and annual CO_2 emissions by more than 500 million tonnes. Taking annual electricity cost saving as perpetual and assuming an interest rate of 5-10 percent would justify investment today to the tune of €240-480 billion.

There seems to be a dichotomy between wind energy and nuclear energy that reinforces rather than solves the wide-ranging debate between pro-nuclear and pro-wind forces.

A key finding is that the low-risk mixes (*P* and *N*) generally reduce fossil shares and increase wind and other non-fossil shares relative to the BAU mix, while the higher-risk/lower-cost mixes (*S* and *Q*) increase primarily nuclear along with gas, wind, and hydro electricity at the expense of coal and oil. There thus seems to be a dichotomy between wind and nuclear, suggesting that our analysis reinforces rather than solves the wide-ranging debate between pro-nuclear and pro-wind forces. However, this debate incorporates numerous additional considerations that are not reflected in our optimisation, including highly uncertain waste disposal management costs. The next section tries to shed more light on the role of nuclear power and other factors influencing the results of our portfolio analysis.

5. An eclectic view on factors influencing optimal electricity mixes

5.1 The role of nuclear power

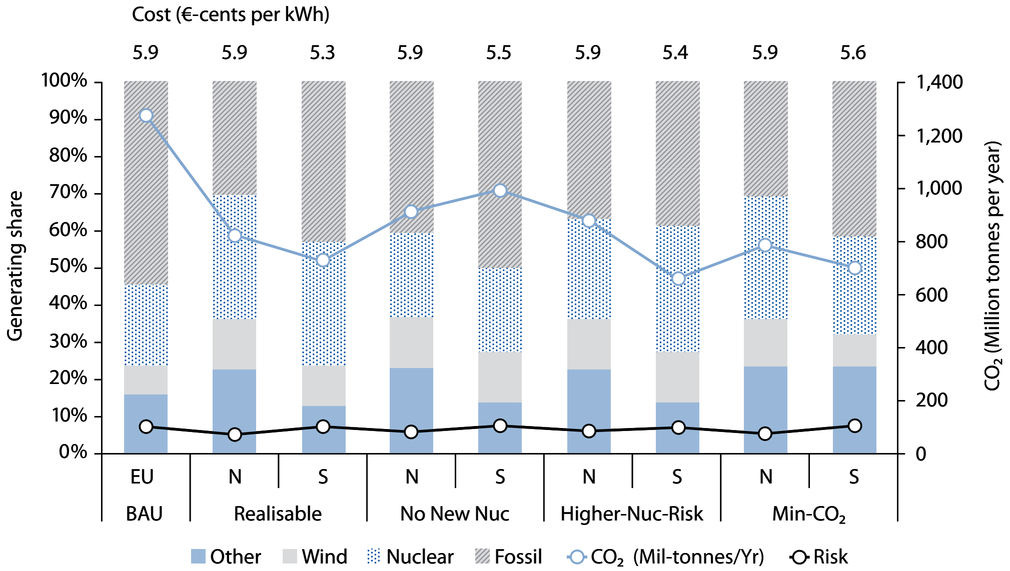
The nuclear cost estimates used for identifying efficient electricity portfolios do not account for the costs and risks of storing nuclear waste, which are essentially incalculable. CORWM (2006) recommends a lengthy, potentially decades-long process, involving interim waste storage in preparation for ultimate geological disposal. Although much of what is risky about nuclear seems to be a matter of expectations and is not necessarily always rational, countries may decide not to build new nuclear power stations – as is currently the case in Germany, for instance. Against this

15 (5.9-5.3) €-cents/kWh x 4,006TWh.

background, it is useful to test a policy of a nuclear moratorium – that is, no new nuclear – to see its effects on cost and risk of the EU portfolio mix. In principle, this can be done for the Baseline case and the Realisable case, but in what follows we will focus on the latter (for ease of comparison, we will call it the 'benchmark' Realisable case). In addition, we concentrate on generating mixes *N* that minimise portfolio risk for the cost of the 2020 EU-BAU mix and on mixes *S* that minimise portfolio cost for the risk of the 2020 EU-BAU mix.

As Figure 11 shows, for mix *S* cost rises from 5.3 €-cents to 5.5 €-cents per kWh. For Mix *N*, risk stays approximately unchanged. The big change is in terms of additional CO_2 emissions, where CO_2 emissions rise from 725 to 993 million tonnes (Mix *S*) and from 825 to 912 million tonnes (mix *N*). This is because for these portfolios, a good part (in mix *N* virtually all) of the drop in the share of nuclear is compensated for by fossil fuel-fired electricity generation.

Figure 11. Technology shares, portfolio risk and cost, and CO_2 emissions – sensitivity analyses



Source: Own calculation.
Notes: Results for €35/t CO_2.

In another sensitivity test, we have examined the impact of a change in risk of constructing and decommissioning nuclear power plants. To recall from Section 3, total generating costs of new nuclear power stations have been estimated at 4.1 €-cents/kWh, including decommissioning costs equivalent to 70 percent of the overnight plant construction cost of €1,710 per kW (see Figure 3). This makes nuclear attractive relative to other alternatives. It can be argued, however, that nuclear risk is understated because construction-period risk was arbitrarily set to the World Bank estimate for the construction-period risk of coal at 23 percent (Bacon *et al.* 1996). To account for this, we re-ran our scenarios several times, gradually increasing nuclear construction risk from 0.23 to 0.38. This raises total technology risk for nuclear from about 16 percent (see Figure 4) to 26 percent.

A higher level of construction cost risk for nuclear has a relatively small effect on the optimal cost-risk combination.

As can be seen from Figure 11, for the Realisable case ('Higher-Nuc-Risk'), a higher risk level for nuclear capital costs has a relatively small effect on the optimal cost-risk combination, that is, mix *N* comes with only a marginal increase in risk relative to the benchmark Realisable case, while mix *S* is associated with only a small increase in portfolio generating costs (5.4 €-cents/kWh instead of 5.3 €-cents/kWh). As expected, both portfolios have a lower share of nuclear – but the change is small because of the already tight upper and lower bounds for most technologies. It is interesting to observe that for the low-risk mix *N*, the share of renewables is virtually constant, with an increase in fossils making up for the drop in nuclear. As a result, CO_2 emissions rise. As for the low-cost mix *S*, the decline in nuclear is associated with a decline in fossils and an increase in renewables, all in all resulting in lower CO_2 emissions. The main reason why renewables become more important in mix *S*, but not in *N*, is that in the benchmark Realisable mix *S*, renewables – biomass in particular – are not as close to their technology upper bounds, whereas they are in the benchmark Realisable mix *N*.[16]

5.2 Efficient electricity portfolios that minimise CO_2 emissions

It is straightforward to illustrate that minimising CO_2 emissions is most likely to be economically inefficient.

We now turn to something that is not so much a sensitivity analysis, but – rather – a change in perspective: we want to identify the combinations of portfolio risk and portfolio generating cost (and the associated technology shares) that minimise CO_2 emissions. For the Realisable case, the results are shown on the very right-hand side of Figure 11. Comparing them to the benchmark Realisable case suggests only a moderate decline in CO_2 emissions: from 825 million tonnes per year to 782 million tonnes for mix *N* and from 725 million tonnes to 700 million tonnes for mix *S*. It is straightforward to illustrate that minimising CO_2 emissions is most likely to be economically inefficient. As Figure 11 shows, for mix *S*, portfolio generating cost increase by 0.3 €-cents/kWh, implying an increase in annual electricity cost of €12 billion and, thus, carbon reduction cost of €480/t CO_2 – a value way above current estimates of global warming damages.

Although not shown in Figure 11, results are very different when taking the Baseline case rather than the Realisable case as a benchmark. As shown in Awerbuch and Yang (2007), moving to the carbon-minimising mix *S* would cut CO_2 emissions by 273 million tonnes, implying carbon reduction cost of €44/t CO_2. Awerbuch and Yang (2007) also show that the risk-cost characteristics of the Baseline carbon-minimising portfolios are very similar to – in fact, slightly better than – those of the Realisable case shown in Figure 9 above. Though it is unlikely that Baseline technology penetration levels could be attained by 2020, this illustrates the significant benefits that could be achieved over a longer period by pursuing deeper penetrations of these technologies.

5.3 The effect of upper limits on technology shares

In Awerbuch and Yang (2007) we investigate in a more rigorous way the economic cost of the constraints that prevent the share of wind, nuclear, and gas to be larger than the upper limit of the Realisable case. Using linear-programming techniques, we show that easing these constraints and, thus, allowing technology shares to move towards the Baseline case, has considerable economic value. More specifically, for the realisable mix *S* we find that increasing the upper limit for the share of nuclear energy by 1 percentage point would result in portfolio cost savings equivalent to 46 percent of the lifetime generating costs of additional nuclear power stations. The comparable results for wind and gas are 21 percent and 8 percent, respectively. The results for wind could significantly and positively impact the current debate regarding development of an EU offshore 'super-grid' to connect diverse offshore wind sites. They also impact on the nuclear debate in a similar fashion.

16 A word of caution is appropriate. The sensitivity of results to changes in underlying assumptions about nuclear energy do not, and are not intended to, resolve the nuclear-renewables debate. Rather, they are meant to quantify and highlight some of the important factors.

All in all, they indicate that failure to fully exploit the EU energy resource potentials needlessly raises generating cost and CO_2 emissions.

5.4 The effect of pricing CO_2 emissions

So far, our analysis assumed a charge of €35 per tonne of CO_2 emitted, which we interpreted as either a market price or a shadow price for carbon emissions. We will now investigate the effect of pricing CO_2 emissions on the cost-risk characteristics of the 2020 EU-BAU mix and of efficient generating portfolios. In addition, we discuss the impact of carbon pricing on CO_2 emissions. To keep things simple, we consider only the effect of moving from a carbon price of zero to one of €35/t CO_2 and we concentrate on the BAU mix and mixes *N* and *S* in the Realisable case.[17]

As Figure 12 illustrates, portfolio risks and costs rise with rising CO_2 prices. This is true for the BAU mix and the efficient electricity generating portfolios. For instance, the cost of the BAU mix increases by 23 percent or 1.1 €-cent per kWh (from 4.8 €-cents to 5.9 €-cents per kWh). The risk of that mix, however, rises a whopping 40 percent (from 5.4 percent to 7.6 percent), illustrating its considerable sensitivity to changing CO_2 (and fossil fuel) prices. By definition, the share of each technology in the BAU mix and, thus, CO_2 emissions do not change with a rise in CO_2 prices. Clearly, it makes little sense to keep technology shares constant when CO_2 prices rise.

On the contrary, with rising CO_2 prices it is optimal to reduce the share of fossil fuels in electricity generation – as indicated by the amount of CO_2 emissions, which is shown by parenthetical values next to the mixes in Figure 12. Since mixes *P* and *N* have lower shares of fossils than mixes *S* and *Q*, they have lower emissions at any given CO_2 price. Absent CO_2 charges, the Realisable mix *N* emits 1,358 million tonnes of CO_2 per year.[18] As the CO_2 price increases, optimal mixes are re-shuffled to minimise portfolios costs and risks. For a carbon price of €35/t CO_2, emissions fall by almost 40 percent to 825 million tonnes per year.

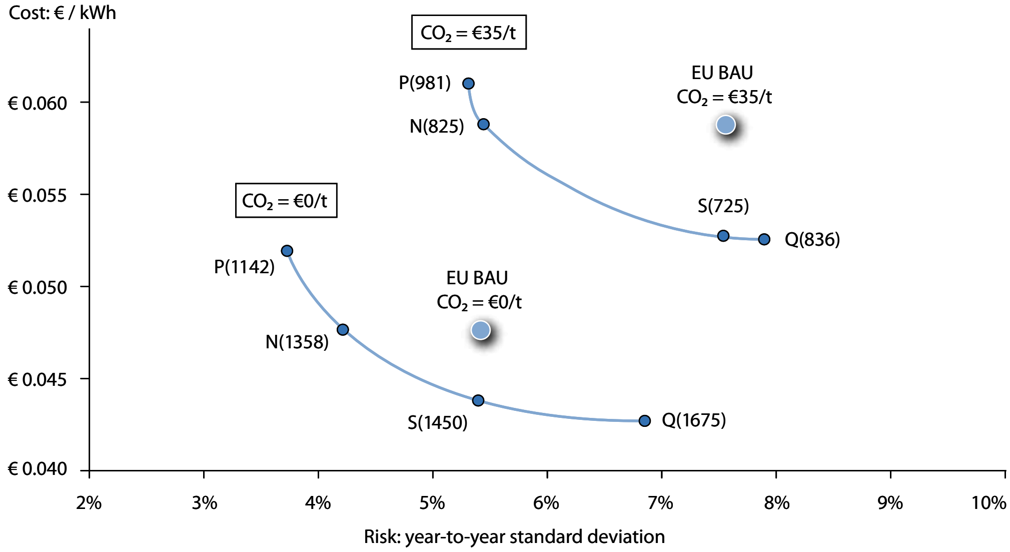
For a carbon price rise from €0 to €35/t CO_2, emissions fall by almost 40 percent to 825 million tonnes per year.

Let us take a closer look at the effect of carbon pricing by considering mix *S*. In general, the change in portfolio costs and CO_2 emissions is the result of two interrelated changes: a rise in CO_2 charges and the re-optimisation of portfolio mixes in response to this rise. Considered in isolation, the increase in the CO_2 price raises the cost of electricity from 4.4 €-cents/kWh (see Figure 12) by about 1.3 €-cents/kWh. This increase reflects the cost of carbon (€35/t CO_2 multiplied by 1,450 million tonnes of CO_2) for a total electricity production of around 4,000 TWh. But as pictured in Figure 12, portfolio generating cost increase only by around by 0.9 €-cents/kWh to a total of 5.3 €-cents/kWh. The cost savings of around 0.4 €-cents/kWh are due to the portfolio re-optimisation triggered by the pricing of carbon. But the associated decline in the share of fossil fuels in mix *S* not only offsets, in part, the increase in electricity costs resulting from the pricing of carbon, it also lowers CO_2 emissions from 1,450 million tonnes to 725 million tonnes.

17 Results for carbon prices between zero and €35/t CO_2 and for other efficient generating mixes (in both the Realisable case and the Baseline case) are discussed in Awerbuch and Yang (2007).

18 It is worth pointing out that without carbon pricing, efficient portfolios that generate electricity at the same or lower cost than the BAU mix are more carbon intensive than the BAU mix (see the points that lie to the southeast of mix N on the 'CO_2 = €0' efficient frontier in Figure 12).

Figure 12. Efficient frontiers (€0/tCO_2 and €35/tCO_2) for 2020 electricity generation mix – Realisable case



Source: Own calculation.

Notes: Values in parentheses next to the mixes show annual CO_2 emissions in million tonnes. The 2020 EU-BAU emits 1,273 million-tonnes per year.

6. Summary and conclusions

Our analysis suggests that greater shares of non-fossil technologies, primarily nuclear or wind, can help reduce the cost and risk of the EU generating portfolio as well as its CO_2 emissions.

This paper has presented a mean-variance portfolio optimisation analysis that develops and evaluates optimal (that is, efficient) EU electricity generating mixes for 2020. The results suggest that greater shares of non-fossil technologies, primarily nuclear or wind, can help reduce the cost and risk of the EU generating portfolio as well as its CO_2 emissions. To illustrate, an efficient generating mix that we consider to be achievable by 2020 is estimated to cut annual EU electricity generating cost by €24 billion and CO_2 emissions by 548 million tonnes. This mix thus produces perpetual annual benefits sufficient to justify current investments of up to €500 billion – which compares to an estimated EU investment of €900 billion in new electricity generation capacity needed by 2030. It is also shown that easing constraints on investment in nuclear and wind energy capacity would lower overall generating cost enough to offset 46 percent and 21 percent of the kWh costs of nuclear and wind generation. Against this background, policies designed to accelerate the deployment of key non-fossil technologies appear to be highly cost-effective.

Perhaps the single most important lesson of the portfolio optimisation analysis is that adding a fuel-less, fixed-cost technology (such as wind energy) to a risky generating mix lowers expected portfolio cost at any level of risk, even if the fuel-less technology costs more when assessed on a stand-alone basis. This underscores the importance of policy-making approaches grounded in portfolio concepts as opposed to stand-alone engineering concepts.

This is a tall order, since quantitative indicators in energy markets are primarily focused on stand-alone performance. In contrast, financial markets provide a *beta* measure to help investors think in terms of portfolio performance. The lack of a similar measure in energy markets prevents some from embracing the energy planning portfolio optimisation approach.

Ironically this issue is akin to the practical problems that initially confronted Harry Markowitz's portfolio approach. The new technique required massive analytic efforts (*sans* computers) to estimate the covariance of returns to each stock in the US market against every other stock. It was not until Sharpe and Lintner developed the Capital Asset Pricing Model (CAPM) to show that a single covariance with the market portfolio is sufficient (Varian 1993). Perhaps with further research, it may be possible to develop energy analogues that will enable a *beta* type measure to index the risk of particular generating technologies against a large generating mix such as the EU mix. This would provide a simple and expedient method for evaluating the costs and risks of individual technologies and their CO_2 emissions.

Today's dynamic and uncertain energy environment requires portfolio-based planning procedures that reflect market risk and de-emphasise stand-alone generating costs. Portfolio theory is well tested and ideally suited to evaluate electricity expansion strategies.[19] It identifies solutions that enhance energy diversity and security and are therefore considerably more robust than arbitrarily mixing technology alternatives. Portfolio analysis reflects the cost-risk relationship (covariances) among generating alternatives. Though crucial for correctly estimating overall cost, electricity-planning models universally ignore this fundamental statistical relationship and instead resort to sensitivity analysis and other ill-suited techniques to deal with risk. Sensitivity analysis cannot replicate the important cost inter-relationships that dramatically affect estimated portfolio costs and risks (Awerbuch 1993), and it is no substitute for portfolio-based approaches. The mean-variance portfolio framework offers solutions that enhance energy diversity and security and are therefore considerably more robust than arbitrarily mixing technology alternatives.

Today's dynamic and uncertain energy environment requires portfolio-based planning procedures that reflect market risk and de-emphasise stand-alone generating costs.

This being said, we must be clear about the purpose and the limitations of the portfolio approach to electricity sector planning. The portfolio optimisation presented in this paper does not point to a specific capacity-expansion plan. Such outputs would require considerably more detailed models. The results presented here are largely expositional, but they demonstrate the value of portfolio optimisation approaches and suggest that capacity planning made on the basis of stand-alone technology costs will likely lead to economically inefficient outcomes.

Moreover, in deregulated markets, individual power producers evaluate only their own direct costs and risks when taking investment decisions. These decisions do not reflect the effects the producers' technologies may have on overall generating portfolio performance. Wind investors, for example, cannot capture the risk-mitigation benefits they produce for the overall portfolio, which leads to under-investment in wind relative to levels that are optimal from society's perspective. Similarly, some investors may prefer the risk menu offered by fuel-intensive technologies such as combined-cycle gas turbines, which have low initial costs. Given sufficient market power, gas generators may be able to externalise fuel risks onto customers. In effect, these investors do not bear the full risk they impose onto the generating mix, which may lead to over-investment in gas relative to what is optimal from a total portfolio perspective (a quantitative treatment of this issue is given in Roques 2006). All this suggests a rationale for economic policies in favour of technologies that bring diversification benefits.

19 Other techniques have also been applied. For instance, Stirling (1996, 1994), develops maximum-diversity portfolios based on a considerably broader uncertainty spectrum. Though radically different in its approach, his diversity model yields qualitatively similar results.

Annex

Table A1. Fuel cost inputs and economic cost of CO_2.

Gas	€4.8/Mbtu
Oil	€41/bbl
Coal	€44/tonne
CO_2	€35/tonne
Uranium	€6/MWh
Biomass	€5.15/GJ

Table A2. O&M correlation coefficients

Technology	Coal	Gas	Nuclear	Oil	Hydro	Wind	Geo	Solar	Bio
Coal	1.00	0.25	0.00	-0.18	0.03	-0.22	0.14	-0.39	0.18
Gas	0.25	1.00	0.24	0.09	-0.04	0.00	-0.18	0.05	0.32
Nuclear	0.00	0.24	1.00	-0.17	-0.41	-0.07	0.12	0.35	0.65
Oil	-0.18	0.09	-0.17	1.00	-0.27	-0.58	-0.06	-0.04	0.01
Hydro	0.03	-0.04	-0.41	-0.27	1.00	0.29	-0.08	0.30	-0.18
Wind	-0.22	0.00	-0.07	-0.58	0.29	1.00	-0.28	0.05	-0.18
Geo	0.14	-0.18	0.12	-0.06	-0.08	-0.28	1.00	-0.48	-0.70
Solar	-0.39	0.05	0.35	-0.04	0.30	0.05	-0.48	1.00	0.25
Biomass	0.18	0.32	0.65	0.01	-0.18	-0.18	-0.70	0.25	1.00

References

Awerbuch, S. (1993). "The surprising role of risk and discount rates in utility integrated-resource planning". *The Electricity Journal*, (6:3), pp. 20-33.

Awerbuch, S. (1995). "New economic cost perspectives for valuing solar technologies", in Böer, K.W. (ed.), *Advances in solar energy: An annual review of research and development.* (10). ASES, Boulder, USA.

Awerbuch, S. (1995a). "Market-based IRP: It's easy!". *Electricity Journal*, (8:3), pp. 50-67.

Awerbuch, S. (1996). "Capital Budgeting, Technological Innovation and the Emerging Competitive Environment of the Electric Power Industry," (with J. Dillard, T. Mouck, and A. Preston), Energy Policy, Vol. 24, No. 2, 195-202.

Awerbuch, S. (2000). "Investing in photovoltaics: Risk, accounting and the value of new technology". *Energy Policy*, Special Issue, (28:14), pp. 1023-1035.

Awerbuch, S. (2000a). "Getting it right: The real cost impacts of a renewables portfolio standard". *Public Utilities Fortnightly*, (February 15).

Awerbuch, S. (2004). "Restructuring our electricity networks to promote decarbonization: Decentralization, mass customization and intermittent renewables in the 21st century". Tyndall Centre Working Paper No. 49. www.tyndall.ac.uk/publications/working_papers/wp49.pdf

Awerbuch, S. (2005). "Portfolio-based electricity generation planning: Policy implications for renewables and energy security". *Mitigation and Adaptation Strategies for Global Change,* in-press.

Awerbuch, S. and Berger, M. (2003). "Energy security and diversity in the EU: A mean-variance portfolio approach". IEA Report Number EET/2003/03. http://library.iea.org/dbtw-wpd/textbase/papers/2003/port.pdf

Awerbuch, S., Beurskens, L., Drennen, T., and Jansen, J.C. (2005). "The cost of geothermal energy in the western US region: A portfolio-based approach: A mean-variance portfolio optimization of the region's generating mix to 2013". Sandia Report SAND 2005-5173.

Awerbuch, S. and Sauter, R. (2006). "Exploiting the oil-GDP effect to support renewables deployment". *Energy Policy* Volume 34, Issue 17, November 2006, pp. 2805-2819.

Awerbuch, S., Stirling, A., and Jansen, J. (2006). "Portfolio and diversity analysis of energy technologies using full-spectrum risk measures", in: Bodde, D. and Leggio, K. (eds.) *Understanding and managing business risk in the electric sector.* Elsevier, forthcoming.

Awerbuch, S. and Yang, S. (2007). "Efficient electricity generating portfolios for europe: Maximizing energy security and climate change mitigation," SPRU Working Paper, (forthcoming).

Bacon, R.H., Besant-Jones, J.F., and Heidurian, J. (1996). "Estimating construction costs and schedules: Experience with power generation projects in developing countries". World Bank Technical Paper No. 325.

Bar-Lev, D. and Katz, S. (1976). “A portfolio approach to fossil fuel procurement in the electric utility industry”. *Journal of Finance*, (31:3), pp. 933-47.

Berger, M. (2003). *Portfolio analysis of EU electricity generating mixes and its implications for renewables*. *Ph.D. Dissertation*, Technische Universität Wien, Vienna.

Brealey, R.A. and Myers, S.C. (2004). *Principles of corporate finance*. McGraw Hill, New York, USA.

Committee On Radioactive Waste Management, CORWM. (2006). “Managing our Radioactive Waste Safely”. July.

Dale, L., Milborrow, D., Slark, R., and Strbac, G. (2004). “Total cost estimates for large-scale wind scenarios in UK”. *Energy Policy,* Volume 32, Issue 17, November 2004, pp. 1949-1956

European Commission. (2001). “Towards a European strategy for the security of energy supply”. Green Paper.

European Commission, EREC. (2005). “European energy and transport: Trends to 2030 - update 2005. SHP Potential”. http://www.erec-renewables.org/sources/hydro.htm

European Commission. (2006). “Commission Staff working Document”. Annex to the ‘Green Paper: A European strategy for sustainable, competitive and secure energy. What is at stake’. Background document SEC(2006) 317/2.

European Commission. (2007). “An energy policy for Europe”. Communication from the Commission to the European Council and the European Parliament, Brussels, 10.1.2007. Com(2007) 1 Final {Sec(2007) 12}.

Fabozzi, F., Gupta, F., and Markowitz, H. (2002). “The legacy of modern portfolio theory”. *The Journal of Investing,* (Fall), pp. 7-22.

Fox, B. and Flynn, D. (2005). “Managing intermittency of wind generation with heating load control”, in collaboration with P. O’Kane. The Queen’s University of Belfast.

German Energy Agency, DENA. (2005). “Grid Study” Presented in detail at the international ‘Conference on the integration of wind energy into the German Electricity Supply’, on May 10, 2005, in Berlin. http://www.ewea.org/fileadmin/ewea_documents/documents/publications/briefings/Dena_Study.pdf

Green, R. (2006). “Carbon tax or carbon permits: the impact on generators’ risks”. Institute for Energy Research and Policy, University of Birmingham.

Grubb, M., Butler, L., and Twomey, P. (2007). “Diversity and security in UK electricity generation: The influence of low carbon objectives”. *Energy Policy*, in press.

Harlow, W.V. (1991). “Asset allocation in a downside-risk framework”. *Financial Analysts Journal,* (47:5), pp. 29-40.

Huber C., Haas R., Faber T., Resch G., Green J., Twidell J., Ruijgrok W., and Erge T. (2001). “Final report of the project ElGreen”. Energy Economics Group (EEG), Vienna University of Technology, Austria.

Humphreys, H.B. and McLain, K.T. (1998). “Reducing the impacts of energy price volatility through dynamic portfolio selection”. *Energy Journal*, (19:3).

International Energy Agency, IEA. (2006). *World Energy Outlook 2006*. IEA, Paris, France.

Jansen, J.C., Beurskens, L.W.M., and van Tilburg, X. (2006). “Application of portfolio analysis to the Dutch generating mix: Reference case and two renewables cases: year 2030 - SE and GE scenario”. Petten: ECN, February.

Mario Ragwitz, Gustav Resch, Claus Huber, and Thomas Faber (2005). “Potential of Renewable Energy Technologies”. Fraunhofer Institute, Presented at Renewable Energy For Europe - Research In Action, Brussels, November 2005.

Roques, F.A. (2006). “The option value of diversity.”, in Grubb M., Jamasb, T., and Pollitt, M. (eds), *Delivering a low carbon electricity system: Ttechnologies, economics and policy.* Cambridge University Press, Cambridge, UK.

Sinden, G. (2005). Wind power and the UK wind resource. Oxford, UK: Environmental Change Institute, http://www.eci.ox.ac.uk/publications/downloads/sinden05-dtiwindreport.pdf

Stirling, A.C. (1996). “On the economics and analysis of diversity”. Science Policy Research Unit (SPRU) University of Sussex, Paper No. 28. www.sussex.ac.uk/spru.

Stirling, A.C. (1994). “Diversity and ignorance in electricity supply - Addressing the solution rather than the problem”. *Energy Policy*, (22:3), pp. 195-216.

UK Energy Research Centre, UKERC. (2006). “The costs and impacts of intermittency: An assessment of the evidence”. Imperial College of London.

Varian, H. (1993). “A Portfolio of Nobel Laureates: Markowitz, Miller and Sharpe”. *The Journal of Economic Perspectives*, (7:1), pp. 159-169.

ABSTRACT

This paper analyses the welfare effects of two policies directed at the security of energy supply: investments in strategic petroleum reserves and a cap on the production of gas from the largest Dutch gas field. *Market failures can justify such policies, in particular failure of individual consumers to account for their impact on energy prices and import dependency and, hence, the vulnerability of a country to geopolitical conflicts. But as the costs of investing in strategic reserves and capping gas production are not negligible, these options are welfare enhancing only in specific circumstances. Generally, measures to improve the functioning of energy markets promise to achieve more than investment-intensive measures or those restricting options of profit-maximising agents. However, policy makers might find it politically expedient to adopt rather than reject inefficient security-of-supply policies.*

Machiel Mulder (m.mulder@ce.nl) is Head of the Economics Division of CE Delft. **Arie ten Cate** (a.ten.cate@cpb.nl) and **Gijsbert Zwart** (g.t.j.zwart@cpb.nl) are researchers at CPB Netherlands Bureau for Economic Policy Analysis. Gijsbert Zwart is also affiliated to Tilburg Law and Economics Center (TILEC) at Tilburg University. This paper draws partly on De Joode *et al.* (2004), Mulder and Zwart (2006) and CPB (2006). The authors thank Douwe Kingma (CPB) for his valuable comments.

The economics of promoting security of energy supply

Machiel Mulder

Arie ten Cate

Gijsbert Zwart

1. Introduction

The security of energy supply has become of increasing concern in recent years. In Europe, important factors contributing to this are the liberalisation of European energy markets and Europe's growing dependence on oil and gas imports from politically less stable countries. As for the debate about policies aimed at ensuring supply security, two different perspectives can be distinguished: a political and an economic one (see CPB 2006).

From a political viewpoint, ensuring security of supply often means that a stable supply of energy needs to be guaranteed at 'affordable' prices, regardless of the circumstances (see, for instance, European Commission 2000). The 2005 Directive on Security of Supply (European Commission 2005, Article 2) states: "security of electricity supply means the ability of an electricity system to supply final customers with electricity, (...) the satisfaction of foreseeable demands of consumers to use electricity without the need to enforce measures to reduce consumption." This definition reflects pre-liberalisation goals as it takes demand as an exogenous factor.[1]

From an economic viewpoint, however, the concept of security of supply is related to the efficiency of providing energy to consumers. Markets will always show variations in supply and demand and, hence, in prices. A reduction in supply puts upward pressure on prices, thereby curbing demand, and an increase in demand raises prices and thus encourages an increase in supply. From an economic perspective, the question of whether or not the degree of supply security is optimal can be rephrased as whether or not the market succeeds in achieving an efficient balancing of supply and demand in the short run and an efficient level of investment in the long run. Efficiency requires that short-run prices fluctuate to reflect changing supply and demand conditions and that market clearing may involve reduction of demand.

The two perspectives, therefore, lead to conflicting goals since from an economic point of view, supplying all demand is bound to be inefficient and prices will have to fluctuate to clear the markets. In this paper, we approach the issue of security of supply from the economic perspective. In Section 2, we explore how energy markets deal with risks and disturbances and we consider possible market failures that could justify government intervention. Section 3 and Section 4 examine the economics of two policy options meant to increase security of supply: investing in strategic petroleum reserves and conserving domestic gas fields, respectively. The final section concludes by answering the question whether such policy measures are efficient responses to risks in the supply of energy.

2. Security of energy supply: market response and market failures

2.1 Market responses to changing supply conditions: historical evidence

Over the past decades, energy markets experienced several disturbances, triggering a response by market participants. To illustrate, we mention a few events in the gas market that demonstrate how energy markets have dealt with changing supply conditions and, more generally, supply risks.

1 That said, the Directive also stresses the importance of removing barriers that prevent the use of interruptible contracts and adopting the use of real-time demand management technologies.

In New Mexico, one of three parallel interstate gas pipelines blew up in August 2000, causing the other two to temporarily shut down. For several weeks, this resulted in a 60-percent drop in the flow of gas (normally 2 billion cubic feet per day) from El Paso to the gas markets of Arizona and California. An EIA study (EIA 2000) into the effects of this disruption concluded that the markets were independently able to adjust and avoid severe gas shortages. This was accompanied by soaring gas prices – at least temporarily. According to the EIA (2000), the system stability was secured by several measures, such as alternate transportation, gas from storage and switching to other fuels to supplement the loss of natural gas supplies.

High prices and high volatility may be viewed as an unsatisfactory but inevitable consequence of liberalised energy markets.

In the United States, regulatory reforms to liberalise the gas market took place in the last three decades, with an initial deregulation of producer prices in response to supply shortages in the 1970s, which were caused by the regulatory price controls then in place (IEA 2004). Gas prices initially dropped during the 1980s and early 1990s as a result of the ensuing gas surplus. This trend was reversed in recent years, however, leading the IEA (2004) to characterise the market as being in a price crisis, meaning that there was no interruption of supply but that the available supply became more expensive. In response to increasing prices, demand fell significantly, both from electricity generators (which switched to alternative fuels) and from industry (which partly relocated production outside North America). On the supply side, capacity was expanded at existing LNG terminals, but many new terminals were planned too. The IEA concluded that – in terms of balancing supply and demand – security of supply was not expected to be a problem in the United States. The IEA expected that the prices would be able to balance supply and demand in North America, both in the short and the long term. However, high prices and high volatility may be viewed as an unsatisfactory consequence of this price mechanism.

The United Kingdom, too, has experienced a prolonged period of low gas prices during the 1990s after liberalisation and the introduction of competition in the first half of this century. Newbery (2000) describes how the abandonment of the centralised role of British Gas led producers to market their gas independently, bringing more gas to the market and leading to sustained low prices for many years because of the overcapacity that had been built up before liberalisation. Moreover, after connecting the British gas market to the continental European gas markets through the Zeebrugge-Bacton interconnector, UK prices typically remained below continental prices, resulting in exports to the continent for most of the time. Recently, this situation has changed, as in the United States. While UK prices remained below continental prices in the summer, some observers argued that the UK gas market was facing a potential shortage in the winters of 2005/06 and 2006/07. In 2005, forward prices for gas deliveries in January 2006 rose steeply to a peak of around 0.5 €/m^3, before tumbling back to some 0.35 €/m^3 (which was still more than twice average spot prices). Global Insight (2005) notes that within a period of a few years, supply and demand will be very tightly balanced. Both winters so far turned out to be relatively mild, except for some isolated cold spells. According to Global Insight, in spite of a drop in imports from the continent during some of these periods, the UK system managed well with the tightness on the market (Global Insight 2005). This was in part attributed to demand responding well to higher prices. Several authors have estimated the possible demand response at almost 10 percent of peak-day consumption. Moreover, in response to the higher gas prices, a large number of investment projects (LNG import terminals and new pipelines) were initiated – some of them have already been realised, others are expected to come on stream before 2008. Global Insight expects that the United Kingdom will see a significant surplus of import capacity by the end of the decade, notwithstanding the fact that the indigenous production continues to decline.

The lessons from these cases are that liberalised markets appear to ensure a balance between supply and demand and to provoke new investments in response to the expectation of higher prices. But it is also true that liberalised markets imply a more substantial demand response to temporary shortages than regulated markets. Short-term prices, in particular, will need to exhibit occasional spikes to trigger a market-clearing demand response. In the absence of market failures, market participants anticipate that disruptions can occur. In addition, prices internalise the impact of the reaction of each energy consumer to tight markets on other consumers. High prices during periods of shortages trigger investments aimed at efficiently managing such shortages. In a perfectly functioning market, if an incident occurs that reduces the supply of energy to a region, market participants immediately respond to the increase in price. Traders will deliver energy from elsewhere to the market while consumers will reduce their consumption, or both. Consumers who bought energy under long-term contracts might reduce their demand and sell the resulting surplus to other users. All in all, there are good reasons to assume that investments in production, storage, and transportation capacity together with a flexible response of energy consumers will take markets to a new equilibrium before too long.

High prices during periods of shortages trigger investments aimed at efficiently managing such shortages.

However, it cannot be denied that energy markets might be subject to market failures that hinder an optimal response to disturbances. Do both producers and consumers really respond efficiently to the challenges arising from changing energy markets? And to which extent do energy consumers take into account the cost of their consumption, in particular in relation to geopolitical risks?

2.2 Market failures

There are different types of market failures, but of particular relevance for the security of energy supply could be externalities, that is, costs or benefits that market agents ignore when making consumption or production decisions. If private costs are smaller than social costs, consumption or production will be higher than the social optimum. Bohi *et al.* (1996), for instance, view the relationship between oil consumption and imports, on the one hand, and the market power of oil-producing countries, on the other, as a clear example of a negative externality. Consumption of energy might cause an externality on security of supply as it raises the import dependence and, hence, the vulnerability of importing countries to geopolitical conflicts. As the Dutch Energy Council *Algemene Energieraad* (2005) argues, geopolitical risks will increase as future gas flows will be increasingly affected by political motives. One could imagine that foreign suppliers who are aware of the negative effects of a curtailment of gas supplies on European economies might use this knowledge to obtain political commitments by affected governments.

To some extent, incidental supply shortages due to political events are comparable to shortages due to technical supply difficulties. Efficient markets will expose market participants importing from unstable regions to higher prices because of the possibility of politically induced supply shortages. A key issue here – distinguishing it from the case of technical supply difficulties – is that the probability of an incident occurring may be endogenous. While the probability of a technical supply incident does not depend on the damage it causes, the attractiveness of politically motivated threats does change with its impact – and this impact rises with dependence on politically imperfectly reliable suppliers. Individual consumers neither take into account the impact of increased individual consumption on overall import dependency nor the price effect of an increasing import dependency. There is thus a negative consumption externality, resulting in too much consumption at too high prices compared to the welfare optimum. This might require a policy intervention, in contrast to a situation where (geo)political supply risks do not depend on the degree of import dependency (political instability in energy-exporting countries, for instance) and are reflected in the price of energy.

Besides negative externalities, positive externalities in energy markets exist too. For instance, market participants will insufficiently invest in the supply, or storage, of energy if they cannot capture the full benefits of the investments. This may happen if prices do not reflect the full scarcity value of energy, which could be a consequence of incomplete markets (such as absence of adequate short-term balancing prices), for instance. But it can also happen if investors fear that the government will interfere by imposing price caps (or creaming off the scarcity rents) when prices soar to extreme levels (see, for example, Bushnell 2005). Conversely, the implicit 'insurance' against high prices of such an intervention leads to a bias in investment towards higher risk alternatives.[2]

In well-developed energy markets, market participants internalise the risk of disturbances in supply and demand.

To conclude Section 2, in well-developed energy markets, market participants internalise the risk of disturbances in supply and demand. In practice, however, imperfect designs of markets and uncertainty about government policies in case of disturbances might result in sub-optimal decisions, such as insufficient investments in flexibility and excessive energy consumption. Too high consumption and prices might also be the result of markets not fully internalising the impact of energy consumption on dependency on a limited number of exporting countries. Market failures and imperfections of this nature might justify government intervention if the benefits of such intervention more than offset its cost. Bearing this condition in mind, we will next examine two policy options aimed at mitigating negative economic consequences of supply disturbances.

3. The economics of investing in Strategic Petroleum Reserves

3.1 Introduction

In this section we analyse the economics of investing in Strategic Petroleum Reserves (SPR). SPR are meant to intervene in the oil market when the oil price has risen sharply due to temporary[3] supply shocks.

Private investors in oil storage ignore in their decisions the benefit of the ensuing lower price for all other oil consumers; this requires government policy.[4] Moreover, the same holds at the level of entire countries: all countries participating in an SPR system contribute to the same decrease of the world oil price and the costs can be spread over the countries. This requires international cooperation.

There are in fact two official SPR systems: the system of the EU and the system of the OECD's International Energy Agency (IEA)[5]. Preceded by an OECD recommendation, the EU issued in 1968 a directive prescribing that member states should maintain strategic reserves of at least 65 days of net consumption. Again, preceded by an OECD recommendation, this level was raised to 90 days in 1972. Following the first oil crisis in 1973, the OECD established in 1974 the IEA. As from 1980, IEA participating countries are obliged to maintain an SPR of 90 days of net consumption. The two systems do not compete with each other but are complementary. Legal and procedural incompatibilities have been smoothed out. In what follows, we refer to both systems together simply as the SPR of the IEA[6].

2 Uncertainty over peak prices is sometimes mentioned as an impediment to efficient investments. Producers would not invest if they perceived the revenues to be too risky. This argument disregards the fact that not investing would be equally risky to those being short in energy: consumers or those from whom they contract energy.

3 Because of the limited size of the SPR, this instrument is of little use against a permanent price rise.

4 Note that for this to be the case, we need to assume that the negative effect of an oil price increase during the build-up of stocks is lower than the positive consumer welfare effect when stocks are released. A rationale for this might be that under normal circumstances, when stocks are built up, oil supplies would be more elastic than in the supply shock situation.

5 This paragraph draws heavily on Willenborg *et al.* (2004).

6 The IEA SPR are held by the United States, the countries of the EU-15, and by Australia, Canada, the Czech Republic, Hungary, Japan, Korea, New Zealand, Switzerland and Turkey. Poland and Slovakia are candidate IEA members. Norway is an IEA member without the SPR obligation.

The past decades have witnessed several temporary disruptions in the supply of oil, creating temporary oil price rises (see Table 1 and Figure 1). In very few cases, oil from the SPR was actually released. The last release from the IEA SPR was related to a natural cause: the Katrina hurricane in 2005.

At present, the SPR of the IEA comprises 1.4 billion barrels of oil, reflecting the required 90 days of net imports. Most of it is in government-controlled stocks. In the EU and in the United States, a unilateral increase of their own SPR has been discussed (see European Commission 2002 and Leiby and Bowman 2000, for instance). In what follows, we analyse the optimal size of the SPR from an economic point of view. We start with a sketch of the model for determining the optimal size of the SPR (Section 3.2), move on to a presentation of key parameter values used for the model (Section 3.3) and results of the model (Section 3.4), and then offer some concluding remarks (Section 3.5).

3.2 A model for optimising the size of the SPR

By computing marginal costs and marginal benefits, one can determine whether the size of the SPR is too small or too large. The optimal level of the SPR is found when marginal benefits equal marginal costs. The marginal costs per barrel of oil are simply the costs per year of buying and storing one barrel. For the sake of simplicity, we assume that these costs are independent of the size of the SPR. The marginal benefit per barrel – that is, the increase in annual benefits when the SPR is increased by one barrel – is the product of two parameters: (i) the benefit of releasing an extra barrel from the stock, thereby moderating a rise in oil prices and its negative effect on the economy; (ii) the number of supply disturbances per year that are large enough to exhaust the SPR and, hence, make this extra barrel necessary.

By computing marginal costs and benefits of the strategic petroleum reserve, one can determine whether its size is too small or too large.

This model can also be applied to other economic activities. For instance, using essentially the same model, the optimal investment in reserve capacity in the electricity supply industry has been studied in the 1970s and 1980s by several authors (see Chao 1983 for example). More generally, in early analyses of optimal investment in power generating capacity (Boiteux 1949), peak electricity demand compares to the large-enough supply disturbance in our SPR model.

Note that this model assumes risk neutrality. For any given size of the SPR, the net benefit of increasing the SPR is a stochastic variable because of the uncertainty of the frequency and timing of extreme supply disturbances. We consider only the expected value (average value) of the stochastic net benefit and ignore both the variance around this average and the risk of an extremely unfortunate outcome.

3.3 Parameter values

The marginal net benefit of the SPR depends on a number of parameters, specifically the probability and the size of oil supply disruptions (relative to the size of the SPR), the effect of these disruptions on the oil price, the negative effect of an oil price rise on the economy, and the costs of buying and storing oil. We will look at each of these parameters in turn.

3.3.1 Probability and size of oil supply disturbances

To assess the probability of future disturbances, it is useful to examine the frequency with which such disturbances have occurred in the past. During the second half of the last century, the world oil market experienced several supply disruptions, as shown in Table 1, which also reports their duration and the size of the loss. The length of these disturbances – which were primarily caused by political events in the Middle East – varied between two months (the Six-Day War between Israel and Arab countries in 1967) and around one year (OPEC Riyadh Pact). The magnitude of the disruptions varied between 0.6 million barrels a day (Nationalisation of oil firms in Algeria in 1971) and 4.6 million barrels/day (Gulf war in 1990). The largest disturbance (OPEC Riyadh Pact) followed from a planned cut in oil output that aimed at raising prices from their trough in the late 1990s rather than a supply disruption. Note that a release of oil from the SPR occurred only in very few cases.

Table 1. Oil market disturbances, 1950-2005

		Period	Duration (months)	Average gross loss (mbd)	Total gross loss (million barrels)	Release of oil from SPR (million barrels)
(1)	Nationalisation in Iran	1951-54	44	0.7	940	
(2)	Suez crisis	1956-57	4	2	245	
(3)	Syrian transit dispute	1966-67	3	1	65	
(4)	Six Day War	1967	2	2	120	
(5)	Libyan price dispute	1970-71	9	1	360	
(6)	Nationalisation in Algeria	1971	5	1	90	
(7)	Oil embargo USA and NL	1973-74	6	6	756	
(8)	Iranian revolution	1978-79	6	6	1,008	
(9)	Iran-Iraq war	1980	3	3	360	
(10)	Gulf war	1990	3	3	378	17
(11)	OPEC Riyadh Pact	1999-2000	12	12	>1,000	
(12)	Venezuelan strikes	2002-03	3	2	200	
(14)	Nigerian unrest	2003	6	0.3	50	
(14)	Iraq war	2003-04	19	1	600	
(15)	Katrina hurricane	2005			?	63
	Size of the present total SPR				1,400	

Sources: De Joode *et al.* (2004); IEA (2006b); EIA (2006a); US Department of Energy, DEO (2006).
Notes: mbd = million barrels per day; the total gross loss in the case of the Venezuelan strikes, Nigerian unrest, and the Iraq war is computed from the reported duration and average gross loss; the numbers in the last column (release of oil) are only rough indications.

One way to look at the data is to compare the size of the total gross loss during a market disturbance to the size of the SPR. Another way is to compare the maximum daily release of oil from the SPR to the size of the average gross loss in million barrels/day, regardless of the duration of the supply disruption. Leiby and Bowman (2000) cite such an analysis of the US Department of Energy. According to this analysis, the probability of exhausting a maximum daily release of 12 million barrels/day is once every century. Table 1 suggest that this has happened once in 1999-2000. This approach fits our model if we assume that adding an extra barrel to the SPR increases the maximum drawdown over the relevant time period also by one barrel; in other words, we assume that the maximum drawdown per unit of time is not a restricting limit.

Any value for the frequency and probability of exhausting the strategic petroleum reserve is highly speculative.

In conclusion, it is worth noting that any value for the frequency and probability of exhausting the SPR is highly speculative. However, to determine the optimal size of the SPR, one needs to make assumptions. As our base case, we use a once-in-a-century frequency; that is, we assume that once in a century, an oil market disturbance is severe enough to exhaust the SPR. As this is a rather arbitrary value, we will also conduct a sensitivity analysis.

3.3.2 Impact of disturbances on the price of oil

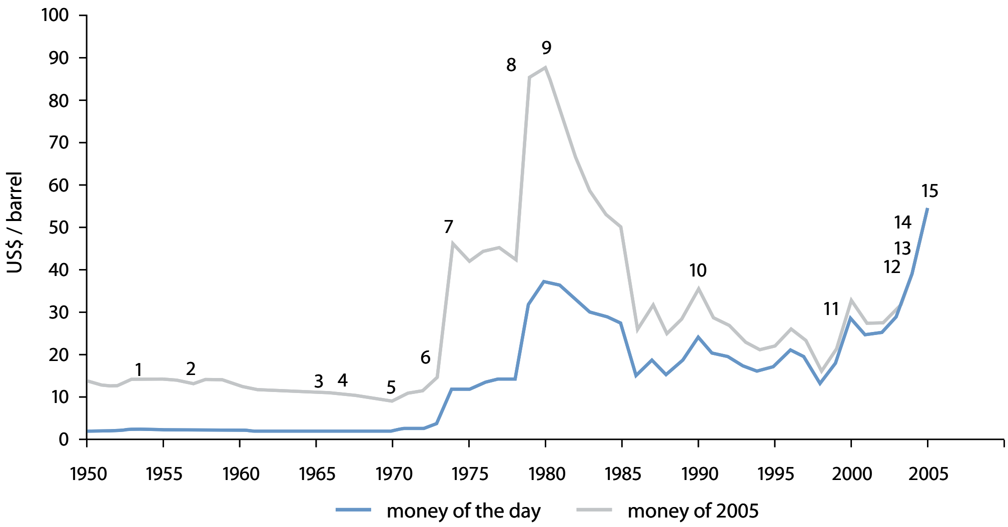
Figure 1 shows that the price impact of various oil market disturbances varied significantly. In the 1950s and 1960s, the impact was negligible due to the organisational structure of the oil market. Between 1973 and the mid-1980s, the influence of disturbances on the oil price increased considerably. In the first oil crisis, in 1973-74, the price of oil surged by approximately 400 percent. Since then, the price of oil has never returned to its pre-1973 level. On the contrary, the price stayed at the new level for the remainder of the decade although the reason (the OPEC embargo on the United States and the Netherlands) that caused the price disappeared. The characteristics of the oil market had altered deeply, with the birth of a powerful oil cartel as the key component.

