Historical Case Studies of Energy Technology Innovation

CASE STUDY 17: ASSESSMENT METRICS.

INPUT, OUTPUT & OUTCOME METRICS FOR ASSESSING ENERGY TECHNOLOGY INNOVATION

Charlie Wilson
Tyndall Centre for Climate Change Research, University of East Anglia

AUTHORS’ SUMMARY

A variety of quantitative methods have been used to analyse energy technology innovation. These range from efforts to identify and isolate the effects of particular policies or innovation processes such as research and development, to studies of the broader social benefits of energy technology innovation. Innovation metrics can thus describe innovation inputs, outputs or outcomes. This case study presents the commonly used metrics in each of these categories, together with the main issues associated with their application. The discussion then identifies some broader issues with the use of quantitative metrics to analyse energy technology innovation including the lack of integrated metrics, problems with quality control, and difficulties with cross-country and cross-technology comparisons.

If referencing this chapter, please cite:
1 QUANTITATIVE ANALYSIS OF ENERGY TECHNOLOGY INNOVATION

A variety of quantitative methods have been used to analyse energy technology innovation. These range from efforts to identify and isolate the effects of particular policies or innovation inputs such as R&D investments, to studies of the broader social benefits of energy technology innovation.

One area of interest has been the relative influence of technology-push and demand-pull factors on innovation outcomes including (as examples) technology costs or market penetration. Some researchers have focused on particular energy technologies including wind power (e.g., Klaassen et al., 2005) and solar PV (Watanabe et al., 2000). Findings have proven sensitive to assumptions about knowledge depreciation and about the time lags between policy actions and the responses of innovation system actors. Both these factors are inherently difficult to estimate empirically.

Other researchers have sought to generalise underlying innovation mechanisms in highly stylized reduced form models (e.g., Kouvaritakis et al., 2000; Miketa and Schrattenholzer, 2004). These include learning or experience curves which describe cost reductions as experience with a technology accumulates (Junginger et al., 2010; Wene, 2000). Resulting models of learning are of direct interest to policymakers concerned with the magnitude and outcomes of public investments or policy initiatives to subsidize the demand for energy technology innovations (Duke and Kammen, 1999). But again, this approach is not without its problems. Learning curve models have been critiqued on the basis of their atheoretical specification (e.g., assumptions of perfect substitutability of R&D and learning-by-doing as sources of technological knowledge) as well as on statistical grounds (e.g., problems of collinearity in model parameter estimation). Moreover, observed discontinuities in learning rates, perhaps resulting from omitted variables, as well as observed high variability of learning rates across technologies and market environments (Nemet, 2009b) limit the reliability and predictive power of models for specific policy interventions.

Shifting from the measurement and modelling of innovation inputs (e.g., R&D investments) and outputs (e.g., cost reductions), cost-benefit analyses have also been used to quantify the broader impacts and outcomes of energy technology innovation. The National Research Council in the US conducted one such study to measure the overall value of federal energy R&D programs (NRC, 2001) and a follow-up study which applied the same methodology prospectively to analyse expected social returns on investment under a range of assumptions including future carbon prices (NRC, 2007). The scope of these cost-benefit analysis is inherently broader than a modelling-based analysis of particular innovation processes or specific technology case studies. Yet insights remain contingent on various assumptions, including those related to counterfactual innovation outcomes in the event that federal R&D investments was not made.

Quantitative analysis is not only retrospective. Decision analytic techniques, commonly used by R&D managers, have been used to elicit the subjective judgment of experts familiar with particular technologies (Clemen and Kwit, 2001; Peerenboom et al., 1989). In the US, this is now recommended practice for helping guide and evaluate public energy technology innovation investments (NRC, 2007). Findings can be used to parameterize models or quantify roadmaps of future technology trajectories. Recent examples include studies of solar PV and carbon capture technologies (Baker et al., 2009).

If referencing this chapter, please cite:
2 INNOVATION METRICS

The most commonly used quantitative metrics of innovation relating to energy technology can be divided into three types. Input metrics describe financial, labour, and other inputs to the innovation system and innovation processes (like R&D). Output metrics describe defined products of the innovation system and innovation processes. These extend from the demonstrated technical feasibility of prototype innovations, the successful scaling-up of technological options either at the level of individual technologies (e.g., larger wind turbines) or of entire industries (e.g., solar PV market deployment), or reaching desired pre-specified technology performance targets (in terms of efficiency or costs). Outcome metrics describe broader energy sector or economy-wide impacts of the successful diffusion of innovations into the marketplace (e.g., carbon emission reductions, jobs created, social returns on innovation investments). These metric proxies of innovation inputs, outputs, and outcomes can be either intangible (e.g., knowledge stock, practical problems and ideas) or tangible and human (e.g., investments needed, scientists, laboratories) (Freeman and Soete, 2000). The following three sections provide an overview of commonly used innovation metrics and their attendant issues, drawing on Gallagher et al. (2006) and Sagar and Holdren (2002) unless otherwise noted.

