

Evaluation and innovation: Challenging the single narrative

Robert Picciotto

The advent of innovation as a privileged policy priority—in New Zealand as in the rest of the world—illustrates the power of the single narrative as a covert advocacy tool for the promotion of an ideology. In the wake of the 2008 financial crisis, a shifting conception of innovation has been used to reinforce the hold of private interests over public policy. This article exposes the artificial simplicity of this mental model through reference to economic theory and historical experience. Next, it unveils the dilemmas posed by intellectual property regimes and highlights the critical role that country institutional and policy contexts play in shaping the impact of innovation on society. Finally, it recommends rigorous assessments of innovation claims and illustrates the risks that the pervasive single narrative has imposed on evaluation practice.

Technique smothers the ideas that put its rule in question and filters out for public discussion only those ideas that are in substantial accord with the values created by a technical civilization. (Merton, 1964, pp. vii–viii)

On November 5, 2011, Eleanor Chelimsky delivered a keynote address at the American Evaluation Association Conference that stressed the insidious threat to society posed by the *single narrative*. Rather than barefaced lies, such a narrative uses misleading simplifications, unproven assertions, and suppression of relevant evidence. It is the careful customising of reality at the service of vested power (Chelimsky, 2012).

The reduction of a complex social reality to a simple, compelling story makes adoption of a predetermined policy agenda irresistible. In such circumstances, it is incumbent on evaluators to unearth the values and motivations that propel the single narrative, to interrogate its unspoken assumptions and to tell it like it is. Thus, evaluation is a much-needed antidote to the single narrative. In what follows I will illustrate how the spin associated with the innovation concept calls for rectification through principled evaluation lest it be allowed to distort policy priorities or undermine the quality of evaluation practice.

The article is structured as follows. First, it defines innovation. Secondly, it describes the single narrative currently associated with the concept. Thirdly, it contrasts its rudimentary logic with the nuanced lessons of development theory and economic history. Fourthly, it probes the policy dimensions of innovation and the need for objectivity in the evaluation of innovation. Finally, it highlights the risks that the single narrative has created for the evaluation discipline.

What is innovation?

As is often the case in the gestation of policy fads (Ellerman, 2007) the meaning of innovation has evolved. Imported from the business sector the term has come to the center of the stage in public policy debates. No longer simply something new or different the concept is now all-embracing as a marker of creativity and social progress.

It evokes positive, transformative, and dynamic change—a self-evident imperative of public policy.

The Organization for Economic Cooperation and Development (OECD) guideline for tracking technological innovations (“the Oslo Manual”) initially related innovation to the introduction of new products and technology processes (Organization for Economic Cooperation and Development, 1992, 1997). Gradually, other enabling conditions reaching well beyond the research and development domain were incorporated in the definition, such as knowledge diffusion, the interactions among innovation actors, interfirm linkages, organisational and marketing innovations, and so forth.

Next the distinctive, incremental, “low tech” nature of innovation in the services sector and the role of new methods were added so that by 2005 the meaning of innovation had expanded to incorporate “the implementation of a new or significantly improved product (good or service), a new process, a new marketing method, or a new organisational method in business practices, workplace organization or external relations” (OECD and Eurostat, 2005, p. 9).

Finally, in 2010, the Innovation Strategy of the OECD further broadened the scope of innovation policy to include non-technological, social, and public-sector innovations. Even *creative imitation*, the use or adaptation of an innovation created elsewhere, is now classified as an innovation. Whereas the management literature had originally defined innovation as the specific function of business entrepreneurship, the term is now an implicit call to arms for decision makers to adopt a wide range of market-oriented policies that are deemed to promote growth through entrepreneurship and technological change.

Innovation as single narrative

Innovation has become a buzzword in the conference circuit. When Bill Gates pressed the G20 to involve the private sector in

development activities he declared that innovation “fundamentally shifts the trajectory of development” and represents “the most powerful force for change in the world” (Bill & Melinda Gates Foundation, 2011, para. 20).