In the 1970s, the characteristics of the oil market changed deeply, with the birth of a powerful oil cartel as its key component.

The second oil crisis, in 1978-79, lead to a further considerable price hike. Although the oil price stayed at that high level for several years, it turned out to be unsustainable because it stimulated production by non-OPEC producers, on the one hand, and energy savings by oil consumers on the other. Consequently, the cooperation among OPEC members was challenged, ultimately leading to a collapse of both the efficacy of the cartel and the price of oil. In 1985, the oil price fell to a level that became the average for almost 15 years.

Since the second oil crisis, the price effect of disturbances has been less clear. An enormous and gradual rise in the real price of oil has coincided with the supply disturbances 12 through 15 – as shown in Figure 1. However, there might be little causal relation between these two observations.

Figure 1. Crude oil price ($/barrel) and the oil market disturbances of Table 1



Source: Oil prices taken from BP, Statistical Review of World Energy, 2006.

As a sudden price spike hardly induces additional supply in the short term, the price effect of a supply disturbance is determined by the price elasticity of the demand for oil. One may then ask by how much prices have to rise (and demand to fall) so that the market clears when the supply of oil unexpectedly drops. The literature offers a number of estimates for this elasticity.

The Energy Information Administration (EIA 2004) of the US Department of Energy uses the following rule of thumb for an oil price of around 40$/barrel: "for every one million barrels per day of oil supply disrupted and not made good by other supplies (i.e. the net disruption size), world oil prices could increase by $4-$6 per barrel. […] At higher prices, $50 per barrel for instance, the rule of thumb would tend more towards the $5-$7 per barrel range". This amounts to a price rise of roughly 12 percent.

Considine (2001) also provides estimates of the price effects of supply disruptions. In a backwardation situation (when spot prices exceed futures prices), a shortfall of one million barrels/day is estimated to lead to a price increase of \$4-\$6 per barrel. In a contango situation (when spot prices are lower than futures prices), the estimated price rise is \$7-\$13. Using a more elaborate model with an imperfectly competitive market structure, Considine (2002) finds much smaller price effects than in the above simple competitive model: a disruption of one million barrels/day rises the oil price by 1\$/barrel – an estimate that has been used by De Joode *et al.* (2004). To illustrate, for an oil price of 30\$/barrel, the expected oil price increase amounts to 3 percent, and it would be even smaller at the currently high oil price.

Kingma and Suyker (2004) suggest that an unexpected decrease in the world oil supply of 1 percent triggers a short-term price increase of around 15 percent. For the current supply of about 84 million barrels/day, this implies a price increase of 18 percent for a supply disruption of one million barrels/day.[7]

Leiby and Bowman (2000), assume the short-term price elasticity of oil demand to range from -0.125 to -0.19, with an average of -0.16. For a supply of 84 million barrels/day, this average means that an unexpected drop in global supply of one million barrels/day would trigger a price increase of about 8 percent.

To summarise the studies reviewed above, the price increase estimated to result from an unexpected decline in global oil supply of one million barrels/day ranges from 3 percent to 18 percent.[8] Somewhat arbitrarily, we choose a value of 10 percent as a baseline assumption for our model, but we will also analyse the sensitivity of the results to this assumption.

3.3.3 Macroeconomic effects of oil price shocks

There has been much debate about the welfare loss incurred by (net) oil-importing countries as a result of higher oil prices. Usually, the welfare loss is defined by the effect of higher oil prices on GDP, with the link between rising oil prices and economic activity also being subject to considerable debate (Box 1 sketches competing views on this issue). The quantitative strength of this relation is summarised in the so-called oil price elasticity of GDP: the percentage change in GDP due to a one-percent change in the oil price.

The oil price elasticity of GDP is a parameter surrounded by considerable uncertainties ...

In general, the negative impact of an oil price increase on the GDP of oil-importing countries is not constant over time. In particular, an oil price hike might or might not trigger a recession (or prevent an economy from getting out of a recession), depending on the current position of the economy in the business cycle, economic and fiscal policies, and other factors. Hence, the oil price elasticity of GDP is a parameter surrounded by considerable uncertainties.[9]

Mory (1993), for instance, estimated an elasticity of -0.055, which is close to value of -0.054 estimated by Mork *et al.* (1994). Leiby and Bowman (2000) used this value in their study on the SPR of the USA. Labonte (2006) is the last of a series of updates on the effects of oil shocks on the economy – prepared for the US Congress; he refers to a study of Jimenez-Rodriguez and Sanchez (2004) that suggests an elasticity of -0.04 to -0.06. The IEA (2004) reports a reduction of 0.4 percent of GDP due to an oil price increase of 10\$/barrel (from \$25 to \$35), implying an elasticity of -0.010 to -0.014 – a value much lower than the figures above.

7 For a total oil supply of 84 mbd, a drop of 1 mbd is equivalent to a drop of about 1.2 percent, thus requiring a price increase of 1.2 times 15 percent to clear the market.

8 Note that these short-term elasticities are significantly above estimated long-term elasticities, that is, those associated with permanent price rises. See for instance IEA (2006a, p. 287).

9 Many studies analyse the effect of a permanent rise of the oil price and not the effect of a temporary rise associated with oil supply disturbances. Note also that for non-US countries, the effect of a change in the oil price in dollars depends on changes in their exchange rates *vis à vis* the US-dollar.

Box 1. The link between the price of oil and economic activity – competing explanations[11]

Several explanations have been put forward for the inverse relationship between oil price and aggregate economic activity. Brown *et al.* (2002) categorise the explanations into four groups.

The classic 'supply side shock' explanation mentions the rising price of a key input factor as the trigger. Increasing costs of production reduce output growth and, hence, productivity growth. Consequently, the growth of real wages declines and consumers save less or borrow more with view to smoothing consumption. As a result, the real rate of interest rises, reducing the demand for money and, thus, increasing inflationary pressure unless – that is – the supply of money is reduced too. If nominal wages are sticky downward, unemployment will grow, further reducing aggregate production.

A totally different perspective follows from the 'income transfer' explanation. This approach stresses that rising oil prices transfer income from oil-importing countries to oil-exporting countries. As the latter have a lower propensity to consume, global spending and, hence, production declines, particularly in the oil-importing countries. This effect is partly offset by the accompanying growth in aggregate savings that reduces the real rate of interest and thus stimulates investment and production.

Another approach – the 'real balance effect' explanation – focuses on the role of money supply. This explanation states that a rising oil price leads to an increase in the demand for money that is not (fully) matched by an increase in money supply. Consequently, interest rates rise and economic growth slows down.

The final approach sees 'the failure of monetary policy' as the major explanation. According to the adherents of this approach (Bernanke *et al.* 1997, for instance), inadequate monetary policies were the major cause for the relationship between the oil price and economic growth. In the past, central banks often tightened the supply of money with a view to restraining oil-priced induced inflationary pressures, thereby exerting a contractionary impact on the economy. This view has been questioned by others (see Hamilton and Herrera 2001 and Hooker 1999, for instance).

According to Brown *et al.* (2002), the 'classic supply side shock' explanation offers the best account for the inverse relationship between oil price and aggregate economic activity.

Using CPB's Athena simulation model of the Netherlands, De Joode *et al.* (2004) computed the effect of an oil price rise on the Dutch economy.[10] The sensitivity of the Dutch economy with respect to the oil price is also discussed in CPB (2004), suggesting that a rise in the oil price of 10$/barrel (from $35 to $45) lowers GDP by 0.4 percent, implying an elasticity of -0.014 to -0.018.

All in all, estimates of the oil price elasticity of GDP cover a wide range (from -0.01 to almost -0.06). For our analysis of the cost and benefits of increasing the SPR, we will use a value of -0.03. For the Netherlands, we choose a value of -0.02, as this country also benefits from higher oil prices through is exports of natural gas, the price of which is positively correlated to the oil price.

… with elasticity estimates covering a range from -0.01 to almost -0.06.

10 CPB is the Netherlands Bureau for Economic Policy Analysis.

11 This Box draws heavily on De Joode *et al.* (2004, p.52).

3.3.4 The costs of buying and storing oil

Oil storage costs depend on geological characteristics of the storage facility and its drawdown capabilities.

The costs of the SPR consist of the costs of buying and storing the oil. In the simplest case, the annual costs of buying the oil is equal to the interest forgone (value of the oil multiplied by the discount rate). Profits may be made when selling high (during a crisis) and buying low (after the crisis), which reduces the costs. Storage costs depend on geological characteristics of the storage facility and its drawdown capabilities.

In the Netherlands, public oil stocks are mainly stored in salt caverns. Based on government information, De Joode *et al.* (2004) estimate the annual costs of buying and storing oil at €2.4/barrel.

Analysing the SPR costs for the United States, Leiby and Bowman (2000, table 4, p. 26) use a net-present-value cost of about 5$/barrel (in 1996 prices). It is unlikely that this very low figure includes the buying of the oil. For the United States and the EU, we simply use the estimated €2.4/barrel per year for the Netherlands.

As noted above in Section 3.1, it pays to spread the costs of the SPR over many participants. For instance, the share of the Netherlands in the SPR of the European Union is 2.9 percent (see De Joode *et al.* 2004), implying an annual cost to the Netherlands of around €0.07 per barrel of the EU SPR.

3.4 Results: efficiency of strategic petroleum reserves

Table 2 shows the parameter values discussed in the previous sub-section, the results of the cost-benefit analysis under these assumptions, and the 'break-even' parameter values that would make marginal economic costs equal the marginal economic benefits of the SPR.

Row [A] indicates the countries that coordinate their SPR activities and row [B] shows the country – or group of countries – for which the cost-benefit analysis is carried out. To illustrate, the second column pictures the cost-benefit analysis for the Netherlands when this country jointly holds strategic petroleum reserves with its EU partners; the first column shows again marginal costs and benefits for the Netherlands, but this time under the assumption that the Netherlands invests in its own SPR. The crucial difference between these two cases is that when the Netherlands – or any other country – jointly holds reserves with partner countries, it carries only a fraction of the costs of increasing strategic reserves (as indicated in row [4]) while fully capturing the benefits – together with its partners – given that the beneficial oil price effect of releasing oil from the SPR has public-good characteristics.

As the first column shows that, for a small country like the Netherlands, the expected marginal benefit of the SPR (€0.02/barrel per year) is negligible compared to the marginal costs of its own SPR (€2.4/barrel per year). Hence, this country on its own should decrease its SPR. Assuming that this does not measurably raise the probability of fully exhausting other SPR (row [5]), the optimal Dutch SPR is zero.

The Netherlands clearly benefits from being part of a much larger SPR (second column), as shown by the much lower marginal costs to the Netherlands of contributing to the SPR of the EU-15 (€0.07/barrel per year). However, the bottom line remains the same: the SPR is too large, as the marginal benefit is still negative (-0.04 €/barrel per year) and the marginal cost-benefit ratio remains larger than one.

Table 2. Cost and benefits of increasing the SPR

Country coverage					
[A]	Countries coordinating their SPR activities	NL	EU-15	EU-15	US
[B]	Perspective of cost-benefit analysis	NL	NL	EU-15	US
Estimated/assumed parameter values					
[1]	Oil price increase resulting from a decline in oil supply by one million barrels/day	10%	10%	10%	10%
[2]	Oil-price elasticity of GDP	–0.02	–0.02	–0.03	–0.03
[3]	Annual GDP of country (region) in row [B]	€450bn	€450bn	€10,000bn	€8,000bn
[4]	Fraction of country/region [B] in the SPR of [A]	1	0.029	1	1
[5]	SPR fully exhausted once in …	100 years	100 years	100 years	100 years
	Marginal costs (per year) of the SPR				
[6a]	Costs to country/region in row [A]	2.4 €/bbl	2.4 €/bbl	2.4 €/bbl	2.4 €/bbl
[6b]	Costs to country/region in row [B] = [6a]x[4]	2.4 €/bbl	0.07 €/bbl	2.4 €/bbl	2.4 €/bbl
Results of cost-benefit analysis under baseline assumptions					
[7]	Expected marginal benefit per year of the SPR = (–[1]x[2]x[3] / [5]) / (365 million barrels per year)	0.02 €/bbl	0.02 €/bbl	0.82 €/bbl	0.66 €/bbl
[8]	Net marginal benefit per year = [7]–[6b]	–2.38 €/bbl	–0.04 €/bbl	–1.58 €/bbl	–1.74 €/bbl
[9]	Marginal cost/benefit ratio = [6b]/[7]	97	2.8	2.9	3.7
Break-even parameter values (resulting in net marginal benefits of zero in [8] and a cost-benefit ratio of 1 in [9])					
[1′]	Oil price increase resulting from a decline in oil supply by one million barrels/day		28%	29%	37%
[2′]	Oil-price elasticity of GDP		–0.06	–0.09	–0.11
[5′]	SPR fully exhausted once in …		35 years	34 years	27 years
[6a′]	Marginal costs (per year) of the SPR		0.85 €/bbl	0.82 €/bbl	0.66 €/bbl

Notes: Estimated/assumed parameter values as described in Section 3.3. Rows [2], [3], [4], [6a], and [6b] depict country- or region-specific parameters; rows [1] and [5] show global oil market parameters. Rows [5] and [5′] are the reciprocals of the probability of fully exhausting the SPR. The crucial calculation is in row [7]; to illustrate, for the second column, the calculation of the expected marginal benefit of increasing strategic reserves in row [7] is: { – (10 percent) x (–0.02) x (€ 450 billion per year) / (100 years) } / (365 million barrels per year) = 0.02 € per barrel per year. End results have been rounded.

From the perspective of the EU-15 (third column), the results of the cost-benefit analysis differ only slightly from the results in the second column, with the cost-benefit ratio increasing to 2.9. Note that the Dutch share in the GDP of the EU-15 of 4.5 percent is larger than its share of 2.9 percent in the EU SPR. If these two shares were equal and if the oil-price elasticities of GDP were equal, too, there would be no difference in the marginal cost-benefit ratio between the second and the third column.

For completeness we note that from the perspective of the United States, increasing its SPR would be economically unviable too (last column). Indeed, the cost-benefit ratio is even less favourable than in the case of the EU because the US economy is smaller than the EU economy. It should be noted, however, that our results for the United States do not agree with the conclusions of Leiby and Bowman (2000). But as we have pointed out in Section 3.3.4, their analysis seems to be based on an estimate of the marginal costs of holding strategic reserves that we consider as far too low.

So far, the overall conclusion resulting from our analysis is that extending the SPR at the EU level or US level is not efficient, as the marginal costs exceed marginal benefits. Even extending the analysis to the entire IEA SPR might not be enough to reach the marginal break-even point. To illustrate, adding the GDP of the major oil-importing IEA members – Australia, Canada, Japan and South Korea – to the GDP of the EU-15 and the United States gives an annual GDP of roughly €25,000 billion. Using this figure in the third (or last) column of Table 2 would raise the expected marginal benefit to €2.1/barrel per year – still slightly less than the marginal cost of €2.4/barrel.

But how robust are these conclusions to changes in the underlying assumptions?[12] To shed light on this question, we proceed in two steps: we first calculate for each of the parameters its break-even value, that is, the value that would result in marginal net benefits of zero (the results of these calculations are shown in the bottom part of Table 2);[13] we then compare these break-even values with the range of values found in the literature and discussed in Section 3.3.

As the probability of fully exhausting the strategic petroleum reserve is particularly uncertain, the balance of marginal costs and benefits might well change considerably under alternative assumptions.

As for rows [1'] and [6a'], the three last columns show parameter values that are outside the range discussed in Section 3.3. Specifically, the impact of a change in the world oil supply on oil prices would have to be significantly larger than what historical evidence suggests while the costs of buying and storing oil would need to be far lower than they currently are. Therefore, our conclusions are fairly immune to the uncertainty surrounding these parameters.

However, the break-even values in row [2'] and row [5'] are reasonably close to some of the estimates reviewed above. As the probability of fully exhausting the SPR (row [5']) is particularly uncertain, the balance of marginal costs and benefits might well change considerably under alternative assumptions. A reasonable conclusion here is that the SPR of the EU-15 might in fact not be too large. This is true, in particular, when we extend the analysis again by including other IEA countries, thus increasing the GDP affected by oil price movements. In fact, in this case the break-even frequency of fully exhausting the SPR is once in about 86 years – not too far off our baseline assumption. Moreover, if one perceives the risk of a supply disturbance to be large, extending the SPR might be viewed as efficient, in particular by risk-averse decision makers.

3.5 Additional observations and conclusions

A few more observations about the economic analysis of the SPR are worth mentioning. The analysis considers expected net marginal benefits as the key criterion for assessing the SPR. Actual benefits

---

12 A general observation might be useful in this context. The effects of alternative assumptions are relatively easy to gauge from Table 2. For instance, a change in the marginal cost of the SPR (row [6b]) feeds directly through to the net benefit row [8] and the cost-benefit ratio [9]. Doubling (halving) the parameters in rows [1] and [2], doubles (halves) the expected marginal benefit shown in row [7]. For the number of years in row [5], the reverse holds. For instance, halving this number from 100 to 50 (thus doubling the probability of fully exhausting the SPR), doubles the expected marginal benefit in all columns, with the net marginal benefit per year (row [8]) amounting to -2.35 €/bbl, -0.02 €/bbl, -0.76 €/bbl, and -1.08 €/bbl. Hence, net marginal benefits remain negative and the conclusions do not change. But with another doubling of expected marginal benefits the conclusions will change, except in NL/NL-case.

13 Break-even values are not computed for the NL/NL column since zero marginal net benefits are entirely out of reach in this case. To illustrate, zero marginal net benefits would require a full exhaustion of the SPR not once in a hundred years (baseline assumption) but almost every year.

might be quite different, however. Moreover, the actual process of an oil supply disturbance and the release of oil from SPR are more complicated than assumed in this model. For example, De Joode *et al.* (2004) pay much attention to the fact that oil can be released only in limited quantities (the drawdown pattern).

Perhaps a more serious concern is that the SPR might replace regular supply. More specifically, the SPR might crowd out private investment in storage capacity and contracts for reducing demand of industrial consumers, both relying on periods of high prices to make such activity profitable. If oil from the SPR is released into the market when there is upward pressure on oil prices (with a view to keeping prices in check), the attractiveness of such investments and contracts will suffer, ultimately reducing the storage supplied by the market. Ironically, as a result of this crowding out, calls for government intervention might grow.

The strategic petroleum reserve might crowd out private investment in storage capacity and contracts for reducing demand of industrial consumers.

As a final comment, there is a possibly strategic role for the SPR: the threat of releasing oil might deter OPEC from trying to raise the price. As EIA (2006) states, most of IEA actions have not resulted in the release of oil, but the mere existence of the SPR might nevertheless have been valuable.

To conclude, it is perhaps not surprising that the cost-benefit analysis presented in this section cannot yield a clear-cut answer to the question of whether or not it makes economic sense to extend the SPR. Ultimately, this question needs a political answer, but the analysis offered here provides a rational economic framework for finding it.

4. The economics of conserving domestic gas reserves

4.1 Dutch gas-depletion policy: capping output of the Groningen field

Since the first oil crisis in the 1970s, the Dutch government has pursued a gas-depletion policy that rests on two related pillars: conserving gas from the Groningen field – a giant, low-cost, and flexible to operate onshore gas field of around 1,100 billion cubic metres (bcm) – and encouraging production from smaller gas fields. In the Dutch Gas Act (Article 55 of the so-called *Gaswet*), the Minister of Economic Affairs has introduced a cap on production from the Groningen field, which replaces the national production cap that had existed before. In a letter to the Parliament, the Ministry of Economic Affairs (2005) has set a cap on Groningen of, on aggregate, 425 bcm for the period of 2006 to 2015. As this cap only limits aggregate production, it allows for the swing function of the Groningen field, that is, larger production when demand is high during the winter season and lower production during the summer.

The purpose of this section is to analyse the welfare effect of this cap. To this end, Section 4.2 examines the impact of the cap on the European gas market. Section 4.3 considers the costs, benefits, and the welfare effect (benefits minus costs) of the cap.

4.2 Effects of the cap on the European gas market

To analyse the welfare effects of the cap on Groningen, we first have to assess its impact on the production from this field and the natural gas market more generally. We use the NATGAS model (described in Zwart and Mulder 2006) to examine the effects of the cap on the European natural gas market under four scenarios, which we have labelled Baseline, Competition, Sellers' Market, and High Price, respectively (Box 2 offers a brief description of these scenarios). The key question is: what is the effect on prices, investment in and use of infrastructure, and on international gas flows of this policy measure?

Box 2. Alternative gas market scenarios

We analyse the welfare effects of the cap on Groningen under four gas market scenarios, which we have called Baseline, Competition, Sellers' Market, and High Price, respectively. There are three criteria that specify each scenario: the degree of competition in the European gas market, demand growth, and LNG prices – which depend on global competition.

Definition of scenarios	Baseline	Competition	Sellers' Market	High Price
Degree of competition in gas market	modest	high	low	very low
Annual growth of gas demand	1.5%	1.0%	1.5%	2.0%
Average threshold price for LNG	0.15 €/m^3	0.13 €/m^3	0.18 €/m^3	0.28 €/m^3

In the baseline scenario, competition in the European gas market is modest because, for instance, of insufficient cross-border transmission capacity and a limited number of suppliers in national markets. Competition in global energy markets is such that LNG would become available for Europe at a price above €0.15 per m^3 (in real terms), which compares to an oil price of about $34/barrel.

In the competition scenario, conditions for competition are more favourable, resulting in fairly fierce competition in European natural gas markets. Moreover, global competition is assumed to be strong too, depressing the threshold price at which LNG becomes available for Europe to €0.13 per m^3 (about $30/barrel of oil).

In the sellers' market scenario, suppliers' market power is stronger than in the baseline scenario, both in the European natural gas market and in global energy markets. The threshold price for LNG to become competitive is significantly above the threshold in the baseline scenario (0.18 €/m^3 and $40/barrel of oil – both in real terms).

In the high price scenario, the degree of competition in European gas markets is assumed to be very low. Economic growth is strong and there is a certain dash-for-gas, in particular by electricity producers, generating buoyant demand for gas. High oil prices, strategic behaviour of non-EU suppliers, and a high price needed to attract LNG to Europe cause gas prices to roughly remain as in 2006.

In estimating the effect of imposing a cap on Groningen production, one has to take into account that Groningen produces low-calorific (low-cal) gas, which is only consumed in a relatively small market.[14] Apart from consuming it directly (mainly by end users in the Netherlands, Germany, Belgium, and northern France), low-cal gas can also be marketed by mixing it with high calorific (high-cal) gas that has a larger-than-average energy content. Norwegian gas is an example of such gas, and given the expected increase in Norwegian imports, there is some scope for growth of the low-cal gas market as a result.

14 See Zwart and Mulder (2006) for a description of the characteristics of the natural gas market.

Lower output from the Groningen field as a result of the cap will have to be compensated for by either an increase from other sources, or a decrease in demand. In the low-cal gas market, the most important other sources are German production and high-cal gas that is quality converted into low-cal gas. For the latter, there is currently a limited capacity in the Netherlands and its utilisation rate is high.[15] Additional conversion capacity is available in Germany. An increased supply of quality-converted gas would mitigate the increase in low-cal gas prices resulting from a decline in the production of Groningen. As a consequence of the larger demand for high-cal gas, prices rise and production (from the Dutch small fields) and imports of high-cal gas are likely to expand – at least in the long run. In the short term, however, the response of high-cal supplies would be limited because of long lead times for new infrastructure investments. This being said, a short-term supply response could come from increased LNG imports in the nearby markets – Belgium and the United Kingdom, for example – which could experience a diversion of LNG ships from non-European markets – the United States, for example – because of higher prices in continental Europe.

Turning to the demand side of the market, in the short run, higher prices are likely to reduce consumption only slightly (perhaps mainly by power stations using low-cal gas). In the long run, however, a drop in low-cal gas supply and higher prices might reduce investments in gas-fired power plants – favouring other technologies, such as coal-fired power plants.

With a cap on Groningen, one would expect a reduction in the output of Groningen because the operator is no longer free to choose the profit-maximising production path. In these circumstances, the shortfall in the cap triggers an increase in output by Dutch small-fields and other European producers, additional imports from outside Europe, or a decline in demand (because of higher prices).[16] However, as the profit-maximising output might not exceed the cap, the latter might be non-binding, implying no shortfall in Groningen output and, thus, no change in the production of other producers, imports, and demand.

With a cap on the production from the Groningen field, one would expect a reduction in its output because the operator is no longer free to choose the profit-maximising production path.

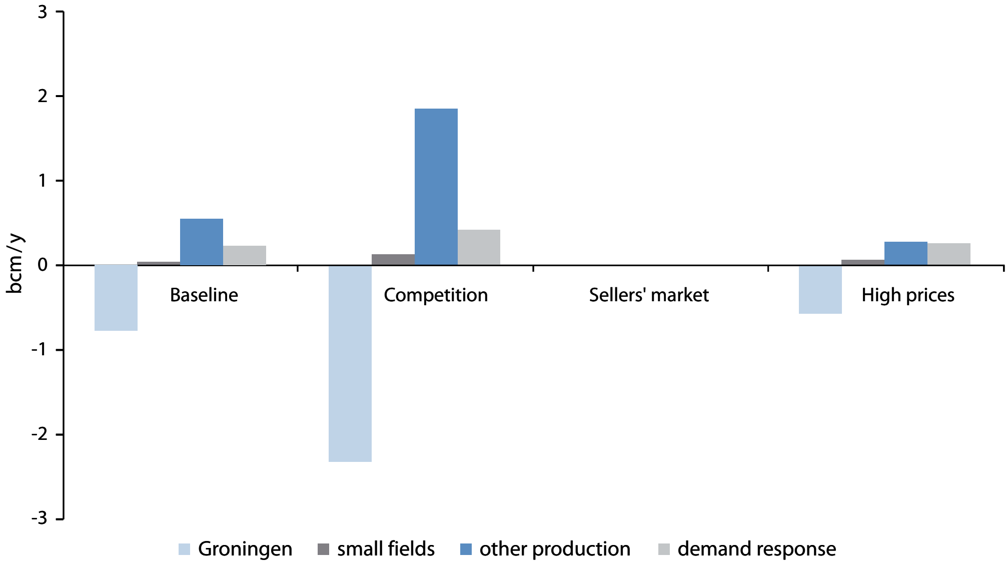
Figure 2 pictures the estimated response to the cap of the Groningen field itself, Dutch small-fields, other European and non-European producers, and of demand under the four scenarios. It turns out that the cap is non-binding in the sellers' market scenario (unconstrained profit-maximising output does not exceed the cap), while it is binding in all other scenarios. This being said, in the baseline scenario and the high price scenario, the cap has only a moderate effect on actual production, amounting to about 0.8 bcm per year and 0.6 bcm per year, respectively. The competition scenario shows the largest impact on Groningen production (an annual average drop of 2.5 bcm over 20 years) and, thus, on other production, imports, and demand. In this scenario, the assumed strong competition in the European gas market reduces the options for strategic behaviour, resulting in a relatively high profit-maximising output, which – however – cannot be attained due to the cap.

The relatively moderate impact on small-fields production shown in Figure 2 is because current capacities are already well utilised and the rise in prices estimated to follow from the cap is too small (0.0007€/m^3 on average over 20 years) to induce significant investment in additional capacity. All in all, the cap on Groningen is estimated to result in only a minor price change on the European natural gas market, mainly because of a fairly elastic response by consumers, other producers, imports and – though to a lesser extent – Dutch small-fields production.

15 See , for instance, DTe (2004) for a discussion of currently available Dutch quality conversion capacity. DTe is the Office of Energy Regulation, a chamber within the Netherlands Competition Authority.

16 It should be noted that these are the effects while the cap is place. When it will be removed in the future (and one enjoys the prolonged life of the field) the effects are reversed.

Figure 2. Effects of the cap on Groningen on the European natural gas market under alternative scenarios (in billion cubic metres per year)



Source: Mulder and Zwart (2006)
Notes : Effects over a 20-year horizon.

4.3 The welfare effect of the cap on Groningen

The welfare effect of the cap on Groningen depends on its impact on the production from this field and the ensuing changes in the European gas market.

Obviously, the costs and benefits – and thus the welfare effect – of the cap on Groningen depend on its impact on the production from this field and the ensuing changes in the European gas market – as discussed above.

As for costs, there are two main cost components to consider when the cap on Groningen binds. First, conserving the Groningen gas field and thus postponing production and sales reduces the present value of the producer's surplus accruing to the operator of the Groningen field. Second, all other things being equal, gas consumers face higher prices today and lower prices in the future, reducing the present value of the consumers' surplus. None of these costs materialise – they are zero – in the seller's market scenario when the Groningen cap does not bind.

As shown in more detail in Mulder and Zwart (2006), for a discount rate of 5 percent (in real terms), the estimated drop in the producer's surplus ranges from €30 million in the high price scenario to €540 million in the competition scenario. The present-value cost to consumers ranges from €140 million in the baseline scenario to €435 million in the competition scenario. Most of the costs to consumers fall on consumers of low-cal gas (that is, the quality of gas produced by the Groningen field), and only a small part on high-cal gas consumers. As shown in Table 3, in present-value terms, total economic costs are estimated to range from €280 million in the high price scenario to almost €1 billion in the competition scenario. For completeness, Table 3 also shows the impact of the cap on output from the Groningen field.

Turning to the expected benefits of the cap, three types of benefits are of particular importance: benefits accruing to small-fields producers, benefits due to a higher reliability of gas supply, and – most important for the topic discussed in this paper – benefits due to an enhanced security of gas supply. We take a brief look at all three categories, and more details can be found in Mulder and

Zwart (2006). Upfront, it is useful to note that the cap brings no benefits when it is non-binding, that is, in the sellers' market scenario.

Table 3. Welfare effect of the cap on Groningen under alternative gas market scenarios

	Baseline	Competition	Sellers' market	High price
	Present value in millions of euros (unless otherwise indicated)			
Costs	500	975	0	280
Benefits				
Additional small-fields producers' surplus	35	145	0	35
Reliability of supply	10	20	0	10
Security of supply	< 200	< 500	0	< 100
Welfare effect = benefits-costs	< –255	< –310	0	< –135
Memorandum item:				
Decline in Groningen output per year (in bcm)	0.8	2.5	0	0.6

Source: Mulder and Zwart (2006).
Notes: Annual cap of 42.5 billion cubic metres on average; discount rate of 5 percent.

Benefits to small-fields producers come in two forms. For one thing, a rise in gas prices due to the cap increases small-fields producers' surplus. Using a discount rate of 5 percent, the present value of this increase is estimated to range from €35 million (baseline and high price scenario) to €145 million (competition scenario) – as shown in Table 3. For another, a cap on Groningen could change its role as a balancing field, thereby allowing small fields to produce at roughly constant rates. However, for all four scenarios, this effect is estimated to be negligible, and it is not shown in Table 3.

Benefits due to a higher reliability of supply reflect the fact that the Groningen field can act as a buffer to make up for temporary supply or demand fluctuations. The cap on Groningen extends the lifespan of this field and thus the lifespan of the buffer. A corollary is that investment in alternative buffers (additional short-term storage, for instance) can be delayed. Hence, the benefit of this measure comes in the form of delaying investments in alternative buffers.

Benefits due to a higher reliability of supply reflect the fact that the Groningen field can act as a buffer to make up for temporary supply or demand fluctuations.

An appropriate alternative buffer would be an expansion of LNG storage capacity that is large enough to cover a one-day supply shortage of 20 million cubic metres. The annual cost of this alternative is estimated at €22 million. We assume that investing in this alternative becomes necessary once the depletion of the Groningen field has progressed to a point where it can no longer act as a buffer. When that point will be reached is hard to predict, but as De Joode *et al.* (2004), we assume this to occur when Groningen reserves have dropped to 400 bcm.

The benefit of capping Groningen is to push this moment further into the future, by several years – depending on the gas market scenario. For these years, one avoids the annual costs of expanding LNG storage capacity. The present value of avoided costs therefore depends on the number of years the investment is postponed and on when the storage needs to be built. Our estimates show that the present value of avoided costs is fairly small, ranging from €10 million in the high price scenario to €20 million in the competition scenario (see Table 3). Indeed, this may still overstate the benefits because the Groningen field cannot act as a buffer against short-term fluctuations during periods of very high demand (winter), whereas alternative storage can.

When discussing investment in strategic petroleum reserves in Section 3, we stressed the possibility of crowding out. The risk that policy measures might crowd out profitable private investment arises, too, in the context of measures to provide a buffer for short-term gas supply and demand fluctuations. However, if we assume that the back-up facility provided by the Groningen field is not used with a view to stabilising prices, but only in emergency situations to avoid forced disconnection of consumers, the effect on private investment is probably low.

This takes us to the third type of benefit: enhanced security of gas supply. As Dutch gas reserves dwindle over the next decades, the Netherlands will turn into a net importer. In this situation, the economy will be increasingly vulnerable to gas price rises. Higher prices, or even price crises, might be incidental, for instance as a result of unusually cold winters, technical supply disruptions, or geopolitical conflicts. There might also be structural reasons for higher gas prices – such as increased market power of a potential future gas cartel. In the remainder of this sub-section, we analyse capping Groningen as a means to provide 'strategic storage' to be used in times of incidental gas supply shortages. In contrast to the reliability benefits discussed above, of interest here are not one-day supply shortages but shortages over a longer period of time.

If market failures leave market participants incompletely exposed to the risk of supply problems, they will invest too little to respond to low-probability supply problems. Solving the market failure itself would be the optimal response here, but if this is deemed unfeasible, the government may step in and reduce the risk of supply shortages – including sharp price increases – by releasing gas from strategic storages.

Leaving the flexibility of the Groningen field looming above the market to interfere when prices rise, reduces the attractiveness of private investment in gas storage.

The flexibility of the Groningen field essentially fulfils this storage function, and conserving Groningen would extend the lifespan during which it can be used to mitigate future price effects of occasional severe winters or technical supply problems. By keeping prices lower in these circumstances, Dutch consumer welfare is increased. However, as with policy measures to reinforce short-term reliability, one should recognise that by effectively capping prices, private-sector investment in alternative storage (LNG import terminals or depleted gas fields, for instance), which relies on occasional price hikes, might be crowded out. Put differently: leaving the Groningen flexibility looming above the market to interfere when prices rise, reduces the attractiveness of private investment in gas storage.

Having said this, we can try to answer what the benefit will be of using Groningen to prevent excessive price hikes under periodic scarcity conditions. Similar to our discussion of short-term reliability, an upper limit for this benefit can be found by considering the cost of other means to achieve this goal – cost that will be avoided due to the cap on Groningen. A natural candidate is strategic storage capacity from which gas can be released in a price crisis. Some countries, Italy in particular, have invested in such storage. As the cost of such storage is estimated to be high (see IEA 2004), the security-of-supply benefits of conserving Groningen might be high, too.

More specifically, we assume that depleted gas fields would be used as strategic storage capacity. As for the size of the capacity needed, we note that the average monthly Groningen production was around 3 bcm below peak production in 2004. If we consider a price crisis of three months (remember that strategic petroleum reserves are required to equal 90 days' consumption), Groningen would be able to supply 9 bcm, on average. Strategic storage in depleted gas fields would have to be in place when the flexibility of the Groningen field is insufficient to produce this additional output, which we again assume to be when the field's remaining reserves have dropped below 400 bcm.

Based on data from ILEX (2005) on storage costs, we estimate the total annual costs for providing storage in depleted fields at €0.05-€0.07 per cubic metre, with the range reflecting alternative discount rate assumptions. The effect of capping Groningen is to postpone the moment when such strategic capacity is necessary to replace the flexibility of Groningen. When this will be depends on the underlying gas market scenario.

As shown in Table 3, our estimates indicate that keeping strategic storage is costly, implying that using Groningen instead of storage entails large benefits, amounting to a maximum of €500 million in the competition scenario (this is for a discount rate of 5 percent). Yet, as other measures dealing with security of supply might be more efficient than investment in strategic storage (measures to influence the demand response, for instance), these calculations overestimate the security-of-supply benefits of Groningen, and they can only be considered an upper bound. This is also because they do not account for possibly crowding out private investment in strategic gas storage.

Considering the economic costs and benefits, the average annual cap on Groningen of 42.5 bcm is estimated to be welfare reducing unless the cap is non-binding. More specifically, the negative welfare effect ranges from (at least) €135 million in the high price scenario to (at least) €310 million in the competition scenario. As shown in Mulder and Zwart (2006), different discount-rate assumptions (3 percent and 7 percent instead of 5 percent) have significant effects on both costs and benefits, but do not change the conclusion that the overall welfare effect is negative if the cap is binding. What is more, this conclusion holds for other reasonable assumptions about the price of LNG, the availability of Dutch offshore gas infrastructure, and the level of tax distortions.

5. Conclusion

The main conclusion from the two policy measures considered in this paper is that security-of-supply measures are unlikely to be welfare enhancing, but they might be in specific circumstances. Considering the economic costs and benefits of such measures, it would often be wiser to accept the consequences of supply disturbances than to avoid them. But it is also true that estimating the benefits of security-of-supply measures is a task surrounded by considerable uncertainties. Governments should thus proceed carefully in taking such measures. If serious market failure is detected, attention should be paid to properly design corrective measure.

Estimating the benefits of security-of-supply measures is a challenge and a task surrounded by considerable uncertainties.

But if markets function reasonably well, prices will give market participants incentives to prepare for and respond to supply disturbances, enabling an efficient balancing of supply and demand not only under normal market conditions but also during periods of stress. It is true that even well-functioning markets might be prone to price spikes, as evidence from both the gas and electricity markets demonstrates. However, the welfare cost of price spikes is often small compared to the cost of policies aimed at preventing them. Given the market's potential to efficiently cope with energy supply disturbances, government action to improve the functioning of energy markets (such as pricing externalities and facilitating transactions between market participant) might be more efficient than measures requiring large up-front investments or measures restricting options of profit-maximising agents.

However, cost-benefit analyses can offer only part of the information needed for decision making. Not all costs and benefits are measurable and those that are not need to be accounted for as *pro memoria* items. Examples include the possible crowding out of private activity geared towards enhancing security of supply by government action. Moreover, the distribution of costs and benefits within society generally plays an important role in the decision-making process. In our analysis, we analysed the distribution effects at a fairly aggregate level only. Another issue concerns the degree of risk aversion underlying cost-benefit analyses. Decision makers might act in a rather risk-averse manner – for instance because they fear for their reputation in times of supply disturbances. But it is also possible that societies as a whole are more risk averse than often presumed. In both cases, the results of cost-benefit analyses are likely to shift in favour of security-of supply policies.

References

Algemene Energieraad. (2005). *Gas for tomorrow*, AER, The Hague, The Netherlands.

Bernanke, B.S., Gertler, M., Watson, M., Sims, C.A., and Friedman, B.M. (1997). "Systematic monetary policy and the effects of oil price shocks". *Brookings Papers on Economic Activity*, (1), pp. 91-157.

Bohi, D.R. and Toman, M.A. (1996). *The economics of energy security*. Kluwer Academic Publishers, Norwell, Massachusetts, USA.

Boiteux, M. (1949). "La tarification des demandes en pointe". *Revue Générale de l'Électricité*, (58), pp. 22-40.

Brown, S.P.A. and Yücel, M.K. (2002). "Energy prices and aggregate economic activity: An interpretative survey". *The Quarterly Review of Economics and Finance*, (42), pp. 193-208.

Bushnell, J. (2005). "Electricity resource adequacy: Matching policies and goals". CSEM Working Paper No. 146.

Centraal Planbureau, CPB. (2004). *Macro Economische Verkenning 2005*, CPB, The Hague, The Netherlands.

Centraal Planbureau, CPB. (2006). "Liberalisation of European energy markets: Challenges and policy options". Centraal Planbureau Document 138.

Chao, H. (1983). "Peak load pricing and capacity planning with demand and supply uncertainty". *The Bell Journal of Economics*, (14), pp.179-190.

Considine, T.J. (2001). "Oil stocks and the price response from stockpile sales: An econometric analysis". Pennsylvania State University Report.

Considine, T.J. (2002). "Inventories and market power in the world crude oil market". Pennsylvania State University Paper.

De Joode, J., Kingma, D., Lijesen, M., Mulder, M., and Shestalova, V. (2004). "Energy policy and risks on energy markets: A cost-benefit analysis". Centraal Planbureau Special Publication 51.

Energy Information Administration, EIA. (2000). "A look at western natural gas infrastructure during the recent El Paso pipeline disruption". EIA Special Report.

Energy Information Administration, EIA. (2004). "Rules-of-thumb for oil supply disruptions". Disruption Analysis (version dated September 2, 2004).
http://www.eia.doe.gov/emeu/security/rule.html

Energy Information Administration, EIA. (2006). "Global oil supply disruptions since 1951". Table. http://www.eia.doe.gov/emeu/security/distable.html.

European Commission. (2002). "The internal market in energy: Coordinated measures on the security of energy supply". Communication from the Commission to the European Parliament and Council COM(2002) 488.

European Commission. (2005). *Directive 2005/89/EC of the European Parliament and of the Council*, EU, Brussels, Belgium.

Global Insight. (2005). "Ensuring effective and efficient forward gas markets", Report to the DTI.

Hamilton, J.D. and Herrera, A.M. (2001). "Oil shocks and aggregate macroeconomic behavior: The role of monetary policy". Department of Economics, University of California, San Diego Discussion Paper 1001-10.

Hooker, M.A. (1999). "Are oil shocks inflationary? Asymmetric and nonlinear specifications versus changes in regime". Board of Governors of the Federal Reserve System Research Paper Series - FEDS Papers No. 99-65

ILEX Energy Consulting (2005). "Storage, gas prices and security of supply", Report to DTI, October 2005.

International Energy Agency, IEA. (2004). *Security of gas supply in open markets; LNG and power on a turning point*. IEA, Paris, France.

International Energy Agency, IEA. (2006a). *World Energy Outlook 2006*. IEA, Paris, France.

International Energy Agency, IEA. (2006b). *Fact sheet on IEA oil stocks and emergency response potential*. IEA, Paris, France.

Jimenez-Rodriguez, R. and Sanchez, M. (2004). "Oil price shocks and real GDP growth: Empirical evidence for some OECD countries". European Central Bank Working Paper No. 362.

Kingma, D. and Suyker, W. (2004). "FAQs about oil and the world economy", Centraal Planbureau Memorandum 104.

Labonte, M. (2006). "The effects of oil shocks on the economy: A review of the empirical evidence". Congressional Research Service Reports for Congress. http://www.ncseonline.org/NLE/CRSreports/06feb/RL31608.pdf

Leiby, P.N. and Bowman, D. (2000). "The value of expanding the US strategic petroleum reserve". Oak Ridge National Laboratory, ORNL/TM-2000/179.

Ministry of Economic Affairs. (2005). "Voorzienings- en leveringszekerheid energie", Tweede Kamer der Staten Generaal, 2005-2006, 29 023, nr. 21, 22 December.

Mork, K.A., Olsen, Ø., and Mysen, H.T. (1994). "Macroeconomic responses to oil price increases and decreases in seven OECD countries". *Energy Journal*, (15:4), pp. 19-35.

Mory, J.F. (1993). "Oil prices and economic activity: Is the relationship symmetric?". *Energy Journal*, (14:4), pp. 151-161.

Mulder, M. and Zwart, G. (2006). "Government involvement in liberalised gas markets; A welfare-economic analysis of Dutch gas-depletion policy". Centraal Planbureau Document 110.

Nelson, J.R. (1964). *Marginal cost pricing in practice*. Prentice-Hall, Englewood Cliffs, New Jersey, USA.

Newbery, D.M. (2000). *Privatization, restructuring, and regulation of network utilities*. Walras-Pareto Lectures, MIT Press, Cambridge, Massachusetts, USA.

OECD. (1999). *Economic Outlook No. 66*, Paris, France.

US Department of Energy, DOE. (2006). "Releasing crude oil from the strategic petroleum reserve". article. http://fossil.energy.gov.

Willenborg, R., Tönjes, C., and Perlot, W. (2004). "Europe's oil defenses: An analysis of Europe's oil supply vulnerability and its emergency oil stockholding systems", *The Journal of Energy Literature* (X:2), pp. 3-49

Zwart, G.T.J. and Mulder, M. (2006). "NATGAS: a model of the European natural gas market". Centraal Planbureau Memorandum 144.

ABSTRACT

This paper provides a quantitative analysis of power relations and strategic investment in the transport system for Russian gas. *First, we analyse how the architecture of the transport system determines Russia's bargaining power vis-à-vis (potential) transit countries. By applying the Shapley value as a solution for multilateral bargaining we find that competition between transit countries is of little strategic importance compared to direct Russian access to its customers in Western Europe. Second, we develop a dynamic model of strategic investment. We find that the failure to include Belarus and Ukraine into a framework for international contract enforcement resulted in underinvestment in cheap pipelines and overinvestment in expensive ones. As capacities are increased to gain leverage over transit countries, customers in Western Europe benefit in terms of lower prices and higher supply security.*

Franz Hubert hubert@wiwi.hu-berlin.de holds the chair for Management Science at the Humboldt Universität zu Berlin. This paper is based on joint work with Svetlana Ikonnikova and Irina Suleymanova.

Strategic investment in international gas transport systems

Franz Hubert

1. Introduction

In late 2005, Russia and Germany signed treaties to build a huge new pipeline through the Baltic Sea, the North European Gas Pipeline, which was renamed to *Nord Stream* shortly afterwards.[1] The project will enable Russia to maintain its position as major supplier of natural gas to Western Europe.[2] Plans for an offshore pipeline to Western Europe have been around for quite a while under names like Baltic Ring and North Trans Gas. However, for a long time Russia's western partners dragged their feet, mainly because of all possible ways to increase the transport capacity for natural gas from Russia to Western Europe, *Nord Stream* is the most expensive option (see Figure 1 for an illustration of the network). With *Yamal 1* already in place and the system in the south still in decay, there are commercially more attractive and technologically less demanding alternatives to *Nord Stream*, such as upgrading the old system in the south, adding a second pipeline to *Yamal 1*, and even building new pipelines in the south. However, cost and technological risk are only part of the picture. As the hostile reactions from Ukraine and new EU members Poland, Lithuania, and Latvia suggest, *Nord Stream* will alter the balance of power in the region. It will also have a long-term impact on the EU's ability to formulate a common energy policy.

Production and transportation of natural gas are characterised by large initial investment in specialised facilities with a long lifetime and low operating costs. Most of the expenditures on project identification, investment planning, and construction are sunk. Once installed, capacities generate large quasi-rents. Hence, it is essential that the countries can credibly commit to grant access to pipelines on agreed terms. Currently, there are no international institutions that could enforce multilateral contracts in case of a dispute if countries outside the EU are involved. If some countries cannot commit *ex ante* to share the rents in long-term contracts, re-contracting after completion of the investment is anticipated. As a result, investment may be distorted to gain leverage in the bargaining process. This distortion may lead to underinvestment in cheap pipelines and/or overinvestment in expensive ones – compared to what would maximise the profits of the supply chain for Russian gas. Given its very substantial markets share, Russia enjoys some market power *vis-à-vis* European customers. Hence, underinvestment tends to increase prices for European customers while overinvestment decreases them. As a result, conflicts along the vertical supply chain have an impact on customers in the West far beyond the rare short-term interruptions that usually attract much public attention.

This paper summarises previous work by Hubert and Ikonnikova (2003), Hubert and Ikonnikova (2004) and Hubert and Suleymanova (2006). In this literature, the interdependencies among the players are captured by a game in value-function form. The 'power' of a player is measured by the Shapley value of this game. A major advantage of this approach is that it does not involve *a priori* assumptions on bargaining power or details of the bargaining process, about which little is known. Instead, it derives the power structure entirely from the geography of the network and the cost differentials between the various pipelines. The value function can be used to model different institutional structures.