2.1 Input Metrics

Table 1 summarises commonly used input metrics applied to energy technology innovation. Common issues associated with the practical application of each metric are also noted.
<table>
<thead>
<tr>
<th>Metric</th>
<th>Description</th>
<th>Issues</th>
</tr>
</thead>
<tbody>
<tr>
<td>R&amp;D Expenditure</td>
<td>• Public, private or total expenditure on R&amp;D (e.g., real $).</td>
<td>• Data on public R&amp;D expenditure typically available; time series data allows trends to be analysed.[1]</td>
</tr>
<tr>
<td></td>
<td>• Can also include expenditure on demonstration projects (i.e., R,D&amp;D)</td>
<td>• Private R&amp;D data, particularly in non-listed companies, can be difficult to obtain; if available, data are often highly aggregated so difficult to isolate expenditures specific to energy R&amp;D.</td>
</tr>
<tr>
<td></td>
<td>• Expressed either in absolute terms or as ‘R&amp;D intensities’ normalised for total output (GDP), total investment, etc.</td>
<td></td>
</tr>
<tr>
<td>Investment</td>
<td>• Public, private or total investment in innovation (e.g., real $).</td>
<td>• Similar issues to R&amp;D expenditures (see above) but broader categorisation of investment can avoid disaggregation issues (e.g., of aggregated investment figures in corporate accounts).</td>
</tr>
<tr>
<td></td>
<td>• Includes R&amp;D expenditure (see above) but also investments in demonstration, early deployment and diffusion.</td>
<td>• Investment data tend to under-represent later stage innovation activities (see text below) and conflate R&amp;D and demonstration stages.</td>
</tr>
<tr>
<td></td>
<td>• Narrow investment metrics can also be normalised (e.g., early stage venture capital as % of total venture capital).[2]</td>
<td>• Some databases compile private investments into specific technology sectors by investor type (e.g., venture capital, private equity); but investment targets are usually start-up companies rather than innovation per se, and databases mainly cover industrialised markets.[3]</td>
</tr>
<tr>
<td>Human Resources</td>
<td>• Number of scientists and engineers engaged in R&amp;D.</td>
<td>• Use as a proxy for ‘tacit’ knowledge embodied in labour input to innovation process.</td>
</tr>
<tr>
<td></td>
<td>• Can be weighted by education (e.g., highest degree attained) or type of training.</td>
<td>• Simplified metric does not account for quality or efficiency, nor differences in research infrastructure and capital equipment (so difficult to assess R&amp;D labour productivity).</td>
</tr>
<tr>
<td></td>
<td>• Expressed either in absolute terms, by sector, or per capita.</td>
<td>• As with investment metrics, difficult to isolate labour input specific to energy innovation, particularly in diversified companies.</td>
</tr>
</tbody>
</table>

[1]. For US R&D trends and analysis, see Nemet and Kammen, 2007; Margolis and Kammen, 1999 and for detailed data, see Gallagher and Anadon, 2009; for selected global data, see Sagar and van der Zwaan, 2006.

[2]. Innovation ratios measure the proportion of high risk private capital invested in early stage (or high tech) technologies as a proxy measure of the extent to which innovation systems support capital constrained entrepreneurial firms through the ‘valley of death’ between innovation and commercial diffusion. See Rin et al., 2006 for an example. In principle, innovation ratios could be used to measure different success factors of an innovation system: for example, total venture capital investments as a proportion of total investments would capture the successful leveraging of risk-taking private capital premised on broad market acceptance and diffusion. [3]. Examples include the European Venture Capital Association yearbooks (for the EU) and New Energy Finance (in the US).

If referencing this chapter, please cite:
2.2 Output Metrics

Table 2 summarises commonly used output metrics applied to energy technology innovation. Again, common issues associated with the practical application of each metric are also noted.