Innovation thus understood has become a litmus test of policy relevance for think tanks. For example, the Center for Global Development has embarked on policy research initiatives focused on investigating incentives to catalyse technology creation, testing, and adoption, and policy options designed to reorient research towards development directions. Equally, social innovation featured prominently at an annual meeting of the World Economic Forum (Schwab Foundation for Social Entrepreneurship, 2014).

Nor have international policy makers resisted the lure of private-sector-led innovation. Specifically, the 2010 OECD Innovation Strategy statement underscored the need for innovation as a way to emerge from the Great Recession triggered by the 2008 financial crisis—along with public expenditure cuts. It made no mention of the need for systemic reform within the banking sector where the crisis originated. Rather, it laid stress on empowering entrepreneurs, unleashing private investment, protecting intellectual property, and developing high-speed broadband networks (OECD, 2010).

By ignoring much-needed systemic reforms in the financial sector and emphasising the need to unshackle private enterprise a single neoliberal narrative has made the concept almost indistinguishable from sound public policy in developed and developing countries alike. Once largely confined to the business firm, innovation has “become the order of the day for the economy as a whole”—as judiciously presaged by Merton (1964, p. vi).

Largely absent from official policy statements is the notion that innovation should be guided by normative objectives grounded in ethical principles and the public interest. The EU-funded civil-society

network that relates socially and environmental responsible innovation to engaged publics, responsive institutions, and ethically acceptable outcomes has been a lone voice (ProGRess, 2014).

Lessons from theory and history

Innovation was not always so central to economic policy thinking. Neoclassical theory relied on capital output coefficients and exogenously determined savings rates to model national economies, with technological change treated as a residual. However, in the mid-1980s, Paul Romer (1986) pioneered the new growth theory which placed knowledge and technology squarely at the core of the development process, thus displacing factor endowments and comparative advantage as key antecedents of national economic dynamism and growth (Romer, 1994).

By now scientific and technological progress has become widely accepted as a major driving force in dynamic economies. It is widely perceived as a capital investment—the indispensable catalyst of economic growth. Unlike land and capital, knowledge is not subject to diminishing returns. In the right policy environment, it is a public good characterised by vast cumulative and spillover effects.

In sum, innovation is an important ingredient of economic policy. Industries cannot survive in a capitalist economy unless they incessantly create better or more effective processes and products. Yet social-research findings about the contribution of new technologies to total factor productivity and output growth vary widely (Kretschmer, Cardona & Strobel, 2012).

The literature has pointed to a *productivity paradox*: a discrepancy between measures of investment in information technology and measures of output at the national level.¹ For example, since the early

¹ Hence, Robert Solow's (1987) quip: "You can see the computer age everywhere but in the productivity statistics" (p. 36).

1970s there has been a major slowdown in economic growth—and in employment creation—in all western economies, even as they spawned new information and communications technologies and as computers became ubiquitous. “Techno optimists” who think information-technology innovations will redefine advanced economies and enhance their productivity have been proven wrong by Robert Gordon (2016).

The reality is that innovation generates costs as well as benefits. It is unruly and induces disequilibrium in the market place. It is fundamentally disruptive since it affects market structures and renders existing processes and products obsolete. It is rife with unexpected and unintended effects. By its very nature innovation makes the existing capital stock and available skills redundant. For example, the advent of the automobile required major investments in paved roads and displaced the railways which had earlier displaced canals and waterways as the primary transport mode (Kogan, Papanikolaou & Stoffman, 2012).

Similarly, at the firm level, innovation involves risks as well as rewards depending on the organisational culture. Complex and iterative feedback loops between research, design, production, and marketing create a wide range of threats that should be carefully weighed and managed. Unbridled pursuit of profits associated with business success in a competitive environment has a dark side: technological change embedded in new capital benefits some workers, but it hurts existing firms and can make workers redundant. From a social-equity standpoint, the fascination with results, the privatisation of knowledge, and the ever-expanding role of technology in society have facilitated the concentration of capital and power in ways that have facilitated the advent of a single narrative implicitly designed to sustain current social arrangements.