1 For ease of reference, we use *italics* for those pipelines that are explicitly shown in the Figures below and/or form part of the model developed in this paper.

2 Throughout this paper we will refer to 'Western Europe' as the market consisting of the EU-15 countries excluding Greece. For ease of reference, we use the names of the countries instead of companies when there is no risk of confusion. Hence, we speak of Russia rather than Gazprom, Ukraine instead of Naftogaz, and so on.

For the Eurasian pipeline system we can quantify the cost of the different pipeline options with reasonable accuracy to calibrate the model and predict the impact of transport capacities on the power structure. This, in turn, allows us to calculate how the players would optimally invest under various assumptions about their ability to make long-term commitments. We consider a 'strategic' investment a player's attempt to influence the power structure – formally: the Shapley value – to its advantage. The resulting investment pattern deviates considerably from 'non-strategic' investment, which would maximise the profits of all players and minimise transportation cost for any given total capacity. The difference between the two is also called the strategic distortion of investment.

In spite of many unresolved conflicts between Russia and gas transit countries, the flow of gas has been interrupted only on two occasions.

In spite of many unresolved conflicts between Russia and transit countries, the flow of gas has been interrupted only on two occasions. During such a crisis, observers tend to focus on the immediate impact of actions. From this shortsighted perspective the power of a player is determined by its control of existing transport capacities. The *status quo*, however, can be changed by adding new pipelines to the existing system. In principle, rational, farsighted players should, therefore, take into account all relevant options to modify the network to obtain a comprehensive assessment of their relative bargaining power.

As a first step, we consider the two borderline cases of shortsightedness and farsightedness to obtain reasonable bounds on what we may expect from a dynamic analysis. By comparing the two scenarios we are also able to assess the strategic relevance of different pipeline options. Our quantitative analysis reveals that some commercially feeble projects are, nevertheless, very important for strategic reasons. Others, which have been deliberately drawn up to alter the balance of power, turn out to be strategically irrelevant.

As a second step, we develop a truly dynamic analysis for a stripped-down version of the model. In every period, the players bargain on the sharing of rents from previous investment. At the same time, however, they can form coalitions for new investment. Additional transport capacities have a long-lasting impact on bargaining power, but they become available only with some delay. In this dynamic, infinite horizon setup, we investigate the incentives for strategically distorting investment.

As in Hubert and Ikonnikova (2004), the distinction between short-term and long-term cooperation is crucial. Short-term cooperation refers to the coordinated use of the existing transport capacities in any given period. It also includes the sharing of current profits if this has not been determined previously in a long-term agreement. Long-term cooperation revolves around the joint determination of transport capacities, ownership or secured access rights, and long-term rent sharing. It requires commitment over time spans of up to forty years. In principal, these commitments can be based on contracts that are enforced by external institutions. If these institutions are not available, long-term cooperation can also be based on dynamic strategies, which support cooperation by the mutual threat of retaliation. In the literature on cartels this informal cooperation is often called collusion.

Our quantitative analysis shows that strategic considerations are of outmost importance in the Eurasian transport network. If the players fail to collude and invest non-cooperatively, all equilibria feature substantial overinvestment to create countervailing power. While cheap investment opportunities are neglected, new expensive pipelines are built, with capacities well ahead of the development of demand. As a result, prices for customers will be lower than if the members of the supply chain for Russian gas would coordinate to maximise their joint payoff. Not surprisingly, there is a large potential for raising profits through dynamic collusion. However, the members of the supply chain largely failed to realise this potential.

The remainder of this paper proceeds as follows. The next section sets the stage by briefly describing the pipeline system for transporting natural gas to Western Europe. Section 3 and 4, respectively, present the static analysis of bargaining power and the dynamic analysis of strategic investment. Section 5 concludes.

2. The supply chain for Eurasian gas

2.1 The current gas transport network

Currently, natural gas has a share of about 25 percent in the fuel mix of the energy consumption of the European Union. This share is likely to grow in the near future, because gas is considered to be an environmentally less harmful source of energy than coal and oil. In 2005, about a quarter of the gas consumption was covered by supplies from Russia, though the share is much larger for France (28 percent), Germany (35 percent), Austria (55 percent), and Poland (53 percent). The magnitude of gas imports from Russia creates concerns of energy dependency and reliability of supplies.

Dependency on Russian gas is difficult to avoid, given that the only alternative producers of pipeline gas, Norway and Algeria, are not able or likely to substantially increase production.[3] However, it is worth stressing that the dependency is mutual. It is often emphasised that Russia honoured contracts in the past and did not abuse its position for political purposes. Hence, reliability refers mainly to the problem of secure transport through transit countries, with which both Russia and the EU failed to establish solid and stable political and economic relations.

Russian natural gas is delivered through a network of pipelines stretching from the industrial centres in Western Europe to the main pipelines in western Russia and further on to fields in permafrost regions of Siberia and the steppes of central Asia. The main features of the transport system have been shaped during the 1970s and 1980s. When the Soviet Union started to supply gas to Western Europe in the late 1960s, it extended an existing pipeline – through which gas was transported from eastern Ukraine to Czechoslovakia – to connect to Austria and Germany. This connection is part of what is called the *Southern System* in Figure 1. As exports increased, additional capacities were established alongside previous routes, which were linked to new fields in the southern Ural (Orenburg). Surprisingly, even when production shifted northwards to Vuktylskoe and fields in western Siberia, the pipelines exporting this gas to the west took a turn towards the old routes in the south. Plans to build a new pipeline through Poland and former East Germany were abandoned, apparently because the Soviets considered occupied Czechoslovakia, through which the southern track went, to be politically more reliable.

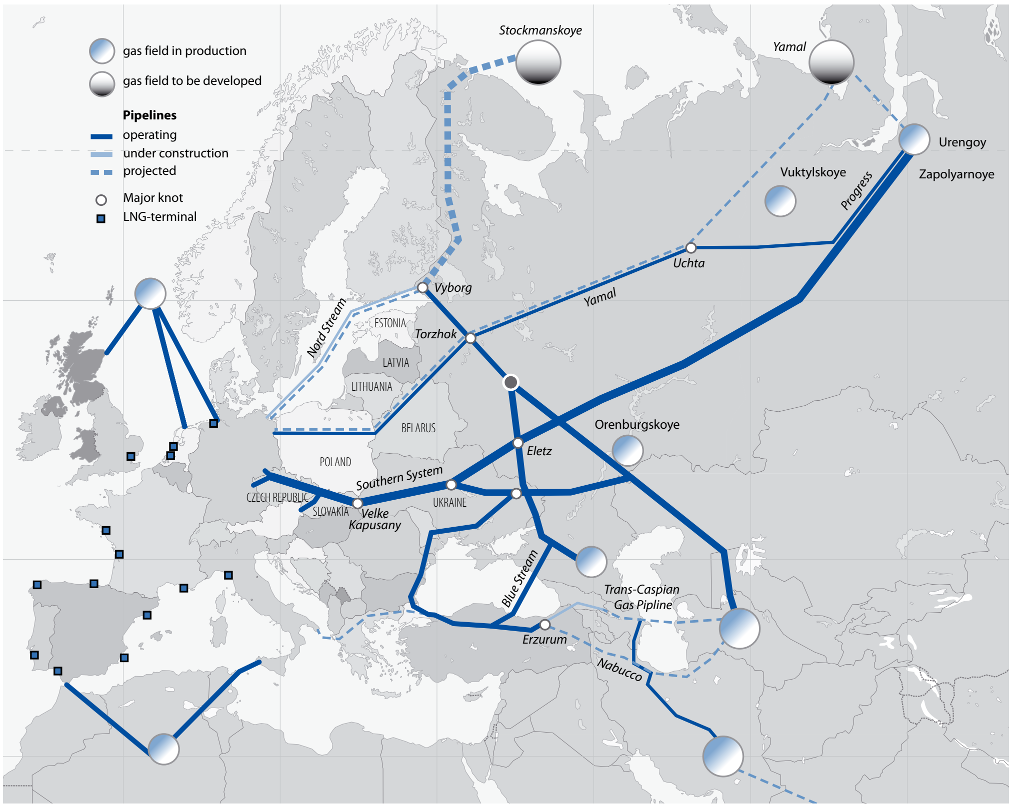
When the Soviet Union collapsed, however, Russia found itself in the uncomfortable position that its only supply route to Western Europe passed through three newly independent states: Slovakia, the Czech Republic, and Ukraine.[4] Looking westward towards integration with the EU, Slovakia and the Czech Republic privatised their pipelines, which were acquired by western importers. The transition to stable commercial gas relations was also eased by both countries' entitlements to large gas deliveries at low cost, inherited from former Czechoslovakia.

When the Soviet Union collapsed, Russia found itself in the uncomfortable position of having its only supply route to Western Europe pass through three newly independent states.

3 The picture may change if the pipeline *Nabucco* directs Iranian and Turkmen gas to Europe, but this project is doubtful for political reasons and not considered in this paper.

4 For detailed and balanced accounts of the ensuing conflicts and Russia's strategy see Stern (1999) and Stern (2005).

Figure 1. Eurasian gas transport network



For many years, Russia and Ukraine failed to find a lasting solution for their gas relationship.

Russia and Ukraine, in contrast, failed to find a lasting solution for their gas relationship. In principle, Russia paid for transit by supplying gas to Ukraine, approximately 26-30 billion cubic metres a year (bcm/a). This payment in kind is sometimes translated into a 'transit fee' by assigning a price to the gas, but as these fees are not actually paid, they have little relevance. The conflicts are essentially over the compensation for additional 25 bcm/a, which Ukraine dearly needs. While Russia claimed average European prices, Ukraine conceded only half of that. In the late 1990s, even the lower figures were not fully paid. Ukraine has also been blamed for syphoning off gas in excess of what it acknowledges officially, a claim that has some credibility, although it is strongly denied by Ukraine.

As a result of non-payments and alleged 'stealing', debts accumulated. In 2002, these amounted to \$1.4bn or \$3.5bn – depending on which side one takes. As a partial solution Ukraine paid \$285m by handing over strategic bombers and missiles in 2002/2003, but both sides could not agree on prices of other components of the proposed barter deal. Meanwhile, due to aging compressors, lack of maintenance, and underinvestment, the capacity of the transport network declined.

In mid-2004, Gazprom and Ukrainian Neftogas apparently reached a comprehensive solution for their long-standing conflict. The agreement converted controversial debt into a formal loan and established a new barter agreement for the use of the transport facilities. Under this agreement, Gazprom was expected to deliver 21-25 bcm/a for the period 2005-09.[5] A Russian-Ukrainian consortium RosUkrEnergo was set up to operate and refurbish the Ukrainian transit system in cooperation with Western partners. By replacing old compressors, the transport capacity could

5 The nominal price for transport was set at \$1.094/tcm/100km, which is fairly standard. The implied price for gas was set at \$50/tcm – about a third of the Western European price at that time (tcm = thousand cubic metres).

have been easily increased by 15 bcm/a. A further expansion still benefits from the established infrastructure but requires complementary investment in pipelines, in particular in Slovakia and the Czech Republic.

However, Ukraine failed to transfer the pipeline system to the consortium. After the Orange Revolution brought a regime change in December 2004, the other parts of the deal quickly unravelled. The new government questioned the debt settlement and opened criminal investigations against RosUkrEnergo, which was declared a failure in summer 2005. The mysterious disappearance of 7.8 bcm from Gazprom's storages in the Ukraine further strained the relation. With gas prices soaring to a record $230/tcm in Western Europe by late-2005, both sides were still haggling over conditions for deliveries in 2006. Russia insisted on international prices while Ukraine offered a third of that amount. With no contract in place, the crisis culminated in January 2006, when Russia shut down deliveries earmarked for Ukraine. However, Ukraine simply continued to take from the export pipelines whatever it needed, and Russia had little choice but to make up for Ukraine's diversions or to default on its obligations to western importers. It took only four days, to find a formal solution to the dispute, but a comprehensive contract was signed only after lengthy negotiations with a new government in Kiev.

Let us then consider how the gas transport network has been extended so far since the break-up of the Soviet Union. After gaining independence, Belarus' ties with Russia remained initially very close. In 1993, both countries agreed on a long-term solution for sales and transit relationships, including the transfer of transit routes from BelTransGaz, the national transmission company, to Gazprom under a 99-year lease. In the case of Poland, a joint stock company, EuroPolGaz, was established in which Polish PGNiG and Russian Gazprom hold equal shares. This encouraged Gazprom to revive old, ambitious plans to develop the huge Yamal field and connect it to internal and external markets with a new massive northern route. However, as demand was weak during the 1990s and the cost of developing the Yamal field turned out to be very high, the project was gradually scaled down. Eventually, attention focussed entirely on the export pipeline, now commonly referred to as *Yamal 1*, which is built 'from the market to the field'.[6] The first pipeline went into operation in 1998. Due to delayed investment in compressor stations, it did not reach its capacity of 28 bcm/a before 2006. A second pipeline, *Yamal 2*, with a potential of another 28 bcm/a has already been laid at major river crossings (see Victor and Victor 2006).

Like Ukraine, Belarus seeks large price concessions for its gas imports using the leverage it gains from its strategic position in the export chain.

After *Yamal 1* started to pump gas, relations between Russia and Belarus deteriorated. Like Ukraine, Belarus seeks large price concessions for its gas imports using the leverage it gains from its strategic position in the export chain. In April 2002, a deal was reached under which Gazprom had to deliver 10 bcm/a at a discount price and accumulated debts were swapped for a controlling stake in BelTransGaz, which manages Gazprom's pipelines in Belarus. However, the second part of the deal, which would have given Gazprom a much more effective control over its export routes, never materialised. When negotiations about new conditions for gas supplies failed, Gazprom stopped deliveries to Belarus at the end of 2003. For a couple of weeks, independent suppliers filled the gap at higher prices, and then Belarus started to divert gas from the export pipeline. Gazprom responded by shutting down gas supplies altogether, deliberately cutting off not only Belarus but also Kaliningrad, Poland, and Germany in February 2004. The immediate crisis was resolved within one day, avoiding any serious impact on customers in the West. Then it took both countries almost five month to agree on a temporary solution until the end of 2006. Formally, Belarus bowed to Russian

6 At that time, the high cost of developing new fields such as Yamal or Shtokman and the availability of low-cost alternatives in old Siberian fields and Turkmenistan cast doubt on the economic viability of grand-scale projects in the near future (Stern 1995). Meanwhile, gas for *Yamal 1* is supplied from fields in the Siberian Basin including newly opened Zapolyarnoye.

demands for higher gas prices, but the net impact was small because Belarus was compensated by an increase in transit fees. In December 2006, both countries found themselves again at the brink of a crisis. The new 5-years supply and transit contract was not signed before late at night on December 31. It raised prices for Belarus from about a quarter of Western European levels to half of that and envisaged a stepwise adjustment to international prices by 2011. What looks like a very substantial price increase is at least partly compensated for by a doubling of transit fees for gas shipped to Poland and by cash payments Gazprom is expected to make for a gradual acquisition of BelTransGaz, Belarus' national gas company. The latter, if fully executed, would give Gazprom a controlling stake by June 2010 (see Yafimava and Stern 2007 for details). In view of Belarus' reluctance to cede control on previous occasions, we see a chance that it may fail to implement the last step. In the meantime, Gazprom's payments would mainly offset the price increase.

2.2 Alternative options for extending the gas transport network

A salient feature of some network extensions is that they would bypass some – or even all – of the current gas transit countries.

A salient feature of the possible network extensions we consider now is that they would bypass some – or even all – of the current transit countries. As a direct threat to Ukraine's strategic position, Russia developed plans for a twin pipeline with a capacity of 60 bcm/a running north-south through Belarus, Poland and Slovakia.[7] Since this link can also be seen as part of the larger Yamal project, it is sometimes referred to as *Yamal 2* (with a planned capacity of 28 bcm/a). However, if realised without additional investment towards customers in the West (and fields in the east) it would mainly serve to bypass Ukraine, hence, we will refer to this project as *Bypass* (see Figure 2). With an estimated cost of €4bn, the *Bypass* has limited commercial value in a narrow sense, because it does little to increase transmission capacities westwards. Nevertheless, offering an alternative to the route through Ukraine, its strategic value is potentially large.

In principle, Belarus can also be bypassed through Latvia and Lithuania, an option to which we will refer to as *Baltic* (see Figure 2). Standing alone, such a pipeline would allow using existing capacities of *Yamal 1* in Poland and Russia without involving Belarus. Beyond the capacity of *Yamal 1*, the same track could be part of a revised *Yamal 2* project. So far, this possibility has attracted little public attention, but it may explain the hostile reaction of the Baltic states towards the announcement of another pipeline even further to the north, to which we turn next.

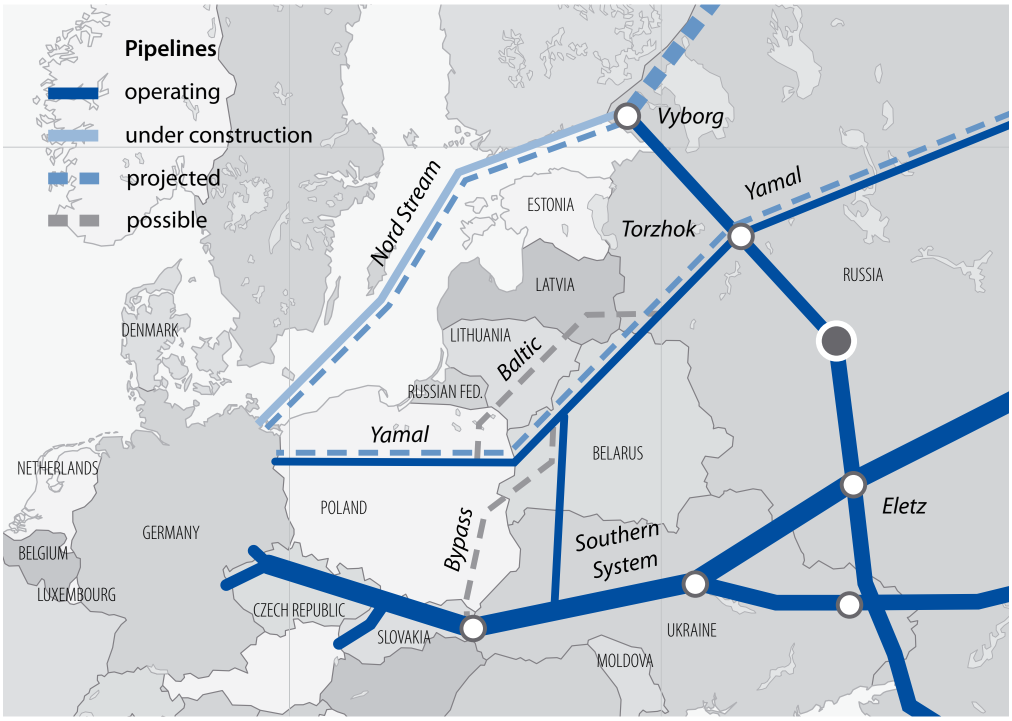
Early plans for a Baltic Ring, connecting Russia through Finland and Sweden to Germany have been abandoned during the late 1990s in favour of a direct offshore connection between Vybourg (Russia) and Germany called North Trans Gas. Initially, Gazprom and Fortum – a Finnish energy company – shared the project, but German Ruhrgas and Wintershall were subsequently invited to join. Planned capacities ranged from 18 bcm/a to 30 bcm/a. Commercially, the link would have looked more attractive if connected to Shtokman, a large field yet to be developed. As with Yamal, the prospects for the development of Shtokman are vague at best. And even if the field is developed, it might be cheaper to liquefy the gas, since the costs of an onshore pipeline appear to be very high due to difficult terrain on the Kola Peninsula. Nevertheless, Russia kept pushing North Trans Gas in international negotiations, while Western partners dragged their feet due to high cost.[8]

7 Gazprom would be in charge of the section in Belarus. For Poland and Slovakia, a consortium was set up including among others Gazprom (18 percent), PGNiG (10 percent), SNAM (29 percent), Ruhrgas (22 percent), GdF (12 percent), and Wintershall (5 percent). The project was pushed in the late 1990s but never got off the ground. Apparently it was shelved with the agreement between Russia and Ukraine in 2003.

8 In the south, another off-shore project, the *Blue Stream* pipeline (see Figure 1) through the Black Sea to Turkey, started operations in 2002 under a long-term agreement with Turkish Botag. It substitutes for pipelines running through Ukraine, Moldavia, Romania, and Bulgaria – where conflicts have been similar to those on the east-west routes.

In 2005, the project took a surprising turn. Under the name North European Gas Pipeline, now called *Nord Stream*, a German-Russian consortium – involving E.ON-Ruhrgas and Wintershall – announced the construction of a twin pipeline with a capacity of 60 bcm/a. Understanding that the project will shift the balance of power in the transport system for natural gas, the announcement triggered a very hostile reaction, in particular from the Baltic states and Poland, which tried to muster their influence as EU members to lobby against the deal. But in spite of this opposition, construction work on the Russian on-shore section started soon after the announcement.

Figure 2. Options for extending the Eurasian gas transport network



3. Static analysis: pipelines and bargaining power

3.1 Modelling bargaining power

The more important a country is for others in a multilateral bargaining situation, the more power it has and the larger will be its profit share.

In this section, we develop a formal model of the supply chain, from which we can derive quantitative results. As countries control only sections of the pipelines, they have to cooperate in order to generate revenues. The interdependencies among the players are represented by the so-called value function, which assigns a payoff, or value, to every possible subset of players. From the value function we can calculate the marginal contributions of any player to any subset of other players. It indicates how much the group can gain by incorporating the player. Hence, a player's marginal contribution reflects the value of his pipelines, or more generally resources, for others. Intuitively, the more important a player is for others, the more power he has and the larger his share of the profits will be. This intuition is nicely captured by the so-called Shapley value (see Box 1 for details). The Shapley value of a player is calculated by taking the weighted average of his marginal contributions. The weights are given by the probability that a subset of players is formed when the players are ordered at random. Hence, the Shapley value can also be interpreted as the expected marginal contribution of a player under random ordering bargaining.

The main players in the network are Russia, Ukraine, Poland, and Belarus. In addition we include Slovakia, Latvia, and Lithuania, which are involved in attempts to bypass Ukraine or Belarus.

As to pipelines, we strife for a fairly comprehensive picture, taking into consideration the old system through Ukraine, which we will refer to as *South*, the possibility to *Upgrade* it, the current *Yamal 1,* its possible enlargement *Yamal 2*, *Bypass*, *Baltic*, and finally *Nord Stream*. However, a player's command over resources also depends on institutional features. For this reason we will not consider the sections of the *Southern System* located in Slovakia and the Czech Republic. They have been sold to importers and cannot be used as a bargaining chip without violating EU laws. We also assume that Poland, as an EU member, cannot obstruct the use of the existing pipeline *Yamal 1*. Poland is bound by long-term agreements, which effectively give Russia assured access to *Yamal 1*. However, like Slovakia, it can veto any new pipeline on its territory. Hence, without Poland, neither *Baltic* nor *Yamal 2* can be built.

The Shapley value describes the contribution of each participant to the overall operation, and the payoff he can reasonably expect.

Box 1. Shapley value

Strategic interdependencies are represented by a game with N players and value function $v(S)$ for all coalitions $S \subseteq N$. The marginal contribution of player $i \in N$ to a coalition $S \subset N$ is given as $v(S \cup i)$ - $v(S)$. The Shapley (Shapley 1953) value solves the game (N, v) by assigning each player i a payoff φ_i , which is equal to his expected contribution to all possible coalitions, assuming that coalitions are formed by adding players at random. It is calculated as

$$\varphi_i = \sum_{S:i \notin S} \frac{|S|!(|N|-|S|-1)!}{|N|!}[v(S \cup i)-v(S)].$$

The Shapley value has a number of desirable features. First, it is unique and always exists. Second, it is the unique solution to the game (N, v), which simultaneously fulfils reasonable criteria such as: the total payoff is distributed (efficiency); interchangeable players receive the same payoff (symmetry); players that contribute nothing to all coalitions receive nothing (dummy player condition); and if one combines coalition games described by different value functions, the players' payoffs of the combined game equal their combined payoffs of the original games (additivity). The last two properties can be replaced with the following (monotonicity): if a player's contribution to all coalitions weakly increases, then his payoff should not decrease (Young 1985). Finally, it is the unique solution featuring balanced contributions, that is, for any two players $i, j \in N$ it is true that i loses as much when j leaves the game as j loses when i leaves the game (Myerson 1980). Hence, if a player objects the Shapley allocation by pointing out the damage he can impose on another player through a boycott of cooperation, his opponent can always counter the argument.

3.2 Calibrating the model

The model is calibrated by calculating the value function for every possible coalition of players. There is considerable uncertainty as to an appropriate estimation of the demand for Russian gas and the cost of producing it in the gas fields of western Siberia.[9] However, these assumptions affect all coalitions in a similar way and have little effect on the relative bargaining power. We use a simple linear function and choose parameters so as to reflect the situation around 2000 and to make existing capacities sufficient (see Hubert and Ikonnikova 2003 for details). Hence, a grand coalition

9 We use a simple linear schedule for the residual demand for Russian gas. Russia cannot raise the price in the short run given its contractual obligations. In the long run, it faces competition from gas suppliers like Norway and Algeria, LNG and other energy sources. Hence, we have opted for a rather flat demand schedule, implying a fairly high price elasticity at current quantities. Our results on the power structure are fairly robust to changes in the elasticity of demand.

of all players would maximise its profit by using the existing capacity of *South* (70 bcm/a) and *Yamal 1* (28 bcm/a), both of which are available at low operating cost. Any additional capacity requires new investment. Up to a limit of about 15 bcm/a, the cheapest option would be *Upgrade*. By modernising compressor stations, the capacity of existing pipelines in Ukraine can be increased at low cost. Beyond that limit, investment in new pipelines is required, for which capital cost are at least twice as high. As for new pipelines, *Yamal 2* is slightly more efficient than further extensions in the south. By far the most expensive option is *Nord Stream*, which requires at least yet another doubling of capital expenditures per unit of capacity. Using this calibration we can calculate how different sub-coalitions would use and extend the system to maximise their payoff.

A grand coalition of all players would maximise profits.

The main results are reported in Table 1. The first column shows the smallest coalition necessary to implement the transport network characterised in the next seven columns. The figures indicate for all pipelines their availability and capacity: "-" means that this link is not available to the coalition, "0" indicates that a link is available, but the coalition chooses not to install capacities. Positive figures indicate usage of existing or investment in new capacities. The last three columns show the resulting price for gas and the payoff, or value, of the coalition. We report the absolute payoff (in million euros) and the relative payoff (in percent), that is, the absolute payoff of a particular coalition relative to the payoff a grand coalition of all players could achieve. In what follows, we will focus on relative payoffs, which largely reflect geography and differences in transportation cost.[10]

Table 1. Coalitions, capacities, prices, and payoffs

Coalition	Pipeline capacity (bcm/a)							Gas price (€/tcm)	Payoff (€ million)	Payoff (% of all players)
	South	South Upgrade	Yamal 1	Yamal 2	Bypass	Baltic	Nord Stream			
R	-	-	-	-	-	-	72	136	3,976	57
R, U	70	15	-	-	-	-	0	132	6,728	96
R, B	-	-	28	-	-	-	45	136	5,039	72
R, B, P	-	-	28	60	-	-	0	131	6,216	89
R, B, U	70	0	28	-	-	-	0	128	6,979	100
R, B, P, S	-	-	28	0	60	-	0	131	6,524	93
R, Li, La	-	-	-	-	-	28	45	136	4,785	69
R, P, Li, La	-	-	-	0	-	88	0	131	5,962	85
All players	70	0	28	0	0	0	0	128	6,979	100

Note: R ≡ Russia, B ≡ Belarus, P ≡ Poland; U ≡ Ukraine; S ≡ Slovakia; La ≡ Latvia Li ≡ Lithuania

On its own, Russia would choose *Nord Stream*, the only pipeline option for which Russia does not need to form a coalition, and install a capacity of 72 bcm/a. This would give Russia an annual payoff of around €4bn, which is equal to 57 percent of the profit of a grand coalition. Russia and Ukraine together would forgo investment in *Nord Stream* and invest in *Upgrade* (that is, the upgrading of the existing *Southern System*). By avoiding the large cost of *Nord Stream* this coalition would achieve a relative payoff of 96 percent. Given that Russia's access to the Polish section of *Yamal 1* is secured by assumption, the coalition of Russia and Belarus would use the existing 28 bcm/a on *Yamal 1* and install a capacity of 45 bcm/a at *Nord Stream*, thereby earning a relative value of 72 percent. If Poland joined that coalition, investment would shift from *Nord Stream* to *Yamal 2*, with a capacity of

10 By contrast, absolute payoffs are sensitive to our assumptions about demand and production costs. Although not shown in Table 1, most coalitions either have equivalent investment opportunities, hence payoffs, or cannot establish a complete supply link and have zero profit. For more details see Hubert and Ikonnikova (2003).

A coalition of Russia, Belarus, and Ukraine would achieve the profits of the grand coalition.

60 bcm/a, yielding a payoff of 89 percent. Not too surprisingly, the coalition of the three major players (namely, Russia, Belarus, and Ukraine) would achieve the value of the grand coalition simply by using the existing capacities.

The *Bypass* would enable Belarus, Poland, and Slovakia to replace the most important transit country, Ukraine, using existing capacities through Slovakia and the Czech Republic. By including Slovakia, the coalition *{R,B,P}* could increase its payoff by 4 percentage points to 93 percent. The difference is modest, but it requires only one additional player. *Baltic* – that is, the link through the Baltic countries – would allow Latvia and Lithuania to replace Belarus in using *Yamal 1* and build a variant of *Yamal 2*. These options would enable the coalition *{R,P,Li,La}* to achieve a payoff of 85 percent. Compared to what Russia can achieve alone, this is an increase of 28 percentage points. The difference is large, but it takes three additional players to achieve it.

3.3 The power structure

In Table 2 we present the players' relative power, as calculated from their Shapley values for various assumptions over the availability of pipeline connections. Since demand for gas and production cost have been chosen to be compatible with current transport capacities, there would be no commercial interest to increase capacity beyond *South* and *Yamal 1*. The available options for investment would not be used. Nevertheless, they have an impact on the power of the players, hence the sharing of profits. Before explaining the details, it is worth mentioning that while we measure the power of a country by its share in total profit from the gas export business, a literal interpretation would be too narrow. Given the complexity of the relations between the countries, it is reasonable to assume that some countries receive their 'share' not in the form of money, but as political concessions on other issues.

Table 2. Countries' profit share (in percent) for alternative pipeline options

	Status quo	Adding one option at a time				
		South Upgrade	Yamal 2	Bypass	Baltic	Nord Stream
Russia	57.1	57.8	60.3	59.2	58.7	79.7
Ukraine	31.8	32.5	22.2	23.2	29.1	15.1
Belarus	11.1	9.6	14.3	13.2	7.5	5.2
Poland	0.0	0.0	3.2	2.1	1.6	0.0
Slovakia	0.0	0.0	0.0	2.1	0.0	0.0
Lithuania	0.0	0.0	0.0	0.0	1.6	0.0
Latvia	0.0	0.0	0.0	0.0	1.6	0.0
	All options	**Excluding one option at a time**				
		South Upgrade	Yamal 2	Bypass	Baltic	Nord Stream
Russia	82.5	82.0	82.0	82.3	82.1	62.7
Ukraine	9.8	9.2	10.6	10.7	10.7	19.2
Belarus	4.3	5.1	3.8	4.1	5.4	10.1
Poland	2.1	2.3	1.6	1.9	1.6	4.9
Slovakia	0.2	0.2	1.1	0.0	0.2	0.2
Lithuania	0.5	0.6	0.5	0.5	0.0	1.4
Latvia	0.5	0.6	0.5	0.5	0.0	1.4

Notes: To recall, the *status quo* is characterised by capacities of 70 bcm/a at *South* and 28 bcm/a at *Yamal 1*. Demand for gas and production cost have been chosen so that these capacities are sufficient and optimal for meeting demand, yielding a total annual profit of €6.979bn, or 100 percent. Figures may not add up due to rounding.

The first column under the heading 'status quo' reports the shares for a situation where the currently existing network cannot be changed, i.e., no new pipelines or upgrading of existing capacity are possible. This corresponds to the shortsighted view on relative power. To interpret the figures, consider the hypothetical situation where all pipelines were to run through Ukraine. In such a bilateral monopoly, Russia and Ukraine would each get 50 percent of the overall profit. With *Yamal 1* in place, however, Ukraine's share is down to 32 percent, while Belarus gains 11 percentage points.[11] However, as the capacity at *South* could not fully replace *Yamal 1* (and *vice versa*), competition between Ukraine and Belarus remains limited. As a result Russia, gains only 7 percentage points compared to a hypothetical bilateral monopoly comprising Russia and Ukraine.

If we adopt the farsighted view of power structure by taking into account the various possibilities to change the transport grid, the picture changes dramatically. The results are reported under the 'all options' heading in the lower part of Table 2. With almost 83 percent, Russia now obtains the lion's share of the profit. Recall that, given our assumption on demand and supply, none of the additional options would materialise. It is the mere possibility to build pipelines – through the Baltic Sea, to increase capacities on *Yamal* and on *South*, or to bypass Ukraine and Belarus – that increases Russia's share by more than a quarter of the total payoff. Given its strength in the current system, it is not surprising that Ukraine's power index suffers most. It is slashed by two thirds, from 32 percent down to 10 percent. With a loss of 7 percentage points, Belarus is also hard hit.

Given their current importance as gas transit countries, Belarus and Ukraine lose most from gas network extensions that can be implemented without them.

The smaller countries – Latvia, Lithuania, and Slovakia – derive their power from making it possible to bypass Belarus and Ukraine. Their payoffs are tiny compared to those of the established transit countries. However, given their much smaller population, the benefits are still very substantial. If measured in per capita terms, the benefits for Lithuania and Latvia fall in between Poland and Ukraine, which may explain their very active lobbying against *Nord Stream*.

3.4 The strategic value of alternative pipeline options

To single out the strategic value of a particular option, we assess its impact on the power structure. One way of doing this is to add one link at a time to the 'status quo' (upper part of Table 2). Alternatively, we can evaluate the strategic value of a particular option in the context of other options by withdrawing one link at a time from the benchmark case 'all options' (lower part of Table 2).

For small additions to the capacity, *Upgrade* is the cheapest, hence, commercially most interesting option. Given that the additional capacity is limited (15 bcm/a), its impact on power is small, whether it is evaluated relative to the status quo or in the context of all other options.

If *Yamal 2* were the only possibility to increase capacity, its strategic impact would be substantial. Adding it to the status quo would cut Ukraine's profit share by more than 9 percentage points. If seen in the context of the other options, however, the impact of *Yamal 2* is small. Comparing the third column to the first in the lower part of Table 2 shows that with *Yamal 2* coming on top of all other options, Russia, Belarus, and Poland would each gain half a percentage point, while Ukraine and Slovakia would lose 0.8 percentage points and 0.5 percentage points, respectively. Slovakia's loss indicates that the strategic value of *Yamal 2* is related to *Bypass*, the only pipeline that requires the involvement of this country.

11 Recall that by assumption Poland cannot obstruct the use of *Yamal 1*, hence it cannot derive any power from threatening to do so. However, in negotiations before the pipeline was built, it had secured a share, which is not accounted for in our model.

Like *Yamal 2*, *Bypass* is important only if it is the sole addition to the status quo. Assessed in the context of all other options, however, its impact on Russia is negligible – and even Slovakia gains very little. This is because all coalitions that can realise *Bypass* (that is, *{R,B,P,S}* and the coalition of all players) can also realise *Yamal 2*, which is just marginally less profitable.

Considering *Baltic,* which inculdes a variation of *Yamal 2* that could be realised without Belarus, we find that Poland, Latvia, Lithuania, and Russia would each gain 1.6 percentage points if *Baltic* were to happen as the only addition to the status quo. These gains would come at the expense of Ukraine and, not surprisingly, Belarus. Examining *Baltic* in the context of all other options shows that it would add around half a percentage point to the profit share of Poland, Latvia, Lithuania, and Russia. As for Russia, this change is marginal. But as argued above, for small countries, like Latvia and Lithuania, the benefits of even a small share in profits are substantial.

By far the strongest impact on the power structure comes from Nord Stream.

By far the strongest impact on the power structure comes from *Nord Stream*. In isolation, it raises Russia's claim on profits from 57 percent to 80 percent. Even when assessed in the context of all other options, *Nord Stream* raises Russia's share from 63 percent to almost 83 percent. It is more important than all other options together. Correspondingly we observe a considerable decline in the bargaining power of Ukraine, Belarus, Poland, Lithuania, and Latvia – as indicated by the sharp change in their profit shares. Only Slovakia remains unaffected. The strategic importance of *Nord Stream* explains Russia's continued interest in a project that from a naïve point of view makes little economic sense due to its high cost.

3.5 The role of geography and cost

To further develop the intuition for the numerical results, it is useful to consider how geography and cost interact in determining bargaining power. For simplicity, we consider only four main countries: Russia, Belarus, Poland, and Ukraine. To isolate the role of geography, suppose gas transport costs are the same for all pipelines, so that total profit is the same whatever pipeline is used. If the only possible transport route was through Ukraine, Russia and Ukraine would each get half of the profit. If we add an equally efficient route through Belarus and Poland, Russia would obtain 7/12 of the profit, Ukraine 1/4, and Belarus and Poland would share the rest equally, obtaining 1/12 each.[12] Ukraine thus suffers a lot from the competing route, but Russia has to share the gain with the other two transit countries. Finally, if Russia establishes a direct offshore link on its own, it obtains the whole profit, as there is no need to share with anyone.

Now, assume that all options are available, but pipelines differ in their cost, either because of different conditions (offshore *vs*. onshore) or because investment costs are already sunk as in the case of *South* and *Yamal 1*. Loosely speaking, Russia would start with the profit it would get if *Nord Stream* were the only connection to Western Europe. While *Nord Stream* looks inefficient compared to options such as *Upgrade* and *Yamal 2*, standing alone, it would be a highly profitable project. Given our calibration, it already yields 57 percent of the profit of the most efficient transportation network – that is, the network resulting from a coalition of all players.[13] In addition, Russia will obtain 7/12 of the increase in profit obtained from switching from the offshore to the next best onshore option, *Yamal*, which increases profits to 89 percent. Finally, Russia would enjoy 1/2 of the increment achieved from using the most efficient solution, which includes the system in the south. Summing

12 With payoffs normalised to one, the value function would be: $\pi(R, B, P) = \pi(R, U) = \pi(R, P, U) = \pi(R, B, U) = \pi(R, B, P, U) = 1$ and zero in all other cases, yielding Shapley values of $\varphi_R = 7/12$, $\varphi_B = \varphi_P = 1/12$, and $\varphi_U = 1/4$.

13 These payoffs are again taken from Table 1.

up, Russia's profit share would amount to 81 percent.[14] Hence, we obtain almost the same figure as the one following from the more complex model reported in Table 2. Enriching the analysis with more players and pipeline options adds to its realism, but the quantitative results change only marginally.

4. Dynamic analysis: hold up, strategic investment, and dynamic cooperation

4.1 Dynamic cooperation in the absence of legally enforceable agreements

So far, we have developed a method for analysing the relation between network architecture and power and explored the strategic relevance of pipeline options. However, provided that all participants understand this relation, there would be no need to deviate from profit-maximising investment. Hence, our analysis cannot explain why *Nord Stream* is built. In this section, we extend the model to analyse the role of commitment and strategic investment within a dynamic framework. For simplicity we consider only those players and pipeline options that turned out to be the most relevant in the previous section. That is, we consider Russia, Ukraine, Poland, and Belarus – countries that control three links: *Nord Stream*, *Yamal* (specifically: the existing *Yamal 1* and a possible *Yamal 2*), and *South* (specifically: the existing infrastructure and a possible upgrade).

Rational players will anticipate the impact of current actions on future bargaining and try to extend their cooperation beyond the current period.

As before, we assume that, in any given period, the players operate the existing transport system efficiently and share profits according to the Shapley value. In doing so, they take the existing capacities as given. In other words, in every period the players play a 'status quo' multilateral bargaining game. In the dynamic setup, however, they can also invest in new capacities, which will become available with some delay. Furthermore, they can grant each other long-term access rights. Both activities together will determine the 'status quo' of future games.[15] Rational players will anticipate the impact of current actions on future bargaining and, hence, try to extend their cooperation beyond the current period. Cooperation spanning over long periods of time requires credible commitments to make payments in the distant future, not to obstruct access to pipelines, and to stick to an investment schedule. In principle, there are two mechanisms to coordinate activities in such a dynamic framework. One rests on explicit contracts, which are enforced by independent institutions. The other is based on agreements, enforced by the mutual threat of the participants to terminate cooperation. We consider both mechanisms in turn.

Since the players are sovereign nations, long-term cooperation based on explicit agreements requires a strict rule of law, independent from political influence, within a country – or even better, international institutions to enforce such contracts. In the absence of such an institutional framework, countries lack the ability to make credible commitments. Since the focus in this section is on investment and not the sharing of profits, we assume that all players who could, in principle, enter explicit long-term agreements form a coalition that determines investment and network access rights to the best advantage of the whole group.[16] All other players remain on their own. In equilibrium, the coalition and all single players maximise their expected payoffs (net of initial

14 57% + 7/12* (89% - 57%) + 1/2* (100% - 89%) = 81%

15 In order to focus on the dynamics of strategic interaction, we assume that the economic environment is stationary, i.e., we abstract from demand growth, depletion of gas fields, technical progress, and so on.

16 Here, the notion 'coalition' has a different meaning than in Section 3.2, where it referred to all possible sub-coalitions in a bargaining situation. Now, we consider a group of players forming a 'pre-coalition' or 'strategic alliance' in advance of such a bargaining situation in order to change the bargaining game to its own advantage. It is straightforward to calculate how such a coalition would share its joint profit using the principles used for calculating the Shapley value (Owen 1977).

investment cost) from future bargaining – given the equilibrium strategies of the other actors (non-cooperative Nash-equilibrium). In Hubert and Ikonnikova (2004) and Hubert and Suleymanova (2006) it is shown that this may involve underinvestment in cheap pipelines as well as overinvestment in expensive ones. The possibility of underinvestment is a variant of the well-known hold-up problem. If the returns on investment are *ex ante* not contractible but shared according to some bargaining rule, the incentives to invest are lower than they would be otherwise. Overinvestment results from an attempt to create countervailing power.

With repeated interaction, cooperation can be supported by mutual threats to retaliate if others defect.

With repeated interaction, cooperation can be supported by mutual threats to retaliate if others defect. Following the literature on industrial organisation, we will refer to this form of cooperation as collusion. The aim of collusion is to raise the profits of the supply chain and avoid the inefficiencies associated with strategic investment. Since the enforcement of collusion does not depend on independent institutions, all players and the coalition can be involved. However, to sustain cooperation, threats have to be credible and the long-term gains from sticking to cooperation have to outweigh the short-term gains from deviation, which would be followed by retaliation from other players. Due to these constraints, collusion is more limited in scope than legally enforceable explicit agreements. We envisage a tacit agreement on a system of transfers and investments between the coalition of players able to commit and all other players. This tacit agreement is supported by the mutual threat to revert to non-cooperative behaviour if one party deviates from the agreement (so called trigger strategies). In practical terms, cooperation is terminated if one player starts bargaining for an increase of his assigned share or if one player deviates from the agreed investment schedule. The former is obvious. It is not possible to increase the share of one party without renegotiating all payments. The latter is because investment in transport capacity is easily observable. Upon observing that a player deviates from cooperative investment, all others anticipate that cooperation will fail once the capacities become available. Backward induction leads them to defect immediately.

While cooperation breaks down immediately, the full impact will be felt only with a delay. Initially, non-cooperative payments reflect players' bargaining power for given transport capacities. At this juncture, a deviating party would make a profit. Once capacities increase to their non-cooperative level, profits decrease. This corresponds to the punishment phase. To sustain collusion, the present value of future income from collusion must not be less than what can be obtained by defecting. Unfortunately, the equilibrium in dynamic strategies is not unique. In the following, we will focus on two borderline cases. At one extreme, players fail to collude, hence, only those players who can make long-term commitments cooperate. At the other, players collude to support the highest profits compatible with the dynamic incentive constraint.

All this leaves us with a fairly complex framework for the analysis. First, we have to decide which countries are able to commit and, hence, form a coalition. Second, we have to determine how the members of this coalition assign access rights so as to enhance their bargaining power as a group in future negotiations. Third, we have to analyse how the players invest, taking into account the impact of capacities on bargaining power for a given access regime. Finally, we have to distinguish between two possible equilibria: (i) a non-collusive one, in which players cooperate only within coalitions and act non-cooperatively *vis-à-vis* outsiders; (ii) a collusive outcome, characterised by the highest total profit that trigger strategies can support.

As to the ability to make credible long-term commitments, we consider four scenarios. As a benchmark (case 1), we assume that all countries can commit and, hence, form a grand coalition that invests so that total profits of the supply chain are maximised. In our standard scenario (case 2), we assume that only Poland can make long-term commitments. As an EU member, Poland can give private companies, even if they were fully owned by a foreign state, considerable legal protection through its and the European legal system. In this case, we thus have a coalition between Russia and Poland *{R,P}*.[17] For comparison, we look at two enlargements of this coalition. In one (case 3), Belarus can commit – in addition to Poland – and thus forms a coalition with Russia and Poland *{R,P,B}*. In the other (case 4), Ukraine can commit – in addition to Poland – and thus forms a coalition with Russia and Poland *{R,P,U}*. The last two cases appear unrealistic under current political circumstances. That said, case 3 might reflect the situation in the mid-1990s, when Belarus' independence from Russia was perceived to be very limited, so that opportunistic re-contracting was not considered a threat. And case 4 could be seen as reflecting a situation in which Ukraine, moving towards the European Union, subjects itself to international arbitration.[18]

4.2 Equilibrium investment in the Eurasian gas transport system

We calibrate the model using the same assumptions on pipeline cost as in Section 3. In addition, we assume a three-year construction period for pipelines, that is, pipelines come on stream three years after investment decisions have been taken. As for gas demand, we envisage substantial growth, such that a capacity increase of about 30 percent (or 30 bcm/a) will become necessary. These assumptions should be taken as an upper bound for reasonable expectations around 2005 (for further details on calibrating the model see Hubert and Suleymanova 2006).

Computed payoffs reveal that strategic considerations are of considerable importance for investment in the Eurasian gas transport network.

The equilibrium capacities under non-cooperative behaviour ('no collusion') and cooperative behaviour ('collusion') are displayed in Table 3. We start from the existing capacities *South* (70 bcm/a), *Yamal* (28 bcm/a), and *Nord Stream* (0 bcm/a) – to which we add the equilibrium investment. The figures reveal that strategic considerations are of considerable importance for investment in the Eurasian gas transport network. To start with the benchmark (case 1), in which all countries can commit and, thus, act like an integrated monopoly against customers in Western Europe. The profit-maximising capacity is 128 bcm/a, which is provided at minimal cost. The optimal network extension is to, first, upgrade *South* by 15 bcm/a and, then, expand *Yamal* by another 15 bcm/a. It follows that *Yamal 2* with its planned capacity of 28 bcm/a (see Section 2.2) would be built only some time after upgrading the *Southern System* and slowly taken to full capacity.

Now consider the most realistic scenario in which only Poland can commit (case 2). In the no-collusion variant, the players fail to upgrade *South* or invest in *Yamal*. Instead, *Nord Stream* is built with a staggering capacity of 80 bcm/a. With 178 bcm/a, the aggregate capacity is much larger than the profit-maximising one. Furthermore, almost one-fifth of the capacity would not be needed for transporting gas to Western Europe. Given our assumption of a fairly elastic long-run demand for Russian gas, the impact on quantities is stronger than the impact on prices.[19] The equilibrium price is 5 percent below the monopoly price, which is the price resulting from a grand coalition, while the quantity of Russian gas is up by 13 percent.

17 In such a coalition Poland would commit to grant access at a rate less favourable than its short-term bargaining position suggests. As a result, Russia's ability to make long-term commitments is not at stake.