**Table 2. Commonly Used Metrics of Innovation Outputs.**

<table>
<thead>
<tr>
<th>Metric</th>
<th>Description</th>
<th>Issues</th>
</tr>
</thead>
<tbody>
<tr>
<td>Publications</td>
<td>• Number of peer-reviewed articles.</td>
<td>• Readily available information, but English language bias.</td>
</tr>
<tr>
<td></td>
<td>• Can be weighted by citations or impact factors.</td>
<td>• Difficult to define clear system boundaries for energy technology innovations: e.g., should articles on complementary technologies such as catalysts, materials, and control systems be included?</td>
</tr>
<tr>
<td></td>
<td>• Can also include other research publications (reports, books, evaluations, etc.)</td>
<td>• Useful metric for program evaluation if quality or impact-weighted.</td>
</tr>
<tr>
<td></td>
<td>• Workshops &amp; conferences</td>
<td></td>
</tr>
<tr>
<td>Patents</td>
<td>• Number of patents filed or granted.</td>
<td>• Similar issues to publications: readily available, but difficult to define system boundaries. Greatest validity if used at low level of aggregation.[1]</td>
</tr>
<tr>
<td></td>
<td>• Can be weighted by citations.</td>
<td>• Biased towards industrialised countries, and towards industrial sectors with higher propensity to patent.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Patents generally relate to R&amp;D rather than later stage innovation activities, and are not necessarily good predictors of successful commercialisation.</td>
</tr>
<tr>
<td>Technologies</td>
<td>• Number of technologies commercialised.</td>
<td>• Most visible measure of ultimate success of innovation process.</td>
</tr>
<tr>
<td></td>
<td>• Can be in terms of plants, production lines, product variants, process improvements, companies, turnover, etc.</td>
<td>• Difficult to define clear system boundaries for what constitutes a technology, particularly for complex multi-component systems (e.g., aircraft).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Fails to capture increases in learning and know-how for technologies based on tacit or non-codified knowledge (e.g., energy efficient building design).</td>
</tr>
<tr>
<td>Technology Characteristics</td>
<td>• Ratios of technical to service characteristics.</td>
<td>• Change in ratios of technical characteristics to performance or service characteristics indicate directionality and variety of innovations, as well as their proximity to the technological frontier.[2]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Technology specific: not possible to use in meta-analyses.</td>
</tr>
</tbody>
</table>

[1]. For an overview of issues relating to patents as an output metric, see Basberg, 1987; OECD, 2009. For studies using patent citations to assess environmental and energy technology innovation, see Nemet, 2009a; Popp, 2005. For a recent application of patent count data to assess the effectiveness of different types of environmental policy on renewable energy innovation, see Johnstone et al., 2010, and for a similar study but on energy efficient building technology, see Noailly, 2012.

[2]. For applications of this approach to helicopters, see Saviotti and Trickett, 1992; to aircraft, see Frenken and Leydesdorff, 2000; to refineries, see Nguyen et al., 2005.
2.3 Outcome Metrics

Table 3 summarises commonly used outcome metrics applied to energy technology innovation with frequent issues associated with the practical application of each metric also noted.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Description</th>
<th>Issues</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market Penetration</td>
<td>• Number or capacity of technologies sold or used.</td>
<td>• Generally available data.</td>
</tr>
<tr>
<td></td>
<td>• Market share (or penetration rate) is alternative measure normalising for size of market or economy, hence describing the extent of substitution into - or capture of - a specific market.</td>
<td>• Extensive empirical literature on diffusion dynamics, market penetration, and substitution effects.[1]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Bias towards ‘successful’ innovations that have diffused widely. Best suited to retrospective historical analysis.</td>
</tr>
<tr>
<td>Learning Rates</td>
<td>• Rate of cost reduction of a technology.</td>
<td>• Learning rates emphasize the commercialisation phase of a technology but substantive cost reductions may also occur in the earlier innovation stages.</td>
</tr>
<tr>
<td></td>
<td>• Conventionally measured as the % reduction in unit cost per doubling of cumulative production.</td>
<td>• Production and cost (or price) data generally available, and mechanisms for learning effects have been extensively researched.[2]</td>
</tr>
<tr>
<td></td>
<td>• Can be measured for production plants, organisations or technologies.</td>
<td>• Learning rates can vary widely between variants of the same technology and between plants producing the same technology; learning rates are also sensitive to timing and data fitting issues.[3]</td>
</tr>
<tr>
<td>Economic Benefits</td>
<td>• Cost benefit analysis.</td>
<td>• (Investment) costs generally easier to quantify than benefits which can include environmental and energy security externalities, knowledge stocks and spillovers, option values of technology portfolios, as well as more conventional net employment, tax, and consumer surplus benefits.</td>
</tr>
<tr>
<td></td>
<td>• Can be aggregated as net social benefit, or left disaggregated in multiple ‘accounts’ (e.g., labour, environment, knowledge, etc.).</td>
<td>• Cost benefit analyses widely used as core component of program evaluations.[4]</td>
</tr>
<tr>
<td>Energy / Emissions Intensity</td>
<td>• Primary energy (GJ), electricity (GWh), or emissions (e.g., tCO₂ tSO₂) per unit of GDP.</td>
<td>• Readily available data; meaningful as part of time series trend.</td>
</tr>
<tr>
<td></td>
<td>• Normalisation can also be more tightly defined, e.g., per sector, or per power plant.</td>
<td>• Aggregate impact of innovation only, and potentially confounded by structural changes to economic activity, non-price induced changes, and inter-fuel substitution.</td>
</tr>
<tr>
<td>Project / Program Evaluation</td>
<td>• Size and number of programs in terms of employees, turnover, investment, outputs, etc.</td>
<td>• Difficult to assess quality, so need to complement with case study or survey research (see text below).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Similar issues with tacit knowledge as for technologies (see above under outputs).</td>
</tr>
</tbody>
</table>
3 DISCUSSION