The enabling environment for innovation

Legislative frameworks have generated abnormally generous pecuniary incentives for private individuals and firms under the shield of patent protection. Yet intellectual-property legislation can have serious deleterious consequences on economic productivity by promoting monopoly power, restricting public access to the benefits of innovation, and discouraging further invention (Azzarelli, 2009).

Technological progress involves the improvement, adaptation, and combination of diverse technologies into new configurations. Patents inevitably raise the cost of such processes: they inhibit knowledge sharing, suppress competition, and discourage further improvement of patented inventions (Chakravarthi, 1990). Hence it is not surprising that there is little robust evidence that patents foster innovation and economic growth (Hu & Png, 2009), or that patented innovation leads to permanent increases in economic growth (Ulku, 2004).

In truth, patents are not strictly necessary as an incentive to innovate. Nor is it obvious that inventors and innovators should be rewarded by granting them monopoly rights that extend over many decades, even though such rights implicitly reduce the public benefits which would have been yielded by open-source, broad-based diffusion of knowledge. Complex technical and social problems are best solved in an open knowledge environment. Personal wealth is not the main motivation that drives researchers and educators involved in the creation and diffusion of knowledge. Professional pride and altruism, more than profit, drive their work.

In the corporate sector the threat of being surpassed by competitors and reputational considerations are often the decisive forces behind innovation (Rothbard, 1970). Confirming this thesis, a study comparing systematically patented and unpatented innovations over the period 1977–2004 across industrial sectors disclosed that only 10% of

technologically significant new products available for sale or licensing were patented (Fontana, Nuvolari, Shimizu & Vezzulli, 2013).

Nor are patents the only market-oriented policy instruments geared to the promotion of innovation. For example, in pursuit of innovation, incentives instruments such as prizes, challenges, and contests have become part of the standard policy toolkit of governments and private foundations.

Vested interests have tirelessly promoted the notion that intellectual property rights (IPRs) are good for everyone and should be protected with vigour even where valuable indigenous knowledge has been displaced, for example, in health, where modern pharmaceuticals have elbowed out alternative remedies. Yet the World Health Organization (WHO) has highlighted the crucial role played by medicinal herbs in the healthcare systems of many developing countries and recommended that governments should give priority to utilising traditional medicine in national drug policies and regulations (WHO, 2002).

At the other end of the policy spectrum, civil-society activists have argued that IPRs inevitably cripple local industry and harm developing countries' populations. To strike the right balance a United Kingdom Independent Commission was established to review the evidence. It concluded that stringent IPRs standards should not be pressed on developing countries without a serious and objective assessment of distinctive development impacts (Commission on Intellectual Property Rights, 2002).

The Commission concluded that the optimal protection threshold (where social benefits exceed the social costs) varies widely by product and sector. In countries where there is substantial research capacity a latent supply of innovative capacity in the private sector waits to be unleashed by IPRs. But in most developing countries local innovation systems are weak, especially in the private sector, so that the

dynamic benefits from intellectual property protection are modest and uncertain.

The country context matters

A prerequisite is the development of adequate local scientific and technological capacities that allow countries to develop their own process of technological innovation and absorb technologies developed elsewhere. The appropriate policy framework for each country should be decided based on what is best in its unique circumstances. In response to strong representations by developing governments and civil-society activists, *special and differential treatment* was adopted as an integral part of the agreement on Trade-related Aspects of Intellectual Property Rights (TRIPS) that took effect on January 1, 1995.

TRIPS intended to provide sufficient policy space to countries to calibrate their legislation in line with their stage of development. But under pressure from industrial lobbies, highly restrictive western patent regimes have been imposed on developing countries as part of bilateral trade agreements (TRIPS Plus). This is neither ethical nor consistent with liberal economic tenets.