18 Hubert and Ikonnikova (2004) analyse in some detail how the different coalitions would optimally modify the access regime. Using results from Segal (2003), they prove that: (i) the coalition *{R,P,B}* would grant Russia access rights to the sections of *Yamal* in Poland and Belarus; (ii) the smaller coalition *{R,P}* would not change the natural access regime (as this would weaken the bargaining power of coalition members because Belarus is complementary to Poland in the presence of Russia); (iii) the coalition *{R,P,U}* would grant Russia access rights to *South*.

19 Specifically, the following linear function has been assumed: $p = 170 - 0.35q$, which is quite elastic at the current quantities.

Given this huge overinvestment, the supply chain could gain much from collusion. In the collusive equilibrium, transit countries voluntarily restrain their claims on profits, and Russia abstains from investing in *Nord Stream*. The credible threat to massively invest in *Nord Stream* is a strong enough deterrent for Ukraine and Belarus not to exploit their bargaining position to the full. The threat is so powerful, that it is even possible to increase the capacity of *South* in the Ukraine, first, by renovating the compressor stations (15 bcm/a). The profit-maximising network extension of case 1, however, is not feasible in the collusive equilibrium. Rather than switching to *Yamal* after exhausting the cheap upgrading option at *South*, the players continue to invest in *South* by installing new pipelines with a capacity of 12 bcm/a. As a result, the collusive equilibrium yields a network capacity that is even lower than in the profit-maximising equilibrium. Tight capacities and slightly higher prices make it the worst-case scenario for customers in the West.

Table 3. Equilibrium gas transport capacity [bcm/a] under alternative scenarios

	South	Yamal	Nord Stream	Total	Used
Case 1: all countries can commit *{grand coalition}*	70+15	28+15	0	128	128
Case 2: Poland can commit *{R,P}*					
No collusion	70	28	0+80	178	145
Collusion	70+15+12	28	0	125	125
Case 3: Poland and Belarus can commit *{R,P,B}*					
No collusion	70	28+85	0	183	145
Collusion	70+15	28+21	0	134	134
Case 4: Poland and Ukraine can commit *{R,P,U}*					
No collusion	70+15+23	28	0	136	136
Collusion	70+15+15	28	0	128	128

Notes: R ≡ Russia, B ≡ Belarus, P ≡ Poland; U ≡ Ukraine; Letters in { } show members of coalition.

That investment in Nord Stream is well under way strongly suggests that Russia and key gas transit countries failed to realise the potential of dynamic cooperation.

How do real-world investment patterns compare to the implications of our analysis? That investment in *Nord Stream* is well under way strongly suggests that Russia and key gas-transit countries failed to realise the potential of dynamic cooperation. For all our assumptions on countries' ability to commit, investment in *Nord Stream* could have been avoided through dynamic collusion. Not surprisingly, the countries also failed to prevent underinvestment in *South* and *Yamal*. However, the magnitude of real-world investment is below our prediction for the non-cooperative equilibrium. We obtain a non-collusive investment of 80 bcm/a in *Nord Stream*. Current plans for this project are 60 bcm/a – of which half will be established in the first step. The difference may indicate that the countries managed to maintain at least a low degree of collusion. Alternatively, the demand assumption underlying our model might be too high compared to actually expected demand.

Finally, we turn to the role of commitment. All non-collusive equilibria feature large overinvestment to create countervailing power. If both Belarus and Ukraine cannot commit, countervailing power is created by investing in *Nord Stream*. If only Ukraine cannot commit (case 3), *Yamal* provides the leverage, and if only Belarus (case 4) is prone to re-contracting, expanding *South* provides countervailing power. In contrast, all collusive equilibria feature capacities close to those following from the profit-maximising benchmark (case 1). Nevertheless, the ability to commit determines which pipeline is expanded.

In the early 1990s, Belarus' independence from Russia was limited. Apparently, the players underestimated the risk from re-contracting. Otherwise, investment in *Yamal 1* cannot be explained in our framework. Currently, the country looks increasingly isolated from the West, which may force it back into Russia's arms. It is difficult to say whether this would make opportunistic re-contracting *vis-à-vis* Russia less likely. In any case, the development of *Yamal 2* has a chance only if Belarus is perceived as a partner able to make long-term commitments. This holds true independently of the type of equilibria in the market.

Although not very likely in the near future, Ukraine may fully implement the European Energy Charter – even move towards closer integration with EU. By providing a framework for international contract enforcement, the charter may enable Ukraine to credibly enter into long-term agreements, which, in turn, is a precondition for investment in *South* in any of the no-collusion equilibria. However, preliminary calculations show that it may already be too late for both countries. Once *Nord Stream* is operating with a capacity of 30 bcm/a, it makes little sense to invest in *South* or *Yamal* unless demand grows well beyond our assumptions.

5. Conclusions

Russian gas, currently pumped through Ukraine and Belarus, makes a very important contribution towards the energy needs of Western Europe, creating concerns about energy dependency, market power, and security of supply.

In the first part of the paper, we develop an analytical framework to analyse power in the supply chain of Eurasian gas. Applying cooperative game theory for multilateral negotiations, we derive the bargaining power of the different players endogenously from the architecture of the transport system and its possible extensions. As a next step, we quantify the strategic importance of each single option to extend the grid, by calculating how it changes the distribution of the profit. The most important lesson from this exercise is that pipelines have to be evaluated in the context of the whole network. *Bypass* – a possible pipeline through Poland and Slovakia – is explicitly designed to shortcut Ukraine. What at first glance may look as a powerful threat to Ukraine's strong position in the current network turns out to have very limited strategic relevance. Slightly more important is the option to extend the capacities on *Yamal*, which is also commercially attractive. However, by far the strongest impact on the bargaining power is exerted by *Nord Stream*. Although this project cannot compete commercially with the other options to increase transport capacity, it strengthens Russia's position more than all other options together. In a nutshell: competition between Belarus and Ukraine is of little strategic importance compared to an option for direct Russian access to customers.

In the past, gas transport through the extended pipeline system has been interrupted occasionally when Russia and transit countries failed to reach agreement on their own gas prices and transit fees. These very rare, very short, but highly publicised events gave the impression that due to conflicts along the transit routes, Russian gas is unreliable and expensive. The dynamic model analysed in the second part of this paper suggests the opposite may be true. Because the members of the supply chain for Russian gas failed to develop a stable long-term cooperation, the pipeline system is expanded and diversified beyond what is in the interest of Russia, Ukraine, and Belarus as group. Investment is partially driven by strategic considerations to increase bargaining power *vis-à-vis* transit countries, rather than consumers. As a result, Western European energy consumers will benefit, both in terms of prices and energy security, from a diversified transport system with substantial spare capacities. At the same time, energy dependency will grow because the fraction of Russian gas in the energy mix becomes larger.

It appears that the pipeline system is being expanded beyond what is in the interest of Russia, Ukraine, and Belarus as a group, with investment partly aiming at increasing bargaining power vis-à-vis transit countries, rather than consumers in the West.

However, potential transit countries in the European Union such as Poland, the Czech Republic, and Slovakia clearly suffer from the lack of an international institutional framework that would allow their eastern neighbours to make credible long-term commitments. If Ukraine and Belarus had developed a strong and commercially sound gas relationship with Russia in time, investment would have gone into *South* and *Yamal* rather than into *Nord Stream*.

References

Hubert, F. and Ikonnikova, S. (2003). “Strategic investment and bargaining power in supply chains: A Shapley value analysis of the Eurasian gas market”. Humboldt-University Berlin Discussion Paper.

Hubert, F. and Ikonnikova, S. (2004). “Hold-up, multilateral bargaining, and strategic investment: the Eurasian supply chain for natural gas”. Humboldt-University Berlin Discussion paper.

Hubert, F. and Suleymanova, I. (2006). “Strategic investment in international gas transport systems: A dynamic analysis of the hold-up problem”. Humboldt-University Berlin Discussion paper.

Myerson, R. (1980). “Conference structures and fair allocation rules”. *International Journal of Game Theory*, (9:3), pp.169-182.

Segal, I. (2003). “Collusion, exclusion, and inclusion in random-order bargaining”. *Review of Economic Studies*, (70), pp. 439-460.

Shapley, L.S. (1953). “A value for n-person games”, in Kuhn, H.W. and Tucker, A.W. (eds.), *Contributions to the theory of games II*. Princeton University Press, Princeton, USA.

Stern, J.P. (1995). *The Russian gas bubble*, The Royal Institute of International Affairs, London, UK.

Stern, J.P. (1999). “Soviet and Russian gas: The origins and evolution of Gazprom’s export strategy”, in Mabro, R. and Wybrew-Bond, I. (eds.), *Gas to Europe. The strategies of four major suppliers*, Oxford University Press, New York, USA.

Stern, J.P. (2005). *The future of Russian gas and Gazprom*, Oxford University Press, New York, USA.

Victor, N.M. and Victor, D. (2006) “Bypassing Ukraine: Exporting Russian gas to Poland and Germany”, in Victor, D. and Jaffe, A. and Hayes, M. (eds), *Natural gas and geopolitics – From 1970 to 2040*. Cambridge University Press.

Yafimava, K. and Stern, J. (2007). “The 2007 Russia-Belarus gas agreement”. Oxford Institute for Energy Studies, Oxford Energy Comment, January 2007

Young, H.P. (1985). “Monotonic solutions of cooperative games”. *International Journal of Game Theory*, (:2), pp. 65-72.

ABSTRACT

Improving energy efficiency is seen as a core strategy for a sustainable energy system, *because it may contribute to cost savings for companies and private households, cost-effectively reduces greenhouse gas emissions and other pollutants, increases security of supply for required energy services. The thrust of engineering-economic analyses suggests that there is a large potential for energy efficiency measures that are also profitable, but – because of barriers to energy efficiency – are not being adopted. This paper presents a taxonomy of these barriers, distinguishing between barriers that would warrant policy intervention and those that do not. As a case study, barriers to energy efficiency in the German higher education sector and measures to overcome those barriers are discussed.*

Joachim Schleich (joachim.schleich@isi.fraunhofer.de) is Senior Researcher at the Fraunhofer Institute for Systems and Innovation Research, Karlsruhe, Germany, and Adjunct Professor at Virginia Polytechnic Institute and State University, Blacksburg, Virginia, USA. This paper draws in parts on Chapter 2 and Chapter 3 of Sorrell *et al.* (2004) and Schleich *et al.* (2001)

The economics of energy efficiency: barriers to profitable investments

Joachim Schleich

1. Motivation

In October 2006, the European Commission published its Action Plan for Energy Efficiency (European Commission 2006a) to help realise savings in energy use in the European Union of at least 20 percent by 2020 compared to the baseline. The Action Plan outlines a framework of policies and measures for all end-use sectors (residential, tertiary, industry, transportation) and the transformation sector to improve energy efficiency. This was previously called for in the Commission Green Paper on "A European Strategy for Sustainable, Competitive and Secure Energy" in March 2006 (European Commission 2006b) and the 2006 Spring European Council Presidency Conclusions (European Council 2006). Improving energy efficiency, that is, obtaining more energy services such as heat, light or mobility for the same or less energy, is seen as the fastest and often most cost-efficient way to achieve a sustainable energy system.

According to the European Commission (2006a), the 20 percent energy savings by 2020 are in addition to savings induced by price effects, structural change in the economy, natural replacement of technology, and measures already in place. Results from technology-based, engineering-economic (bottom-up) modelling analyses suggest that in total, the energy savings compared to the baseline scenario would mean annual savings of around 390 million tonnes of oil equivalent (Mtoe), most of it in the end-use sectors. The residential and commercial buildings sector exhibits the largest relative cost-effective potentials – 27 percent and 30 percent, respectively. The most important measures include retrofitted wall and roof insulation for residential buildings and improved energy management systems in commercial buildings. Energy savings in industry amount to 25 percent, where measures not specific to the industry concerned – such as high efficient motors, fans, and lighting – offer the most important savings potential. Finally, the estimated savings of 26 percent in the transport sector are – to a large extent – the result of shifting to other modes of transportation.

These savings in energy use correspond to direct energy cost savings of more than €100 billion per year by 2020[1], which the Commission estimates to more than compensate the additional costs for the required investments in energy efficiency over time. Energy cost savings would then translate into improved competitiveness for companies and lower expenses on energy for households, making companies and households less vulnerable to energy price hikes in the future.

The 20 percent potential in 2020 corresponds to reductions in CO_2-emissions of 780 million tonnes, which would be more than twice the reductions required by the Kyoto Protocol by 2012 for the European Union (European Commission 2006a). In general, most other studies also find improved energy efficiency to be the single largest source of fossil fuel-related greenhouse gas emissions savings, at least until 2050. Typically, energy efficiency tends to account for 30-50 percent of all emission reductions in such technology-based models (HM Treasury 2006, IEA 2006a, IPCC 2001).

Ongoing geopolitical crises such as conflicts in the Middle East and recent disputes over gas supply from Russia to Ukraine and Belarus have lead to an increased focus on the volatility of international

1 At $ 48/barrel.

Higher energy efficiency will reduce import dependency and increase security of supply.

energy markets and the security of supply. By 2030, more than 80 percent of the natural gas and more than 90 percent of the oil used in the EU will be imported, most of it from politically sensitive regions (European Commission 2006b). Obviously, increased energy efficiency will reduce import dependency and increase security of supply for required energy services.

Investments in energy efficiency will bring other benefits, notably a decline in other, local, pollutants such as nitrogen oxides and sulphur. Also, substituting fuel imports for investments in energy efficiency, tends to increase domestic production and employment, in particular in the construction, electrical, and mechanical engineering sectors. Finally, improved energy efficiency at home may also lead to increased export opportunities for new, energy-efficient technologies via first-mover-advantages.

The modelling results, on which the Action Plan for Energy Efficiency is based, suggest that there is a large potential of energy-efficient measures – an energy efficiency gap – which may be realised at low or even negative costs (so called 'no-regret' potential). This raises several questions (see also Sorrell *et al*. 2004).

First, do individuals and organisations really 'leave money on the floor' by neglecting cost-effective measures to improve energy efficiency? Second, what is the nature of the 'barriers' to energy efficiency, that is, the mechanisms which inhibit a decision or behaviour that appears to be both energy efficient and profitable under existing (and expected) economic conditions? Third, do these barriers hinder an efficient resource allocation? And, if so, can these barriers be overcome by adequate policy intervention?

These questions lie at the heart of recent and current policy debates over energy and climate policies and are a focus of continuing dispute within energy economics, with purely technology-based bottom-up modellers on one side of the spectrum, and rather aggregate economic top-down models on the other side. To verify the claims of bottom-up type modelling, it would first be necessary to show that barriers explain the lack of investment in cost-effective energy-efficient technologies; second, that these barriers should be overcome because they inhibit economic efficiency; and third that they could be cost-effectively overcome by non-price measures, which would be the case if the benefits from implementing those measures outweigh the costs.

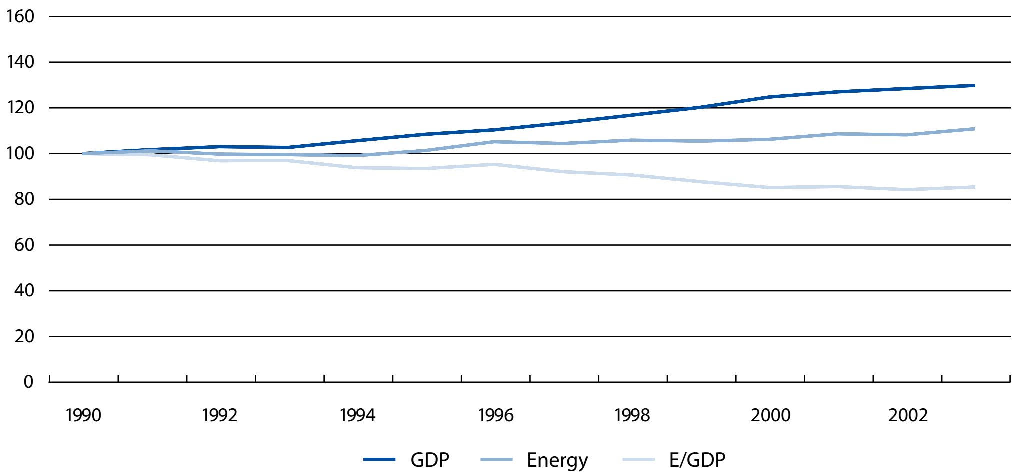
The remainder of the paper is organised as follows. Section 2 briefly portrays recent historic trends in energy efficiency and the underlying reasons. It also includes a short description of the main differences between bottom-up and top-down models and their implications. Section 3 offers a taxonomy of barriers to energy efficiency based on neoclassical, institutional, and behavioural economics. Barriers that hinder economic efficiency and thus might be addressed by policy intervention will be identified. As an illustration, Section 4 presents results from a case study on barriers to energy efficiency in the German higher education sector and includes suggestions for policies to overcome those barriers. The concluding section points to limitations and future research.

2. Trends in energy efficiency and approaches to energy modelling

A common indicator to portray and compare energy performance over time and across regions at a rather aggregate level is energy intensity, i.e., the quantity of energy use divided by gross domestic product (GDP). Since improvements in energy efficiency in processes and equipment will translate into observed changes in energy intensity, energy intensity is often used as a proxy for energy efficiency. When there is structural change in the economy – for instance, an increase in the share of the (less energy-intensive) services sectors such as banking or insurance services and a decrease in (energy-intensive) manufacturing sectors such as steel or cement production – the observed energy intensity would change, even if the quantity of energy used to produce one tonne of steel or cement remains unchanged. This should be kept in mind when looking at the development of energy intensities in the EU and the United States in Figure 1 and Figure 2. In both regions energy intensity decreased in the 1990s, but has almost been stagnating since then. For the EU-25, the average annual reduction in energy intensity since 1990 is 1.2 percent compared to 1.8 percent for the United Sates and 1.6 percent for all OECD countries. Lower reduction rates in the EU may be rationalised by the lower starting level in 1990 and the fact that additional reductions become increasingly more difficult to achieve.

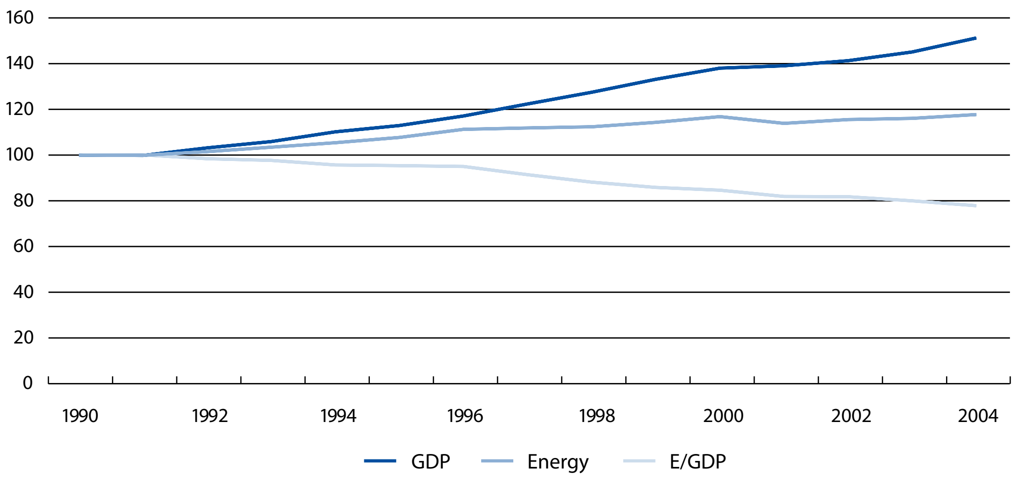
In the 1990s, energy intensity declined in the EU and the United States, but has stagnated since then.

Figure 1. Energy intensity in the EU 25 (1990=100)



Source: Own calculations based on data provided by EUROSTAT

Figure 2. Energy intensity in the United States (1990=100)



Source: Own calculations based on data provided by Energy Information Agency

The targeted reduction in energy intensity under the EU Action Plan for Energy Efficiency appears to be quite ambitious …

In its latest World Energy Outlook, the IEA (2006b) assumes that until 2030 the average annual reduction in energy intensity in the EU will be 1.3 percent in the baseline scenario, which is based on existing policies to enhance energy efficiency. With an implementation of additional policies currently considered in those countries, the annual average reduction is expected to be 1.6 percent. The estimates by the European Commission as given in the Action Plan for Energy Efficiency (European Commission 2006a) are more optimistic: in the baseline scenario, the rate is 1.8 percent; but with an implementation of the Action Plan and, thus, energy savings of 20 percent compared to the baseline, energy intensity is expected to fall at an annual average rate of 3.3 percent. Thus, in light of historic developments and of estimates by other organisations, the implied reduction in energy intensity in the Action Plan for Energy Efficiency appears to be quite ambitious.

Against this background, it is useful to sketch different approaches to energy modelling. Energy and climate policy makers are typically interested in the influence of proposed policies (for instance, taxes on energy or CO_2 emissions) on individuals' and companies' decisions. In particular, they want to know the effectiveness and the costs of these policies. Historically, two types of models were developed to address these questions: 'bottom-up' models and 'top-down' models.

Conventional bottom-up models are engineering-economic models that describe current and future energy demand and supply technologies in detail. They simulate the ageing and replacement of these technologies, thereby assuming that the cost of meeting the demand for energy services (from all end-use sectors) are minimised. Bottom-up models allow for 'no-regret' opportunities and are able to portray the possibilities for a radically different technology stock in the future. However, bottom-up models have been criticised for their lack of adequately describing microeconomic decision-making behaviour of companies and individuals (for example, they do not allow for transaction costs) and their lack of macroeconomic feedback – such as income effects, price effects, or international trade[2].

In contrast, macroeconomic top-down models are able to model microeconomic behaviour and macroeconomic feedback mechanisms, but are more aggregate. Rather than including specific technologies, top-down models indirectly reflect production possibilities via production or cost functions, elasticities of substitution (between capital, labour, and energy), and parameters reflecting technological progress over time. As a result, they allow for the decoupling of GDP and energy as shown in Figure 1 and Figure 2. Computable General Equilibrium (CGE) models, which have come to dominate the top-down approach in recent years, typically imply that agents react perfectly rational to prices and also that markets are always in equilibrium (abstracting, for example, from unemployment). By design, CGE models do not allow for no-regret opportunities, which is one of the reasons why top-down models tend to show higher costs of climate policy than bottom-up models.

As summarised by Hourcade *et al.* (2006), conventional bottom-up models perform well in terms of technological explicitness, but they lack microeconomic realism and – in particular – macroeconomic completeness. In contrast, conventional top-down models perform well in terms of macroeconomic completeness, but to some extent lack microeconomic realism and entirely lack technological explicitness. In recent years, substantial modelling efforts have been made to reconcile the bottom-up and top-down approaches via hybrid models. Some top-down models now specifically represent energy-supply and energy-demand technologies, and allow for technological change to be included explicitly – rather than through a time trend or a fixed coefficient. Some bottom-

2 For a brief overview see also Hourcade *et al.* (2006).

up models have made considerable progress towards including behavioural parameters to better portray microeconomic decision making.[3] Likewise, top-down models and bottom-up models may be linked to get the best of both worlds. For example, the results of a bottom-up model (for example, investments in and prices for energy) enter a top-down model as an input. Ideally, the results of a top-down model are then fed back to the bottom-up model; this process is iterated until convergence is achieved.

Achieving the ambitious targets of the European Commission's Action Plan for Energy Efficiency and the considerable potential for energy savings often suggested by bottom-up models is perhaps less daunting than one may think at first glance if there are easy to remove barriers to energy efficiency. But what is the nature of such barriers, and are they easy to remove? We shall see next.

… but achieving it is perhaps possible if there are easy to remove barriers to energy efficiency.

3. Taxonomy of barriers to energy efficiency[4]

This section develops a taxonomy of barriers to energy efficiency, drawing on concepts from neo-classical economics, institutional economics (principal-agency theory and transaction cost economics), and behavioural economics. These barriers and their implications are described – without formally developing their grounding within the various strings of economic theory, but noting that there is a great deal of overlap between these concepts.[5] The barriers to be discussed represent potential answers to one or more of the following questions:

- Why do organisations impose very stringent investment criteria for projects to improve energy efficiency?
- Why do organisations neglect projects that appear to meet these criteria?
- Why do organisations neglect energy-efficient and apparently cost-effective alternatives when making broader investment, operational, maintenance, and purchasing decisions?

Table 1 presents an overview of barriers to energy efficiency as developed by Sorrell *et al.* (2004).

In any case though, such barriers are more likely to be found in organisations where the share of energy costs in total production costs is low – such as in the services sectors, public administrations, or in industries like mechanical engineering and the food sectors. In contrast, the considerable importance of energy costs in energy-intensive industries – electric power, iron and steel, and mineral processing, for example – provides a strong economic incentive to find and realise efficiency potentials. Likewise, since investing in energy efficiency directly affects the core production processes in energy-intensive companies, energy use is automatically considered in investment decisions. Bearing this in mind, let us examine various barriers one by one.

3 A Special Issue of The Energy Journal (October 2006) is devoted to hybrid modelling of energy-environment policies. Likewise, the Special Issue of The Energy Journal (March 2006) presents several innovative ways to model technological change in (long-run) climate policy analyses (Endogenous Technological Change and the Economics of Atmospheric Stabilisation, Special Issue, The Energy Journal, March 2006).

4 Parts of this section summarise the main concepts and arguments presented in Sorrell *et al.* (2004).

5 See, for example, Chapter 2 in Sorrell *et al.* (2004) for a thorough development of these concepts in relation to barriers to energy efficiency.

Table 1. Taxonomy of barriers to energy efficiency

Barrier	Claim
Risk	Short paybacks required for energy efficiency investments may reflect a rational response to higher technical or financial risk and business and market uncertainty.
Imperfect information	Lack of information on energy efficiency opportunities may lead to cost effective opportunities being missed.
Hidden costs	Engineering-economic analyses may fail to account for either the reduction in utility associated with energy efficiency technologies, or the additional costs associated with them. As a consequence, the studies may overestimate the energy efficiency potential. Hidden costs (to observer!) include overhead costs for management, disruptions to production, staff replacement and training, and the costs associated with gathering, analysing, and applying information.
Access to capital	If organisation cannot raise sufficient external funds, energy-efficient investments may be prevented from going ahead. Investment could also be inhibited by internal capital budgeting procedures, investment appraisal rules, and the short-term incentives of energy management staff.
Split incentives	Energy efficiency opportunities are likely to be foregone if actors cannot appropriate the benefits of the investment. For example, if individual departments within an organisation are not accountable for their energy use, they will have no incentive to improve energy efficiency.
Bounded rationality	Owing to constraints on time, attention, and the ability to process information, individuals do not make decisions in the manner assumed in classical economic models. As a consequence, they may neglect energy efficiency opportunities, even when given good information and appropriate incentives.

Source: Based on Sorrell *et al.* (2004).

3.1 Risk

***Although high discount rates are often observed for investments in energy efficiency, they are not a source of the energy efficiency gap* per se.**

High (implied) discount rates are often observed for investments in energy efficiency. In essence though, high discount rates are merely a restatement and not a source of the energy efficiency gap *per se*. Instead, stringent investment criteria and the rejection of particular energy-efficient technologies may represent a rational response to risk. In particular, they may result from financial risks such as business-specific risk, general economic risk (business cycle, inflation, interest rates, exchange rates, and so on), potential changes in government policy, trends in input and output markets (fuel and electricity prices, for example) or financing risk (an anticipated reaction of capital markets to increases in borrowing, for example). Also, there may be technical risks – unreliability, for instance – associated with individual technologies.

This being said, what matters for the debate about investment barriers is whether or why investments in energy efficiency would carry higher risks than other investments, and therefore would be systematically overlooked. If energy-efficient technologies are unreliable, the risk of breakdowns and disruptions could outweigh any potential benefits from reduced energy costs. Ignoring these technologies is not only perfectly rational but also avoids inefficient outcomes. However, if these – often new and unfamiliar – technologies are wrongly perceived to be unreliable, government funded demonstration programmes aimed at increasing confidence and disseminating information and awareness among potential adopters might be justified. Interestingly enough though, many of the technologies that are included in engineering-economic models and recommended in energy efficiency publications are well proven, reliable, and widely used. Examples include energy-efficient lighting and motors, condensing boilers, thermal insulation, thermostatic radiator valves, and lighting controls.

Next, since energy efficiency investments are normally embedded within buildings and equipment, costly to remove and with limited scope for subsequent resale, they may carry a higher financial risk (see Sutherland 1991, for instance). For example, they may require higher hurdle rates because compared to stock and bonds they are 'illiquid' and irreversible, with limited scope for diversifying risks. While this argument may account for the differing treatment of 'liquid' and 'illiquid' assets, it fails to account for the differing treatment of comparable assets. For example, it could not explain why cost-saving energy efficiency investments should be subject to more stringent investment criteria than investment in a new production plant, when the latter is equally illiquid and irreversible.

Postponing irreversible investments in energy efficiency may be optimal if future energy prices are uncertain.

Finally, postponing irreversible investments in energy efficiency may be optimal if future energy prices are uncertain (Hasset and Metcalf 1993; van Soest and Bulte 2001). For example, investing in a more energy-efficient technology may turn out to be unprofitable if energy prices fall after the new technology has been implemented. Hence, there is an option value associated with postponing investments (McDonald and Siegel 1986; Dixit and Pindyck 1994). However, this approach fails to account for the potential costs of delaying energy efficiency investments. For example, it is much more costly to retrofit heat recovery systems than to include them when a plant or building is designed. Since most decisions relevant to energy efficiency involve a choice between efficient and inefficient options within an investment that is being made for other purposes, the scope of this approach to explaining why allegedly viable investments in energy efficiency are not made may be limited.[6]

To conclude, the argument that high discount rates can be considered a rational response to risk for all types of energy efficiency investment does not seem plausible. However, depending on the application, business, technical, or regulatory risks might be relevant barriers. If high discount rates are due to business or actual technical risks, they can be taken to reflect an economically efficient response of private decision makers and, thus, there does not seem to be a relevant barrier. If they are due to regulatory or misperceived technical risks, however, there might be relevant barriers that merit a policy intervention.

3.2 Imperfect information

If individuals lack adequate information on either energy efficiency opportunities or the energy performance of technologies, they may underinvest in energy efficiency. Conceptually, information problems possibly hindering energy efficiency investment can be grouped into three broad categories.

6 See Howarth and Sanstad (1995) for further criticism of the option-value approach to explaining the energy-efficiency gap.

First, there could be inadequate information on the level and pattern of current energy consumption. The availability of such information depends on the information content of utility bills, the level of sub-metering, the availability of relevant benchmarks, the use of computerised information systems, the time devoted to analysing consumption information, and so on. Gathering and analysing information on current energy consumption is associated with investment, operational, and staff costs, which can be seen as a particular category of transaction costs. Typically, these costs are not taken into account in engineering-economic models.

Second, there could be imperfect information on specific energy-saving opportunities – such as the retrofit of thermal insulation. Information on energy-specific investment opportunities within an organisation consists of two components. One concerns the extent to which organisations have evaluated energy efficiency opportunities, for example through energy audits. Since the value of an audit becomes known only after the audit has been carried out, it can only be judged with hindsight whether an audit was actually useful or not. In this sense, it is unlikely that the market produces an efficient outcome (Goldstone 1995). The other component refers to the availability of information on the costs and performance of specific energy-saving technologies. Since the search costs for energy-efficient technologies are likely to be much greater than those for energy commodities (electricity, gas, fuel oil, and so on), individuals' and organisations' choices may be systematically biased against energy efficiency. For example, the performance of technologies such as control systems, motors, and variable-speed drives may be difficult to evaluate even after purchase because detailed metering is not feasible. Thus, feedback on the performance of the energy-efficient technology is not available. Similarly, information on the performance of new energy-efficient technologies rests with the investor, but would be of value to others too. In this case, markets undersupply such information (because of the public-goods character of information), potentially justifying publicly funded information programmes and demonstration schemes.

Third, information on the energy consumption of new and refurbished buildings, process plants, and equipment and machinery could be asymmetric, resulting in adverse selection and thus inefficient outcomes. For example, among many aspects, the value of a house should reflect its energy efficiency. While this information may be available to the seller, potential buyers have difficulty in recognising and evaluating the potential energy savings. As a consequence, their bids on the house will be too low. In the end, only energy-inefficient houses (or other technologies) may be on the market and investment in improving the energy efficiency of houses is lower than it would be with symmetric information.

Problems of imperfect information are likely to pervade energy service markets and could potentially explain a substantial proportion of the efficiency gap.

To sum up, problems of imperfect information are likely to pervade energy service markets and could potentially explain a substantial proportion of the efficiency gap. If private markets do not provide adequate information, policy interventions (such as energy labelling) could be justified. Information programmes appear to be the most obvious policy approach, but minimum energy efficiency standards might be more effective in some instances. If information programmes are to be employed, both the manner in which information is presented and the credibility of the source need to be taken into account. There is some overlap between the barrier 'imperfect information' and 'hidden costs' – the barrier to be addressed next.

3.3 Hidden costs

If engineering-economic studies fail to account for either the reduction in utility associated with energy-efficient technologies or the additional costs associated with their use, the cost-efficient potential may be overestimated (Nichols 1994). Hidden costs are only hidden to the observer, but not to the decision-making individual or organisation. As Table 2 shows, three broad groups of hidden costs can be distinguished.

Table 2. Hidden costs of energy efficiency investment – main categories and examples

Sub-category	Examples
Loss of utility associated with energy-efficient choices	Problems with safety, noise, working conditions, service quality, and so on (e.g., lighting levels). Extra maintenance, lower reliability.
Cost involved in individual technology decisions	Cost of: (i) identifying opportunities; (ii) detailed investigation and design; (iii) formal investment appraisal. Cost of formal procedures for seeking approval of capital expenditures. Cost of specification and tendering for capital works to manufacturers and contractors. Additional staff cost for maintenance. Cost for replacement, early retirement, or retraining of staff. Cost of disruptions and inconvenience.
General overhead cost of energy management	Cost of employing specialists (e.g., energy manager). Cost of energy information systems (including gathering of energy consumption data; maintaining sub-metering systems; analysing data and correcting for influencing factors; identifying faults; and so on). Cost of energy auditing.

Source: Sorrell *et al.* (2004)

A rational choice of energy efficiency measures needs to account for a loss of utility possibly associated with them.

The first category concerns the potential loss of utility associated with energy-efficient choices. In essence, there are costs that result from inferior performance of energy-efficient technologies with respect to dimensions other than energy services. For example: an energy-efficient production process may lead to increased noise; the installation of cavity wall insulation in an old building may encourage damp; a variable-speed drive may require extra maintenance and training for new skills and tools. While these considerations clearly apply to energy-specific investment opportunities, they are likely to be even more important for investments where energy efficiency is only one of many attributes to consider. Incorporating these costs in engineering-economic models is feasible, in principle, but difficult to achieve in practice. Since taking into account costs related to inferior performance is necessary for a rational technology choice, they do not justify policy interventions.

Hidden costs in the second category can be considered part of the production costs of energy efficiency. They are specific to an individual investment in energy efficiency or the choice of an energy-efficient option. Examples include design fees for large items of a plant, the civil engineering costs of installing a combined-heat-and-power (CHP) unit and of connecting it to the grid, the costs of re-routing pipework, the costs of new light fittings to accommodate compact fluorescents, and the cost of production interruptions during equipment installation. In principle, these costs can be included in engineering-economic models (Ostertag 2003), but as they are site-specific and difficult to estimate, they may be easily overlooked. Arguably, they represent real costs, and organisations can be expected to take them into account when appraising investment opportunities. In any case, these types of barriers do not result in inefficient choices and, consequently, do not provide a rationale for policy intervention.

A third group of hidden costs relates to general overhead costs of energy management, and it corresponds to the search costs discussed in the context of the 'imperfect information' barrier. On the one hand, these costs partly depend on factors outside the control of the organisation adopting or considering the energy efficiency investment – such as the existence of standardised labelling schemes. On the other hand, these search costs also depend on factors internal to the organisation such as organisational procedures for information gathering, specification, purchasing and procurement. Within the broader category of transaction costs, they include all the organisational costs associated with establishing and maintaining an energy management scheme, investing in specific energy-saving technologies, and implementing specific energy-efficient options within broader investment programmes (for example, choosing an energy-efficient motor rather than a standard one). In contrast to the production costs and loss of utility discussed above, transaction costs depend closely on organisational and contractual structures, procedures, incentives, and routines. This makes them much more difficult to incorporate in models that represent costs purely in relation to individual technologies (Ostertag 2003). Nevertheless, these types of market or organisational transaction costs could be reduced through public policy or changes in the internal organisational structure.

All in all, claiming that hidden costs can explain the entire efficiency gap seems to be a tautology, but asserting that hidden costs are unimportant seems to be equally wrong. Sorrell *et al.* (2004) conclude that the truth lies somewhere between the two and the relative importance of different categories of cost is likely to vary between technologies and between organisations.

3.4 Access to capital

Imperfect access to capital may prevent the implementation of profitable energy efficiency projects.

The literature on barriers to energy efficiency usually discusses lack of access to capital in the context of private households. If low-income households have limited access to credit and can only borrow at high interest rates, this may prevent energy efficiency projects with a high rate of return from being undertaken. From the perspective of neoclassical economics, the inability to access capital may well constitute a barrier, but it need not imply a failure in capital markets. If low-income households are considered high-risk borrowers, potential lenders may demand a high risk-adjusted rate of return (Sutherland 1996). In this case, the market outcome is efficient and policy interventions are not justified. From the perspective of transaction costs economics, it may be that transaction costs necessary to investigating the creditworthiness of individual households are sufficiently high to diminish the economic viability of such loans (Golove and Eto 1996). Thus, overall efficiency may be improved if transaction costs to assess households' creditworthiness could be lowered.[7] In practice, policy interventions to overcome this barrier in the household sector are usually justified primarily on equity grounds.

As for the enterprise sector, lack of access to capital as a barrier to energy efficiency is more complex. Here, the 'access to capital' problem has an external and an internal dimension. To start with possibly insufficient access to external finance, in principle, firms should invest in all projects that have a rate of return exceeding the average cost of capital. However, there may be several reasons why a firm might fail in raising additional debt or equity. For example, external funding for a highly profitable investment in energy may be denied simply because of business risk.

The internal 'access to capital' problem stems from neglect of energy efficiency within internal capital budgeting procedures, combined with other organisational rules such as strict requirements

7 A similar argument could be made for small and medium-sized companies.

on payback periods. Two observations are commonly made in this context. First, since energy efficiency investments tend to be classified as discretionary maintenance projects, they are usually given a lower priority over essential maintenance projects or strategic investments. Second, energy efficiency projects tend to be evaluated based on payback periods rather than discounted cashflow analyses. The (implicitly) required rate of return implied by short payback periods exceeds those for business development projects. In practice, such short payback periods may be required for all (not just energy efficiency) projects by central or upper management as a safeguard to managerial slack at the lower management levels, because they cannot perfectly observe or assess the lower management's abilities or the project's profitability. Similarly, relatively high hurdle rates may be required for smaller projects – and many energy efficiency projects fall into this category – since the transaction costs of determining the profitability of such investments are likely to represent a larger portion of the expected savings.

In addition, top management does not consider energy-cost savings as a strategic priority. Thus, given the constraints on time and attention (see also sub-section 3.6 on bounded rationality), it may be overlooked by top management. This bias towards strict investment criteria may be worsened by individual managers' incentives to favour large, strategic projects, which are more prestigious than energy management activities.

In sum, there may be good reasons for imposing strict investment criteria or restricting capital budgets for energy efficiency investments within organisations. Empirical research would have to identify the rationale for such behaviour, the extent to which this is reproduced in other comparable organisations, and whether such behaviour is a contingent feature of particular organisational arrangements, which may be altered.

3.5 Split incentives and appropriability

Asymmetric information and transaction costs are at the root of the split incentives problem.

If a company is renting office space, neither the landlord nor the company (tenant) may have an incentive to invest in energy efficiency because the investor cannot appropriate the energy-cost savings. On the one hand, the landlord may not invest in energy efficiency if the investment costs cannot be passed on to the tenant, who will benefit from the investment through lower energy costs. On the other hand, the tenant may not invest if he is likely to move out before fully benefiting from the energy-cost savings. In principle, this so-called investor/user or landlord/tenant dilemma could be avoided if the investor were able to credibly transmit the information about the benefits (that is, future cost savings) arising from the investment and to enter into a contract with those benefiting from the investment. Such a contract would have to secure the appropriation of cost savings so that the investor can cover the investment costs. However, the costs of verifying energy-cost savings and the costs for the contractual arrangements are often prohibitive. Thus, asymmetric information and transaction costs are at the root of this investor/user dilemma problem (Jaffe and Stavins 1994).

Similarly, if managers – because of job rotation – remain in their post only for a short time, they may not have an incentive to invest in energy-efficient projects that have a longer pay back time. Further, if departments (in larger organisations) are not accountable for their own energy costs, they may have no incentive to invest in energy efficiency because the benefits in terms of cost savings accrue elsewhere. Finally, the purchaser of equipment may have a strong incentive to minimise capital costs, but may not be accountable for operating costs (including energy costs). This type of problem may also arise with users of buildings, operators of process equipment, and designers and contractors in the case of construction projects.

To conclude, various types of split incentives are likely to explain part of the energy efficiency gap. Policies that can be implemented at low costs – labelling, for instance – are likely to be economically efficient.

3.6 Bounded rationality

When faced with a complex decision structure, agents may not be able to optimise because of lack of time, attention, or the ability to adequately process information. Instead, bounded rationality may result in using routines or rules of thumb (Simon 1957, 1959), thus neglecting opportunities for improving energy efficiency – even when information is perfect and incentives are appropriate. For example, small motor end-users tend to consider only delivery time or price instead of life-cycle costs when buying a new motor to replace an old one (de Almeida 1998). Similarly, when making decisions about investment priorities, firms are likely to focus on the core production process rather than on ways to save energy costs. Likewise, in cases where investments in energy-efficient technologies are being considered, the same profitability or payback criteria may be required as for the core production technologies although the economic risks associated with the former are much lower.

4. Case study on German higher education sector

4.1 Overview

The German higher education sector (HE) consists of about 370 institutions for about 1.8 million students (Federal Ministry of Education and Research 2005), but only a few are private. Operating expenses for higher education institutions are largely financed through the budgets of the federal states (*Länder*). Investment costs for large equipment, new buildings and the building extensions are evenly split between the state and federal budgets.

When energy expenditures account for only a small share of energy users' overall spending, there is little financial incentive to reflect on energy efficiency.

Total energy consumption in the German higher education sector is significant and accounts for about 0.4 percent of German final energy consumption. The majority of this is for generic uses, notably heating, ventilation, air conditioning, and lighting (see Table 3). Since individual institutions only spend around 2 percent of their budget on energy, there is little financial incentive to pay attention to energy efficiency in university decision making, which is dominated by research and teaching concerns.

Table 3. Energy use in the German higher education sector

Share of electricity consumption	40 percent
o/w Ventilation and air conditioning	30-50 percent
o/w Lighting	20-40 percent
o/w Office equipment	20-30 percent
Share of thermal energy consumption	60 percent
o/w Space (and process) heating	> 90 percent
o/w Hot water	< 10 percent
Total energy costs	€500 million
Share of electricity costs in total energy costs	60 percent
Share of energy costs on total budget	2 percent

Source: Sorrell *et al.* (2004)

The technical and economic potential for energy efficiency in the German higher education sector is estimated to be substantial. Government analyses suggest that organisational measures in the public sector may save 5-15 percent while technical measures could reduce thermal energy consumption by 25-60 percent and electricity consumption by at least 10 percent (*Umweltbundesamt* 1999, *Energieverwertungsagentur* 1999). Energy efficiency opportunities in the higher education sector are typically of a general nature (that is, they are not specific to the higher education sector, but may be implemented in other sectors too) rather than process specific (see Table 4). These represent established and low-risk technologies widely recommended in the best-practice literature.

Energy efficiency opportunities in the higher education sector are typically of a general nature rather than process specific.

Table 4. Selected measures for the rational use of energy

Space heating
Thermostatic radiator valves
Programming heating and ventilation controls to match occupancy patterns and/or temperature
Use of building energy management system (BEMS)
Lighting
Replacement of 38mm fluorescents with 26mm
Use of high frequency electronic ballasts
Use of compact fluorescents
Use of photocell, acoustic or movement sensors
Plant room
Insulation of pipes, valves and flanges
Use of boiler sequencing controls
Replacement of oversized boiler plant
Installation of condensing boilers
Installation of CHP
Building fabric
Draught proofing of windows and doors
Retrofitting insulation to walls and roofs
Use of secondary or double glazing on refurbishment
Electrical equipment
Specification of high efficiency office equipment
Specification of high efficiency motors
Use of variable speed drives in pumps, fans and other applications
Automatic switch off of fans & pumps

Source: Sorrell *et al.* (2004)

Like almost all publicly funded projects and institutions, higher education in Germany has suffered from tight federal and state budgets. Since many institutions need new equipment for research and teaching and many buildings need to be refurbished, the budget situation constitutes a major challenge. Further challenges arise from current reforms of the higher education sector, including increased autonomy from the state authorities, the introduction of global budgeting and business accounting, new funding schemes, tuition payments in some states, rankings and evaluation, increased competition from other public or private universities, new forms of teaching such as virtual universities, and new designs for bachelors' and masters' degrees. Since the German constitution grants authority over education largely to the *Länder* (rather than the federal government), these challenges vary across institutions. As will be described below, some of these reforms may be beneficial for energy efficiency.

In this section, barriers to energy efficiency in the German higher education sector will be explored and some recommendations on how these could be overcome will be identified. The section draws on results from six in-depth case studies of individual universities and a number of additional interviews carried out under the EU research project "Barriers to energy efficiency in public and private organisations", published in Sorrell *et al.* (2004). At these universities, up to six people (top administrators, energy managers, technical managers, finance managers, purchasing officers, buildings officers, and so on) were interviewed in person and, where necessary, follow-up telephone interviews were conducted. The interviews were semi-structured and used detailed protocols based on the theoretical framework developed in the previous section. The next sub-section briefly summarises the energy management practices observed in the German higher education sector.

4.2 Summary of energy management

A variety of factors shape energy management practices in Germany's higher education sector.

The energy management practices at the six case-study universities in German higher education may be summarised as follows.

Organisation: universities are large institutions with complex decision-making structures, where energy consumption is influenced by several departments and institutes, individuals and groups from within and outside the university; the university administration is primarily responsible for measuring and controlling energy consumption and costs, setting investment priorities, purchasing equipment, space planning, and to a limited extend also for construction planning and maintenance; typically, there is no energy manager; instead a department for technical services is responsible for the supply of heat and electricity services, while a construction department participates in the planning of buildings and is in charge of their maintenance; many universities exhibited a lack of coordination, clear delegation, and clear responsibilities for energy management.

Energy/environmental policy: formal environmental or energy policies are an exception; no university had a certified environmental management scheme in place; teaching and research are the top priorities.

Energy costs and specific energy consumption figures: the share of energy cost in the budget ranged from 1.7 percent to 2.4 percent (including third-party funding); final thermal energy consumption was between 0.56 GJ_{th}/m^2 and 1.5 GJ_{th}/m^2, while final electricity consumption was between 0.28 GJ_{el}/m^2 and 0.66 GJ_{el}/m^2; annual energy costs per student ranged from €110 to €350.

Energy information systems and energy consumption control: energy consumption is measured and controlled regularly at the level of individual buildings, but no targets exist; energy management systems for buildings are usually in place, but mostly only for some of them and often outdated; energy costs are paid out of a central budget and unknown to individual departments or institutes.

Capital budgeting and investment criteria: separate budgets for energy efficiency measures do not exist; typically a share of 6 percent of the maintenance budget is required by law to be spent on energy saving measures, but this is not sufficient to realise all measures; quantity and quality of profitability and risk analyses for energy efficiency measures vary considerably across universities; for investments, payback periods of 5 years tend to be required, but they may be longer for equipment with a long lifetime.

New buildings and refurbishment: for construction of buildings and most refurbishment, a state construction agency is responsible for planning and carrying out the work, with limited influence of the universities; often, prestige and design dominate; there is lack of coordination between planners, various engineering firms, and trades.

Purchasing and policy integration: specifications for new equipment are provided by individual departments, and equipment is ordered by central purchasing offices, where energy efficiency is taken into account; but usually initial outlays (and not life-cycle costs) are most important; existing laws and guidelines for integrating environmental performance into the purchasing specifications are not very powerful and usually not enforced.