3.1 Lack of integrated metrics for energy technology innovation

With respect to high-tech companies, Hagedoorn and Cloodt (2003) discuss the absence of a generally accepted metric of innovation despite the long tradition of using single metrics to study innovation, including R&D, patents, and new product announcements. In response they propose a composite innovation metric to describe companies’ overall ‘innovative performance’

It is similarly salient that no single metric dominates assessments of energy technology innovation. As shown in Tables 1, 2, and 3 above, this can be explained partly by issues of data availability, reliability, and definition which mean each metric has its shortcomings. Defining what constitutes an energy technology is also far from straightforward. Energy assessments typically focus on the supply-side only, but many activities in the industrial, agricultural, and residential and commercial buildings sectors are also fundamentally connected to energy technology innovation.

A more general observation, however, is that no single metric captures the overall success of the energy technology innovation endeavour, linking inputs to outputs and broader outcomes. Learning rates are one exception as they are a hybrid metric that links inputs (investments in production) to outputs (cost per unit production), but they are typically confined to the diffusion phase of technologies that have succeeded in the market. As such, they omit the uncertain and complex inter-relationships among innovation processes during the early innovation stages which generate technological novelty.

Other fields of innovation - e.g., agriculture, pharmaceuticals, aerospace - are similarly characterised by a mixed, staged involvement of both public and private resources, and a complex set of time-lagged processes linking innovation inputs to outputs. In these fields, innovation assessments commonly use econometric or regression techniques to evaluate either net social returns on R&D, or productivity gains. In agriculture, for example, such techniques have long been used to examine the effectiveness of particular structures and roles in the innovation process (Evenson et al., 1979), and to account for the often long time lags between innovation inputs and outputs (Alston et al., 2000).

Certainly, the use of econometric methods to evaluate innovation is not without its issues (Griliches, 1987). However, innovation assessments in the energy field have not similarly converged on a broadly

If referencing this chapter, please cite:
accepted set of analytical approaches. Many, if not all of the data and methodological limitations raised in Table 1 above equally hamper the assessment of innovation in other fields. Neither is energy unique in the capital intensiveness of innovation outcomes (cf. pharmaceuticals) and the shape of returns in high volume, low margin regulated markets (cf. agriculture).

With respect to the evaluation of research, “problems arise because not everyone perceives the same purpose” (p14, Alston et al., 2000). Is this a particular problem with assessments of energy technology innovation? Agricultural innovation may emphasize productivity outcomes; pharmaceutical innovation may emphasize health outcomes. In contrast, there are many, varied, and oft-competing objectives for energy technology innovation. Does successful innovation reduce emissions in absolute terms? reduce energy or carbon intensity? reduce unit costs? reduce the cost of final service provision? increase option values in diverse portfolios? improve security, reliability, flexibility? ensure backstops are available? increase accessibility? And so on. The most appropriate dependent variable of an energy technology innovation assessment may also be technology-specific. This creates problems for meta-analyses and dampens convergence on a set of common assessment techniques and indicators.