Within distinctive factor endowments the rate of development “catch up” varies depending on the capacity of individual countries to access and make effective use of the technological knowhow of advanced nations. Distinct policy packages are needed for individual countries to benefit from technological and institutional innovations generated in developed countries. This equitable, context-dependant policy message contrasts sharply with the absolute certainties of the single narrative.

The global context matters too

Against this background, innovation strategies have been formed around two sets of critical uncertainties (Rockefeller Foundation & Global Business Network, 2010). The first refers to the evolving

state of the global system. At one extreme of this continuum, sound and predictable governance, open trade and investment regimes, and connectivity to problem-solving networks will favour the spread of innovation. At the other end, weak international governance, rising protectionism, civil strife, and fragility throttle innovation.

The second continuum refers to the capacity of individual countries, communities, and organisations to innovate and cope with change. Resilience of institutions is the capability to adapt and adjust. It is correlated with the level of education and trust in a society, its tolerance of diversity and diversity, and the freedom it affords individuals to pursue their aspirations. The extent of innovation and the capacity to make effective use of it hinges on how these two sets of driving forces are combined (see Table 1).

Table 1. Forces influencing extent of innovation

	Low adaptive capacity	High adaptive capacity
Open, secure global environment	High risks with high rewards for capacity building innovations	Low risks and high rewards to technological and institutional innovation
Closed, insecure global environment	High risks low rewards: dangerous innovations spread in an insecure world	Low rewards and low risks for local innovations

Implications for public policy

The single narrative is on firm grounds when it ascribes high priority to research and development to keep the innovation engine of a diverse and complex modern economy humming. For example, the New Zealand government is significantly increasing its investment in science and innovation through the Budget 2016 Innovative New Zealand package that will provide a further \$410.5 million in operating funding over the next 4 years. It gives pride of place to government involvement in research and development.

This goes against the market magic myth of the single narrative which leaves private entrepreneurs to their own devices to drive innovation. In fact, it is the state not the private sector that funded space exploration and spawned three-fourths of all new medicines. It is not young entrepreneurs tinkering in their garages free of government interference that laid the foundations for the information technology revolution. Neither the internet, nor smartphones, nor touchscreen displays, nor powerful search engines would have materialised without government vision and funding.

The stark reality is that venture capitalists cannot afford to take the time or the risks involved in the exploration of new scientific frontiers. Only the state can do so and create new markets in the process. The main contribution of private finance to innovation has been to help adapt, disseminate, and market new technologies that would not have come into being without government sponsorship.

In that sense the private-innovation myth has been a convenient advocacy tool for vested interests. The same flawed narrative is now being put to work to denigrate government contributions to social innovation and to rely on processes that simulate market mechanisms to achieve social objectives. Yet the state has a key role to play in piloting and incubating social innovation, and all social interventions whether sponsored by private, public, or voluntary organisation actors—or by multisector coalitions—should comply with the same evaluation standards to protect the public interest.

Evaluation of innovation

As decision makers' interest in innovation increased, the demand for evaluation of innovation policies and programs grew. The Manchester Institute for Innovation Research confirms that the demand for evaluation in science and innovation policy has continued to grow alongside demands on policy makers to deliver a broad range of

economic and societal goals against financial constraints (University of Manchester, 2016).

The growth in demand is likely to cross national borders and come into New Zealand given (a) the higher policy profile of innovation in development policy; (b) the diverse forms that innovation takes depending on the country and sector context (OECD, 2012); and (c) the diminishing returns of policy research relying on cross country correlations. For example, the European Commission Competition and Innovation Program commissioned 49 evaluations and related reports during 2007–2012 (European Commission, 2014).