Awareness and culture: a general lack of awareness for energy performance was prevalent; where educational programmes aimed at increasing energy efficiency awareness exist, research and teaching staff tend to abstain; support for rational-use-of-energy measures from the top administration is important, but the personal motivation of those in charge of energy management seems to be crucial for finding and realising rational-use-of-energy potentials; some of the interviewees would like to see universities to lead by example (that is, do what they preach); but the concept of 'sustainability' had only started to enter the curricula.

4.3 Barriers to energy efficiency in the German higher education sector

The case-study interviews confirmed that there are numerous opportunities for improving energy efficiency in a cost effective way, but many are not being realised. Most interview partners agreed with the following statement from a pre-interview questionnaire:

> "There is a wide range of energy-efficiency measures that could be implemented in my university that would yield paybacks of less than four years at current energy prices".

This section assesses the importance of various barriers to energy efficiency in the German higher education sector using the theoretical framework presented earlier. When interpreting the results, it is important to note that the barriers represent perspectives that highlight particular aspects of a complex situation and that there is much overlap and interdependence between the different categories. The barriers found to be of particular importance are split incentives, lack of capital, hidden costs, and imperfect information. These are discussed in turn below. Table 5 specifies how these barriers operate, with the final column identifying potential policy measures to remove or lower them.

The barriers found to be of particular importance are split incentives, lack of capital, hidden costs, and imperfect information.

To start with split incentives, the case studies suggest that these are the most important barriers to energy efficiency in Germany's higher education sector. Three key split-incentives problems are worth highlighting.

First, at the university level, there was no incentive to save energy costs since the savings were generally not allowed to be used for other purposes and – even worse – may lead to reduced budgets in the future. At the root of this problem are public accounting principles in Germany (*Kameralistik*) that limit the transferability of funds across the separate budgets for capital expenditures and administrative expenditures as well as within those budgets. Unused funds designated for a particular purpose can neither be spent on other purposes nor transferred across budgeting periods. In particular, savings in energy costs must not be spent on investment in energy efficiency, buildings maintenance, or office equipment. Historically, this was the case in all universities. Recently, most federal states have started to change budgeting principles, but

Incentives for German universities to save energy are likely to improve, but will remain constrained.

the policies and future plans on how to allocate savings differ across states. For most universities, partial transferability of unspent funds to other purposes within the same budgeting period is being planned for the near future in connection with the introduction of 'global budgeting'. Similarly, for savings at the university level, some type of cost-sharing arrangements with the state administration is being considered, but universities will not be allowed to fully keep the savings made in one period for future periods. Thus, incentives for the university to save energy costs are likely to improve, but remain constrained.

Table 5. German higher education sector – barriers to energy efficiency and policies to overcome them

General category	Specific instance	Policies
Split incentives	Departments not accountable for energy costs.	Devolved budgeting with new business accounting system.
	Limited transferability of funds.	Global budgeting (degree of implementation varies by Federal State).
	State construction agency responsible for the planning of new buildings and refurbishment.	Increase universities' planning and financial authority; put operating/facility/space management company in charge; privatisation of all facility/space, building codes.
	Contractors, etc. for buildings not accountable for operating costs.	Targets for energy performance and operating costs to be included in tender; integrated design process, building codes.
Access to capital	Availability of capital to university.	Energy service contracting.
	Allocation of capital within university.	Conduct profitability analyses; life-cycle costing; make university funding a function of energy performance or energy audits; environmental/energy management schemes; subsidies for energy audits; raise awareness at top administration level via voluntary agreements, etc.
Hidden costs	Lack of time, management costs.	Energy service contracting; full-time energy manager; targeted information programmes; co-operative procurement; subsidies for energy audits.
	Complex and time-consuming decision-making process.	Change responsibilities and shorten process through laws; devolve financial and decision-making responsibility to individual institutions; involve ESCOs.
Imperfect information	Information on energy use and needs.	Invest in information systems and Business and Environment Management Schemes; energy manager to improve co-ordination of energy management; improved communication with top management; energy committee; best practice programmes.

Second, energy consumption was not always metered for individual departments, and historically there had been no arrangements for decentralised accountability of energy costs. Instead these costs were paid out of the university-wide budget for operating costs. Since energy budgets and the responsibility for energy management were not devolved to individual departments, these had no incentive to save energy through purchasing efficient equipment, office and space management, or behavioural measures. Again, these are typical split-incentives problems and they were encountered in all the universities studied. With the introduction of the business accounting system in most universities, it is planned to allocate energy costs in future, most likely using some proxy such as floor space. Although this represents an imperfect second-best to directly charging for energy consumption, it should nevertheless create stronger incentives than the existing mechanisms to manage office and room space. One interviewee noted that:

"… so far the use of space is considered free and not associated with any costs, ... and the status of a professor is also determined by the office space he manages to seize,... Institute managers will be stunned once they actually have to face costs, especially energy costs."

Third, as already pointed out, for all new construction and to some extent for building maintenance too, a state construction agency and not the university was responsible for planning and implementing projects. Thus, those planning the projects were not the same as those who had to pay the operating costs. Except for one university, which for historical reasons was also in charge of construction planning, all universities complained that the state construction agency did not adequately take into account future energy and other operating costs. Instead, other motives like prestige and design appeared to be more important. Quite often, this situation caused friction between those responsible for energy management at the university level and state planners, as indicated by the following statements from interviewees:

"... decision makers have the power, but they do not have the relevant knowledge and information about the actual needs at the university level."

"… planners at the state construction agency are utterly incompetent and haven't a clue about what is going on."

"Often, one hand does not know what the other hand is doing, or why it is doing what it is doing. … input from the universities is not always appreciated at the state construction agency."

"… it is not important to save energy or energy costs, but rather to comply with regulations."

Until recently, efforts on the part of building designers and contractors to improve energy efficiency were neither incentivised nor rewarded.

In the past, these problems were exacerbated by external designers and contractors, who were commissioned by the state construction agency and reimbursed according to the Ordinance on Fees for Architects and Engineers (HOAI). Since their fees depend on the financial size of the projects, designers and contractors have an incentive to choose larger heating, ventilation, and cooling equipment than necessary, leading to needlessly high energy consumption. Similarly, until recently, efforts on the part of designers and contractors to improve energy efficiency were not incentivised or rewarded. In 1994, the German Federal Government modified the HOAI to provide financial incentives for improved energy efficiency or the inclusion of renewable energy sources. Although higher fees can now be charged if the services provided exceed those usually required, the reimbursement may not fully cover the costs.

Even with these additional incentives, split-incentives problems remain since architects, engineers, sub-contractors, and consultants are accountable for capital cost but not for future operating costs. For them, meeting deadlines and staying within the budget are vital. Lack of coordination between different trades tends to aggravate this problem[8].

8 For details, see Chapter 7 in Sorrell *et al.* (2004).

In contrast, other potential forms of split incentives – landlord-tenant relationships in leased buildings and the lack of long-term incentives created by rapid job rotation, for example – were not found to be significant in the German higher education sector.

Turning then to the 'access to capital' problem, for the case-study universities, this was a major barrier to improving energy efficiency – reflecting both lack of funds for universities and an allocation of funds within universities that shows low priority for energy efficiency investments.

The access-to-capital problem has been a major barrier to improving energy efficiency in Germany's higher education sector.

Financing for larger investments comes from the state and federal budgets, which have been very tight for years. The situation is seriously aggravated by the fact that, historically, universities have not been allowed to borrow on the capital market. Thus, borrowing for investments in energy efficiency has not been an option, no matter how profitable the investment may have been. Likewise, in some states, third-party financing through an energy services contract was considered illegal. Hence, in universities, the funding situation for investments in general is difficult, which tends to restrict funding to small projects with payback period of less than five years. Investments in energy efficiency will at best be delayed, as in the case of one university where the implementation of a highly cost-effective lighting system took eleven years! This delay resulted from the combination of lack of capital in the public sector as a whole and the complex decision-making structures in the higher education sector.

While investment in energy efficiency suffers from aggregate capital restrictions, the lack of funding for these investments is also a consequence of priority setting inside and outside the university. The primary criteria are quality of research, quality of teaching, urgency, and design prestige. Energy efficiency – and in some cases even profitability – are only of minor importance. Hence, it does not come as a surprise that none of the universities had a specific budget for energy efficiency investment. Instead, small investments had to be financed through unused funds from other areas, in particular through the maintenance budget, where laws require that a certain percentage be spent to improve energy efficiency. But in most cases the amount of money available from this source is both small and insufficient.

As for hidden costs, the case studies considered all three main categories discussed above: general overhead costs of energy management; costs specific to a technology investment; and loss of utility associated with an energy-efficient technology.

The overhead costs of energy management (including expenditures on staff salaries, energy audits, information systems, sub-metering, and so on) were usually not included in profitability analyses of possible energy efficiency projects – if such analyses were conducted at all. Typically, only investment costs, direct personnel costs, and maintenance and fuel costs were considered. Accounting for overhead costs of energy management in profitability assessments would require calculating their magnitude. This is no easy task, and leads to additional costs of its own. The prevalence of severe time constraints on university employees was suggestive of the importance of salary costs, in that it may be uneconomic to increase staff resources for energy efficiency improvements. However, not only were such calculations not made, but hiring new staff was often prohibited by rigid public employment schemes (*Stellenplan*) that fix the number of university staff.

As outlined above, the costs associated with complex and time-consuming decision-making processes constituted a major barrier to all types of investment, including energy efficiency investments. For example, the installation of a CHP-plant would have to be planned 5 to 7 years in advance – a period during which energy markets and CHP-technology may change significantly.

By contrast, hidden costs specific to a technology investment – such as disruption, hassle, and inconvenience – were found to be much less important. Major construction or equipment replacement can always be carried out during the relatively long winter or summer breaks. Likewise, loss of utility associated with energy-efficient technologies was not considered to be important. However, in many cases the costs for additional energy services were not considered when buying new equipment. For example, at one university, a new transmission microscope was estimated to increase annual operating costs by €1,600. However, the microscope significantly added to cooling and ventilation loads, and when the fixed, operating and maintenance costs for cooling and ventilating the equipment were taken into account, the additional costs exceeded €6,000 a year. Hence, when indirect costs are neglected, as is typically the case, total costs may be seriously underestimated, with the result that inefficient investment decisions are made.

In summary, hidden costs associated with identifying savings potentials, finding appropriate energy efficiency measures, conducting profitability analyses, and preparing public procurement processes were important barriers to energy efficiency in the case-study institutions. But the evidence does not support the hypothesis that these hidden costs are significantly higher for energy efficiency measures than for standard (i.e., non-energy efficiency) measures.

This takes us to imperfect information – the last barrier to energy efficiency investment found to be of particular importance in case-study universities. Two categories of information were considered: information on the volume and pattern of universities' energy use and on opportunities for dedicated energy efficiency investments.

Information on current energy use at the level of individual buildings or departments was poor and the end-use split of energy consumption was generally unknown, particularly for electricity. Data on energy costs, energy consumption, and user needs were collected by individual departments rather than a central organisational unit, and there was no evidence of comparison with either generic or sector-specific benchmarks. Similarly, the needs of end users were often neither known nor well communicated within the institution. The case-study results thus suggest problems of decentralised information and lack of both knowledge and coordination of user needs throughout the higher education sector. But they also suggest that these problems are less pronounced in universities with a full-time energy manager.

While lack of information on the volume and pattern of energy use was considered to be an important barrier to energy efficiency, the quality and quantity of information available on energy-specific investment opportunities was generally judged to be good. Information on investment opportunities appeared to depend on the competence and motivation of the staff in charge of energy management. With regard to sources of information on energy efficiency measures, the media used included the internet, informal networks, meetings of energy managers, the information system for the higher-education sector (HIS-Higher Education Information Systems GmbH, Hannover), special seminars and workshops, and energy service companies (ESCOs). If used at all, informal networks were considered to be excellent and trusted sources of information. By contrast, there was considerable suspicion of ESCOs.

Lack of information on the volume and pattern of energy has been an important barrier to energy efficiency.

4.4 Policy implications

The empirical research demonstrates a wide range of barriers to energy efficiency in the German higher education sector that also inhibit economically efficient outcomes. It follows that significant improvements in energy efficiency are likely to require a similarly wide range of initiatives at the organisational, sectoral, and national level. A general observation from the case studies is that successful energy efficiency policy has to encompass more than just fiscal measures. It must also take into account the communicative, organisational, and cooperative challenge that energy efficiency creates for individual institutions (Ostertag 2003).

Energy service contracts may provide an effective route for overcoming barriers such as lack of capital, time, staff, and expertise …

In addition to public policy measures, energy service contracts may provide an effective route for overcoming barriers such as lack of capital, time, staff and expertise. Likewise, the involvement of ESCOs (see Box 1) may avoid the time-consuming and complex decision-making structures within the higher education sector. Since the potential – and limitations – of ESCOs in the German higher education sector have been discussed in detail elsewhere (Schleich *et al.* 2001), the remainder of this section will focus on potential measures at the organisational, sector and national level (see the last column in Table 5). It should be noted, however, that it is beyond the scope of this paper to assess the efficiency of these measures, that is, whether the benefits of implementing them outweigh their costs.

At the organisational level, creating the position of an energy manager would have considerable potential to reduce information deficits, improve the coordination of energy management and user needs, and encourage vertical and horizontal communication – thereby reducing barriers to energy efficiency. Along with other tasks, an energy manager could assume the role of 'product champion'

Box 1. The energy services concept with application to the higher education sector

Energy services represents a new and rapidly growing business model in which suppliers offer a single contract to minimise the total bill for the services that energy provides – such as heating, lighting, and air conditioning. This contrasts with the traditional approach in which energy consumers contract separately for energy commodities (fuel oil, natural gas, electricity, and so on) and for a range of conversion equipment that delivers energy services (heating, lighting, air conditioning, and so on). In its simplest form, an energy services contract guarantees supplies of heat and power at reduced cost, but in a more sophisticated form the contract may guarantee particular levels of service provision (for example, lighting levels and room temperatures).

By focussing on better performance and solutions to customer needs rather than commodity sales, energy service companies (ESCOs) have a strong incentive to improve energy efficiency. ESCOs typically offer energy management, energy information systems, energy audits, installation, operation and maintenance of equipment, competitive finance, and fuel and electricity purchasing. The contract allows the host organisation to lower risk, avoid capital expenditure, reduce energy costs, and concentrate attention on core activities. Energy services contracts could provide a cost-effective route to overcoming barriers to the diffusion of both established and innovative low-carbon technologies in the public, commercial, industrial, and household sectors. The model is applicable to both energy-use and energy-supply technologies and is, in many countries, the primary mechanism for the diffusion of CHP technology. It is also consistent with energy market liberalisation and the broader trend towards the outsourcing of non-core activities.

As for ESCOs' potential in the higher education sector, it is useful to point out that they may be reluctant to offer their services in this sector. One of the typical reasons in the higher education sector is the long negotiation process before a contract can actually be made: the decision making at universities involves administrative procedures and budgeting laws, is highly complex, and rather inflexible. Also, some universities compare new offers from ESCOs with the variable costs of providing the same services internally. This incorrect procedure renders offers by ESCOs, which calculate on for energy efficiency, which has proven to be instrumental for improving energy efficiency in the industrial sector (Ramesohl 1998). Likewise, bundling information and competence for energy use and consumption within a single department saves coordination costs and helps identify and solve internal conflicts of interest. This means that responsibility for planning, maintenance, technical services, space management, heat and electricity supply, and buildings energy management systems should be under the same roof. In any case, clear delegation and responsibility for energy management is crucial.

Certified environmental management schemes might also be an effective measure, especially for poor performers. These schemes not only organise environmental management, but are also designed to motivate staff and students and to get energy and environmental issues high on the administration's agenda. For large universities, implementing such schemes at the level of the institution requires considerable effort in coordination and hence may be difficult to achieve. Environmental management schemes for individual schools or institutes should therefore be considered as an alternative. Getting key administrators' attention may also be achieved via voluntary or negotiated agreements, either at the level of the institution or for the entire higher education sector. Given the crucial role of the state administrations, voluntary agreements may be most effective for organisations at the state level. Instead of using payback periods, profitability

... and certified environmental management schemes might also be an effective measure.

the basis of full costs, (seemingly) unprofitable. Outsourcing or privatisation is often associated with the loss of jobs in the public sector, which tends to create a political problem. Furthermore, at least historically, because of public budgeting laws, the budget for energy costs could not easily be used for financing contract energy management. In particular, universities had problems getting the investment part of the contracting fee 'reimbursed' by the state administrations. Likewise, some state administrations have denied approval of contract energy management because it was regarded as a type of unauthorised 'hidden credits'.

On the other hand, ESCOs may find conditions in the higher education sector particularly beneficial. Since energy consumption (and costs) of universities is rather high, projects for contract energy management are of sufficient size for ESCOs to recover overhead and transaction costs. Moreover, energy supply and demand technologies in the higher education sector are fairly homogenous (generic), which encourages strategies by ESCOs to focus on particular customer groups and to realise economies of scale and scope. What is more, since institutions of higher education usually belong to the public sector, they carry a very low financial risk. Since contract energy management for investments in energy efficiency typically requires long contract periods, a low financial risk is crucial for such projects. At the state level, several administrations have pushed contract energy management as a strategy to overcome financial restrictions in the public sector. Furthermore, ESCOs are expected to benefit from more recent developments towards increased autonomy in terms of decision making and financial resource allocation for universities – developments that are expected to speed up the decision process and reduce uncertainty stemming from hierarchical administrative structures. Likewise, the introduction of global budgeting and business accounting systems (rather than separate budgets for capital and administrative expenditures under a *Kameralistik* system) should facilitate contract energy management via ESCOs and help to correctly assess the costs of contract energy management vis à vis other alternatives.

Thus, ESCOs may help overcome important barriers to energy efficiency in the higher education sector such as lack of sufficient internal or external capital to finance profitable measures, know-how, manpower, and time to realise such measures internally.

analyses using life-cycle costs should be carried out, so the net benefits of energy efficiency measures can be demonstrated. With modern software tools, the transaction costs of this should not be prohibitive. Similarly, when deciding on new equipment, the indirect costs for additional energy services should be considered. Since indirect costs can sometimes be a multiple of the direct investment costs, neglecting them might result in inefficient investment decisions.

Introducing global budgeting with unrestricted transferability of funds within and across budgeting periods has considerable potential to encourage energy savings.

Creating positive incentives for universities to invest in energy efficiency implies that institutions should be able to keep the cost savings from these investments and use them for other purposes now or later. Reforms to this end would constitute one of the crucial measures at the sectoral level. In essence, universities should be able to shift funds between budget headings and they should not be punished for lower energy costs by receiving less funding overall in subsequent financial periods. Thus, global budgeting with unrestricted transferability of funds within and across budgeting periods should be introduced in all organisations. To make incentives for planning and operating buildings compatible, individual universities should be given more planning (and financial) authority at the expense of the state planning agencies. Alternatively, the operation and management of facilities and estates could be outsourced to private companies.

If a portion of university funding were to be made a function of prior performance (such as achieved energy savings or performance against some kind of benchmark), the incentive to realise energy efficiency measures within the university would be much stronger. This kind of funding system would also force energy costs onto the top administration's agenda. Alternatively, a portion of state funding could be dependent on the universities having carried out an energy audit within the last year or two or on having an environmental/energy management system in place.

Similarly, individual departments should have individual budgets and be held accountable for their energy costs as far as possible. Clearly, allocation of costs based on the space used would only be a first step. At the same time, the business accounting system, which is currently being introduced in many universities, is expected to lead to a better allocation of financial resources and to provide better incentives for energy-saving measures than the obsolete *Kameralistik* system. Business accounting systems are also expected to provide opportunities for incorporating data on environmental performance.

Future operating costs should be an integral part of the procurement specifications for new buildings and major refurbishments. Likewise, integrated planning should be implemented as a rule. Since integrated planning involves all the relevant actors (notably architects, specialist engineers, sub-contractors, and the energy manager), the content of the individual work-packages can be better coordinated and designed so that split incentives are accounted for and a more efficient solution emerges.

Other possibilities at the sectoral level include training seminars for those in charge of energy management and targeted information programmes on specific topics (for instance, contract energy management). Both are expected to reduce information and other transaction costs. In the same way, informal networks can be a cost-effective tool to reduce information-related barriers.

Benchmarking at the sectoral level is an important tool to get the attention of top administrators and it may help improve the status of energy management. However, for meaningful comparisons across different institutions, proper benchmarking would have to differentiate between a wide range of buildings types and uses.

Finally, cooperative procurement of energy-efficient equipment by several higher education institutions provides another route to reduce transaction costs and obtain price reductions on equipment through bulk purchases. For example, primary and secondary schools in the city of Hamburg have successfully used such a procurement process for more energy-efficient lighting.

To conclude with key measures at the national level, it should be stressed that the effectiveness of measures at the organisational and sectoral levels will be enhanced if they are embedded in a broad-based, long-term national programme to address the climate-change challenge. In particular, allowing energy prices to reflect external environmental costs will render measures improving energy efficiency more cost-effective and raise awareness among higher administration. In Germany, such policies include the continuation of the ecological tax reform, which was implemented in 1998. Increasing tax rates on fuels and electricity reduces the financial risk associated with investments in energy efficiency and allows for long-term planning. The introduction of the EU CO_2-emissions trading scheme (EU ETS), which started in 2005 for about 11,500 installations from the power and most energy-intensive industry sectors in the EU, has lead to a substantial increase in the costs of electricity. Companies subject to the EU ETS receive more than 95 percent of allowances for free, but in particular power producers were able to pass on a large part of the full (opportunity or marginal) costs to consumers.[9] If the cap on CO_2 emissions is chosen so that climate change targets are met, energy prices could increase significantly, creating additional incentives for energy efficiency.

A national programme addressing climate change challenges will enhance the effectiveness of energy efficiency measures at the organisational and sectoral levels.

National policies in support of CHP plants may affect universities either directly, as operators of CHP plants, or indirectly as end users who bear the costs of national subsidy programmes. Other policies at the national level include the continued tightening of the standards in the Energy Conservation Ordinance; the re-examination of technical standards for heating, ventilation, air conditioning, and cooling services (to avoid over-sizing); and the introduction of minimum efficiency standards and labels for energy-consuming equipment such as personal computers. Moreover, public programmes subsidising energy audits and the implementation of energy management systems could be extended to make public institutions eligible. Finally, the recently formed Federal Energy Agency could initiate, coordinate, or develop information and education programmes targeted at the higher education sector, together with best-practice programmes, pilot projects, support networks (such as eco-campus net – a network for an environmentally sound development of universities), and 'energy-cocktails'. At such energy cocktail, top university administrators would be invited for food and drinks and to listen to a short keynote presentation on the importance of energy costs and energy efficiency, similar to the Swiss RAVEL programme (Bush 1996). Top administrators may have strong incentives to participate in such events, since – as a side benefit – they also provide a stage for networking among top administrators where other relevant topics may also be discussed.

5. Conclusions

The case-study results presented in this paper for the higher education sector in Germany confirm the notion that there are barriers to energy efficiency. That is, there are mechanisms that inhibit the adoption of profitable energy-efficient measures. The main barrier found were various forms of split incentives, which – apart from preventing higher energy efficiency – lead to economically

9 Note that the logic of emissions trading requires opportunity costs (rather than actual costs) to be passed on to consumers. Otherwise, product prices would not reflect true environmental costs. Of course, a free allocation of allowances may result in substantial windfall profits for companies. Auctioning off allowances would reduce those windfall profits.

The case study on Germany's higher education sector supports – at least to some extent – the presumptions of technology-economic type modelling and the call for policy interventions.

inefficient outcomes. The findings support – at least to some extent – the presumptions of technology-economic, bottom-up type modelling and the call for policy interventions. For example, policy measures such as global budgeting at the level of universities and devolved budgeting at the level of departments may be implemented at relatively low costs. Barriers that would not have justified policy interventions, like hidden costs of production interruption, were not found to be relevant for the German higher education system. For other sectors, however, such as the brewing or mechanical engineering (see Sorrell *et al.* 2004) hidden costs such as production interruption and loss of quality were found to be significant, but did not provide a rationale for policy intervention either. This illustrates that barriers to energy efficiency will vary across applications, and judgment on whether policy measures should be implemented is likely to be case specific. The case studies also suggest that multiple, possibly reinforcing, policies may be necessary to address the different types of barriers.

However, since the number of observations in a case-study analysis – as the one presented for the German higher education sector in this paper – tends to be small (by definition), the findings cannot be generalised in a statistical sense. Nevertheless, case studies are well suited to gain insights into complex decision-making processes and structures within organisations – even if their findings are usually limited to an analytical generalisation, where observed outcomes of decision-making processes are explained by identifying relevant causal mechanisms (Yin 1994). Although all these qualifications are sensible, existing econometric analyses based on large samples tend to support the general findings of analyses based on case studies (see for example, Scott 1997, Schleich 2004, and Schleich and Gruber 2007).

Finally, cost-benefit analyses ought to be conducted with a view to assessing the economic efficiency of the proposed policies. Likewise, thorough methodologically sound *ex-post* evaluations of existing energy-efficiency programmes are vital. With increased data availability, using econometric techniques to evaluate such programmes may become more popular. Such analyses could also help reduce the uncertainty about just how many kilowatt hours and CO_2 emissions were saved by a particular policy intervention.

References

Bush, E. (1996). "Energie-aperos - Energiesparmöglichkeiten lustvoll vermitteln". Impuls No. 19, Zeitschrift für IP BAU, RAVEL und PACER, Bundesamt für Konjunkturfragen, Bern.

de Almeida, E.L.F. (1998). "Energy efficiency and the limits of market forces: The example of the electric motor market in France." *Energy Policy,* (26), pp. 643-653.

Dixit, A.K. and Pindyck, R.S. (1994). *Investment under uncertainty,* Princeton University Press, Princeton, USA.

Energieverwertungsagentur (E.V.A.). (1999). *Energieeffiziente Universitäten*, Endbericht, Vienna, Austria.

European Commission (2006a). "Action plan for energy efficiency: Realising the potential". COM(2006)545 final.

European Commission (2006b). "A European strategy for sustainable, competitive and secure energy". COM(2006)105 final.

European Council (2006): Presidency Conclusions of 23/24 March 2006. 7775/1/06 REV1 of 18.05.2006.

Federal Ministry of Education and Research (2005). *Basic and structural data 2005*, Bonn, Berlin, Germany.

Goldstone, S. (1995). "Restructuring: stimulus to improving utility DSM: how economists might help". Western Economic Association 70th annual conference, San Diego, July 5-9.

Golove, W.H. and Eto, J.H. (1996). "Market barriers to energy efficiency: A critical reappraisal of the rationale for public policies to promote energy efficiency"., University of California, Lawrence Berkeley Laboratory report LBL-38059.

Hassett, K.A. and Metcalf, G.E. (1993). "Energy conservation investment: Do consumers discount the future correctly?" *Energy Policy,* (23:1), pp. 710-716.

Howarth, R.B. and Sanstad, A.H. (1995). "Discount rates and energy efficiency". *Contemporary Economic Policy,* (13), pp. 101-109.

Hourcade, J.-C., Jaccard, M., Bataille, C., and Ghersi, F. (2006). "Hybrid modeling: New answers to old challenges". *The Energy Journal,* (Special Issue), pp. 1-11.

International Energy Agency (2006a). *Energy technology perspectives report*. IEA, Paris, France.

International Energy Agency (2006b). *World Energy Outlook*. IEA, Paris, France.

Intergovernmental Panel on Climate Change (2001). *Third assessment report*. IPCC, Geneva, Switzerland.

Jaffe, A.B. and Stavins, R.N. (1994). "The energy-efficiency gap: What does it mean?". *Energy Policy,* (22), pp. 804-810.

McDonald, R. and Siegel, D. (1986)."The value of waiting to invest". *Quarterly Journal of Economics*, (101), pp. 707-727.

Nichols, A.L. (1994). "Demand-side management: Overcoming market barriers or obscuring real costs?". *Energy Policy,* (22), pp. 840-847.

Ostertag, K. (2003). *No-regret potentials in energy conservation: An analysis of their relevance, size and determinants*. Physica, Heidelberg, Germany.

Ramesohl, S. (1998). "Successful implementation of energy efficiency in light industry: A socio-economic approach to industrial energy policies". Industrial energy efficiency policies: Understanding success and failure. Proceedings of a workshop organized by the international network for energy demand analysis in the industrial sector. Utrecht, The Netherlands, June 11-12, 1998. (LBNL-42368).

Schleich, J., Böde, U., Köwener, D., and Radgen, P. (2001). "Chances and barriers for energy services companies? A comparative analysis for the German brewery and university sectors", in European Council for an energy-efficient economy (ed.), *Proceedings of the 2001 ECEEE summer study*. ADEME, Paris, France.

Schleich, J. (2004). "Do energy audits help reduce barriers to energy efficiency? An empirical analysis for Germany". *International Journal of Energy Technology and Policy* , (2), pp. 226-239.

Schleich., J. and Gruber, E. (2007). "Beyond case studies: Barriers to energy efficiency in commerce and the services sectors". *Energy Economics*, forthcoming.

Scott, S. (1997). "Household energy efficiency in Ireland: A replication study of ownership of energy saving items". *Energy Economics,* (19), pp. 187-208.

Simon, H.A. (1957). *Models of man*. Wiley, London, UK.

Simon, H.A. (1959). "Theories of decision-making in economics and behavioural science." *The American Economic Review*, (49), pp. 253-283.

Sorrell, S., O'Malley, E., Schleich, J., and Scott, S. (2004). *The economics of energy efficiency - Barriers to cost-effective investment*. Edward Elgar, Cheltenham, UK.

HM Treasury (2006). *Stern Review: The economics of climate change*, Cambridge University Press, Cambridge, UK.

Sutherland, R.J. (1991). "Market barriers to energy efficiency investments". *The Energy Journal*, (12:3), pp. 15-34.

Sutherland, R.J. (1996). "'The economics of energy conservation policy". *Energy Policy*, (24:4), pp. 361-370.

Umweltbundesamt (1999), *Umweltcontrolling im Bereich der öffentlichen Hand*, UBA Text No. 8/99.

van Soest, D.P. and Bulte, E.H. (2001). "Does the energy-efficiency paradox exist? Technological progress and uncertainty". *Environmental and Resource Economics,* (18), pp. 101-112.

Yin, R.K. (1994). *Case study research: Design and methods*. Sage Publications, London, UK.

ABSTRACT

Following a brief review of the rationale for promoting renewable energy sources, this paper compares alternative policies to promote the production of renewable electricity. *The focus is on feed-in tariffs (used in Germany, Spain, and France – for instance) and tradable green certificate (TGC) systems (United Kingdom and Italy, for instance). Considering economic theory and practical experience, the criteria for comparing these two alternatives are: cost-effectiveness, environmental effectiveness, and compatibility with market liberalisation. The paper argues that economic theory does not suggest a clear-cut advantage of one instrument over the other and it emphasises that, in any event, the choice of instrument depends on the relative importance attached to these criteria and on cultural factors such as faith – or lack thereof – in markets to help solve environmental problems. In this context, the paper questions the practical usefulness of a European-wide TGC system.*

Dominique Finon (finon@centre-cired.fr) is Senior Research Fellow in Economics at the Centre International de Recherche sur l'Environnement et le Développement, France. He is the present chair of the French Association of Energy Economists.

Pros and cons of alternative policies aimed at promoting renewables

1. Introduction

The promotion of renewables re-started in Europe – and elsewhere in the world – during the first half of the 1990s to align with the objective of reducing greenhouse gas emissions and with energy efficiency policies. This follows an earlier, though temporary, boost in renewables after the oil shocks of the 1970s. Efforts have been directed at developing renewables, in general, and using them for the production of electricity in particular. Specifically, the European Directive 2001/77/C on the promotion of renewable energy sources aims to increase the share of renewables in the energy supply of the EU from 5.6 percent in 2000 to 11.8 percent in 2010. At the same time, the share of renewable energy sources in the production of electricity is targeted to increase from 14 to 21 percent. Moreover, the European Council of March 2007 endorsed a binding target of a 20-percent share of renewable energies in overall EU energy consumption.

Dominique Finon

Reflecting earlier experience with, and limitations of, investment subsidies for renewables, efforts made as from the mid-1990s were initially based on two instruments: feed-in tariffs and tendering systems for long-term contracts at guaranteed prices. Traditional instruments (soft loans, investment grants, tax allowances, tax exemptions, and so on) to encourage the diffusion of near-market innovative technologies complement these two principal instruments. More recently, conventional wisdom seems to suggest that moving towards market-based instruments – such as a green certificate trading scheme – is necessary to reduce the costs of promoting pre-commercial-stage technologies.

In fact, there is the notion of an optimal sequencing of instruments: preferential prices (feed-in tariffs or tendering of contracts at guaranteed prices, for instance) in the early pre-commercial phase of promoting renewables followed by a trading scheme based on green certificate quotas. Moreover, adopting a European-wide system of such quotas is considered a means for minimising the cost of increasing the share of renewables. This is because a trading scheme based on green certificate quotas would shift the development of renewables to those EU member states that are relatively well endowed with renewable energy sources and could thus supply them at low costs.

Against this background, this paper examines the pros and cons of alternative policies to promote renewable energy technologies that have not yet reached commercial maturity. Section 2 briefly reviews the rationale for promoting renewable energy and introduces criteria for comparing alternative policy instruments. Using these criteria, Section 3 shows what economic theory tells us about the pros and cons of alternative policies. Section 4 moves beyond theoretical considerations by offering practical policy lessons. Section 5 concludes.

2. Promoting renewable energy: rationale, instruments, and criteria for assessing alternative instruments

2.1 The rationale for policies in support of renewable energy

Governments traditionally support research and development and demonstration projects in a variety of sectors given well-known market failures in the creation and diffusion of product and process innovation. To the extent that such market failures arise in the area of renewable energy, too, there is an argument for promoting renewable energy technologies. However, more important in the case of renewables is their contribution to replacing fossil fuels associated with environmental externalities – in particular climate change.[1]

Focussing on environmental aspects, it is nonetheless pertinent to ask whether there is a need to promote renewables in situations where economic policies internalise the environmental externalities of polluting energy technologies. In principle, solutions exist that internalise the environmental effects of using fossil fuels and, thus, get round the need for specific instruments in favour of renewables. As for climate-change externalities, these instruments are a tax on CO_2 emissions (with the tax equalling the climate-change externality) or a quota-and-trade system (with the quota on CO_2 emissions set at the socially optimal level). Indeed, a criticism of specific policies in favour of renewables is that the costs of avoiding CO_2 emissions through these policies (€100-€150/t CO_2) exceed by far the estimated social damages of CO_2 emissions ($20-$30/t CO_2) (see, for instance, Newbery 2003, and Fischer and Newell 2004).

Imposing a sufficiently high CO_2 tax, which would foster the replacement of fossil fuels with renewables, encounters a variety of difficulties.

But imposing a sufficiently high CO_2 tax (or stringent CO_2 quota), which would foster the replacement of fossil fuels with renewables and technological progress, encounters a variety of difficulties. First, estimating an optimal CO_2 tax or CO_2 quota is surrounded by considerable uncertainties. Second, both instruments – notably when set at the right level – will have visible distributional effects and, thus, encounter problems of acceptability. This is evidenced by the refusal of a European eco-tax at the beginning of the 1990s and the small number of EU member states (Denmark, Sweden, the Netherlands, and Finland) that imposed a significant eco-tax on energy. Third, there is no guarantee that even a high price of CO_2 emissions will trigger more substitution of new clean technologies for fossil fuels than policies directly supporting clean technologies (see, for instance, the literature reviewed in Jaffe *et al*. 2002). This is mainly for two reasons: regulatory uncertainty as to the price of CO_2 (which is liable to follow, for instance, when the period during which the quota applies is too short, as in the EU Emission Trading Scheme) and entry barriers for renewable technologies.

To elaborate on the last point, even when internalising the environmental costs of using fossil fuels, renewable technologies might face entry barriers. For instance, technologies at an early stage of development might be expensive now, but their costs are likely to fall considerably as and when they gain commercial maturity. Kolev and Riess (this volume) discuss the underlying rationale for promoting new renewable technologies in greater detail. Another entry barrier stems from constraints and costs of integrating decentralised, renewable technologies into an existing centralised infrastructure. The cost of integrating renewable electricity into the network (comprising network investment cost, balancing cost for intermittent production, the cost of regulating voltage and frequency, and so on) are indeed among the most important obstacles for developers and producers of renewable electricity. And then, there are constraints of land use by, for instance, renewable power units with landscape impacts and projects that rest on the development of energy crops. Easing these constraints requires specific siting and land planning procedures, and not having such procedures increases the risk for the developers of new technologies.

The arguments presented so far highlight environmental reasons for promoting renewable energy. But replacing fossil fuels with renewables is expected to bring non-environmental benefits, too. For instance, it can be argued that increasing the share of renewables in the energy mix of a country brings diversification benefits and enhances security of supply – an issue that Awerbuch and Yang

1 For simplicity, we ignore possible environmental effects of renewable energy resources – impacts on the landscape, for instance.

(this volume) address in detail. Furthermore, some governments consider the support for new, clean energy technologies as a means to foster competitive export industries, employment, and regional development.

In sum, the existence of entry barriers suggests a need for renewable energy deployment policies alongside efforts to internalise the environmental costs of using fossil fuels. And the nature of these barriers suggests a need for a successive set of policies rather than a one-for-all second best policy. In addition, there is a variety of non-environmental reasons that could justify instruments in support of renewable energy. Let us then look at possible instruments.

2.2 Instruments and their link to the development stage of new technologies

The choice of policy instrument needs to reflect the stage of development of the renewable technology.

A first fundamental observation is that the choice of policy instrument needs to reflect at what stage the development of a renewable technology is. Considering the sequence of development stages from the R&D phase, to the demonstration phase, pre-commercial phase, and – finally – to the commercial maturity of the technology, the choice and sequence of instruments could be broadly described as follows (Foxon *et al*. 2005): R&D subsidies in the R&D phase; investment grants and public procurement in the demonstration phase; purchase obligations, quotas, or fiscal incentives in the pre-commercial phase; full reliance on the energy price effect of internalising the environmental cost of fossil fuels at the stage of commercial maturity. What this indicates is, in essence, a shift from directly subsidising investment in renewable technologies during the R&D and demonstration phases to subsidising the production of renewable energy in subsequent stages of developing technologies.

One reason for this shift is that supporting investment beyond the demonstration phase often leads to productive inefficiencies and is exposed to the risk of stop-and-go policies. Direct investment support is prone to two problems. First, there is limited concern about long-term performance and maintenance. Experience with direct subsidies and tax credits for investment in renewable electricity capacity – largely used in the 1980s, notably in United States – has indeed shown that producers soon neglected maintenance after the capacity went on stream and stopped it at the first operational incident (Sawin 2004). Second, when investment-tax-credit regulation expires, new projects dry up and the industry producing renewable energy equipment tends to collapse. Against this background, the shift from supporting investment to supporting production needs to be carefully designed, taking into account the development phase of a technology.

Let us then focus on the three main instruments in support of the production of renewable energy (rather than direct support of investment in the underlying production capacity): feed-in tariffs, tradable green (i.e., renewable) certificate systems, and tendering systems – all coming ideally after direct investment support has pushed technologies beyond the demonstration phase. These instruments have common characteristics. To begin with, they effectively subsidise renewable energy production over a long period of time (10-15 years), covering the economic lifespan of the equipment. As a result, the return on investment is increased and the pay-back period reduced by rewarding producers for the actual amount of energy produced. All this assumes that the support will indeed be available (and sufficiently high) over the investment lifetime – an issue that will be discussed in more detail in Section 3.

Second, the instruments rest on an obligation to buy renewable energy, green certificates, or both. To be effective, the economic agents mandated to buy must be clearly designated, and mandated buyers are generally the suppliers of electricity. For completeness, it is useful to point out that there may also be voluntary buyers of renewable energy or green certificates. While this contributes to the promotion of renewable energy, it is bound to be inefficient given the well-known free rider problem.

Third, these instruments do not involve public expenditure. Typically, the higher financial cost of producing renewable energy is passed on to consumers through the electricity price. For European countries, the wholesale price effect of this currently ranges from €1-€3/MWh – which is very little compared to a wholesale price of around €50-€55/MWh in 2006-2007. Although it is true that the instruments do not require public expenditure, it is possible that the costs are shared between energy consumers and tax payers. In the United States and some European countries, for instance, policies existed that combined a purchase obligation (raising energy prices) with a tax credit on the production (lowering tax revenue). The advantage of fully passing on costs to consumers and not making the support of renewables dependent on budgetary decisions is that this provides a more reliable investment framework.

2.3 Criteria for assessing alternative policy instruments

Instruments for fostering renewable energy should aim at achieving a social optimum.

In principle, the choice of instruments for fostering renewable energy needs to rest on the objective of achieving a social optimum, and many dimensions of a social optimum need to be taken into account – notably the estimated value of negative environmental externalities. Given the difficulty of reliably estimating such externalities, there are – in practice – implicit or explicit quantitative 'renewable' targets to be achieved over a specific time span, maximum acceptable costs of these instruments, and distributional considerations.

In general, when designing environmental policies in the presence of uncertainty about the costs of environmental damages, one cannot reason simply in terms of cost-benefit analyses or second-best optimal tax policies. Rather, it is more appropriate to conceive policies that achieve a targeted reduction in pollution in a cost-effective manner (Baumol and Oates 1988). This is also true when it comes to designing policies in support of renewable electricity, mainly because of the enormous difficulty of reliably estimating the benefits of such policies, i.e., the economic value of emissions avoided and other benefits of using renewables for electricity generation.

If we thus take the objective of raising the share of renewables to a certain level (without trying to assess the social benefits of meeting this objective), one could think of three criteria to compare alternative policies. First, the criterion of social efficiency – here largely defined as cost-effectiveness. Obviously, one would like to achieve the objective at least cost. A variety of issues need to be considered in this context. Uncertainty and how it influences the cost-effectiveness is one. Another is whether alternative instruments differ in their impact on technology development and, thus, the cost of renewables in the future. Does the instrument incite the deployment of a variety of renewable technologies of different degrees of maturity with a technology-specific design, or does it play by encouraging indistinctly the set of renewable energy technologies? In a perspective of dynamic efficiency, the cost increase resulting from a technology-specific support can be beneficial in a long-term perspective: indeed, it opens the way to large-scale deployment of new technologies before the exhaustion of the resource potential of front-runner technologies in order to avoid an undue cost increase in the transition from front-runner technologies to the next generation of renewable technologies (Neuhoff 2005, Huber, *et al.* 2001). And then, there is the question of how the cost-effectiveness of an instrument might change if it is jointly applied by a group of countries, such as the European Union, rather than by each country individually.

It is worth noting that defining social efficiency in this way leaves open the possibility to examine the distributional effects of alternative policies, notably the rents accruing to producers of renewable energy located in regions with a natural advantage in using renewables or – more generally – producers generally benefiting from learning effects. In sum, a legitimate government concern is to avoid producers' rents and the cost for consumers to become too high. This means, in particular, that policies must be flexible, including the possibility of revising downward or even terminating the support in light of the progress that renewable technologies make towards becoming commercially competitive without further public support.

The second criterion for assessing alternative policies is environmental effectiveness, which is measured by the additional renewable energy capacity installed due to the policy, taking into account that the support should be set as low as possible to achieve the desired result. The capacity impact of a policy depends on the size of the incentives it offers and on how long investors can count on them. The size of the incentive is determined by the additional remuneration per unit of renewable energy produced. To successfully increase capacity, the additional remuneration must cover all the extra costs and risks of energy production, taking into account that these vary across different technologies. As for how long investors can count on the incentives, two factors are of importance: the long-term predictability of the policy itself and the efficiency of the trilateral relationship between the government, mandated purchasers, and developers/producers of renewable energy (Langniss and Wiser 2003; Finon and Perez 2006). Obviously, producers of and investors in renewable energy production need stable and predicable cashflows, which are heavily influenced by the indirect subsidy offered by alternative policies. All this implies that even when there is a change in policies for future investments, the long-term commitment of the government and the relationship between parties must remain intact for investments already made.

The impact of policies promoting renewables depend on the size of the incentives they offer and on how long investors can count on these incentives.

The third criterion guiding the comparison of alternative policies in support of renewables concerns their conformity with the market regime of the energy sector in question. An instrument to support renewable electricity should be compatible with the market regime of the electricity industry. But this does not mean that an instrument must rely as much as possible on a market mechanism. Rather, the design of a policy in favour of renewables should be coherent with the market principles by not distorting competition.

3. Comparing alternative policies in support of renewables

To work out as clearly as possible how alternative policies in support of renewables score against the criteria introduced in the previous section, this section will concentrate on two key policy alternatives: feed-in tariffs and tradable green certificate (TGC) systems. We will begin with a brief description of these instruments, move on to analysing them in-depth in light of the criteria set above, and then propose an overall assessment.

3.1 Salient features of feed-in tariffs and tradable green certificate systems

To start with feed-in tariffs, their main characteristics are: an obligation to purchase electricity based on renewable energy at a fairly high price – with both obligation and price guaranteed over a long period of time (8 years in Spain, 15 years in France, and 20 years in Germany, for example). The purchase obligation is with the distributors-suppliers in their service areas and it applies to all new renewable power generation units. To promote the development of a diverse set of renewable technologies, feed-in tariffs differ across technologies. They reflect the generating costs of a typical renewable electricity unit (including some risk premium) and are not set on the basis of the avoided generating cost of the distributor-supplier subject to the purchase obligation. Unless the supply curves for renewable electricity are known, the quantity of renewable electricity production resulting from setting feed-in tariffs is not known *ex ante*.

The recovery of the extra cost of renewable electricity that mandated buyers incur can be organised in three ways: an increase in the price of every kWh sold by the distributors subject to the purchase obligation when such distributors have a legal monopoly as in the former monopoly regime; a compensation between competing distributors-suppliers (given that they are obliged – irrespective of their own sales – to buy all the renewable electricity produced in the area of their distribution networks); or reimbursements financed by a tax on all electricity transmitted via the national grid. In the latter case, the extra cost of renewable electricity is paid by all electricity consumers. An alternative, or complement, to passing on the extra cost to electricity consumers is budgetary support to mandated buyers. Budgetary support could also be given to producers of renewable electricity to limit the level of the feed-in tariffs and, thus, cost for consumers; this could be done either through eco-tax and/or VAT exemption, as in the Netherlands and Denmark, or tax credits on the renewable electricity production – as in the United States.

Turning to tradable green certificate systems, the main features of this instrument are the following. It designates economic agents subjected to a rising renewable, or green, electricity quota (normally, these agents are electricity suppliers or distributors/retailers) and eligible technologies and installations (typically including only new installations and possibly excluding new large hydro plants and waste incineration). Designated agents – suppliers for short – can fulfil their quotas (expressed in percent of each supplier's annual sales of electricity and rising over time) in different ways. They can produce renewable electricity, purchase it under long-term contracts from specialised producers, or purchase green certificates, which originate from suppliers that exceed their quotas or from specialised producers that choose to sell part of their renewable electricity in the market rather than directly under long-term contracts. The quota is complemented by a penalty to be paid in case of non-fulfilment of the quota. But this penalty could be seen as a safety valve rather than a threat to force suppliers to meet their quotas. Rather than fulfilling his quota, a supplier may opt to pay the 'buy-out price' (its name in the UK system) for not meeting the quota, which could in extreme cases represent the full quota. In essence, this buy-out price puts a ceiling on the cost of renewable certificates. A last trait of the TGC design is the reallocation of penalty revenues to the agents who strictly respected their quotas, which is an incentive to respect the quota.

The price of green certificates acts like a premium for the production of renewable electricity.

A number of conceptual differences between feed-in tariffs and TGC systems are worth highlighting. In contrast to feed-in tariffs, TGC systems directly specify the targeted quantity of renewable electricity. This being said, as suppliers can also fulfil their quotas by purchasing renewable electricity or green certificates, TGC systems create competition that encourages the production of renewable electricity at least cost. Another difference is that TGC systems do not impose a contractual arrangement on price and quantity between producers and buyers of renewable electricity and green certificates. The remuneration for producing renewable electricity essentially has two components. One is the price of electricity as determined in the electricity market where renewable electricity is sold. The other is the price of certificates as determined in the market for green certificates. Thus, the price of green certificates acts like a premium for the production of renewable electricity. Finally, as the renewable electricity quota for a country as whole is allocated in an equitable way to competing suppliers, there is no need for a specific financing mechanism to compensate suppliers for the extra cost of fulfilling their quotas. This facilitates the acceptance of this type of support for renewables by electricity regulators and large users of electricity – both always eager to limit the cost of promoting renewables.