### 3.2 Under-representation of demonstration and deployment activities

Although the overall innovation process is characterised by multiple feedbacks between stages, a general distinction can be made between research, development, demonstration, market formation and diffusion activities. Early deployment during the market formation stage is a precursor (but overlapping) stage to diffusion in which barriers to widespread commercialisation are targeted (e.g., through ‘demand-pull’ type policies) (Sagar and van der Zwaan, 2006). R&D expenditures, investment data, as well as patent and publication data, tend to relate to the earlier research and development phases. For the financial metrics, this is driven by data availability associated with government’s more central role in these higher risk activities with public good characteristics (e.g., non-appropriable knowledge generation). Outcome metrics inherently relate to the diffusion phase. This leaves the demonstration and early deployment phases under-represented by innovation metrics (Sagar and Gallagher, 2004). Both phases are critical to avoiding the so-called “valley of death” between innovation and successful diffusion if technologies still carry preventively high risks for private sector capital. Metrics of demonstration activities can be extracted from demonstration project databases for specific technologies (Harborne and Hendry, 2009). However, available statistics (e.g., IEA (2011) data on public R,D&D expenditures) are extremely limited with respect to technology demonstration activities. Assessing deployment activities is also difficult to do outside of specific program evaluations (see Table 1 above). For examples on clean coal and buildings, see Sagar and Gallagher, 2004.

### 3.3 Cross-country comparison of innovation metrics

Cross-country comparisons of technological capabilities and innovation processes are complicated by the absence of standardised data. For a detailed discussion of issues, see Archibugi and Coco, 2005. As noted in Table 1 above, both R&D expenditure and investment data vary as to the system boundaries or definition of what is included (e.g., sources and targets) and to the coverage of the innovation lifecycle (Sagar and Holdren, 2002). Comparisons of innovation metrics across countries also need to control for confounding factors. For example, comparisons of human resource metrics need to account for labour cost differentials, and comparisons of peer-reviewed papers may be biased by different cultural attitudes towards the importance of publications, and by access to English language journals (Archibugi and Coco, 2005). The same is broadly true of patents as the institutional nature of intellectual property.
rights varies widely between countries both in terms of propensity to patent, as well as the quality and accessibility of the patent system (Basberg, 1987). Cross-country comparisons of energy and emissions intensity are confounded by numerous factors including the structural weighting of economic activity towards certain types of production (e.g., Alcantara and Duro, 2004), relative energy prices (Miketa, 2001), climatic factors, urban form, resource endowments, and so on. Comparison of within-country time trends are informative through decomposition analyses to isolate the influence of technological change (IEA, 2004).

3.4 Quality control of quantitative metrics

As with any quantitative metric, quality control is important to ensure results are meaningful. Some examples are included in Tables 1, 2, and 3 above. Numbers of publications or patents can be weighted by subsequent citations or some other measure of impact. Changes in the ratios between certain technical and performance characteristics offer another indication of qualitative changes in a technology’s design (Saviotti and Metcalfe, 1984).

In general, however, complementary qualitative assessments of innovation processes are needed. Methods include interviews, process-tracing, surveys of innovation process participants, and detailed case studies of innovation programs (as in this book). Important qualities of energy technology innovations to assess include contribution to portfolio diversity and innovation system management. Creating and maintaining diverse knowledge stocks can enhance system resilience and preserve option value with respect to long-term climate change mitigation goals (Sanden and Azar, 2005). Management capabilities, as well as sufficient infrastructure and collaborative relationships to support knowledge generation, are important structural features of innovation systems and influence their successful functioning. (For an overview of innovation systems at a sectoral level, see Malerba, 2002; for applications to renewable energy, see Hekkert and Negro, 2009, and to biofuels, see Jacobsson, 2008).

4 CONCLUSION

The growing body of quantitative analysis and modelling using input, output and outcome metrics of energy technology innovation has greatly improved our understanding of innovation processes. Systematic renderings of the whole energy technology innovation system are, however, rare. Wilson et al. (2012) compare a range of different metrics of both inputs, outputs and outcomes into global energy technology innovation (see also Figure 1). They find directed innovation efforts to clearly privilege energy supply technologies even though end-use technologies dominate innovation system outputs and desired innovation outcomes in terms of climate change mitigation or social rates of return. This is but one potential application of a compound set of metrics to assess the overall functioning of the energy technology innovation system.

However, as a final cautionary note, it is important to recognise the inherent limitations of quantitative metrics which emphasise particular elements of the energy technology innovation system and risk offering overly mechanistic explanations of causality. Quantitative assessments are also limited by data availability, as is evidenced by the relative dearth of empirical work on energy end-use technologies compared to energy supply technologies, and on developing country transitions compared to developed country innovation systems. Descriptively as well as analytically rich case studies of energy technology innovation can help to counter some of these limitations.

If referencing this chapter, please cite:


5 FURTHER READING
For a good overview of issues associated with assessing innovation in energy technologies, see: Gallagher et al., 2006. For an interesting recent application of one widely used metric, patent counts, see: Johnstone et al., 2010. For a comparison of innovation input and output metrics in the context of the energy system, see: Wilson et al., 2012.

6 REFERENCES

If referencing this chapter, please cite:


*If referencing this chapter, please cite:*