Evaluation sponsors often use prestigious events, the pronouncements of eminent persons, and glossy reports by fee-dependant consultants in lieu of objective and rigorous assessments. The field is dominated by “pseudo-evaluations”, that is, evaluations which that are carried out for political or public-relations reasons (Stufflebeam & Webster, 1980). Market-led initiatives have often been treated as if competition is an iron-clad guarantee that social value will be generated. For example, in the prizes and challenges domain, a recent report by Doblin, Deloitte Consulting provided assessment guidance for prize designers that did not take account of the hidden costs borne by losers or compare the prizes and challenges option to other ways of promoting discoveries (Goldhammer, Mitchell, Parker, Anderson & Joshi, 2014).

Evaluators should resist decision makers’ capture of evaluation processes and help dispel the lure of the single narrative. They should live up to the ethical strictures of their craft and protect the integrity and independence of their assessments of innovation policies and programs. They should educate their clients about the generic uncertainties of innovation, stress the need to weigh its costs alongside its benefits and help their clients steer free of the illusion that only the private sector generates innovation.

Evaluating innovation at the country level

At the country level, innovation league tables have emerged with metrics shaped by the single narrative. For example, the Global Innovation Index (Dutta, Lanvin & Wunsch-Vincent, 2014) includes data on 143 countries. It adopts an inclusive notion of innovation based on the Oslo manual and the OECD strategy. Its conceptual framework combines innovation input measures and output measures and establishes a correlation between the two sets of indicators, but it does not establish that together they generate equitable and sustainable development.

Input measures include (a) institutions and country policy dimensions—good governance indicators, private sector friendliness, security of property rights, labour market flexibility, ease of starting a business, and so forth; (b) human capital and research—expenditures on education; educational quality indicators, tertiary education enrolment, number of researchers and R&D expenditures; (c) infrastructure—ICT, general infrastructure and ecological sustainability indicators; (d) market sophistication—availability of credit, investor protection, size of equity market, trade and competition openness; (e) business sophistication—availability of knowledge workers, innovation linkages, and knowledge absorption indicators.

Output measures cover all the variables traditionally viewed as fruits of innovation and invention and include (a) knowledge creation—patent applications, utility model applications, articles published in peer reviewed journals, and so forth; (b) knowledge impact—new businesses, software spending, International Organization for Standardization (ISO) quality certificates, and so forth; and (c) knowledge diffusion—royalty and licence fees, high-tech exports, foreign direct investment outflows, and so forth. Output measures also cover creativity measures: (a) intangible assets—trademark applications, business model creation, and so forth; (b) creative

goods and services—cultural services, feature films, and so forth; and (c) online creativity—average monthly Wikipedia edits, YouTube video uploads.

These are plausible indicators for tracking innovation at country level, but it is far from obvious that policy makers should systematically privilege them over social criteria or that they should ignore the state of knowledge regarding intellectual-property policies. Innovation is not a panacea. Its effects vary widely, depending on the economic and social context.

Equitable access to knowledge, increased productivity, and technological development should be facilitated. Policy should remain subservient to democratic processes and reflect the universal policy objectives legitimised by the international community. The enabling environment, including intellectual property systems, should be assessed at arm's length taking account of the special needs of individual countries.

Innovation in evaluation

Evaluation has not been immune to the seductive power of the single narrative. And just as in society, innovation in evaluation has proved to be a mixed blessing. Three major shifts in evaluation practice have swept over the world over the past decade: (a) experimental methods; (b) big data; and (c) social-impact assessment in the private sector. All three have been touted as major innovations in the evaluation domain, and yet all three have generated costs and risks—along with potential benefits.

Experimentalism

The new evaluation experimentalists are idealists embarked on a crusade consistent with the single narrative. They equate evaluation with social research and social research with laboratory research.

They subscribe to the heady vision famously enunciated by the MIT Poverty Action Lab's co-founder Esther Duflo: "Just as randomized evaluations revolutionized medicine in the 20th century they have the potential to revolutionize social policy during the 21st" (Duflo & Kremer, 2003). They have been remarkably successful in convincing influential policy makers that experimental methods are the gold standard in evaluation.