3.2 Cost-effectiveness

3.2.1 Static and dynamic cost-effectiveness – a closed-economy perspective

As in other areas of environmental policy, price-based instruments (here: feed-in tariffs) and quantity-based instruments (here: TGC systems) lead to similar results if the cost of renewable electricity is known with certainty, transaction costs are zero, and no dynamic learning effects are considered. In these circumstances, setting a feed-in tariff at level *p* will result in an overall quantity of renewable electricity *q* and, conversely, fixing that quantity upfront under a TGC system will result in a price *p* for renewable electricity.[2] In other words, with perfect information and zero transaction cost, whether the government fixes the price – as in the case of feed-in tariffs – or the quantity, as in the case of TGC systems, makes no difference. Moreover, it does not matter whether or not the price or quantity is the same for all technologies.

However, it is equally well known that when information is incomplete and when the shape of cost curves is uncertain, price-based and quantity-based instruments lead to different results (see Cropper and Oates 1992 and Weitzman 1974). In fact, depending on the shape of cost curves for renewable electricity, feed-in tariffs may be better than TGC systems – and *vice versa*. Another important issue here is whether the total cost of producing renewable electricity turns out to be higher or lower compared to the level anticipated when choosing between feed-in tariffs and TGC systems.

When information is incomplete and when renewable electricity costs are uncertain, price-based and quantity-based instruments lead to different results.

Let us illustrate all this with the help of Figure 1. If marginal costs are known with certainty, as represented by *MC* in Figure 1, setting a feed-in tariff *P* will result in an output of renewable electricity *Q*. Given this tariff and quantity, consumers pay for renewable electricity an amount indicated by the area *OQXP*. This amount can be broken down into a part covering actual production costs (*OQXW*) and a part representing producers' surplus (*WXP*). With certainty, a TGC system leads to the same outcome.

Consider now the situation where, contrary to expectations at the time when deciding either in favour of feed-in tariffs or a TGC system, the marginal cost curve turns out to be *MC'* rather than *MC*. In case a feed-in tariff is the policy instrument, renewable electricity production turns out to be *Q'*, the cost for consumers is *OQ'YP*, and producers' surplus increases to *WYP***.** Depending on the size of the difference between expected and actual costs, the impact on output and the cost for consumers can be substantial. In case a TGC system is the instrument, output reaches the expected level *Q*, the price of renewable electricity is *P'*, cost for consumers is *OQZP'*, and producers' surplus is *WZP'*. All in all, when the marginal cost of renewable electricity turns out to be lower than expected by policy makers, feed-in tariffs deliver a higher output at the pre-determined price *P* whereas a TGC system delivers the targeted output at a lower price.

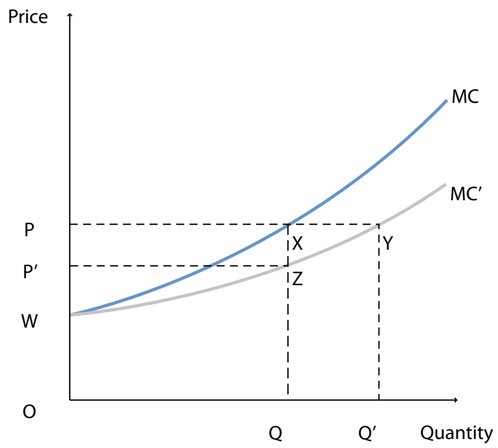
Further insights can be gained when we compare both instruments for the same output level, notably *Q*, that is, the output explicitly or implicitly targeted by both instruments. As Figure 1 indicates, for this output, consumers would save an amount equal to *P'ZXP* if the instrument is a TGC system and not a feed-in tariff. It is worth noting that what consumers pay less in the case of a TGC system comes fully at the expense of the producers' surplus. The conclusion so far is that when the cost of producing renewable electricity turns out to be lower than expected, the TGC system is more

2 It is useful to note that in equilibrium and under perfect information, the feed-in tariff is equal to the shadow price of the quantity objective of the TGC system, which is equal to the sum of the price of green certificates and the price of conventional electricity.

cost-effective than a feed-in tariff and it limits the risk of an excessively high output and burden on consumers.

The results are reversed, however, when the marginal costs of producing renewable electricity prove to be higher than anticipated (in Figure 1 the marginal cost curve would lie above *MC*). With a feed-in tariff, a smaller-than-expected amount of renewable electricity will be produced at the pre-determined feed-in tariff. With a TGC system, the targeted quantity will be generated, but at a higher price than expected – and higher costs for consumers. What is more, the producers' surplus under a TGC system will be larger than the surplus under a feed-in tariff. In sum, when reality shows that assumptions about the cost of producing renewable electricity were too optimistic, a price-based instrument (i.e., feed-in tariffs) will be more cost-effective than a quantity-based instrument (i.e., TGC systems). In this situation, feed-in tariffs set a ceiling for the marginal cost of each renewable technology; conversely, a TGC system – while directly controlling output – leaves the setting of prices to the workings of the system, possibly resulting in surprisingly high prices.

Figure 1. Cost-effectiveness and cost for consumers: feed-in tariffs *vs.* tradable quotas with lower-than-expected marginal costs of producing renewable electricity



Hybrid solutions, combining price-based elements with quantity-based elements, can limit the costs for consumers.

Both types of instruments can limit the costs for consumers through hybrid solutions, which combine price-based elements with quantity-based elements (Roberts and Spence 1976). To illustrate, as mentioned above, TGC systems usually include the option that electricity suppliers pay a penalty instead of (fully) meeting their quotas. This option provides a safety valve in case renewable electricity turns out to be much more expensive than anticipated. In the price-based approach, two approaches are developed. Under the first, successive downward adjustments to the feed-in tariffs can be made at certain intervals if marginal costs turn out to be lower than anticipated. Under the second approach, downward adjustments of prices for new capacities are programmed upfront, reflecting – among other things – anticipated cost declines due to learning and experience effects, technical progress, and the increasing use of renewable resources.

Reflecting the possibility of overestimating marginal costs and, equally important, successive future cost declines, feed-in tariffs are often considered less suited than TGC systems to let market forces play their role in controlling the cost for consumers. While producers of renewable electricity can, and probably will, exert competitive pressure on the producers of equipment used for generating renewable electricity, any decline in the cost of equipment will boost the profits of producers – as long as the level of feed-in tariffs remains unchanged for new capacity. By contrast, TGC systems have the potential to encourage *ex ante* competition not only between equipment producers, but also between renewable electricity producers. The main reason for this is that obligated suppliers have a keen interest in minimising the cost of complying with their quotas.

However, all this does not mean that we can necessarily expect TGC systems to deliver renewable electricity at lower cost for consumers than feed-in tariffs. This is because TGC systems come with far more uncertainties than feed-in tariffs, and this higher degree of uncertainty affects the relationship between obligated suppliers, producers, and financiers. Ultimately, this results in higher risk premiums that suppliers, producers, and financiers take into account when embarking on renewable electricity projects, thereby raising the cost of such projects.

One should not necessarily expect a tradable green certificates system to deliver renewable electricity at lower cost than feed-in tariffs.

To illustrate this point, let us look at the revenue characteristics of a renewable electricity project under each of the two instruments. In the case of feed-in tariffs, revenues are fairly certain given that there is a guaranteed price at which production can be fed into the network. In the case of TGC systems, revenues depend on the uncertain market price of electricity and the uncertain price of green certificates: electricity-price risk combines with green-certificate-price risk. To the extent that some suppliers do not fulfil their quotas and pay a penalty instead, there is an additional source of revenue for suppliers that meet their quotas because total penalties paid are allocated *ex post* to them. But this source of revenue is uncertain too.

When the production of renewable electricity is difficult to schedule – as in the case of wind energy, for instance – the electricity-price risk is exacerbated by uncertainties concerning balancing costs, which – in TGC systems – are entirely borne by producers of renewable electricity. In TGC systems, the generation of renewable electricity needs to observe all electricity market rules, including those pertaining to the balancing market mechanism that aims at ensuring the reliability of the whole power system (Mitchell *et al*. 2004). By contrast, under feed-in tariffs, renewable power plants do not need to supply a certain load profile and the balancing costs fall on obligated suppliers.

Revenue risk also arises from uncertainty as to how the quota will increase over time and, in particular, at which level it will not be raised any further. When the quota approaches its ultimate level, investment in additional renewable electricity generating capacity may create an oversupply of green certificates and, thus, a drop in their price. This adds to the risk of renewable energy projects in a TGC system and, thus, their costs.

Both instruments considered here also differ in their exposure to political and regulatory risks. In general, the impact of regulatory changes is more difficult to anticipate under a TGC system than under a system of feed-in tariffs. For instance, although it is clear that a decrease in the penalty for not observing quotas will reduce the price of green certificates and, hence, the amount of penalties that can be passed on to complying suppliers, the size of this effect is difficult to anticipate. This creates uncertainty as to the overall profitability of investment in renewable electricity projects. Moreover, in a feed-in tariff system it is easier than in a TGC system to insulate existing renewable projects from the effects of changes in rules governing the system. For instance, although a cut in feed-in tariffs could apply, in principle, to both new and existing projects, the latter are usually protected by appropriate legal or institutional arrangements.[3] Things are different in a TGC system. For instance, broadening the scope of eligible technologies (e.g., making wood co-firing eligible) and changing the way technologies are certified (i.e., the quantity of certificates associated to a

3 For instance, in Germany, where feed-in tariffs are backed by two successive laws (EFL law of 1990 and EEG law of April 2000 – amended in 2004), legal and constitutional principles effectively protect investors from the regulatory change (Langniss and Wiser 2003). In France, developers of renewable electricity under feed-in tariffs are less well protected because the implementation of such tariffs is a matter of decree and not of law (decree of the 6.12.2000).

technique) could increase abruptly the number of certificates. The ensuing drop in the price of green certificates will adversely affect the profitability of existing renewable projects, with investors having no legal protection against this form of partly alienating the value of 'green' capacity installed in the hope that the regulatory framework remains unchanged.

The considerable uncertainties surrounding renewable projects in TGC systems has two main consequences. First, other things being equal, the relatively high risk will make finance for renewable electricity more costly than it would be under feed-in tariffs, thereby increasing the cost of renewable electricity. Empirical support for this hypothesis will be provided in Section 4, where we look at the experience with alternative policy instruments.

The second consequence of TGC systems is that they tend to reduce competition among producers of renewable electricity. In addition, there are built-in tendencies that undermine competition in the market for green certificates, thereby distorting the price signal that this market is expected to send. To see why, consider first that the risks and transaction costs associated with supplying renewable electricity under a TGC system will encourage large, independent suppliers to become part of vertically integrated electricity companies, or – alternatively – buy electricity and trade green certificates under long-term contracts. Indeed, long-term contractual arrangements rather than spot transactions on the market of green certificates and/or vertical integration could become necessary to buttress the profitability of renewable electricity projects and thus ensure their financing.[4] All this implies that only the small obligated suppliers with changing loads will make permanent use of the certificates market, implying limited competition and liquidity in that market. This creates an opportunity for large obligated suppliers – though not really relying on the certificate markets – to exert market power to increase certificate prices and thus the cost of renewable electricity to consumers.

Let us now move beyond static aspects of cost-effectiveness and briefly consider it from a dynamic efficiency viewpoint. The main point to develop here is that the relatively large producers' surplus generated by feed-in tariffs (thereby making this instrument not look particularly cost-effective from a static viewpoint) is the very means of fostering technological progress and, thus, cheaper renewable electricity over time. In terms of Figure 1, the hypothesis is that feed-in tariffs lead to a faster rightward shift of the marginal cost curve over time than a TGC system would .

There are three channels through which feed-in tariffs encourage technological progress. First a generous producers' surplus spurs the deployment of renewable technologies which, in turn, gives rise to learning and experience effects that lower marginal production cost.

A generous producer surplus allows greater research and development efforts by equipment manufacturers.

Second, a generous surplus allows greater research and development efforts by equipment manufacturers. This is because a high surplus and attractive feed-in tariffs allow electricity producers to offer equipment manufacturers better prices than they could under the more competitive TGC instrument. In fact, the experience of German, Danish and Spanish producers of wind energy equipment indicates the favourable impact that feed-in tariffs can have on the development of nascent technologies. That said, this favourable impact and, thus, the advantage of feed-in tariffs weakens once a performing international industry emerges.

4 In the United Kingdom, since the implementation of the Renewable Obligation Certificates (ROC) system in mid-2002, most of the investments have been carried out by subsidiary companies of the five large suppliers, and a minor part of the investment has been undertaken by independent producers protected by long-term contracts with a minimum contractual timespan of 13 years. This has also been observed in Texas where, despite the possibility to exchange certificates, all the distributors-suppliers who carry the obligation have negotiated bilateral, long-term contracts (10 to 25 years) to reach their quotas of renewable electricity (Langniss and Wiser 2003).

Third, compared to the TGC instrument, feed-in tariffs are easier to design with technology-specific tariffs so that electricity producers do not only select the most mature renewable technologies, but also promising technologies that are at an earlier stage of technological and commercial maturity. Some authors argue that market-based incentives of TGC systems stimulate risk-taking and innovation too (see Egenhofer 2005, for example). Although true, the stimulus to innovation is unlikely to match that of feed-in tariffs, which allow firms to adopt innovative technologies while being sure of a reasonable stream of revenues.

Overall, there is good reason to believe that feed-in tariffs are better than TGC systems in fostering technological progress. Somewhat paradoxically, supporters of TGC systems argue that if technology-specific investment grants for projects based on second-ranked technologies complement TGC systems, their drawback in fostering technological progress will be mitigated. Alternatively, it is sometimes proposed to create technology-specific TGC systems; however, instead of one big and possibly efficient market for exchanging certificates, there would then be a number of different – but smaller – exchanges.

We will now broaden the analysis by examining the cost-effectiveness of feed-in tariffs and TGC systems in a European policy context.

3.2.2 Cost-effectiveness – a European perspective

An alleged advantage of an EU-wide tradable green certificates system over feed-in tariffs is that it would foster a cost-effective development of renewables across the EU.

An alleged advantage of a TGC system over feed-in tariffs is that it would – if applied at the European level – foster a cost-effective development of renewables across the EU. The underlying economic logic is a straightforward extension of the one underpinning the case for TGC system at the national level. As argued above, at the national level, a firm subject to a renewable electricity quota can meet its obligation by producing renewable electricity, purchasing it (directly from other firms, or indirectly from the green certificates market), or a combination of the two. Applied to the EU – where each member state will have to observe its quota – this would mean that a country with comparatively high costs of renewable electricity buys renewable certificates from other countries for as long as its marginal cost of producing renewable electricity exceeds the sum of the certificate price and the market price of electricity. And *vice versa*: a country with low costs of producing renewable electricity produces such electricity in excess of its quota, earns green certificates for this excess, and sells them in the certificates market – all this making sense as long as its marginal cost of producing renewable electricity is lower than the sum of the certificate price and the market price of electricity. In equilibrium, supposing a common price of electricity across EU member states, there would be least-cost production of the targeted amount of renewable electricity, a common price of green certificates, and the same long-run marginal cost of producing renewable electricity in all member states.

Notwithstanding the useful role a European-wide TGC system could play, in principle, in efficiently achieving renewable electricity targets, it must be emphasised that for such a system to be fully effective, two conditions must hold: first, there must be an integrated European electricity market and, second, the national institutional and regulatory frameworks supporting the green certificates market must be harmonised. In fact, establishing such a system in a situation where these conditions are not fulfilled could lead to considerable inefficiencies. The remainder of this sub-section explains why.

To start with the need for an integrated electricity market, note that in equilibrium, the certificate price is the difference between the marginal cost of renewable electricity and the market price of electricity (see Box 1). Consider now a situation where EU certificate markets are fully integrated but electricity markets are not. There will then be a common price of certificates, but different electricity prices. It follows that the marginal costs of renewable electricity differ across EU countries – contrary to what would happen in the idealised world of fully integrated electricity markets. Moreover, there will be undue rents to producers of renewable electricity in countries with higher electricity price.

This being said, introducing a common TGC system in Europe in lieu of national TGC systems could nonetheless lead to a more efficient allocation of renewable electricity production across Europe. For instance, with perfectly integrated certificates markets, but imperfectly integrated electricity

Box 1. The benefits of a European TGC system in imperfect market conditions

Considering a tradable green certificate (TGC) system, the purpose of this Box is to set out the equilibrium relationship between the market price of electricity, the price of green – or renewable – certificates, and the marginal costs of producing renewable – or green – electricity. We will illustrate this relationship by considering three cases, involving two countries.

A key observation to make upfront is that renewable electricity is a joint product, comprising two sub-products: 'normal' electricity (i.e., electricity as consumed, comprising renewable electricity and electricity from polluting plants) and – assuming that renewable technologies replace polluting ones – better environmental quality (which results from establishing property rights on the environment). The price of green certificates is linked to the marginal cost of renewable electricity and the market price of electricity. An important corollary is that the sub-products are sold on independent markets (the green certificates market and the electricity market), which have different structures and demand functions.

Case 1 – each country has its own TGC system

In equilibrium, the relationship between the market price of electricity (P_E), the price of green certificates (P_C), and the marginal costs of producing renewable electricity (MC_R) is:

$$P_C^A = MC_R^A - P_E^A$$

$$P_C^B = MC_R^B - P_E^B$$

The first equation shows this relationship for country *A* and the second equation for country *B*. Let us assume that both countries produce the same amount of renewable electricity, but that country *B* produces at lower marginal cost than country *A* ($MC_R^B < MC_R^A$). In these circumstances, both countries could achieve the same overall amount of renewable electricity at lower costs if country *B* increased its production and if country *A* reduced its production. Assuming upward-sloping marginal cost curves, no further efficiency gains can be made once marginal cost are the same in each country.

Case 2 – countries introduce a common TGC system

Variant 1: perfect integration of certificate markets and of electricity markets

If both countries move to a common TGC system and succeed in establishing a fully integrated electricity market, the price of certificates and the price of electricity will be the same in both countries, that is, $P_C^A = P_C^B$ and $P_C^A = P_C^B$. It follows that, in equilibrium, marginal costs are also the markets, the production of renewable electricity could fall in countries with initially high marginal cost of renewable electricity and it could increase in countries with low marginal costs. But the opposite could happen, too, as set out in detail in Box 1, when a common TGC system would reduce rather than increase the cost-effectiveness of producing renewable electricity. The risk of this happening is high if countries with costly renewable resources have high electricity prices and countries with low-cost renewable resources have low electricity prices.

same in both countries, which – in turn – implies that renewable electricity production has risen in country *B* and fallen in country *A*. In these circumstances, a common TGC system results in an optimal allocation of renewable electricity production across countries.

Variant 2: perfect integration of certificate markets, but imperfect integration of electricity markets

Things are different when the certificate markets are fully integrated ($P_C^A = P_C^B$) – which is a reasonable assumption as certificates are tradable financial assets – but the electricity markets are not ($P_E^A \neq P_E^B$) – because of physical constraints to exchange electricity between countries, for instance. In these circumstances, marginal costs of producing renewable electricity are not equalised across countries ($MC_R^B \neq MC_R^A$) and, thus, a common TGC system does not lead to an optimal allocation of renewable electricity production across countries. Whether or not the ensuing allocation of renewable electricity is better than the one under separate TGC systems depends on the constellation of P_C^A, P_C^B , P_E^A and P_E^B before establishing a common TGC system. In principle, three outcomes are possible.

First, the allocation of renewable electricity production between the two countries moves in the right direction, but falls short of reaching its optimum. With $MC_R^B < MC_R^A$ before creating a common system, this means that production and marginal costs increase (fall) in *B* (*A*) – but MC_R^B remains smaller than MC_R^A.

Second, the allocation of renewable electricity production between the two countries moves in the right direction, but exceeds its optimum. In other words, production and marginal costs increase (fall) in *B* (*A*) beyond the optimal level, resulting in $MC_R^B > MC_R^A$. This outcome could be less, equally, or more efficient than the situation before introducing a common TGC system.

Third, the allocation of renewable electricity production between the two countries moves in the wrong direction: country *B*, which produces renewable electricity at lower marginal costs than country *A* before the creation of a common TGC system, reduces its renewable output while country *A* raises its output. As a result, the positive difference between MC_R^A and MC_R^B becomes even greater. This outcome is clearly less efficient than the situation before introducing a common TGC system.

In sum, without fully integrated European electricity markets, a common TGC system does not ensure an optimal allocation of renewable electricity production across countries. The reasoning sketched here for given capacities to generate renewable electricity applies, too, to investments in additional capacity: investments in additional renewable power capacity might not take place where it is cheapest.

What are then the chances for a well-integrated EU-wide electricity market? While some EU governments and the European Commission continue to aim at creating such a market, a truly functioning EU-wide market with common electricity prices is very unlikely to materialise. A more realistic scenario is the emergence of smaller regional markets, each possibly comprising a number of EU countries. For instance, the markets of the Nordic countries are fairly integrated, and there are indications of interaction between continental European markets (France, Germany, Austria, Belgium, and the Netherlands) during large periods of the year. Efforts – led by the Commission, regulators, and transmission system operators – to harmonise access to national grids and trans-border interconnections will improve this nascent integration (European Commission 2005). Although true, because of limited interconnection capacities between markets, it is unlikely – neither over the short nor medium term – that the so-called electric peninsulas (i.e., Italy, Iberic countries, Nordic countries, the United Kingdom, and Greece) will be integrated in a pan-European market.[5] In the long run, electricity prices could possibly converge even if the physical separation of markets endures thanks to the convergence of technology mixes across countries that will happen with the replacement of existing generating capacity. However, differences in technology mix will persist because of differences between countries in, for instance, coal policies, hydro resources, and acceptance of nuclear technology.

A necessary condition for an effective common tradable green certificates system is that EU members harmonise all rules governing the system.

A second condition for an effective common TGC system is that EU members harmonise all rules governing the system (eligible technologies, type and duration of the certificate, certificate exchange rules, and so on).[6] In addition, they must harmonise – or abolish – all other instruments in support of renewable electricity (tax credits, investment grants, preferential indirect taxation, and so on). Obviously, a country that maintains such instruments creates an artificial cost advantage for its producers, thereby distorting competition between projects at the European level and preventing an efficient allocation of new investment in renewable energy.

Advocates of an EU-wide TGC system appear to minimise the distortion that could arise if these two conditions are not met and/or the intrinsic difficulties of harmonising the rules of the game and of integrating Europe's electricity markets. It should also be pointed out that the economic advantage of a common TGC could be reached through a burden sharing agreement among member states. Based on the estimated potential of renewable resources in each member state, members' obligations (that is, a targeted percentage of renewable electricity in national electricity production) could be set to approximate an equalisation of marginal costs. Indeed, this is the philosophy of the 2001 Directive, which proposes voluntary objectives for developing renewable electricity – objectives defined on the basis of experts' studies on the renewable resources potential of each EU member state (see European Commission 1996, for instance). Adopting renewable electricity targets for each member state aligned with its specific renewable resource potential seems to be more promising than an EU-wide TGC system. Under such an arrangement, member states would be free to choose the instrument (feed-in tariffs, national TGC system, or competitive tendering) to achieve their national targets.

5 In normal years, wholesale prices in the 'electric peninsulas' without capacity surplus (Spain and Italy) have amounted to €60/MWh and more, which is considerably above prices in 'continental markets', the Nordic market, and the UK market (around €30-€50/MWh) – where surplus of capacity has remained. Without increasing trans-border transmission capacity, these differences will continue to persist.

6 Although evident, it merits repeating that renewable electricity production existing prior to the possible introduction of a common TGC system must be excluded from the certificates system to avoid the transfer of rents from one country to another without there being any investment in new renewable capacity.

3.3 Environmental effectiveness

The environmental effectiveness of an instrument is measured by its success in stimulating investment in renewable power generating capacity and electricity production from this capacity. A variety of factors influence investment and production, notably the level of support and its reliability and predictability. As far as reliability and predictability are concerned, an important aspect is the vulnerability of an instrument to external shocks and changes in the political balance of power after elections. Obviously, the higher the level of support and the more reliable and predictable investment revenues, the greater the impact on investment and production will be. But there is also a trade-off in the sense that with relatively low revenue reliability and predictability, expected revenues need to be high to induce investment. Bearing this in mind, let us consider now both instruments.

The environmental effectiveness of an instrument is measured by its success in stimulating investment in renewable power generating capacity and electricity production from this capacity.

Feed-in tariffs have a good chance to be environmentally effective if they grant sufficient financial support for a long enough time span while minimising transaction costs of producers in their relation with the obligated purchasers of renewable electricity. Indeed, such transaction costs are normally minimal as there is no need to establish any contract between producers and purchasers, except for agreements on technical conventions governing the secondary duties of producers (maximum annual production covered by the tariff, conditions of connection, technical tuning, and so on). Provided feed-in tariffs are sufficiently high and guaranteed for the whole economic life of the investment, potential producers should have no difficulty in sourcing finance for their projects. In sum, although feed-in tariffs do not directly target the quantity of renewable electricity production but its price, they can be expected to perform well in terms of environmental effectiveness without being more costly than TGC systems.

In principle, one would expect TGC systems to be more effective than feed-in tariffs in increasing investment in and production of renewable electricity – after all, this instrument directly targets the quantity of renewable electricity. However, real-world TGC systems are in fact hybrid instruments, controlling quantity and price. This is because they allow suppliers to pay a penalty rather than fulfilling their quota. Obviously, too low a penalty – for instance one very close to the electricity generating cost of the marginal project needed to respect the general quota – would induce mandated suppliers to pay the penalty for part of their quota rather than to comply with them. In these circumstances, investment in new renewable capacity might not be considered attractive and, as a result, although fixed *a priori*, the quantitative target might be missed.

The UK experience clearly shows how too low a penalty – the buy-out price – makes suppliers decide to disrespect their quotas. In 2003 and 2004, the buy-out price was €43/MWh. The market price for electricity ranged between €30/MWh and €45/MWh. The lower end of this range together with the buy-out price suggest a reference price of renewable electricity of €73/MWh – compared to a long-run marginal cost of wind energy of €90/MWh. In these circumstances, many suppliers decided to pay the penalty for part of their obligation rather than observe their quota. More specifically, in 2002-04, between 41 percent and 45 percent of the overall quota was 'met' by penalty payments. In 2004, the buy-out price was raised to €47/MWh, reducing the gap between the targeted quota and what was actually achieved in 2005 to 31 percent.

All in all, meeting an increasing target for renewable electricity with a TGC system presents intrinsic difficulties. The challenge is to adequately raise the quota, which influences the certificate price, and the penalty that is meant to cap the cost for consumers. In other words, the aim of limiting cost for consumers makes it difficult to achieve ambitious renewable energy targets. And experience so far suggests that feed-in tariffs have been more successful than TGC systems in spurring the production of renewable electricity.

3.4 Conformity with the underlying market regime

Over recent years, EU member states have been liberalising their electricity sectors – albeit at different speeds and to different degrees. This raises the question of how alternative instruments for promoting renewables conform to a market regime that might be characterised by a vertical separation of network business, the abolishment of regional monopolies, and – in general – by more competition.

As a matter of principle, a tradable green certificates system seems to readily conform with liberalised electricity markets.

As a matter of principle, TGC systems seems to readily conform with liberalised electricity markets – for a variety of reasons. First, TGC systems do not distort competition between suppliers. This is because the ability of a supplier to meet his quota obligation does not depend on whether or not additional renewable electricity can be produced at reasonable cost in his supply area. The obligation can be met by purchasing green certificates, which can be bought at a market price from any plant generating renewable electricity. Second, as long as the renewable electricity quota is the same for all competing suppliers, they all carry the same cost for supplying renewable electricity quota and there is, thus, no need for a specific financing mechanism. Third, as explained in Box 1, the value of the 'greenness' of renewable electricity is linked to the market price of electricity, given that producers of green electricity receive the electricity price and the certificate price.

By contrast, with feed-in tariffs, the production of renewable electricity is not governed by market forces given the obligation to purchase at a fixed price. But there is a way to strengthen the role of market forces – a solution applied in Spain: to define a price premium that is added to the wholesale electricity price. Each year, this premium is calculated on the basis of the cost of renewable electricity and the average wholesale electricity price during the preceding year. This solution links the revenue (per kWh) of renewable electricity producers to the market price of electricity.

Feed-in tariffs could distort competition when there is clear-cut unbundling between distribution (a physical network activity) and supply (a commercial activity) – a model the European Commission tries to promote – by imposing unequal obligations on competing suppliers. To avoid distorting competition, it is necessary to entrust an agency with the responsibility for buying renewable electricity at guaranteed tariffs because suppliers are no longer regional monopolists. This agency then needs to auction the renewable electricity it has bought, or reallocate it to suppliers in proportion to their market shares.

However, making feed-in tariffs compatible with liberalised electricity markets does not really pose problems in countries where the activities of local, regional, or national incumbent operators have been 'unbundled' only moderately (e.g., when only the accounting of distribution and supply activities has been unbundled). And then, the spatially prescribed obligation to purchase renewable electricity raises few problems when incumbents remain dominant suppliers with *de facto* captive customers – as it is the case in France, Germany, Spain, Ireland, and Portugal where feed-in tariffs are in place. However, to ensure conformity with market principles it is necessary that the incumbents subject to the purchase obligation are compensated in a transparent and fair way from a fund financed by a special tax on every transported kWh. In sum, there is considerable scope for making feed-in tariffs compatible with a liberalised electricity market regime.

3.5 Which instrument is preferable?

Overall, economic reasoning does not suggest a clear-cut advantage of feed-in tariffs over TGC systems – and *vice versa*. Each system has its strengths and weaknesses.

Economic reasoning does not suggest a clear-cut advantage of feed-in tariffs over a tradable green certificates system – and* vice versa*. Each system has its strengths and weaknesses.

Feed-in tariffs promise greater environmental effectiveness and they are relatively easy to design so that they foster the development of a diverse set of renewable technologies. They can also be fairly cost-effective – at the national level and in the European policy context – provided they account reasonably well for countries' underlying renewable resource potential, cost differences across technologies, cost differences when a given technology is used in different locations, and anticipated cost decreases because of technological progress. Tariffs that take these factors into account help limit producers' rents and, thus, costs to consumers. In line with expected technological progress, they can also be phased out as and when technologies have matured or targeted capacities have been installed. This constitutes now the reference design for feed-in tariffs. But, admittedly, feed-in tariffs could result in too much renewable capacity in certain technologies and they could be too costly if there is no timely adaptation and eventual phasing out of tariffs.

TGC systems make good use of market forces with a view to minimising the cost of meeting renewable electricity targets. Almost by design, there is no risk that this instrument inadvertently leads to excessive investment in renewable capacity. If the targeted amount of renewable electricity is set at too high a level – relative to underlying marginal cost curves, that is – buy-out prices offer a safety valve that caps the cost for consumers. All in all, the use of market forces and the setting of quantitative targets in combination with buy-out prices all promote a cost-effective supply of renewable electricity. This being said, potential investors in renewable electricity might perceive the stream of revenue resulting from their investment as too uncertain. This tends to stifle investment in renewable capacity, increase financing cost, or both. In addition, large transaction costs might offset the downward pressure on cost resulting from competition. What is more, competition itself might not develop as hoped for if producers and suppliers strive for vertical integration or long-term contracts with a view to limiting risks and reducing transactions costs. A further drawback of TGC systems is that they effectively concentrate their support on the least costly technological solution and, thus, they do not stimulate technological diversification. To mitigate this shortcoming, there could be complementary support mechanisms, such as tax credits and investment grants for promising, though less advanced, technologies. However, this would lessen the beneficial impact of competition – one of the presumed strengths of TGC systems.

Having compared the economics of feed-in tariffs and TGC systems and concluded that neither outperforms the other on theoretical grounds, it needs to be pointed that the choice of instrument is largely political in any case. It reflects the preferences of governments and citizens, notably preferences pertaining to environmentalism and the respective role of the state and free markets in an economy. With faith in markets and a strong preference for controlling costs, governments are likely to choose market-based instruments, such as TGC systems. Conversely, with less of a free-market culture and a strong preference for attaining quantitative environmental goals, the choice is more likely to be in favour of feed-in tariffs.

Besides economic reasoning and political considerations, guidance in choosing instruments for promoting renewable electricity should also come from good – or bad – practice. A point we address next.

4. Lesson from the application of alternative policy instruments

The experience of EU member states with promoting renewable electricity production is now sufficiently documented to draw some lessons on how various instruments have worked in practice. Insights follow from experience with designing and applying various policy instruments and from what they have achieved in meeting policy objectives. As before, we will concentrate on feed-in tariffs and TGC systems.

One lesson is that the influence of a particular instrument cannot be isolated from other factors that foster, or hinder, the development of a country's renewable electricity resources.[7]

Key obstacles to developing renewable electricity generation include fragmented planning procedures and local acceptability problems.

Specifically, how successful an instrument is depends as much on the level of support it provides as on the planning procedures and rules that govern the recovery of balancing costs and the cost of connecting renewable power plants to the network. To illustrate, although France adopted in 2000 feed-in tariffs as generous and predictable as those in Germany, investment in renewable generating capacity and its performance fall well short of what has been achieved in Germany (e.g., in 2005, installed wind energy capacity amounted to 530 MW in France and 15,000 MW in Germany). Key obstacles to developing renewable electricity generation in France include fragmented planning procedures and local acceptability problems. There is thus no doubt that effective planning procedures and network integration rules can help reduce project costs and risks and they must therefore be an integral part of a successful renewable energy policy. But like the support for renewable energy in general, they reflect the political backing of the underlying renewable technology. In addition, they reflect social preferences for global environmental protection and energy security, on the one hand, and local environmental concerns on the other.

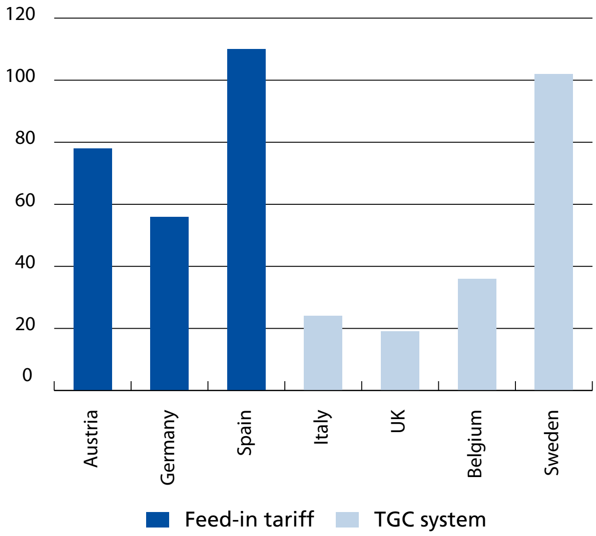
Another lesson is that differences in the stability and predictability of the support for renewable electricity largely explain why some European countries were more successful in increasing the share of renewable electricity than others. Let us take the case of onshore wind, so far the most successful renewable electricity technology.

Figure 2 shows, for a selected number of EU countries, the annual average per capita output (in kWh) produced by new wind energy installations in 2000-04. The results come from Ragwitz *et al.* (2006). In analysing the efficiency of various support mechanisms for renewable electricity, they have estimated expected revenues to new producers of onshore wind energy and the induced renewable electricity production in Austria, Germany, and Spain – all offering feed-in tariffs – and in Belgium, Italy, Sweden, and the United Kingdom – countries with TGC systems. Using the level of output to measure the environmental effectiveness of the underlying policy, it is fair to say that countries with feed-in tariffs (Austria, Germany, and Spain) performed better than countries with TGC systems – the exception being Sweden. But in the Swedish case, results follow from the specificity of this country's TGC system, which was only adopted in 2003 (replacing a system that offered large tax credits and investment subsidies) and includes existing installations in the portfolio of eligible technologies. Austria, Germany, and Spain applied feed-in tariffs in 1998-2005. Combined with low administrative barriers, this stimulated a strong and continuous growth in wind energy. By contrast, in the United Kingdom, Italy, Belgium and Sweden, the change from a

7 An abundant literature discusses the causal links between the diffusion of renewable electricity and variations in the design and strength of policy instruments. Examples include Reiche and Becherger (2004), Reiche (2005), Meyer (2003), and van Djik *et al*. (2003).

tendering system or feed-in tariffs to a TGC system created much uncertainty during the transition period. Moreover, one can consider the findings summarised in Figure 2 as evidence for the hypothesis that TGC systems do not create an environment secure enough to invest in renewable electricity generation.

Figure 2. Annual average wind energy output per capita (in kWh) of onshore units installed in 2000-2004



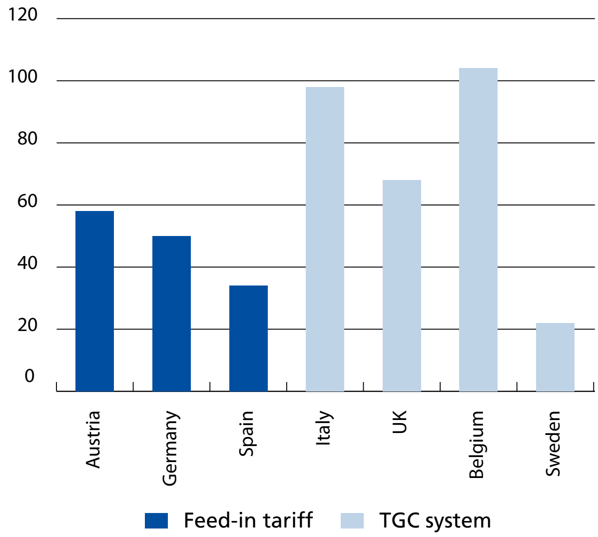
Source: Ragwitz *et al.* (2006)

Feed-in tariffs are often presumed to offer more generous support for renewable electricity than TGC systems, and this – rather than the predictability of the support – could explain their greater environmental effectiveness. But as we have argued in Section 3, from the perspective of potential producers and investors, TGC systems are surrounded by considerable uncertainties, and one consequence of this could be that the revenue required to induce investment in renewables is higher under TGC systems than under feed-in tariffs. Empirical support for this hypothesis comes from Butler and Neuhoff (2004). They show that the remuneration of wind energy is higher under the UK TGC system than under the German feed-in tariffs, which are often portrayed as excessively generous. More specifically, they show that the remuneration of wind energy ranges from around €77/MWh to €100/MWh in the British mechanism, which compares to a figure of €70/MWh under Germany's feed-in tariffs. Similar evidence comes from Ragwitz *et al.* (2006), who have estimated expected revenues to new producers of onshore wind energy. The estimates – pictured in Figure 3 – show that expected revenues are much higher in the group of countries using the TGC systems than in those relying on feed-in tariffs.[8] All in all, a fair conclusion is that feed-in tariffs, in practice, do not offer exceptionally high revenues to producers and that reliability and predictability of the policy and investment environment is key for successfully developing the market for renewable electricity.

From the perspective of potential producers and investors, a tradable green certificates system is surrounded by considerable uncertainties.

8 This being said, the difference could probably decrease in the future as and when institutional experience with the relatively new instrument of TGC systems accumulates. But even if this were to happen, it would not reduce the risk premium associated with the production of renewable electricity under TGC systems.

Figure 3. Expected revenues (€/MWh) of onshore wind energy: feed-in tariffs *vs.* tradable green certificate (TGC) system



Source: Ragwitz *et al*. (2006)
Notes: Estimates reflect levelised expected revenues.[9]

The third and final lesson is that governments often offer complementary investment support (soft loans, tax allowances, and so on) – in addition to feed-in tariffs or TGC systems, for example. The need for such additional support scheme seems to be higher in the case of TGC systems, in particular when the objective is to foster not only the technologically and commercially most advanced renewable option but also those lagging behind.

In designing policies to promote renewables, policy makers need to be aware of the complexity of the innovation process driving renewable electricity technologies.

In conclusion, given the experience gained with competing instruments to promote renewable electricity, it may not come as a surprise that erstwhile strong supporters of TGC systems have become more cautious, as evidenced by the evolving position of the European Commission in the debate (European Commission 2005). Policy makers need to be aware, and increasingly are, of the complexity of the innovation process driving renewable electricity technologies. Once they have decided on the instrument, they must be aware of the necessity to clearly signal that the support mechanism will remain in place long enough to ensure an acceptable return to the producers of renewable electricity.

5. Conclusion

Economic reasoning does not provide an unambiguous answer to the question which of the two instruments – feed-in tariffs or TGC systems – is best for promoting renewable electricity. One reason is that there is a range of criteria for assessing the pros and cons of alternative policies, and while one instrument might be strong when measured against one criterion, it might be weak when measured against others. There are then possible trade-offs to consider – such as a trade-off

9 The expected levelised revenues were calculated for 2004. Calculations are based on the effective support conditions in each country. Tax exemptions and soft loans are also taken into account in estimating revenues. For countries with TGC systems, certificate prices of 2004 have been extrapolated for the entire active period of the support system for a new equipment. The low estimate for Sweden results from the specificity of the Swedish TGC system, which includes existing installations in the portfolio of eligible technologies.

between good performance of an instrument in terms of cost-effectiveness and possibly less-than-satisfactory performance with respect to environmental effectiveness.

Needless to say that such trade-offs become more relevant when moving from the principles of a particular instrument to its practical application. But it is also true that each of the two instruments examined in this paper could be designed so that its weaknesses are mitigated without compromising its strength too much. This being said, experience with the two instruments in various countries seems to suggest that, in practice, feed-in tariffs are easier to adapt to real-world situations than TGC systems – a finding that holds when considering a common European approach to promoting renewable electricity. What is more, feed-in tariffs seem to offer greater success than TGC systems in providing a predictable revenue planning horizon, boosting investment in renewables, and fostering technological diversification. If society values this more than minimising cost to electricity consumers, feed-in tariffs are a good choice. Conversely, if society considers cost minimisation under market pressures the norm in public policies, TGC systems might be the route to follow. That said, in practice, TGC systems imply a less reliable investment environment, which raises risk premiums and thus capital costs, thereby making cost for consumers higher than what simplified theoretical reasoning suggests.

Feed-in tariffs and tradable green certificate systems could be designed so that their weaknesses are mitigated without compromising their strengths too much.

References

Baumol, W.J. and Oates, W.E. (1988). *The theory of environmental policy*, 2nd edition, Cambridge University Press, Cambridge, USA.

Butler, L. and Neuhoff, K. (2004). "Comparison of feed-in tariff, quota and auction as mechanisms to support wind power development". University of Cambridge Department of Applied Economics Working Paper.

Cropper, M.L. and Oates, W.E. (1992). "Environmental economics: A survey". *Journal of Economic Literature*, (30), pp. 675-740.

Egenhofer, C., Jansen, J.C., Gialoglou, K. (2005). "Market stimulation of renewable electricity in the EU. What degree of harmonisation of support mechanisms is required?" Centre for European Policy studies, CEPS Task Force Report n°. 56, October 2005

European Commision. (2005). "The support of electricity from renewable energy sources". Com(2005) 627.

Finon, D. and Menanteau, P. (2003). "The static and dynamic efficiency of instruments of promotion of renewables". *Energy Studies Review*, (12:1), p. 53-83.

Finon, D. and Perez, Y. (2006). "The social efficiency of instruments for the promotion of renewable in the electricity industry: A transaction cost perspective". *Ecological Economics*, (62), pp.77-92.

Foxon, T.J., Gross, R., Chase, A., Holwes, J., Arnall, A., and Anderson, D. (2004). "UK innovation systems for new and renewable energy technologies drivers, barriers and system failures". *Energy Policy*, (33), pp. 2123-2137.

Huber, C., Haas, R., and Faber, T. (2001). "Action plan for a green European electricity market". Compiled within the project "ElGreen". International Energy Workshop (IEW).

Jaffe, A.B., Newell, R., and Stavins, R.N. (2002). "Environmental policy and technological change". *Environmental and Resource Economics*, (22:1-2), pp. 41-69.

Langniss, O. and Wiser, R. (2003). "The renewable portfolio standard in Texas: An early assessment". Energy Policy, (31:6), pp. 527-535.

Meyer, N. (2003). "European schemes for promoting renewables in liberalised markets". *Energy Policy*, (31), pp. 665-676.

Mitchell, C., Bauknecht, E., and Connor, P.M. (2004). "Effectiveness through risk reduction: A comparison of the renewable obligation in England and Wales and the feed-in system in Germany". *Energy Policy*, (32), pp. 1935-1947.

Neuhoff, K. (2005). "Large scale deployment of renewables for electricity generation". *Oxford Review of Economic Policy*, (21:1), pp. 88-110.

Newbery, D. (2003). "Government intervention in energy markets". BIEE conference, 25 September 2003.

Fischer, C. and Newell, R. (2004). "Environmental and technology policies for climate mitigation". Resources for the future Discussion Paper No. 04-05. http://www.rff.org/rff/Documents/RFF-DP-04-05.pdf

Ragwitz, M., Held, A., Sensfuss, F., Huber, C., Resch, G., Faber, T., Haas R., Coenraads, R., Morotz, A., Jensen, S.G., Morthorst, P.E., Konstantinaviciute, I., and Heyder, B. (2006). "Assessment and optimisation of renewable support schemes in the European electricity market". OPTRES interim report.

Reiche, D. and Bechberger, M. (2004). "Policy differences in the promotion of renewable energies in the EU member states". *Energy Policy,* (32), pp. 843-849.

Reiche, D. (2005). *Handbook of renewable energies in the European Union - Case studies of the EU-15 states*, Peter Lang Publishers, Berlin, Germany.

Roberts, M.J. and Spence, M. (1976). "Effluent charges and licenses under uncertainty". *Journal of Public Economies*, (5:3-4), pp. 193-208.

Sawin, J. (2004). "National policy instruments for the promotion of renewables". Worldwatch Institute Report to International Conference for Renewable Energies.

van Dijk, A.L., Beurskens, L.W.M., Boots, M.G., Kaal, M.B.T., de Lange T.J., van Sambeek, E.J.W., and Uyterlinde, M.A. (2003). "Renewable energy policies and market developments". Energy Research Centre of the Netherlands, Policy Studies Report ECN-C--99-072.

Weitzman, M.L. (1974). "Prices vs. quantities". *The Review of Economic Studies,* (41:4), pp. 477-491.

ABSTRACT

Recognising that environmental and technology externalities affect the development of renewable energy technologies, *this paper illustrates how environmental policies induce technological change and how market failures that hinder technological progress weaken the impact of environmental policies on technological change; examines the rationale for and type of policies in support of renewables at an early stage of their commercialisation; analyses to what extent so-called experience curves enlighten the debate on the rationale of such policies; and – assuming that early-stage renewables cannot establish themselves in the market – develops a method for assessing the economics of renewable energy projects based on new technologies.*

Atanas Kolev (a.kolev@eib.org) and **Armin Riess** (a.riess@eib.org) are, respectively, Economist and Deputy Head in the Economic and Financial Studies Division of the EIB. The views expressed are strictly personal.

Environmental and technology externalities: policy and investment implications

Nothing is so simple that it cannot be misunderstood.
Freeman Teague, Jr.

1. Introduction

The production and use of energy is characterised by a variety of negative environmental externalities, that is, environmental cost to society normally ignored by energy producers and users in their decision making. Notorious is the emission of airborne pollutants associated with heavy industry, transport, and electricity generation. Without policies aimed at making external environmental cost influence private decision making – through a tax on emissions, for instance – the use of energy is likely to exceed its social optimum and the energy mix is biased in favour of fossil fuels and against renewable sources of energy – renewables, for short.

Atanas Kolev

Armin Riess

A variety of renewables is available to partly replace fossil fuels. Some of them – like onshore wind energy in good locations – use fairly mature technologies and allow supplying energy at attractive cost compared to fossil fuels, provided the cost of the latter include their external environmental cost. Other renewables – solar thermal electric power, for instance – use less developed, new technologies and remain economically too costly even when accounting for the negative environmental externalities of fossil fuels.

However, renewables not yet economically competitive might become so in the future – for a number of reasons. For a start, the external environmental cost of fossils might rise over time, changing relative cost in favour of renewables. In addition, one could envisage an increase in the cost of mature renewables as and when low-cost options – such as good locations for onshore wind farms – become scarce. This would lower the cost of new renewables relative to that of mature renewables. And then, one might expect an absolute decline in the cost of new renewables. By definition, they are at an early stage in the lifecycle of developing technologies, and future technological progress might reduce their cost.