Consequently, disproportionate resources have been invested worldwide in experimental impact evaluations that are not always appropriate. They are redundant when no other plausible explanation for the results observed is available. They are often not feasible, for example, it is not possible to randomise the location of an infrastructure project. Nor are experimental methods useable when no untreated target group can be identified, for example, for full-coverage policy interventions.

Experimental designs are not the *only* scientific basis of ascertaining causation or attribution. Biology, geology, astronomy, epidemiology, the forensic sciences, and so on all testify to the proposition that causation can be established without randomised control trials. Similarly, investigatory techniques, contestability protocols, and rules of evidence are widely perceived to be adequate to reach evaluative judgments.

Experimental methods on their own only tackle two of the evaluative questions that are within the remit of evaluation: (a) do the intervention effects meet its goals (efficacy)? and (b) do they result from the intervention (attribution)? They fail to deal with issues of social relevance, efficiency, and sustainability that are often at least as important. Nor do experimental evaluations provide estimates of the distinct contributions of partners in the collective enterprises that design and implement social programs. Yet not all partners contribute equally to the success or failure of policy interventions

which confirms that experimental methods do not contribute to the accountability mandate of evaluation.

Experimental evaluations are very hard to do well and they require large populations of subjects to meet statistical validity tests. Only if the treatment group and the control group and the process that affects each of the two groups are strictly identical (except in terms of cause and effect) can inferences be established with confidence. Validity may be jeopardised by latent and unobserved causal factors that are not built into the process when constructing the treatment and control groups or when the intervention design is flexible and adaptable to changed circumstances. Last but not least, depriving members of the control group of a useful treatment can be unethical, discriminatory, or even illegal.

From a social-learning perspective, frequent claims that the randomised trial procedures which made their mark in the health and agricultural research sectors hold the key to evaluation rigor in the economic and social domain are invalid. Experimental evaluation allows formative conclusions to be drawn only for simple interventions implemented in stable environments. Since other evaluation methods are available to answer the wide range of questions that policy makers as well as citizens wish answered the single narrative that has equated experimentalism with a gold standard has been costly to society.

Big data

Big-data analytics have also generated huge and, so far, unmet expectations. Big data examines large amounts of data to uncover hidden patterns, correlations, and other insights. Given modern computing it can provide just-in-time management information. But the illusion that the data collected speaks for itself can be costly: correlation is not causation.

The risks associated with limited contextual comprehension are considerable. Thus, big-data approaches need to be complemented with full understanding of the context and valid insights about the issues being scrutinised. There is no substitute for theory-based approaches that help experienced evaluators identify meaningful correlations. Addressing the wrong questions can induce strategic errors and high costs.

The lion's share of big data is privately owned. Access to data for use in evaluation is restricted and regulated. Retrieving and analysing active driven data from social media is challenging since it is unstructured and rarely fact based. Passive driven data (i.e., data about "what people do"), and machine-to-machine, algorithm-driven data are equally hard to collect and assess. Tools to capture big data for assessment of social interventions are still untested.

A major shortcoming of big data for evaluation is the lack of transparency around how it has been generated so that its reliability and validity can be ascertained. Are agencies that own the data likely to disclose all the relevant data of interest, and can they vouch for its reliability? Do commercial interests inhibit timely and transparent disclosure? Privacy issues need special attention too since the users of services and devices generating data are often unaware that they are doing so. These are early days: ethical protocols about the use of big data have yet to be defined and agreed, let alone implemented.

Social-impact assessment

The single narrative about innovation does not acknowledge that the ideal conditions under which invisible hand guides private actors towards the public good is absent in the real world. Markets are volatile and do not automatically generate an optimal equilibrium. They frequently overshoot and crash. Absent public action, they induce growing inequality. Absent regulation, they favour monopolies and

deplete natural resources. Absent social norms, they exploit labour and produce pollution.