This takes us to the second type of externality in the title of this paper. If technological progress were to proceed at an optimal pace (and if environmental externalities were fully internalised), society could simply wait for the new technologies to mature and then use them. However, technological progress is fraught with market failures and externalities, too, but in contrast to those affecting the environment, they are 'positive' so that free markets left to themselves might generate too little technological progress. A particular aspect concerns learning and the accumulation of experience of firms embarking on new technologies. When firms start using a new technology, they increasingly learn how to use it better and, as a result, with an increase in output they experience a decline in production costs. The trouble is, however, that various market failures and externalities might prevent learning and experience to go as far as it should from society's viewpoint. If true, there is an economic case for public support in favour of new technologies with a view to increasing their use and, thereby, allowing firms to benefit from learning and experience effects.

Interactions between environmental and technology policies complicate the design of such policies.

From an economic policy perspective, environmental and technology externalities raise a variety of questions. There are the 'usual suspects' of whether environmental challenges are best addressed by market-based policies or command-and-control policies and whether promoting technological progress is best achieved by non-selective measures fostering the creation and diffusion of new knowledge in general or by targeted R&D support for specific sectors, firms, or technologies. Interactions between environmental and technology policies make this question more difficult to answer. To illustrate, an emissions tax implicitly rewards clean technologies, thereby fostering not only renewable energy production, but also research directed at improving these technologies. Does this imply that directly supporting technological progress becomes less pressing because of the technological push induced by environmental policy? Or are the costs of environmental policies lower than they appear because environmental policies kill two birds with one stone – apologies to animal rights defenders – by tackling not only environmental problems but also technology externalities? Along similar lines, is there an argument for making environmental targets more ambitious than environmental externalities alone advise because of the favourable impact of environmental policies on technological progress? And – to end a non-exhaustive list of questions – what is the rationale for promoting new renewable energy technologies given that we have mature ones?

Similar questions arise from an investment perspective – more specifically: an economic cost-benefit perspective of renewable energy projects based on new technologies. In addition, there is the issue of how to account for both environmental and technology externalities in the appraisal of energy investments. And then, does it matter whether or not real-world policies fully internalise external environmental cost? More heretically: should the environmental cost of fossil-fuel-based energy affect decisions on investments in new-technology renewables?

Trying to address all these questions in one paper would certainly be far too ambitious. Rather, we will concentrate on some of them and promise not to shy away from the heretical one. To this end, the remainder of this paper unfolds as follows. The next section examines interactions between environmental policy and technological change, notably the links between policies aimed at internalising the external environmental cost of producing energy, technological progress, and policies aimed at promoting technological progress. Section 3 zooms in on the rationale for promoting new-technology renewables, that is, technologies that are known and do not need to be invented but that are at an early stage of their commercialisation. As this section will show, the rationale for promoting them largely rests on market failures and externalities possibly associated with learning and experience effects. Against this background, Section 4 reviews the empirical literature on learning and experience effects and discusses how well – rather, how poorly – it informs on the extent to which learning and the accumulation of experience suffer from market failures and externalities. In Section 5 we change tack: leaving behind the policy-oriented presentation of the previous sections, we will develop on the basis of a welfare-maximising model a cost-benefit rule for assessing energy projects based on new-technology renewables. Section 6 concludes.

2. Environmental policy and technological change

The main purpose of this section is fourfold: first, to examine how environmental policies induce technological change; second, to discuss how policies directed at fostering technological progress contribute to achieving environmental targets (notably when environmental policies are sub-optimal); third, to explain how market failures that hinder technological progress weaken the impact of environmental policies on technological change; and fourth, to outline how market failures that hinder technological progress might affect the choice of environmental policy targets and instruments. The survey article by Jaffe *et al.* (2003) and the paper by Jaffe *et al*. (2005) discuss these and other issues in greater detail. In what follows, we condense and illustrate the insights from this literature that are relevant for our paper.

2.1 Technological change induced by environmental policies

To discuss how environmental policies induce technological change, let us consider a tax on the emission of airborne pollutants – such as SO_2, NO_x, particulates, CO_2 and other greenhouse gases. The purpose of an emission tax is to make polluters account for the environmental damage of their emissions. For now, we assume that the tax is set so that it fully internalises the environmental damage, that is, the economic cost of emissions.

Internalising the economic cost of emissions fosters renewable energy in a direct and indirect way.

An emission tax has two main effects (see Box 1 for details). For one thing, by putting a price tag on emissions, the tax penalises pollution and thus encourages its abatement. We may call this the static effect of internalising the cost of emissions. For another, penalising emissions encourages efforts to improve on existing abatement technologies or to invent new, cheaper ones. Provided such efforts are successful, the cost of abatement falls and abatement increases. We may call this the dynamic effect of internalising the cost of emissions. The technological progress leading to a fall in abatement cost is aptly thought of as induced by environmental policies.

Given the theme of this paper, let us make things a little more concrete by considering the production of electricity – one of the main sources of airborne pollutants together with transport and industry – and the emission of air pollutants by fossil-fuel-fired power plants. Moreover, let us focus on a particular abatement option, that is, the replacement of fossil fuels with renewable electricity, and we ignore negative environmental externalities caused by renewables. The static effect of taxing emissions is an increase in renewable electricity output for a given level of technological development of renewables. The dynamic effect resulting from the induced technological progress implies that output increases further.

In sum, policies to internalise the economic cost of emissions raise the production of renewable electricity directly and indirectly. In line with the notation used in Box 1, let A^*_1 denote this dynamic production optimum. The direct, static effect is due to making the cost of fossil-fuel-fired electricity reflect its negative environmental impact, thereby lowering the cost of renewables relative to the cost of fossils. The indirect, dynamic effect is due to the economic rent that producers of renewable electricity can earn if they succeed in lowering their production cost.

Box 1. Interaction between environmental policy, induced technological progress, and market failures hindering technological progress

This Box offers a graphical presentation of the static and dynamic effects of policies to internalise the economic cost of emitting pollutants. As in the main text, we assume that emissions are caused by fossil-fuel-fired power plants.

At the core of the presentation is the comparison between marginal abatement benefits (*MAB*) and marginal abatement costs (*MAC*). Marginal abatement benefits equal the (avoided) marginal economic costs of emissions. Marginal abatement costs reflect the marginal cost of reducing emissions from fossil-fuel-fired power plants (with the help of appropriate technologies – flue-gas desulphurisation, for example) or the marginal cost of producing electricity on the basis of zero-emission electricity-generating technologies (renewables and nuclear, for instance) and, thus, of partly replacing fossil fuels with zero-emission electricity. In contrast to the main text, the presentation here considers emission abatement in general rather than only the abatement associated with an increase in renewable electricity. For simplicity, we assume that abatement technologies do not give rise to negative environmental externalities.[1]

The sequencing of the presentation parallels the structure of Section 2.

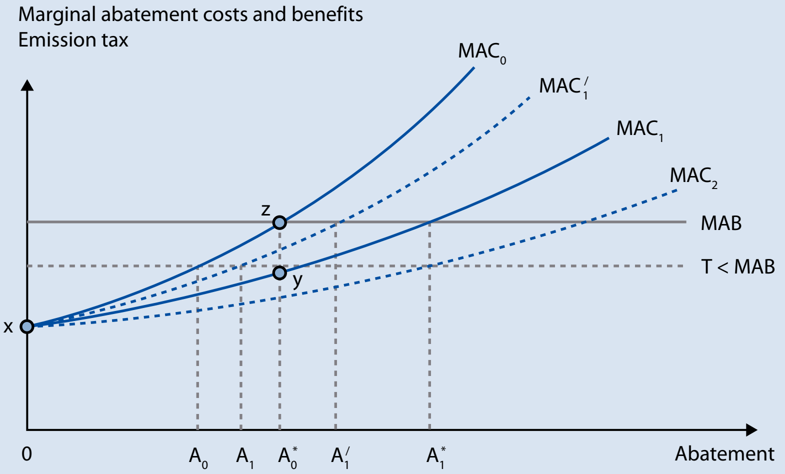
Technological change induced by environmental policies

In the figure below, the horizontal line *MAB* shows marginal abatement benefits – assumed to be constant for simplicity although they probably fall with the level of abatement. MAC_0 shows marginal abatement costs, which typically increase with the level of abatement. Starting from no abatement at all, it is economically efficient to increase abatement – and thus reduce emissions – as long as $MAB > MAC$. The optimal level of abatement is A^*_0.

Without policies to internalise the economic cost of emissions, there would be no abatement. To reach A^*_0, one could impose a tax on emission equal to *MAB*. Alternatively, under an emission cap-and-trade system, emissions could be capped so that A^*_0 is achieved. In a perfect world – notably with perfect information on marginal abatement costs and benefits – the cap-and-trade system would yield a price of emission rights equal to the optimal emission tax. The move from zero abatement to A^*_0 along the curve MAC_0 captures the static effect of taxing emissions.

The dynamic effect – that is, the fall in marginal abatement cost induced by taxing emissions – is illustrated by the rightward rotation of the marginal-abatement-cost curve from MAC_0 to MAC_1. Two points are worth stressing. First, at the 'static' optimal abatement level A^*_0, the economic rent accruing to suppliers of abatement technologies (including producers of renewable electricity) increases by *XYZ*. In fact, it is the prospect of higher economic rents that stimulates efforts to reduce abatement cost. Second, given the drop in abatement cost, $MAB > MAC$ at the 'static' optimum A^*_0, making it worthwhile to further curb emissions and increase abatement to A^*_1. For completeness, note that with a downward-sloping *MAB*-curve, the further increase in abatement would be smaller, but equilibrium marginal abatement cost would be lower too.

1 Clearly, the technologies labelled 'zero emission' also emit air pollutants (for instance, in the manufacturing of wind farms and photovoltaic electricity-generating equipment) and cause other environmental externalities – such as noise pollution and visual intrusion in the case of renewables or the risk of radioactive contamination in the case of nuclear energy. Still, there is broad agreement that the external environmental cost of fossil fuels by far exceeds that of other sources of primary energy (ExternE 2004).



Promoting technological change to achieve environmental targets

Consider now a situation where policies fail to fully internalise the economic cost of emissions. This is the case, for instance, if the emission tax is lower than the marginal abatement benefit – as pictured by the line $T < MAB$. In these circumstances, abatement reaches A_0, but it remains below the static optimum A^*_0. The static effect of taxing emissions is thus smaller. As a result, the dynamic effect is likely to be smaller, too, as illustrated by the less pronounced rightward rotation of the marginal-abatement-cost curve from MAC_0 to MAC'_1. Although abatement increases further from A_0 to A_1, it remains below the dynamic optimum A^*_1. In principle, A^*_1 can be attained by policies that directly promote technological progress, as illustrated by the considerable further twist in the *MAC*-curve from MAC'_1 to MAC_2. But as discussed in the main text, the economic cost of achieving A^*_1 in this way is liable to be higher than the cost of a policy that fully internalises emission costs.

Market failures hindering technological progress and the environment

Assume again an emission tax equal to *MAB* that leads to the static optimum A^*_0. In the process of achieving this level of abatement, producers and users of abatement technologies learn and, thus, marginal abatement costs fall. To fix a benchmark, take MAC_1 as picturing the situation without market failures that stifle learning and experience effects. Thus, with such market failures, the rightward rotation of the marginal-abatement-cost curve is less pronounced, reaching only MAC'_1, for instance, implying abatement of A'_1. As a result, even with an emission tax high enough to fully internalise the economic cost of emissions, abatement remains below A^*_1. To achieve A^*_1 nonetheless, policies are needed that would ensure that the marginal-abatement-cost curve moves to MAC_1.

A tax on emission that tries to achieve more than just internalising emission cost

Sub-section 2.4 alludes to the idea of an emission tax higher than marginal abatement benefits when market failures stifle technological progress. It is easy enough to picture the apparent logic of this idea. To recall, with full internalisation of emission cost in a situation where market failures stifle technological progress, the abatement level (A'_1) will be determined by the intersection of MAC'_1 and *MAB*. It seems that the dynamic optimum A^*_1 could be induced by an emission tax $T > MAB$ so that the tax line (not shown in the figure) and MAC'_1 intersect above A^*_1. The shortcomings of this idea are discussed in the main text.

2.2 Promoting technological change to achieve environmental targets

To illustrate how policies directly fostering technological change contribute to the achievement of environmental targets, we now assume that the emission tax is not high enough to fully internalise the economic cost of emissions. In these circumstances, renewable electricity output increases too, but not as much as with full internalisation of emission costs. It is fair to presume that this nonetheless stimulates efforts to reduce the cost of renewable electricity, but that they are not as big as in the case of fully internalising the economic cost of emissions. All in all, compared to the case of optimal environmental policies, both the static and the dynamic effect of taxing emissions are weaker. Reflecting the notation used in Box 1, let A_1 denote this level of renewable electricity output, which is lower than the dynamic optimum A_1^*.

Given this sub-optimal outcome and assuming that policy makers shy away from raising emission taxes, one could ask whether directly supporting technological progress could not lead to the optimal level of renewable electricity output. In principle, this is possible if such support sufficiently reduces the cost of renewable electricity. It is important to stress that this cost reduction is not induced by policies aimed at correcting environmental market failures. Rather, it results from policy measures such as public R&D in favour of renewables, or it might be triggered by preferential prices offered to producers of renewable energy.

Second-best policy packages that combine the partial internalisation of environmental cost with direct technology support are economically less efficient than the first-best policy of fully internalising environmental cost.

It thus appears that the optimal outcome can be reached either by a first-best policy that fully internalises the economic cost of emissions or by a second-best policy package that combines partial internalisation with direct technology support. Although this is true as far as the optimal amount of renewable electricity is concerned, it would be an erroneous conclusion from a welfare-maximising viewpoint since the technology support component of the second-best policy package is not for nothing. For a start, there are opportunity costs of promoting technological advances in renewable energy. Take public R&D in support of renewables, for instance. Research and development resources committed to this undertaking cannot be used to accelerate technological advances in other fields – biotechnology for example. Moreover, in contrast to emission taxes that correct a distortion in the economy, mobilising the public finance needed to directly foster technological change is distortionary. In addition, the cost of administering technology support is probably higher than the cost of administering emission taxes. What is more, technology support inevitably comes with the challenge of picking winners – or the risk of choosing losers. In sum, the second-best policy package that combines the partial internalisation of environmental costs with direct technology support cannot outperform the first-best policy of fully internalising the environmental costs.

2.3 Market failures hindering technological progress and the environment

Technological progress – whether or not induced by environmental policy – can be thought of as comprising two broad components. One reflects the creation and diffusion of new technology, that is, product and process innovation. Typically, this type of technological progress follows from research and development, and it occurs at the pre-commercialisation stage in the lifecycle of technology developments (Foxon *et al.* 2005).

There are various reasons why markets might fail in stimulating the creation and diffusion of new technology as much as is desirable from society's viewpoint. We have critically reviewed them elsewhere (Riess and Välilä 2006). Suffice it to note here that firms are liable to underinvest in the creation of new technologies if they cannot fully appropriate the fruits of their innovations – and whatever innovation there is might not disseminate through the economy as much as it could because innovators deny the use of their innovations to others, notably competitors, or overcharge them for using their innovations. Riess and Välilä (2006) conclude that this type of market failure is not as grave as often feared, markets are quite innovative in trying to overcome their own failures, and that policies most appropriate for addressing remaining failures are support for basic research and development, protection of intellectual property rights that strikes the right balance between promoting innovation and not hindering too much its diffusion, and measures to strengthen markets for technologies.

The second component of technological progress does not concern the creation and diffusion of new technologies. Rather, it concerns improvements to new technologies resulting from so-called learning and experience effects. In contrast to the technological progress due to the creation and diffusion of new technologies, technological progress due to learning and experience effects happens at the commercialisation stage in the lifecycle of technology developments. The nature of learning and experience effects and the rationale for economic policies possibly following from them will be the focus of Section 3. For now, we simply note that as and when firms start using a new technology – be it the manufacturing of new products or the use of new production processes – they increasingly learn how to use this technology better and, as a result, experience a decline in production costs. The trouble is, however, that various market failures might prevent learning and experience to go as far as it should from society's viewpoint. What does this imply for induced technological change?

Market failures that hinder technological progress weaken the impact of environmental policies on technological change.

To find the answer, assume as in sub-section 2.1 an emission tax high enough to fully internalise emission cost. As argued above, this triggers the optimal static supply response by producers of renewable electricity. By increasing supply, producers of renewable electricity and manufacturers of equipment for the production of renewable electricity learn and, thus, production costs fall. As a result, renewable electricity production increases further. However, this indirect, dynamic effect of making producers of fossil-fuel-fired electricity account for emission cost is sub-optimal if market failures hinder the learning and experience process. The combined static and dynamic effect is thus sub-optimal although the emission tax is high enough to fully internalise the economic cost of emissions. In other words, market failures that hinder technological progress weaken the impact of environmental policies on technological change.

Against this background, arguments are made in favour of so-called strategic deployment policies, that is, measures helping a known technology at its early stage of commercialisation to achieve greater market penetration. The underlying rationale for such policies will be addressed in Section 3. But before, let us briefly turn to some other intriguing policy issues arising from the interaction between environmental policies, on the one hand, and market failures and externalities affecting technological change on the other hand.

2.4 Environmental policy instruments and targets when market failures stifle technological change

Further to the interaction between environmental policies, on the one hand, and market failures and externalities affecting technological change, on the other hand, discussed so far, four interactions are worth stressing.

First, when considering negative environmental externalities in isolation and a situation of certainty, economists broadly agree that market-based policy instruments (emission taxes and tradable emission permits, for instance) are economically more efficient in addressing environmental externalities than command-and-control measures (quantitative emission targets and imposing the use of specific technologies, for instance), but that once uncertainty is introduced, the superiority of market-based instruments might not hold in all circumstances (Perman *et al.* 2003). Rivers and Jaccard (2006) investigate how learning and experience in the process of technology development might affect the ranking of policy instruments. They find that the advantage of market-based instruments remains but could be small. Given political-economy obstacles to market-based instruments stringent enough to fully internalise environmental externalities, the efficiency loss of choosing second-best command-and-control measures instead could thus be small as well.

Second, when market failures and externalities stifle technological progress, it might be tempting to argue that emission taxes need to be above the level suggested by environmental considerations alone (the apparent logic of this idea is graphically illustrated in Box 1). To put it differently, a higher emission tax might substitute for direct technology support aimed at lowering the cost of renewables.

While this idea seems appealing, there are reasons to consider it flawed. For a start, it runs against the 'Tinbergen rule' (Tinbergen 1955, 1956), suggesting that the number of independent policy tools must be at least as high as the number of policy objectives. In other words, two independent policy instruments are needed to simultaneously internalise the economic cost of emissions and correct market failures affecting technological change. In fact, in their survey, Jaffe *et al.* (2003) emphasise the empirical work of Katsoulacos and Xepapadeas (1996), who found that simultaneously taxing emissions and subsidising environmental research and development promises greater success in correcting environmental and technology externalities than using either instrument alone. In any event, if policy makers find it politically impossible to impose an emission tax that fully prices in emission cost, they will not levy a tax even higher than that.

Efforts to spur innovation in the area of renewables do normally not rest on the use of unemployed resources and, thus, come at the expense of creating new knowledge of a different kind.

Third – and related to the previous point – with environmental policies inducing technological change that is itself fraught with market failures and externalities, one could argue that the net social benefit of environmental policies is larger than the difference between gross environmental benefits and the costs of such policies, or – to put it differently – that these costs are offset not only by environmental benefits but also by benefits related to mitigating technology market failures and externalities. Although this argument has some charm at first glance, two caveats are worth making (see Jaffe *et al.* 2003 and the literature reviewed there). First, while it is true that technological progress induced by environmental policy reduces the cost of renewables for a given level of renewable electricity output, it increases the level of output and, thus, total cost. Second, as pointed out above, the production of technological progress does normally not rest on the use of otherwise unemployed resources. On the contrary, creating new knowledge that eventually lowers the cost of renewables needs highly skilled labour – scientists, for instance – and thus comes at the expense of creating new knowledge of a different kind. Settling this issue would then require a comparison of economic rates of return to competing research and development expenditures.

Fourth, the work of Goulder and Mathai (2000) suggests that while induced technological progress reduces the cost of renewables, it might increase their near-term cost relative to their cost further into the future. As a result, while induced technological progress raises the optimal level of renewable energy, it might be optimal to have less of it today and more tomorrow. Although not linked to technological progress induced by environmental policies, the proper timing of environmental action is one of the key issues in the global warming debate. The *Stern Review* (Stern *et al*. 2006), for instance, strongly argues for near-term measures to tackle global warming whereas others – such as Nordhaus (2006) and Jaccard (2006) – find that societies are probably better served by climate-change policies that "tighten or ramp up over time" (Nordhaus 2006, p.3).

3. The rationale for promoting new renewable energy technologies

3.1 Defining the perspective: obstacles to the commercialisation of new renewables

New technologies are central to economic growth and prosperity and are, therefore, very desirable for any society. Developing, producing, and using them, however, faces a host of difficulties. Some of these are overcome through the workings of the market, while others are not as they originate from the existence of barriers and market failures. Difficulties of the first type are necessary and desirable since they are part of the market quest for the best candidate technology. Problems of the second type do not contribute to this selection process. Rather, they hinder the appearance and diffusion of superior new technologies and products and might thus call for pubic policy intervention.

As sketched in the previous section, there are different types of market failures that could hinder technological progress. In this paper, we focus on market failures associated with economies of learning. An important point to recall is that this type of market failure is an obstacle to the commercialisation rather than the creation of new technologies. But what is the nature of learning and why might it be prone to market failures?

Producers of new technologies crucially rely on learning and experience in the course of production to reduce costs, and learning and experience might spill over to firms that do not pay for the benefits they obtain.

Producers of new technologies crucially rely on learning and experience in the course of production to reduce costs. More specifically, when a firm produces its first units of a product based on a new technology, marginal production cost are relatively high, but are expected to fall with cumulative output due to learning (Wright 1936). The problem is that even if the firm could fully appropriate the benefits of its learning (that is, learning is private to the firm), market failures might still prevent it from taking on this new technology. The problem gets bigger if the firm cannot fully appropriate the learning benefits but if they accrue to other firms too. In other words, the knowledge and experience acquired by producers may spill over to firms that do not pay in any way for the benefits they obtain (Arrow 1962). In practice, market failures hindering private learning and those discouraging firms to embark on new technologies because of learning spillovers might occur together. From a policy perspective, however, it is important to consider them one by one because there is not a single policy that would fit all situations. Borrowing from the infant-industry literature (Corden 1985), we discuss the two cases under the heading internal economies of learning and external economies of learning, respectively.[2]

3.2 Internal economies of learning

We start with the situation where learning is private to firms and thus ignore learning spillovers for now. Normally, a new technology is characterised by production costs that are higher than the costs of a comparable mature technology. In fact, costs may be so high that production is not commercially viable. Existence of a competitor mature technology sets an additional hurdle since the technology needs to be priced so that it can compete with the mature one. Indeed, the price of a new technology needs to be the closer to the price of a mature one, the less potential buyers can differentiate between the two technologies. Different market strategies could be designed, as we discuss below, to allow new technologies to develop and gain market share. Since our interest is in

2 More generally, the analysis in this section is inspired by the infant-industry-protection debate that pre-occupied development economists in the 1960s and 1970s. Indeed, new renewable technologies are infant industries. Interestingly enough, many economists viewed infant industry protection with a fair dose of scepticism (Johnson 1965, Corden 1985, Krueger 1985, Baldwin 1969, and Bliss 1989), but there have also been more optimistic assessments (Dasgupta and Stiglitz 1988).

energy technologies, we illustrate our discussion with a stylised example from the energy industry – more specifically the production of renewable electricity.[3]

The marginal cost of producing a good based on a new production technology tends to decrease as producers become more experienced.

Consider two different renewable electricity-generating technologies: a new one (solar thermal power, for instance) and a mature one (onshore wind energy, for instance). The new technology is characterised by high marginal production cost while the cost of the mature one has reached a long-run floor well below the marginal cost of the new technology. Figure 1 illustrates the situation. It plots marginal production costs against cumulative output produced with a given technology.[4] The downward-sloping marginal cost curve of the new technology reflects the observation that marginal costs tend to decrease as producers accumulate production experience – the so-called learning curve. Since the mature technology is assumed to have reached its long-term cost floor, it is plotted flat.[5] The state of affairs with the new technology being considerably more expensive than the mature one is pictured to the left of Q^* in Figure 1.

In this situation, producers of the new technology should target early technology adopters and sell their product at a price that is initially higher than the price of the mature technology. As producers move along the learning curve by expanding output over time, production costs fall, making the new technology increasingly competitive and gaining market share. However, a necessary condition for such a strategy is that users can clearly distinguish between the two technologies and are willing to pay a premium for the distinguishing features of the new technology. Selling new products that are similar to but sufficiently better than mature ones is possible in a number of industries – consumer electronics, for instance. But does it work in the market for electricity?

Although it could, electricity generation is an industry in which producers using new technologies face great difficulties in finding demand for their product at a price higher than that of mature technologies. Electricity producers' efforts to introduce higher tariffs for those consumers willing to pay for 'green' electricity have largely failed although surveys reveal such willingness among approximately one third of electricity consumers (see EWEA 2004, for instance). This has a simple explanation: services provided by green electricity to consumers are indistinguishable from those provided by electricity based on other energy sources. It is for the cleaner environment that consumers are required to pay a premium. A cleaner environment, however, is a pure public good, implying that even if some consumers are willing to pay a premium, demand for green electricity will fall short of its social optimum.

The case considered here is even more disheartening as electricity users would need to be willing to pay yet more for 'new' green electricity compared to 'mature' green electricity although the latter already yields the clean-environment gain. Some users are perhaps willing to pay for greater diversification of clean technologies – for instance, as an insurance against the possibility that the cost of further extending mature renewables will become prohibitive at some point in the

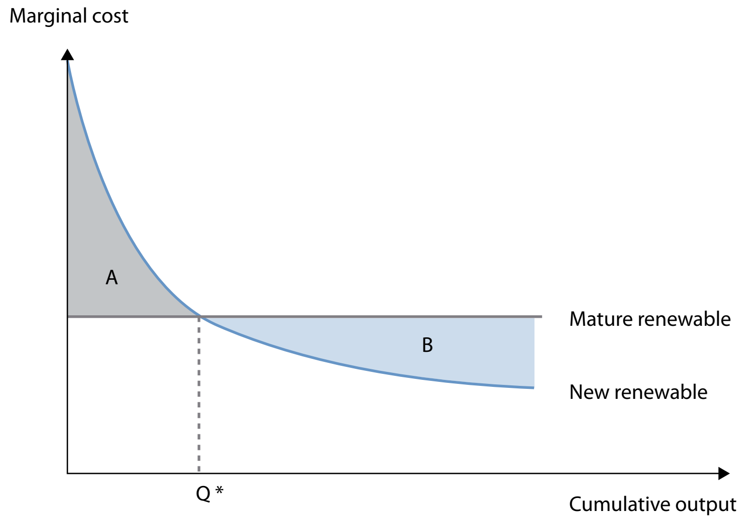
3 Learning economies possibly shape all stages of the production chain. As for renewable electricity, two stages are of particular importance: the manufacturing of equipment for generating renewable electricity (wind turbines suitable for offshore wind farms, for instance) and the production of renewable electricity itself. For the purpose of this paper, there is no need to explicitly distinguish between different stages of the production chain and we thus simply talk about learning economies in the supply of renewable electricity based on new technologies – new renewables, for short.

4 Note that this marginal cost schedule is plotted against cumulative output produced over time, i.e., allowing for variation of all inputs and for all sorts of improvements in the production process. It should be distinguished from the short-run upward-sloping marginal cost schedules of the figure in Box 1. In fact, the link between that figure and Figure 1 is as follows: falling marginal cost along the curve for the new technology in Figure 1 is one factor contributing to the downward rotation of the *MAC* curve in the figure of Box 1.

5 One could argue that the mature technology will eventually exhaust its potential for further expansion and, as a result, marginal production costs may start sloping upwards. Although not reflected in Figure 1, an upward-sloping marginal cost curve for the mature technology does not change the gist of the arguments presented here. But as will be discussed below, rising marginal cost of the mature technology tend to weaken the case of promoting new technologies.

future. But again, it is fairly unlikely that altruism is strong enough to push new renewables down their learning curves. To conclude: in contrast to flat TV screens, for example, a new renewable electricity-generating technology cannot bank on consumers' choice to help it into the market. But this does not preclude the possibility that profit-maximising firms nonetheless embark on the new technology, expand output over time, and thus reap economies of learning.

Figure 1. Marginal cost of a new renewable technology with internal economies of learning – firm perspective



Using Figure 1, it is easy to see why. Suppose the marginal cost of the mature renewable technology set the price for clean electricity. Potential electricity producers using the new technology then face a standard investment problem, that is, to spend upfront an amount equal to *A* in return for future profits *B*. They would thus use dynamic pricing, that is, sell at a price below cost as long as cumulative output is below Q^* and sell at a price above cost once cumulative output is above Q^*. If the net present value of *B* is greater than that of *A*, it is financially profitable to use the new technology and there would be, in principle, no need for policies to promote it.[6]

In practice, however, the way to market success of new technologies is fraught with market failures and barriers. For instance, financial markets might fail in providing the finance needed to help the new technology to sustain initial losses, or the finance they provide might be too expensive. It is true that this applies to the financing of investment in general, but when the success of that investment depends on uncertain learning effects, it might be particularly relevant. Given that this market failure originates in financial markets, policy measures directly addressing the causes of financial market failures are first best, but subsidised funds for financially constrained developers of new technologies could be an effective second best.

The way to market success of new technologies is fraught with market failures and barriers.

6 It is useful to add that this conclusion holds even if environmental policies do not fully internalise the economic cost of emissions associated with fossil fuels. What is crucial is that there is a price for renewable electricity – regardless of whether this price equals a feed-in tariff, the sum of the market price of electricity and the price of 'renewable' certificates in a tradable renewable certificates system, or the market price of electricity with fully or partly internalised emission cost, or a price following from some other renewables support scheme. In our stylised presentation of Figure 1, this price must be at least as high as the marginal cost of the mature technology. If not, only the new technology could be profitable. If the economic cost of emissions are not fully internalised, there is a case for promoting renewables in addition to what is achieved through influencing the relative price of renewables. But this is an argument for additional support to all renewables, new and mature.

The presence of dominant incumbents using well-established mature renewable technologies might pose a formidable barrier to the entry of firms using new ones.

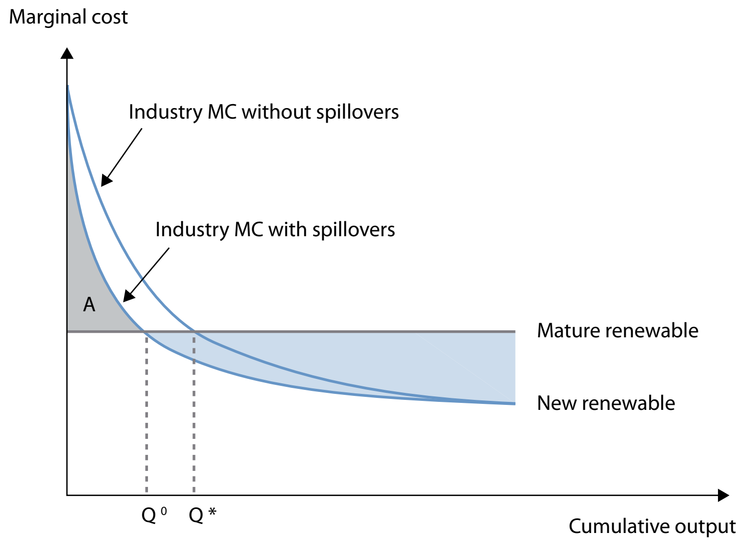
Another reason why new renewable technologies might fail to establish themselves can be found in the structure of electricity markets: they are everything but perfectly competitive. To illustrate, the presence of dominant incumbents using well-established mature renewable technologies might pose a formidable barrier to the entry of firms using new ones. These firms face the risk that once they have moved down the learning curve, incumbents use their power to instigate a decline in the price of renewable electricity, thereby foiling the backloaded part of the dynamic pricing strategy, that is, to sell at a price (sufficiently) above marginal cost. Indeed, this risk might give financial markets good reason to be cautious when considering finance for new renewable technologies. In these circumstances, the first-best policy would be competition policy that lowers barriers to market entry and exit and ensures that dominant incumbents do not abuse their power to keep new technologies away. Again, direct support for new technologies can be considered a practical second best. In this context, it is interesting to note that support in the form of feed-in tariffs that credibly promises a sufficiently high long-term renewable electricity price are relatively immune to the misbehaviour of incumbents.

The dynamic pricing strategy necessary for a new technology to establish itself also fails when firms that have not invested in the learning process nonetheless benefit from it. Not having incurred initial losses *A*, they can sell electricity produced on the basis of the new technology at a price below the marginal cost of the mature one. Obviously, the possibility that firms free ride on the learning acquired by other firms assumes that learning spills over, which takes us to the rationale for promoting new renewable technologies when economies of learning are external.

3.3 External economies of learning

From the viewpoint of a firm trying to establish a new technology, learning spillovers are not desirable because the firm makes an investment whose return it cannot fully appropriate. From the viewpoint of society, however, learning spillovers are beneficial as they represent a positive externality from the activity of a particular firm. Figure 2 illustrates this effect. The setup is similar to that in Figure 1, but now the marginal cost curve is for the whole industry, and there is a distinction between the case with and without learning spillovers.

Figure 2. Marginal cost of a new renewable technology with internal and external economies of learning – industry perspective



Without spillovers, all firms in the new-technology industry develop it on their own, each making progress in different directions and being able to prevent other firms to benefit from its progress – unless they pay for it. With spillovers in the industry, all benefiting firms have lower marginal costs than in the case with no spillovers for the same level of cumulative output. Therefore, the industry marginal cost curve in the presence of learning spillovers lies below the one without spillovers for any level of cumulative output. As Figure 2 suggests, from society's viewpoint, learning externalities lower the upfront cost associated with establishing the new technology (area *A* shrinks) and increase its future benefits (area *B* increases). This means that the economic return to the new technology is larger than its financial return and that the new technology would become competitive with the mature one at a lower level of cumulative output ($Q°$ instead of Q^*), thereby bringing forward other benefits possibly associated with the new technology – such as its contribution to a diversified set of clean-energy generation capacities.

The trouble is that learning spillovers discourage firms from establishing the new technology as they can no longer appropriate all the gains. This is liable to delay the commercialisation of economically profitable new technologies or completely prevent them from entering the market. Learning spillovers thus represent a clear market failure justifying policy intervention. The first-best policy is one that directly addresses the learning externality. Suppose that learning is embodied in the labour force of firms that choose the new technology and consider that this labour force might move on to free-riding firms. A first-best policy would be to subsidise on-the-job training. Another first-best policy candidate is support for demonstration plants on condition that the learning and experience gained in this endeavour is made available to other firms in the industry. By contrast, a long-term output subsidy to all firms does not seem to be first-best unless it is well targeted to the early movers in trying out new technologies.

Learning spillovers discourage firms from establishing new technologies as they can no longer appropriate all the gains.

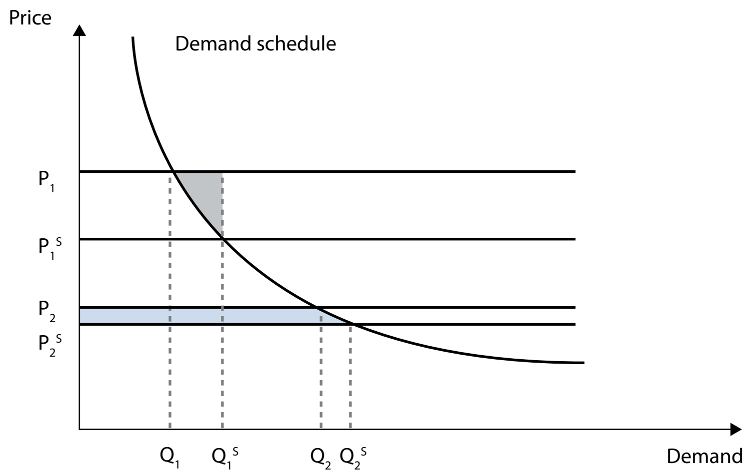
3.4 Strategic deployment

The first-best policies we have alluded to above aim at the supply side of establishing new technologies. First-best policies are notoriously difficult to implement both for political and practical reasons. As for renewable energy, many studies – notably Duke and Kammen (1999), Duke (2002), and Neuhoff (2005) – argue for demand-oriented policies in addition to supply side measures, in particular when policy makers shy away from environmental policies that are strong enough to fully internalise the negative externalities of producing and using energy. Strategic deployment or buy-down is one such policy. The thrust of it is to boost demand for new, near-market technologies so as to help producers move down the learning curve until they become competitive with existing technologies.

Figure 3 illustrates strategic deployment policies with a simple two-period analysis. Without a policy intervention, first-period sales by firms using a new technology are assumed to amount to Q_1 at a unit price of P_1. In the next period, marginal production costs will be lower because of learning effects, allowing firms to cut their price to P_2 and sell Q_2 units. Consider now a policy to help deploy this new technology and assume that this policy consists of subsidising buyers in the first period. Such a subsidy reduces the price buyers have to pay to P^S_1, thereby increasing demand to Q^S_1. This higher first-period output helps firms go even further down the learning curve, allowing second-period sales of Q^S_2 at a non-subsidised price of P^S_2.

Does this policy raise or reduce economic welfare? The factors shaping the answer to this question are pictured in Figure 3. The net cost of subsidising first-period purchases is equal to the grey area. This needs to be compared to the increase in the second-period consumer surplus, which is equal to the blue area. Obviously, for a given demand schedule and first-period subsidy, deployment policies are the more likely to be welfare enhancing the bigger the learning effect triggered by the increase in first-period sales from Q_1 to Q_1^S. At least two caveats should be made. First, the costs of strategic deployment policies pictured in Figure 3 do not include opportunity costs. The economic cost of deployment could thus be higher than what the grey area suggests. Second, benefits could be higher if the additional use of new-technology renewables were to replace polluting energy whose environmental externality is not fully internalised. But there will be no environmental benefit if new-technology renewables replace mature-technology renewables.

Figure 3. Welfare effects of strategic deployment



3.5 Summary and qualifications

Establishing new renewable technologies makes sense only if their costs are expected to fall below the costs of mature renewables.

An obvious conclusion following from our analysis is that establishing new renewable technologies makes sense only if their costs are expected to fall below the costs of mature renewables. This is true as long as new technologies do not have other advantages compared to mature technologies. As obvious as it seems, support in favour of new renewable technologies is often justified on the grounds that they are expected to become competitive with mature technologies. But just becoming competitive is clearly not good enough. Why should firms or societies invest in a learning process that is anticipated to achieve nothing more than eventually making new technologies just as good as mature ones? To make it concrete, suppose the marginal cost of generating electricity on the basis of a mature renewable technology, say, onshore wind is €50 per MWh. Establishing a new technology, say, solar thermal power that currently generates electricity at €200 per MWh makes sense only if there is hope that due to economies of learning, the cost of solar energy will fall below €50 per MWh.

In this context it is sometimes observed that today's cost of mature technologies (€50 per MWh in our example) is the wrong benchmark. Mature technologies might become more expensive in the future because – to remain with the example of onshore wind energy – favourable locations for onshore wind farms become scarce, forcing additional wind farms into marginal sites with higher if not prohibitive production cost. Against this background, helping to commercialise currently expensive modes of producing renewable electricity could be seen as a means to ensure that affordable alternatives are available as and when mature renewables become costly or, worse, cannot contribute at all to further raising the share of renewables in the overall energy mix.

This argument is appealing at first glance, but it is flawed nonetheless. Although it is true that the marginal cost of mature renewables might rise in the decades to come, this does not strengthen the case for supporting the commercialisation of new renewables. If anything, the opposite holds. A glance at Figure 2 shows why. An expected increase in the cost of mature renewables would imply an upward-sloping marginal cost curve for this technology and, in fact, one could imagine the curve to become vertical for a very high level of cumulative output. All other things being equal, this increases the return to learning associated with establishing new renewables (in Figure 2 area *A* shrinks while *B* expands), thereby encouraging firms to start using new renewables. It also tends to mitigate the market failures that could prevent new technologies from establishing themselves. Take the financial market failure for instance: while a firm might find it hard to convince financiers of its learning potential – there definitely is an asymmetric information problem – they surely recognise the potential for new technologies if the deployment of mature technologies is widely believed to become increasingly constrained. In sum, an expected increase in the cost of mature renewable technologies will encourage a market-driven transition from mature to new renewables similar to the gradual development and commercialisation of other 'backstop' technologies, such as unconventional oil as a substitute for conventional oil.

Another conclusion worth stressing is that when arguing in favour of policies to promote new renewable technologies, it is not sufficient to observe that their future benefits will outweigh today's cost. What needs to be shown is that new renewable technologies cannot establish themselves or – if they can – that social returns to investing in learning economies are larger than private returns. There is then a dilemma: while it is intellectually fairly easy to contemplate market failures that could hinder the commercialisation of economically viable technologies, it is much harder to find out how relevant these market failures are in practice and how much support new technologies need to overcome them. In this context, guidance is often sought from learning and experience processes that today's mature technologies went through in the past. Against this background, the next section will turn to empirical learning and experience curves and discuss what they tell us about the market failures that might hinder the commercialisation of promising new technologies.

When arguing in favour of policies to promote new renewable technologies, it is not sufficient to observe that their future benefits will outweigh today's cost.

4. Empirical experience curves – what they tell, and what not

The empirical observation that many technologies have become cheaper with increasing market penetration is one of the main arguments of proponents of policies in support of new renewables (see Duke 2002, IEA 2000, and Stern *et al*. 2006). At the centre of this observation are empirical estimates for the learning curves, introduced in stylised fashion in Section 3. The purpose of this section is to review key empirical findings about learning and experience curves, assess how much they help in deciding whether or not to promote new renewables technologies, and to illustrate the pitfalls if they are used to gauge the scope of policies in favour of new technologies.

Surveying the research literature on learning curves, Dutton and Thomas (1984) find that, on average, unit costs decline by approximately 20 percent each time production doubles. As Box 2 sets out in greater detail, the percentage decline in cost associated with a doubling of output is called the learning rate. Closely related to the learning rate is the so-called progress ratio, which is 100 (percent) minus the learning rate. The sheer scope of the survey, covering studies of more than one hundred different technologies in a wide range of industries, seems to lend credibility to the claim that the link between a rise in cumulative output and a decline in cost is a fact rather than just a coincidence, but we will see below that coincidence cannot be ruled out.

Box 2. Learning and experience curves, progress ratio, and learning rate

Economists have defined learning and experience curves – and the difference between the two – in a more precise way than is needed for most of the points made in this paper. Learning curves are meant to capture the process of improving labour productivity as workers learn to work faster and more efficiently. Specifically, learning curves plot unit labour costs as a function of cumulative output of a firm. The fundamental assumption here is that cost reductions are driven by cumulative output. In the empirical literature, learning curves assume the following conventional formulation:

$$c_t = c_0 \left(\frac{Y_t}{Y_0} \right)^b$$

In this equation, c_0 and Y_0 , respectively, represents the unit labour cost and production in period *0*, c_t is the unit labour cost in period t, Y_t is the cumulative production up to period t (but not including the production of period t), and b is the learning parameter. With this particular functional form, two indicators have gained popularity – the progress ratio (PR) and the learning rate (LR). The progress ratio is the ratio of unit labour costs after production has doubled to unit costs before production doubles. In other words, multiplying unit costs associated with a given level of cumulative output by the progress ratio yields unit costs after a doubling of output. The learning rate is just 1-PR. Thus, multiplied by 100, LR gives the percentage change in unit costs when cumulative output doubles. Mathematically, if period T, $Y_T=2Y_t$, then,

$$PR = \frac{c_T}{c_t} = \frac{c_0 \left(\frac{2Y_t}{Y_0} \right)^b}{c_0 \left(\frac{Y_t}{Y_0} \right)^b} = 2^b, \text{ and } LR=1-2^b.$$

The concept of experience curves – related to learning curves, but broader – was introduced by the Boston Consulting Group (1972). They argued that total unit costs would come down rapidly not only because workers learned (the learning curve) but also because experience would lead to optimisation of research, development, production, marketing, and so on. Experience curves, therefore, plot total unit costs as function of cumulative output. In practice, they are formulated using the same functional form as that of learning curves. The only change is that c_t now denotes total unit costs rather than only unit labour costs.

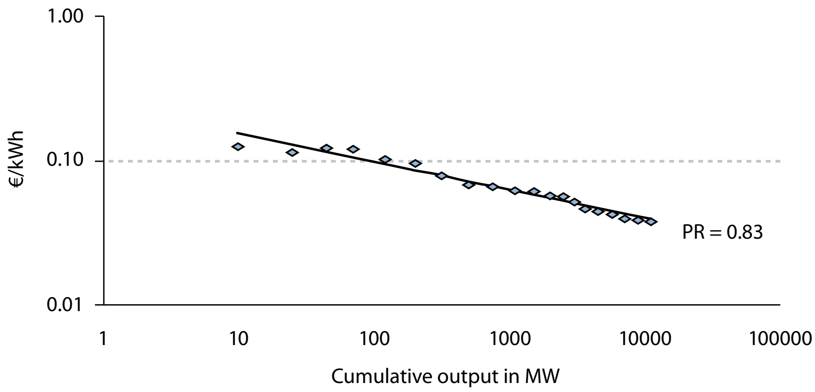
Empirical work, constrained by the availability of data on unit cost, typically assumes that prices equal marginal cost, as under perfect competition, so that the former can substitute for the latter in the specified functional form. Replacing unit cost with price (p), normalising Y_0 to one, and applying logarithmic transformation, the estimated relation becomes:

$$\log(p_t) = c_0 + b \text{ x } \log (Y_t) + v_t$$

where v_t is a random error term. Such relations have been fitted to both firm-level and industry-level data. Different experience curves have also been fitted for producers of energy equipment (wind turbines, photovoltaic panels, and so on) and for energy producers using this equipment. Experience curves estimates exist for particular countries or regions as well as for the whole world.

Figure 4 shows an experience curve for the production of wind energy based on wind turbines produced by Danish manufacturers in the period 1981-2000 (Neij *et al.* 2003). The authors have estimated a progress ratio of 0.83, implying that, on average, electricity generating cost decreased by 17 percent each time wind turbine sales doubled in the period 1981-2000. Wind energy cost declined from around €130/MWh in 1981 to around €40/MWh in 2000.

Figure 4. Experience curve for wind energy, 1981-2000



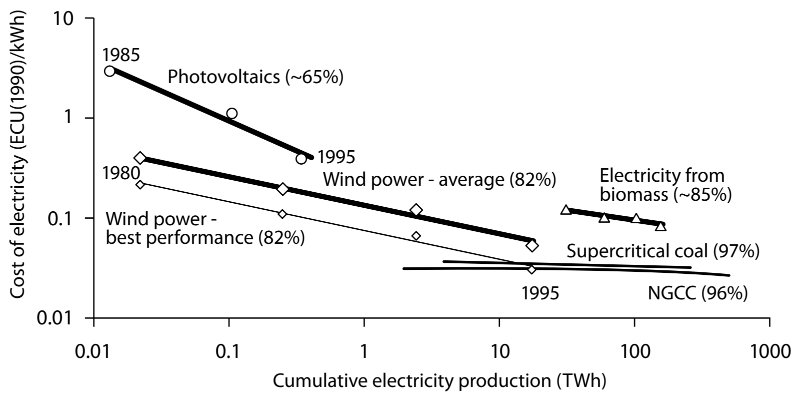
Source: Neij *et al.* (2003).

Notes: Both axes in logarithmic scale; levelised electricity generation cost (in 2000 prices) using wind turbines made by Danish manufacturers; PR ≡ progress ratio.

A key conclusion following from the survey of Dutton and Thomas (1984) is that progress ratios vary considerably across technologies. Figure 5 – taken from IEA (2000) – shows that this also applies to experience curves and progress ratios for various electricity-generating technologies. More specifically, experience curves range from the very steep one for photovoltaics (with a progress ratio of 65 percent) to the almost flat curve for supercritical coal (97 percent). Against this background, it could be misleading to assume that progress ratios known from one technology apply to others or to extrapolate historical progress ratios into the future (Box 3 illustrates the scope for error when empirical progress ratios are used to quantify the policy support for a currently new technology). But from a policy perspective there are even more fundamental problems.

Experience effects vary considerably across technologies and it could thus be misleading to assume that experience effects known from one technology apply to others.

Figure 5. Experience curves for different renewable technologies in the EU, 1980-95



Source: IEA (2000).

Notes: Both axes in logarithmic scale; electricity generation cost (in 1990 prices); in contrast to Figure 4, cumulative output is not measured in generating capacity (MW) but in electricity produced (TWh); numbers in parentheses are estimates of progress ratios.

Empirical learning curves merely show a correlation between cumulative output and cost rather than a causal link …

One is that empirical learning curves merely show a correlation between cumulative output and cost rather than a causal link. The estimated decline in total unit costs, for example, is likely to reflect other factors in addition to experience: economies of scale, general technological progress, standardisation, change in input prices, and so on. All these factors would contribute to a downward-sloping experience curve. Nemet (2006), for instance, finds that plant size, module efficiency, and the cost of silicon are the key determinants of cost reductions in the photovoltaic industry. His analysis reveals that learning is only weakly related to these factors. Papineau (2006) has shown that if one accounts for the elapsed time, cumulative output is no longer a significant determinant of cost reductions in photovoltaics, thermal solar power, and wind. This time trend most probably captures the effect of important omitted variables such as economy-wide productivity improvements. A final caveat as to the reliability and interpretation of estimated progress ratios: they might be distorted because assumptions made in estimating them might be false. For example, estimates assume that prices approximate well unobserved marginal unit costs. If that is not true, estimated progress ratios would be incorrect and misleading.

Box 3. Using empirical progress ratios for policy-making purposes can be costly

The purpose of this Box is to illustrate the ambiguities that arise when empirical progress ratios are used to gauge the scope of policies in support of new technologies. For this illustration, we use some of the findings of Neij *et al.* (2003) as a starting point. In this study the authors have estimated a variety of experience curves, using data for different producers, different measures of cumulative output, or different measures of costs. Figure 4 in the text presents one of these estimates, for which unit costs are measured as levelised electricity production costs. For the purpose of this illustration, we take another experience-curve estimate that uses market price of wind turbines as measure of unit costs. Considering the period 1981-2000, the estimated progress ratio (PR) for onshore wind turbines produced in Denmark is 0.92.

Imagine now that policy makers are considering whether to subsidise the purchase of a new renewable technology similar to onshore wind turbines. In order to make a good decision about how much this technology should be supported, policy makers take into account that production costs for onshore wind turbines decreased by 8 percent each time production doubled (PR = 0.92) and that the long-run marginal cost of the new technology is expected to be around €750 per kW (in 2000 prices), which was approximately the average market price for onshore wind turbines back in 2000. Further, the current new-technology renewable is approximately as expensive as onshore wind was in 1981, i.e. €1,500 per kW in 2000 prices with a cumulative output of 10 MW.

Following the argument of Duke (2002) that buy-down support should be extended until a technology reaches its long-run marginal costs and based on previous experience with onshore wind, policy makers readily calculate that the technology will be supported until cumulative output reaches 4.7 GW.

The trouble is that, even if we agree that the new technology will repeat the progress of onshore wind, the PR of 0.92 is only a statistical estimate that is surrounded by some uncertainty. In order to account for this uncertainty, policy makers need to quantify it. A proper way to do so is to calculate the so-called confidence interval, which gives an upper and a lower bound for

Even if ones assumes that cost savings are indeed caused by cumulative output and that estimated progress ratios reflect reality, a key problem remains: empirical experience curves do not provide any information on market failures that are believed to hinder new technologies in establishing themselves. Take the wind energy experience curve of Figure 4, for example. We do not know the counterfactual, that is, how cumulative output would have increased and cost declined in the absence of support for wind energy. And even if we knew the counterfactual, we could not tell whether the then observed experience curve pictures a learning process that suffered from market failures, and how severe these failures were.

… and they do not provide any information on market failures that are believed to hinder new technologies in establishing themselves.

To conclude, empirical experience curves show first and foremost correlations between total unit cost (approximated by market prices) and cumulative output of a formerly new technology. In contrast to what proponents of policies in support of new technologies implicitly assume, they do not vindicate such policies. But this should not really come as a surprise if one considers the origin of experience curves as an underpinning of profit-maximising firm behaviour. The concept

the estimated progress ratio and informs policy makers that the estimate will lie within these bounds with a certain probability. For a probability of 95 percent, the confidence interval ranges from 0.913 and 0.936. Thus, policy makers should consider these alternative values to see how uncertainty affects the estimate of 4.7 GW. In other words, they should consider two more scenarios: a low PR of 0.913 (suggesting faster progress) and a high PR of 0.936 (suggesting slower progress); the former will give a lower bound for the cumulative output needed to reach the targeted unit cost whereas the latter will give an upper bound. Recalculating, policy makers find that the warranted cumulative output could be as low as 2 GW and as high as 15 GW!

This wide range for the warranted increase in cumulative output will most probably result in large bounds for the cost of the policy. To illustrate this we need to make an assumption about the demand for this new technology. In general, for less elastic demand, subsidies have to be larger to induce potential users of the technology to buy it while more elastic demand results in smaller policy cost. This brings us to the second important source of uncertainty – the estimate of the price elasticity of demand, which like progress ratios, comes with statistical errors.

For the sake of simplicity, let demand have a constant price elasticity both for a given level of output and over time. Further, let the price elasticity of demand have a very small statistical error resulting in a 95 percent confidence interval of +/- 0.5 percent around the central estimate. We then calculate demand under the same three scenarios as for the progress ratios – a central scenario that uses the point estimate of price elasticity, and a 95 percent confidence interval for this estimate. Assuming a five percent discount rate and a twenty-year horizon for the policy, we obtain the following results. If both the progress ratio and demand coincide with the central estimate, the present value of the public subsidy is €776m. If progress is fast and demand is more elastic, the present value is only €202m. However, if progress turns out to be slow and demand less price elastic, the present value of the subsidy becomes €1,440m!

Put simply, with a probability of 95 percent, the present value of the deployment subsidy falls in the interval €202m and €1,440m, with the upper bound seven times bigger than the lower bound. Evidently, even taking the central estimate, policy makers bear a non-trivial risk of the deployment programme being twice as expensive as foreseen.

of experience curves was introduced by the Boston Consulting Group (1972), a management consultancy, in the 1970s. BCG advised their customers to strategically increase production of new products, even though they might encounter losses in the beginning. The argument was that total unit costs would come down rapidly, giving a firm riding down its experience curve a strategic advantage over competitors. Business strategies based on experience curves are reported to have sometimes ended in spectacular failures. But this should not be too surprising either if advice is given on the basis of statistical correlations rather than a good understanding of cause and effect.

5. Cost-benefit rules for new-technology renewable energy projects

The previous sections stressed that the rationale for promoting new technologies rests on two conditions: first, the cost of new technologies falls as and when the use of them spreads and, second, new technologies cannot establish themselves at all, or not as fast as they should, because of various market failures – notably learning spillovers. Leaving the policy-oriented discussion behind and assuming that both conditions are fulfilled, this section turns to the question of how to assess the economic costs and benefits of investments in new-technology renewables.

When assessing the economics of projects, one needs to be clear about the decision situation. In the parlance of cost-benefit analyses, the 'with' and 'without' project scenarios need to be correctly specified. In the case at hand, we assume that a power plant is needed either to satisfy a growing demand or to replace an obsolete plant for a constant demand.[7] We define the new-technology renewable as the 'with' project scenario and consider two alternative 'without' project scenarios. One is a fossil-fuel-fired power plant, the other a mature-technology renewable.

When assessing a renewable energy project that uses a new technology, one needs to bear in mind that the energy output of this project is as 'green' as renewable energy based on mature technologies …

Which economic costs and benefits need to be taken into account when comparing these three options? They all produce the same amount of electricity and, thus, the economic value of electricity can be ignored. And as they do not generate any other benefits, choosing among options depends only on their costs, that is, the cost-benefit analysis simplifies to a least-cost analysis. The cost of the fossil fuel alternative comprises the private cost of generating electricity and the external environmental cost associated with fossil fuels. By contrast, we ignore negative environmental externalities that both renewables might have and, therefore, consider only their private electricity generating costs. As both renewables would come in lieu of fossils, one could treat the avoided environmental cost of fossils as a benefit of renewables. It is crucial to note, however, that both renewables would generate the same (relative) environmental benefit. In other words, the new renewable is as 'green' as the mature one.

To further narrow down the decision situation, we assume that while the fossil-fuel option has lower private generating cost than the mature renewable, its economic costs are higher because of its negative environmental effects. Thus, in a comparison of the mature renewable with the fossil fuel option, the latter is discarded. This leaves a choice between the mature renewable and the new one. In sum, when considering a new-technology renewable energy project, the 'without' project scenario is not a fossil-fuel-fired power plant, but one based on the mature renewable technology. It follows that environmental aspects should not influence the decision for or against the new renewable. But what, then, determines the choice between the mature and the new renewable?

7 The economic viability of meeting electricity demand is thus taken for granted, or – to put it differently – leaving demand unmet is not considered a relevant 'without' project scenario.

By virtue of the problem we want to analyse, the new renewable currently costs more than the mature one. Let MC^{N0} and MC^{M0} denote, respectively, today's levelised marginal electricity generating cost of the new renewable and today's levelised marginal electricity generating cost of the mature renewable. Today's situation is thus characterised by $MC^{N0} > MC^{M0}$. But the new renewable could become cheaper in the future if it is more widely used. Let us presume that the new renewable cannot establish itself due to the market failures discussed in Section 3, but that choosing it despite its current cost disadvantage pushes it down its experience curve, triggering a cost decline in the future. Let MC^{Nt} and MC^{Mt} denote, respectively, the levelised marginal electricity generating cost of the new renewable and the mature renewable in all future periods t = 1, ...n.

… implying that environmental aspects should not influence the decision for or against the new-technology renewable.

We could then imagine two alternative future trajectories, one in which the mature renewable is used and another trajectory in which the new renewable is used. As shown in detail in Kolev and Riess (2007),[8] getting on the new-renewable trajectory would make society better off if and only if

(1) $$MC^{N0} + \delta^1 MC^{N1} + \cdots + \delta^n MC^{Nn} \quad < \quad MC^{M0} + \delta^1 MC^{M1} + \cdots + \delta^n MC^{Mn}$$

with $\delta = 1/(1+r)$ being the one-period discount factor and r the discount rate. If inequality (1) holds, the present-value generating cost of the new-renewable trajectory is smaller than the present-value generating cost of the mature-renewable trajectory.

Because economies of learning are assumed to reduce the cost of the new renewable, $MC^{N0} \geq MC^{N1} \geq \cdots \geq MC^{Nn}$. As indicated above, a crucial assumption embedded in the decision rule (1) is that the consecutive decline in the cost of the new renewable materialises only if society embarks on the new-renewable trajectory, and (1) suggests when this is better than staying on the mature-renewable trajectory. Needless to say: this is a very favourable assumption from the perspective of the new renewable.

To use (1) in applied project appraisal, it is necessary to specify the size and time profile of the expected decline in the cost of the new renewable. Moreover, to the extent that the cost of the mature renewable is envisaged to rise in the future, for reasons discussed in Section 3, the magnitude and timing of this increase needs to be accounted for in (1). In the remainder of this section, we consider a special case of (1) and offer a numerical illustration.

One feature of this special case is that the hoped-for decline in MC^N materialises in period $t = i$ $(1 < i < n)$ and that there is no further decline thereafter. The relationship between cost before (MC^{N0}) and after ($MC^{N\,t \geq i}$) the cost decline is $MC^{N\,t \geq i} = \alpha\, MC^{N0}$ with $\alpha < 1$ and 1- α (multiplied by 100) indicating the percentage decline in the cost of the new renewable. The other feature of this special case is that the cost of the mature renewable remains unchanged and can thus be expressed as a constant fraction $\beta < 1$ of the cost of the new renewable before the new renewable becomes cheaper, that is, we can write $MC^{Mt} = \beta\, MC^{N0}$. In essence, β captures the current cost disadvantage of the new renewable, with this disadvantage being the bigger, the smaller β.

Assuming an infinite planning horizon, we can then calculate a critical value α^*:

(2) $$\alpha^* = \frac{1}{\delta^i}\left[\beta - \left(1 - \delta^i\right)\right] \quad with \quad \frac{\partial \alpha^*}{\partial \beta} > 0, \quad \frac{\partial \alpha^*}{\partial i} < 0 \quad and \quad \frac{\partial \alpha^*}{\partial \delta} > 0.$$

8 In this paper we also show why environmental aspects are irrelevant for the decision rule although they have been taken into account in the welfare-maximising model that leads to (1).

If the hoped-for cost decline is such that $\alpha < \alpha^*$, embarking on the new-renewable trajectory results in present-value electricity costs that are lower than those associated with the mature-renewable trajectory, and *vice versa*.

The often-made claim that promoting new renewables is worthwhile provided they become competitive would be correct if society had virtually no time preference.

A few implications of (2) are useful to point out. First, imagine there is virtually no time preference, which implies that the discount rate r approaches zero and the discount factor δ approaches 1. In this case, $\alpha^* = \beta - \varepsilon$, with ε being a number very close to zero. That is, to decide in favour of the new renewable, it would need to become only marginally cheaper than the mature renewable at some point in the future.[9] Hence, the often-made claim that promoting new renewables is worthwhile provided they become competitive would be correct if society was virtually indifferent between income today and income tomorrow.

Second, as the positive sign of the partial derivative $\partial\alpha^*/\partial\delta$ shows, the smaller δ, the lower the critical threshold α^*. In words: the higher the time preference (small δ), the larger the required cost decline (small α).

Third, as the positive sign of the partial derivative $\partial\alpha^*/\partial\beta$ shows, the smaller β, the lower the critical threshold α^*. In words: the larger the cost disadvantage of the new renewable today (small β), the larger the required cost decline (small α).

Fourth, as the negative sign of the partial derivative $\partial\alpha^*/\partial i$ shows, the larger i, the lower the critical threshold α^*. In words: the longer it takes for cost to decline (large i), the larger the required cost decline (small α).

Let us illustrate this with a numerical example. Table 1 shows by how much the cost of the new-technology renewable must decline (in percent) to justify investing in this technology today despite the fact that society can use a currently cheaper mature technology. The required cost decline is shown for alternative values of the new technology's current cost disadvantage (β) and for an alternative number of years it takes for the cost decline to materialise (i). Recall that a lower β signals a greater cost disadvantage of the new renewable.

Suppose the current cost disadvantage of the new renewable is such that the mature renewable offers electricity at 70 percent of the cost of the new renewable ($\beta = 0.7$). Further assume that the new renewable will experience its cost decline after ten years – conditional on being chosen today. For this choice to be economically beneficial, the hoped-for cost decline would need to amount to at least 49 percent. The table also illustrates that the required decline in the cost of the new renewable rises with the number of years for learning to reduce costs and with the initial cost disadvantage. For instance, keeping the current disadvantage unchanged ($\beta = 0.7$), but assuming that the cost of the new-technology renewable declines after 15 years, yields a required cost decline of 62 percent. And then, fixing the number of years at ten, but assuming $\beta = 0.5$, the required cost decline would be 81 percent; in other words, if the current cost disadvantage of the new renewable is such that the mature renewable offers electricity at half the cost of the new renewable and if the cost decline happens after 10 years, the hoped-for cost decline needs to be at least 81 percent.

As can be seen from the table, the current cost disadvantage of the new renewable could be so big and/or the hoped-for cost decline could lie so far in the future that cost would need to decline by more than 100 percent (reflecting $\alpha^* < 0$ in (2)). Obviously, this is not feasible, suggesting that on the basis of present-value generating cost, the new renewable cannot catch up with the mature one.

9 'Marginally' cheaper and 'at some point' in the future are sufficient as the alternative trajectories continue forever.

Table 1. Required cost decline (in %) of new-technology renewable to make it viable

		Years i to cost decline of new renewable		
		$i = 5$	$i = 10$	$i = 15$
Cost disadvantage of new renewable today	$\beta = 0.9$	13%	16%	21%
	$\beta = 0.7$	38%	49%	62%
	$\beta = 0.5$	64%	81%	>100%
	$\beta = 0.3$	89%	>100%	>100%

Notes: The figures in the table show the percentage decline in the cost of the new renewable relative to today's level. Moving down the β-column indicates a greater cost disadvantage of the new-technology renewable. Calculations are based on equation (2) for a discount rate of 5 percent. Values larger than 100% (α*<0) suggest that the new renewable cannot outperform the mature renewable on a present-value cost basis.

All in all, inequality (1) seems to offer a sensible rule to assess the economics of new-technology energy projects. Obviously, its purpose is to inform decision-making, but as with any rule its intention is not to indisputably distinguish between the good, the bad, and the ugly. It is relatively easy to turn into a hands-on project appraisal tool – as (2) and its numerical illustration shows. What is more, as shown in Kolev and Riess (2007), it is straightforward to refine the approach in developing (2). For instance, instead of considering a one-off drop in the cost of the new renewable after a certain time, one can model a gradual cost decline in line with the notion of learning effects. Furthermore, it is easy to account for the possibility that the cost of the mature technology increases over time. In any event, for project appraisal purposes, one would need to compare the required cost decline (for instance that shown in Table 1) with estimates of the decline and its timing. Arguably, arriving at such estimates is a challenge, and as we have pointed out in Section 4, experience curves associated with formerly new technologies could be very misleading.

Our rule for assessing the economics of a new-technology energy project is useful only if one believes that using this technology today causes the hoped-for future cost decline – if not, there would be no need for a rule in the first place.

As a final and perhaps most important point: our rule for assessing the economics of a new-technology energy project is useful only if one believes that using this technology today causes the hoped-for future cost decline. If that is not the case, our rule is meaningless – but then there would be no need for a rule in the first place.

6. Conclusions

In this paper, we have focused on two questions: first, what is the rationale for promoting new renewable energy technologies given that society already has mature ones and, second, how can cost-benefit analyses of energy projects based on new renewables account for this rationale – assuming there is one?

As to the first question, one conclusion – rather, a lesson re-learned – is that it is relatively easy to think of possible market failures that could justify a policy intervention – support for new renewables in our case – but that it is much harder to ascertain the practical relevance of such failures and to decide on the proper type, size, and duration of policy measures. This being said, the environmental market failures that bias the energy mix against renewables – new and mature – are arguably very relevant. In addition, considerable progress has been made in recent decades to value environmental externalities in monetary terms and the policy measures most suitable to internalise them are reasonably well understood – though not necessarily applied. Alas, this cannot be said for the technology externalities and market failures that could justify public support for new, but known renewable energy technologies.

To recap the main reasons for our scepticism, take the financial market failures and the behaviour of dominant incumbents that could prevent new renewables from penetrating the market and thus from generating internal economies of learning. In practice, it is virtually impossible to find out how severe these obstacles are and how much public support is needed to remove them. The same is true for external economies of learning since nobody knows how to value learning spillovers between firms with reasonable accuracy. In this context, reference is frequently made to learning and experience effects that came with formerly new technologies. The trouble is that empirical learning and experience effects do not inform about possible market failures and learning externalities. And even if they did, using them to devise support for currently new technologies could result in costly errors. As an aside: many schemes in support of new renewables – such as output or investment subsidies – normally do not directly address the presumed market failures, as subsidising on-the-job training would in the case of learning spillovers, for instance.

But why, then, is support for establishing new-technology renewables so popular? One reason could be that our analysis is flawed and our scepticism misplaced. We think, political-economy considerations offer a better explanation. As discussed in Section 2, policy measures that lower the cost of renewables have some logic if environmental policies are not stringent enough to bring about the socially optimal level of renewable energy. From a purely efficiency viewpoint, society would gain from policies that encourage renewables by fully pricing in the environmental cost of producing and using energy. However, such a policy has winners and losers, making it difficult to implement in practice. In fact, policy makers might prefer direct support for renewables for a simple reason: the benefits of direct support for renewables are visible, signalling that policy makers care about the problem at hand, while its costs are not. In sum, support for new-technology renewables could be seen as part of a second-best policy package that tries to raise the share of renewables in the overall energy mix to its optimal level in a situation where policy makers shy away from policies that would fully internalise the environmental cost of energy, for instance a sufficiently high tax on environmentally damaging emissions.

Policy measures that lower the cost of renewables have some logic if environmental policies are not stringent enough – but this does not justify supporting new renewables more than mature ones.

Even with this explanation, a snag remains: why specifically and more generously supporting new-technology renewables and not supporting renewables in general – new, mature, and in between? One reason examined in this paper is that helping to commercialise currently expensive modes of producing renewable electricity could be a means to ensure that affordable alternatives are available as and when mature renewables become costly or, worse, cannot contribute at all to further raising the share of renewables in the overall energy mix. A political-economy explanation could be that policy makers probably find it more rewarding to be seen as pushing the new rather than the established – at least in a field like renewable energy. Another reason has an industrial-policy flavour in that new renewables are seen as promising new industries that could conquer world markets and create employment. Of all the reasons in favour of promoting new renewables, this is perhaps one of the weakest unless, that is, one believes in the capacity of governments to pick winners or assumes that establishing new renewables will draw on resources that would have been unemployed otherwise. This leaves the reason we started with, namely that new-technology renewables indeed cannot establish themselves in the market as much (and as fast) as economic efficiency suggests they should.

This takes us to the second question. Assuming that new renewables cannot establish themselves, we have developed decision rules for assessing the economics of new-technology renewable energy projects. What these rules tell is hardly surprising for aficionados of cost-benefit analyses, but two of the messages they contain run against conventional wisdom. First, conventional wisdom has it that choosing a new-technology renewable makes economic sense if the new technology is expected to become competitive with mature technologies. Our decision rule shows that just becoming competitive is not good enough. On the contrary, the new technology needs to become cheaper than mature ones. The intuition is obvious: society should invest in a learning process only if that process yields something better than what society already has, otherwise the return on this investment would be negative. Second, conventional wisdom typically points to the environmental benefits that new renewables have relative to fossil fuels. Although it is certainly true that new renewables avoid the adverse environmental impact of burning fossil fuels, they do that no better than mature-technology renewables. The long and short of this is that environmental aspects are largely irrelevant for a rational decision on new-technology renewable energy projects when equally clean mature renewables are available.

References

Arrow, K. (1962). "The economic implications of learning by doing". *Review of Economic Studies* (29), pp. 155-173.

Baldwin, R. (1969). "The case against infant-industry protection". *Journal of Political Economy* (77), pp. 295-305.

Bliss, C. (1989). "Trade and development", in Chenery, H. and Srinivasan, T.N. (eds.), *Handbook of development economics*, Vol. 2, Elsevier Science, Amsterdam, The Netherlands.

Boston Consulting Group (1972). *Perspectives on experience*, Boston Consulting Group Inc., Boston, Massachusetts, USA.

Corden, M.W. (1985). "The normative theory of international trade", in Jones, R.W. and Kenen, P.B. (eds.), *Handbook of international economics*, Vol. 1, Elsevier Science, Amsterdam, The Netherlands.

Dasgupta, P. and Stiglitz, J. (1988). "Learning-by-doing, market structure, and industrial and trade policies". *Oxford Economic Papers* (40), pp. 256-268.

Duke, R.D. (2002). *Clean energy technology buydowns: economic theory, analytic tools, and photovoltaics case*, *Ph.D. Dissertation,* Woodrow Wilson School of Public and International Affairs, Princeton University.

Duke, R.D. and Kammen, D. M. (1999). "The economics of energy market transformation programs." *Energy Journal* (20:4), pp. 15-64.

Dutton, J.M. and Thomas, A. (1984). "Treating progress functions as managerial opportunity". *Academy of Management Review* (9:2), pp. 235-247.

EWEA (2004). *Wind Energy – The Facts*, European Wind Energy Association. http://www.ewea.org/fileadmin/ewea_documents/documents/publications/WETF/WETF.pdf

ExternE (2004) *Final report,* Externalities of Energy, European Commission Project. http://www.ier.uni-stuttgart.de/forschung/projektwebsites/newext/newext_final.pdf

Foxon, T.J., Gross, R., Chase, A., Holwes, J., Arnall A., and Anderson D. (2005). "UK innovation systems for new and renewable energy technologies: drivers, barriers and system failures". *Energy Policy* (33), pp. 2123-2137.

Goulder, L.H. and Mathai, K. (2000). "Optimal CO_2 abatement in the presence of induced technological change". *Journal Of Environmental Economics and Management* (39), pp. 1-38.

IEA (2000). *Experience curves for energy technology policy*, International Energy Agency, Paris, France.

Jaccrad, M. (2006). *Sustainable fossil fuels: the unusual suspect in the quest for clean and enduring energy*, Cambridge University Press, Cambridge, UK.

Jaffe, A.B., Newell, R.G., and Stavins, R.N. (2003). "Technological change and the environment", in Mäler, K.-G. and Vincent, J.R. (eds.), *Handbook of environmental economics*, Vol. 1, Elsevier Science, Amsterdam, The Netherlands.

Jaffe, A.B., Newell, R.G., and Stavins, R.N. (2005). "A tale of two market failures: technology and environmental policy". *Ecological Economics,* (54) pp. 164-174.

Johnson, H.G. (1965). "Optimal trade intervention in the presence of domestic distortion", in Caves, R., Johnson, H.G., and Kenen, P.B. (eds.), *Trade, Growth and the Balance of Payment*, Rand McNally.

Katsoulacos, Y. and Xepapadeas, A. (1996). "Environmental innovation, spillovers and optimal policy rules", in: Carraro, C., Katsoulacos, Y., and Xepapadeas, A. (eds.), *Environmental policy and market structure*, Kluwer Academic Publishers, Dordrecht, The Netherlands.

Kolev, A. and Riess, A. (2007). "Cost-benefit rules for new-technology renewable energy projects." *Economic and Financial Reports*, 2007/04, European Investment Bank.
http://www.eib.org/efs/publications/economic-and-financial-reports/index.htm

Krueger, A.O. (1985). "Trade policies in developing countries", in Jones, R.W. and Kenen, P.B. (eds.), *Handbook of international economics*, Vol. 1, Elsevier Science, Amsterdam, The Netherlands.

Neij, L., Andersen, P.D., Durstewitz, M., Helby, P., Hoppe-Kilper, M., and Morthorst, P.E. (2003). "Experience curves: a tool for energy policy assessment", *IMES/EESS Report No. 40,* Department of Technology and Society, Environmental and Energy Systems Studies, Sweden.

Nemet, G.F. (2006). "Beyond the learning curve: factors influencing cost reductions in photovoltaics". *Energy Policy* (34:17), pp. 3218-3232.

Neuhoff, K. (2005). "Large scale deployment of renewables for electricity generation". *Oxford Review of Economic Policy* (21:1), pp. 88-110.

Nordhaus, W. (2006) "The Stern review on the economics of climate change". Manuscript, http://nordhaus.econ.yale.edu/SternReviewD2.pdf

Papineau, M. (2006). "An economic perspective on experience curves and dynamic economies in renewable energy technologies", *Energy Policy* (34:4), pp. 422-432.

Perman, R., Ma, Y., and McGilvray, J. (2003). *Natural resource and environmental economics*, Third edition, Pearson Adison Wesley, Harlow, UK.

Riess, A. and Välilä, T. (2006). "Industrial policy: a tale of innovators, champions, and B52s." *EIB Papers*, (11:1), pp. 10-34.

Rivers, N. and Jaccard, M. (2006). "Choice of environmental policy in the presence of learning by doing". *Energy Economics* (28), pp. 223-242.

Stern, N., Peters, S., Bakhshi, V., Bowen, A., Cameron, C., Catovsky, S., Crane, D., Cruickshank, S., Dietz, S., Edmonson, N., Garbett, S.-L., Hamid, L., Hoffman, G., Ingram, D., Jones, B., Patmore, N., Radcliffe, H., Sathiyarajah, R., Stock, Taylor, M.C., Vernon, T., Wanjie, H., and Zenghelis, D. (2006). *Stern Review: the economics of climate change*, HM Treasury, London.

Tinbergen, J. (1955). *On the theory of economic policy*, Second Edition, North-Holland Publishing Company. Amsterdam, The Netherlands.

Tinbergen, J. (1956). *Economic policy: principles and design*, North-Holland Publishing Company. Amsterdam, The Netherlands.

Wright, T. (1936). "Factors affecting the cost of airplanes". *Journal of Aeronautical Sciences* (3), pp. 122-128.

Economic and Financial Reports*

Report 2006/03 **Rien Wagenvoort,** *"Comparing Distributions: The Harmonic Mass Index: Extension to m Samples"*

Report 2006/02 **Rien Wagenvoort,** *"Does the hedge fund industry deliver alpha?"*

Report 2006/01 **Frédéric Blanc-Brude, Hugh Goldsmith and Timo Välilä,** "Ex Ante *Construction Costs in the European Road Sector: A Comparison of Public-Private Partnerships and Traditional Public Procurement"*

Report 2005/03 **Gerti Dudkin and Timo Välilä,** *"Transaction Costs in Public-Private Partnerships".*

Report 2005/02 **Eric Perée and Timo Välilä,** *"Fiscal Rules and Public Investment".*

Report 2005/01 **Timo Välilä and Aaron Mehrotra,** *"Evolution and determinants of public investment in Europe".*

Report 2004/02 **Alfred Steinherr, Ali Tukel and Murat Ucer,** *"The Turkish Banking Sector: Challenges and Outlook in Transition to EU Membership".*

Report 2004/01 **Federico Galizia,** *"Systemic credit risk in the presence of concentration".*

Report 2003/02 **Federico Galizia,** *"Measuring the "financial gap" of European corporations: an update".*

Report 2003/01 **Rien Wagenvoort,** *"Bank Survey evidence on bank lending to SMEs in the European Union".*

Report 2002/02 **A. Rossi, G. Bichisao & F. Campolongo,** *"Hedge accounting within IAS39".*

Report 2002/01 **Luisa Ferreira de Almeida & Patrick Vanhoudt,** *"Catching the Celtic tiger by its tail".*

Report 2001/03 **Eric Perée & Alfred Steinherr,** *"The euro and capital markets: A new era".*

Report 2001/02 **Federico Galizia & Dermot O'Brien,** *"Do capital expenditures determine debt issues?"*

Report 2001/01 **Federico Galizia with Thomas Steinberger,** *"The "savings gap" of European Corporations: A first look at the available data".*

Report 2000/02 **Christopher Hurst & Kristian Uppenberg,** *"Investment issues in south-eastern Europe".*

Report 2000/01 **Dermot O'Brien & Rien Wagenvoort,** *"Some empirical findings on the characteristics of cost-efficient credit institutions".*

Report 99/P3 **Carsten Sprenger,** *"Direct investment and the eastern enlargement of the EU".*

Report 99/P2 **Daniel Piazolo,** *"The credibility and growth effects of EU institutions for Eastern Europe".*

Report 99/P1 **Hendrik Hassheider,** *"Private sector investment as a means of financing Trans-European Networks".*

Report 99/02 **Rien Wagenvoort & Paul Schure,** *"The recursive thick frontier approach to estimating efficiency".*

Report 99/01 **Paul Schure & Rien Wagenvoort,** *"Economies of scale and efficiency in European banking: New evidence".*

Report 98/05 **Daniel Ottolenghi,** *"Walking a tightrope: Financial liberalisation in Mediterranean countries".*

Report 98/04 **Christian Keuschnigg,** *"Venture capital: A case for investment promotion".*

Report 98/03 **Roman Arjona, Pascale Viala & Rien Wagenvoort,** *"Small business finance and the European financial environment: Some empirical evidence".*

Report 98/02 **Alfred Steinherr & Roman Arjona,** *"The impact of European Monetary Union on the development of the Mediterranean zone".*

Report 98/01 **Marco Buti, Lucio R. Pench & Paolo Sestito,** *"European unemployment: Contending theories and institutional complexities".*

Report 97/05 **Mireille Fischbach & Maria Scattaglia,** *"Financial flows to Eastern Europe: Determinants, volatility and sustainability issues".*

Report 97/04 **Mirela Keuschnigg,** *"Banking in Central and Eastern Europe".*

Report 97/03 **Gregorio Impavido,** *"Pension funds and stock market development in Eastern Europe: Is there a link?".*

Report 97/02 **Agnès Belaisch & Kristian Kjeldsen,** *"Value-at-risk models for determining capital adequacy: Does allowance for correlation within broad risk categories matter?".*

Report 97/01 **Agnès Belaisch,** *"Why are long-term interest rates so high in Europe?".*

Economic and Financial Reports are preliminary material circulated to stimulate discussion and critical comment.
*As from 2004, EFRs are available in electronic format only.

The above publications are available free of charge from:

European Investment Bank
Economic and Financial Studies

100, boulevard Konrad Adenauer
L-2950 Luxembourg

infoefs@eib.org (+352) 4379-8895

www.eib.org/efs

EIB Papers 2006 - 2007

Volume 12, Number 2, 2007: An efficient, sustainable and secure supply of energy for Europe. Meeting the Challenge

Shimon Awerbuch & Spencer Yang, *"Efficient electricity generating portfolios for Europe: maximising energy security and climate change mitigation"*
Machiel Mulder, Arie ten Cate & Gijsbert Zwart , *"The economics of promoting security of energy supply"*
Franz Hubert, *"Strategic investment in international gas transport systems"*
Joachim Schleich, *"The economics of energy efficiency: barriers to profitable investments"*
Dominique Finon, *"Pros and cons of alternative policies aimed at promoting renewables"*
Atanas Kolev & Armin Riess, *"Environmental and technology externalities: policy and investment implications"*

Volume 12, Number 1, 2007: An efficient, sustainable and secure supply of energy for Europe. Global and European policy perspectives

Atanas Kolev & Armin Riess, *"Energy – revival of a burning matter"*
Dieter Helm, *"European energy policy: meeting the security of supply and climate change challenges"*
Coby van der Linde, *"The art of managing energy security risks"*
Mark Jaccard, *"Fossil fuels and clean, plentiful energy in the 21st century: the example of coal"*
Juan Alario, *"EU policy objectives and energy investment decisions"*
Bassam Fattouh, *"The drivers of oil prices: the usefulness and limitations of non-structural models, supply-demand frameworks, and informal approaches"*

Volume 11, Number 2, 2006: An industrial policy for Europe? From concepts to action

Timo Välilä, *"No policy is an island – on the interaction between industrial and other policies"*
Andreas Strohm, *"Competition policy at war with industrial policy?"*
Otto Toivanen, *"Innovation and research policies: two case studies of R&D subsidies"*
Jordi Gual & Sandra Jódar, *"Vertical industrial policy in the EU: an empirical analysis of the effectiveness of state aid"*
Ha-Joon Chang, *"Industrial policy in East Asia – lessons for Europe"*
Christopher Allen, Didier Herbert & Gert-Jan Koopman, *"The European Commission's new industrial policy"*

Volume 11, Number 1, 2006: An industrial policy for Europe? Context and concepts

Armin Riess & Timo Välilä, *"Industrial policy: a tale of innovators, champions, and B52s"*
James Foreman-Peck, *"Industrial policy in Europe in the 20th century"*
Olivier Debande, *"De-industrialisation"*
Elie Cohen, *"Theoretical foundations of industrial policy"*
Charles Edquist & Cristina Chaminade, *"Industrial policy from a systems-of-innovation perspective"*
Richard Baldwin & Philippe Martin, *"Coordination of industrial policy in the European Union"*

The above publications are available free of charge from:

European Investment Bank
Economic and Financial Studies

100, boulevard Konrad Adenauer
L-2950 Luxembourg

infoefs@eib.org (+352) 43 79-8895

www.eib.org/efs

EIB Papers 2004 – 2005

Volume 10, Number 2, 2005: Innovative financing of infrastructure - the role of public-private partnerships: lessons from the early movers

Armin Riess, *"Is the PPP model applicable across sectors?"*
Paul A. Grout, *"Value-for-money measurement in public-private partnerships"*
Patricia Leahy, *"Lessons from the Private Finance Initiative in the United Kingdom"*
Rui Sousa Monteiro, *"Public-private partnerships: some lessons from Portugal"*
Andreas Brenck, Thorsten Beckers, Maria Heinrich, Christian von Hirschhausen, *"Public-private partnerships in new EU member countries of Central and Eastern Europe"*
Campbell Thomson, *"Public-private partnerships: prerequisites for prime performance"*

Volume 10, Number 1, 2005: Innovative financing of infrastructure - the role of public-private partnerships: infrastructure, economic growth, and the economics of PPPs

Timo Välilä, Tomasz Kozluk, & Aaron Mehrotra, *"Roads on a downhill? Trends in EU infrastructure investment"*
Ward Romp & Jakob de Haan, *"Public capital and economic growth: a critical survey"*
Christophe Kamps, *"Is there a lack of public capital in the European Union?"*
Timo Välilä, *"How expensive are cost savings? On the economics of public-private partnerships"*
Mathias Dewatripont & Patrick Legros, *"Public-private partnerships: contract design and risk transfer"*

Volume 9, Number 2, 2004: The internationalisation of production in Europe: Case studies of foreign direct investment in old and new EU members

Frank Barry, *"Export-platform FDI: the Irish experience"*
Thierry Mayer, *"Where do foreign firms locate in France and why?"*
Magdolna Sass, *"FDI in Hungary: the first mover's advantage and disadvantage"*
Gábor Hunya, *"FDI in small accession countries: the Baltic states"*

Volume 9, Number 1, 2004: The internationalisation of production in Europe: Causes and effects of foreign direct investment and non-equity forms of international production

Armin Riess & Kristian Uppenberg, *"The internationalisation of production: moving plants, products, and people"*
Zbigniew Zimny, *"Long-term trends in international production"*
Kristian Uppenberg & Armin Riess, *"Determinants and growth effects of foreign direct investment"*
Jozef Konings, *"The employment effects of foreign direct investment"*
Ari Kokko & Patrik Gustavsson, *"Regional integration, FDI, and regional development"*

The above publications are available free of charge from:

European Investment Bank
Economic and Financial Studies

100, boulevard Konrad Adenauer
L-2950 Luxembourg

infoefs@eib.org (+352) 4379-8895

www.eib.org/efs

EIB Papers 2002 – 2003

Volume 8, Number 2, 2003: Europe's changing financial landscape: The financing of small and medium-sized enterprises

Rien Wagenvoort, *"SME finance in Europe: introduction and overview".*
Rien Wagenvoort, *"Are finance contraints hindering the growth of SMEs in Europe?".*
Ulrich Hommel & Hilmar Schneider, *"Financing the German Mittlelstand".*
Michel Dietsch, *"Financing small business in France".*
Luigi Guiso, *"Small business finance in Italy".*

Volume 8, Number 1, 2003: Europe's changing financial landscape: Recent developments and prospects

Eric Perée & Armin Riess, *"The transformation of finance in Europe: introduction and overview".*
Thorsten Beck, *"Stock markets, banks, and economic development: theory and evidence".*
Graham Bishop, *"The role and development of EU capital markets".*
E. Philip Davis, *"Institutional investors, financial market efficiency, and financial stability".*
Arnoud Boot, *"Restructuring in the banking industry with implications for Europe".*

Volume 7, Number 2, 2002: The financial integration of an enlarged EU: Macroeconomic interlinkages

Robert A. Feldman & Nancy Wagner, *"The financial sector, macroeconomic policy and performance".*
Paul De Grauwe & Marianna Grimaldi, *"Exchange rate regimes and financial vulnerability".*
Willem Buiter & Clemes Grafe, *"Anchor, float or abandon ship: exchange rate regimes for the Accession countries".*

Volume 7, Number 1, 2002: The financial integration of an enlarged EU: Banking and capital markets

Armin Riess, Rien Wagenvoort & Peter Zajc, *"Practice makes perfect: A review of banking in Central and Eastern Europe".*
Dana Hájková, Jan Hanousek & Libor Němeček, *"The Czech Republic's banking sector: Emerging from turbulent times".*
Éva Várhegyi, *"Hungary's banking sector: Achievements and challenges".*
Peter Zajc, *"Slovenian banks a decade later".*
Erich Hampel, *"Successful banking in an enlarged EU".*
Jens Köke & Michael Schröder, *"The future of Eastern European capital markets".*

The above publications are available free of charge from:

European Investment Bank
Economic and Financial Studies
100, boulevard Konrad Adenauer
L-2950 Luxembourg
infoefs@eib.org (+352) 43 79-8895

www.eib.org/efs

EIB Papers 2000 – 2001

Volume 6, Number 2, 2001: The 2001 EIB Prize

First Prize; Erich Gundlach & Luder Wößmann, *"Better school for Europe"*.
Second Prize; Jan Fidrmuc, *"Democracy in transition economies: Grease or sand in the wheels of growth"*.
Third Prize; Gilles Duranton & Vassilis Monastiriotis, *"The evolution of the UK North-South divide: Should we mind the gap?"*.

Volume 6, Number 1, 2001: European economic growth: The impact of new technologies

Christopher Hurst & Kristian Uppenberg, *"Wonders will never cease: Prospects for a new economy in Europe"*.
Paul David, *"Productivity growth prospects and the new economy in historical perspective"*.
Patrick Vanhoudt & Luca Onorante, *"Measuring economic growth and the new economy"*.
Danny Quah, *"ICT clusters in development: Theory and evidence"*.
Hermann Hauser, *"Nothing ventured, nothing gained"*.
Andrew Gillespie, Ranald Richardson & James Cornford, *"Regional development and the new economy"*.
Luc Soete & Bas ter Weel, *"Computers and employment: The truth about e-skills"*.
Harald Gruber, *"The diffusion of information technology in Europe"*.

Volume 5, Number 2, 2000: Regional convergence in Europe: Theory and empirical evidence

Daniel Mouque, *"A survey of socio-economic disparities between the regions of the EU"*.
Angel de la Fuente, *"Convergence across countries and regions: Theory and empirics"*.
Jacques-François Thisse, *"Agglomeration and regional imbalance: Why? And is it bad?"*.
Philippe Martin, *"The role of public policy in the process of regional convergence"*.
Patrick Vanhoudt, Thomas Mathä & Bert Smid, *"How productive are capital investments in Europe?"*.

Volume 5, Number 1, 2000: Regional development in Europe: An assessment of policy strategies

Christopher Hurst, Jacques-François Thisse & Patrick Vanhoudt, *"What diagnosis for Europe's ailing regions?"*
Yannis Ioannides & George Petrakos, *"Regional disparities in Greece: The performance of Crete, Peloponnese & Thessaly"*.
Rodolfo Helg, Giovanni Peri & Gianfranco Viesti, *"Abruzzo & Sicily: Catching up and lagging behind"*.
Andrés Rodríguez-Pose, *"Economic convergence and regional development strategies in Spain: The case of Galicia & Navarre"*.
Margarethe Quehenberger, *"Lessons from Eastern Germany"*.
Bertrand Rossert, *"Contributing to regional development through project selection"*.

The above publications are available free of charge from:

European Investment Bank
Economic and Financial Studies

100, boulevard Konrad Adenauer
L-2950 Luxembourg
infoefs@eib.org (+352) 43 79-8895

www.eib.org/efs

EIB Papers 1998 – 1999

Volume 4, Number 2, 1999: The 1999 EIB Prize

First Prize; Philippe Martin, *"Are European regional policies delivering?"*.
Second Prize; Rudi Dornbusch & Francesco Giavazzi, *"Hard currency and sound credit: A financial agenda for Central Europe"*.
Third Prize; Jordi Gual, *"Deregulation, integration and market structure in European banking"*.
Special Topic Prize; Jens Verner Andersen & Jørn Henrik Rasmussen, *"Solving the unemployment puzzle in Europe"*.

Volume 4, Number 1, 1999: European banking after EMU

Christopher Hurst & Rien Wagenvoort, *"European banking after EMU"*.
Martin Brookes, *"The impact of EMU on portfolio management"*.
Graham Bishop, *"New capital market opportunities in Euroland"*.
Colin Mayer, *"European capital markets: Competition between systems"*.
Luigi Spaventa, *"Comments on the future of the EMU capital markets."*
Daniel Gros & Karel Lannoo, *"The structure of financial systems and macroeconomic instability"*.
Clas Wihlborg, *"Supervision of banks after EMU"*.
Christopher Hurst, Eric Perée & Mireille Fischbach, *"On the road to wonderland? Bank restructuring after EMU"*.
Rien Wagenvoort & Paul Schure, *"Who are Europe's efficient bankers?"*.
Philip Molyneux, *"Increasing concentration and competition in European banking: The end of anti-trust?"*
Jean Dermine, *"The case for a European-wide strategy"*.
Ingo Walter, *"Financial services strategies in the euro-zone"*.

Volume 3, Number 2, 1998: International financial institutions in the 21st century

Christopher Hurst & Eric Perée, *"Only a mid-life crisis? The future for IFIs in an integrated world"*.
Michael Klein, *"One hundred years after Bretton Woods: A future history of the World Bank Group"*.
Jannick Lindbaek, Guy Pfeffermann & Neil Gregory, *"The evolving role of multilateral development banks: History and prospects?"*.
Jean-François Rischard, *"Multilateral development banks and global public networks: Speculations for the next century"*.
Pasquale Scandizzo, *"The purpose, management and governance of IFIs: A case study of ambiguity"*.
Nicholas Stern & Hans Peter Lankes, *"Making the most of markets: The role of IFIs"*.
Joseph Stiglitz, *"IFIs and the provision of international public goods"*.
A discussion between Alfred Steinherr & Jacques Attali.

Volume 3, Number 1, 1998: Employment in Europe

Christopher Hurst, *"A discussion of the papers and an overview of the role of investment"*.
Jørgen Elmeskov, *"The unemployment problem in Europe: Lessons for implementing the OECD Jobs Strategy"*.
Gilles Saint-Paul, *"Does technical progress create unemployment?"*.
Michael Burda, *"The consequences of EU enlargement for Central and Eastern European labour markets"*.
José Viñals & Juan Jimeno, *"The impact of EMU on European unemployment"*.
Jacques Drèze, Alain Durré & Henri Sneessens, *"Investment stimulation, with the example of housing"*.
Ole Rummel & Thomas Schröder, *"The employment effects of investment projects"*.
Pascale Viala, *"Financing young and innovative enterprises in Europe: Supporting the venture capital industry"*.

The above publications are available free of charge from:

European Investment Bank
Economic and Financial Studies
100, boulevard Konrad Adenauer
L-2950 Luxembourg
infoefs@eib.org (+352) 4379-8895

www.eib.org/efs

EIB Papers 1996 – 1997

Volume 2, Number 2, 1997: The 1997 EIB Prize

First Prize; Bernadette Frédérick & Peter Vanden Houte, "*What is the optimum strategy for freezing exchange rate parities at the start of EMU*".
Second Prize; Jean Dermine, "*Eurobanking, a new world*".
Joint Third Prize; Daniel Gros, "*The dynamics of European integration*".
Joint Third Prize; Martin Kolmar, "*On the efficiency of national social security systems in the European Union*".
Awards for Authors under 30 for essays by Pablo de Andrés Alonso & Félix J. Lopez Iturriaga; Volker Köllmann & Andri Kopperschmidt; and Jérôme Vacher.

Volume 2, Number 1, 1997: Banking in an uncertain world

Edward Gardener & Philip Molyneux, "*The too-big-to-fail doctrine revisited*".
Agnès Belaisch, "*High real interest rates: A long-term risk?*".
Pier-Luigi Gilibert, "*Expected loss as a credit risk measure*".
Jacques Girard & Christopher Hurst, "*The capital structure of private infrastructure projects and the risk of default*".
Kristian Kjeldsen, "*Value-at-risk and capital adequacy: The challenge for financial regulation*".

Volume 1, Number 1, 1996: Special Issue on EMU

Daniel Gros, "*The steeple chase towards EMU*".
Ole Rummel, "*On the feasibility of EMU with diverse European economic transmission mechanisms*".
Bernhard Winkler, "*Is Maastricht a good contract?*".
Paul de Grauwe, "*How to fix conversion rates at the start of EMU*".
Luis Gonzalez-Pacheco & Alfred Steinherr, "*A proposal to stabilise the value of the ECU*".

The above publications are available free of charge from:

European Investment Bank
Economic and Financial Studies

100, boulevard Konrad Adenauer
L-2950 Luxembourg

infoefs@eib.org (+352) 4379-8895

www.eib.org/efs

ISSN 1830-3676

EIB PAPERS

European Investment Bank

Economic and Financial Studies

100, boulevard Konrad Adenauer
L-2950 Luxembourg
www.eib.org/efs

ISBN 92-861-0596-8