Yet, the myth which holds that innovation and competition automatically produce optimum social outcomes persists. This is ultimately why private sponsors of social interventions often view traditional evaluation as redundant and intrusive. Instead, they have created new assessment methods principally designed to serve intervention design and monitoring. To be sure, social-impact assessment is intended to focus on all the direct and indirect effects that an intervention may have on the way of life, the culture, the community, the political system, the environment, the health and wellbeing, and the rights of people. But it does not advocate comprehensive field verification of results, that is, it fails to embrace independent, no-holds-barred, ex post reviews of what ultimately happened.

Social-impact investors put great store in the use of indicators for impact measurement and management. Theoretically anything can be quantified and measured. But attempting to capture all relevant effects of a social intervention through indicators inevitably leads to unmanageable complexity. Auditing of the relevance and accuracy of data used for performance tracking does not constitute evaluation. Merely checking on predetermined indicators at project closure falls well short of ascertaining the intended and unintended effects of the intervention on the ground, draw pertinent lessons of experience, or provide a reliable economic and social justification for the intervention.

Social-impact practitioners in the private sector tend to select success indicators that privilege intervention goals defined at the outset. This may generate illusions of success. Efficacy—the extent to which relevant goals have been achieved—does not establish that the results observed were due to the intervention rather than other factors. Nor does it establish whether the results were achieved economically

relative to alternative ways of achieving the same goals.

Here again it emerges that what is presented as an evaluation innovation may induce erroneous conclusions about the economic and social justification of a social intervention. What if the goals were overambitious, trivial, or irrelevant? What about positive and negative unintended and secondary effects not envisaged when the indicators were selected? What if the selected indicators did not reflect the aspirations and concerns of stakeholders other than those of program sponsors and managers?

Finally, managers and staff who oversee social-impact assessments in the private sector are routinely subject to pressure to deliver good news. They may be tempted to select evidence that supports their preconceptions. Their interests and concerns are not always aligned with those of beneficiaries and they do not enjoy the arm's length relationship needed to deliver uncompromising assessment reports. The pliant, fee-dependent external consultants and auditors that managers retain are not independent either.

Conclusions

The advent of innovation as a privileged goal of public policy illustrates the power of the “single narrative” as a covert advocacy tool at the service of a distinctive neoliberal ideology. In the wake of the 2008 financial crisis a shifting conception of innovation has been put to work to reinforce the hold of vested interests on the policy discourse in New Zealand as elsewhere. This has generated a mental model that largely ignores the disruptive effects of innovation and that takes for granted that tough intellectual property legislation is a necessary ingredient of a dynamic knowledge economy.

These propositions fly in the face of economic theory and historical experience. Hence, evaluators should be skeptical about the broad generalisations of the single narrative and embrace a sober

and independent approach to the costs and benefits of innovation. Innovation yields huge rewards, but it also involves commensurate risks. This highlights the need to carefully weigh alternative policy options and to balance the benefits of innovation against the costs.

Evidently the innovation process has often been socially disruptive and a major cause of wealth concentration and monopoly power. To be sure, rigid government regulation and obsolete public-sector institutions have often hobbled entrepreneurship. Creation and diffusion of new technologies have contributed to development of emerging market economies. But the effect on overall productivity in developed countries has yet to be proven and the social impact of robotisation looms as a major policy concern.

Given their monopoly features, intellectual-property regimes have dampened competition and slowed down further innovation at the expense of the broader public. They have not always served the public interest. In the design of effective policy regimes as well as in the standards used to evaluate development interventions the country context matters as does the global policy framework. In sum, innovation is an important policy vector but it is not a panacea and it has a dark side.

The ideology underlying the single narrative regarding innovation has also affected evaluation. It has propelled new approaches to evaluation that have favoured experimental methods focused on evaluation questions aligned with the politics of government austerity: “whether the intervention works”, instead of “what explains performance shortfalls”, and “how the intervention can be improved in the public interest”. It has privileged experimental methods over all others despite their well-known limitations. In any event the very concept of a gold standard in is a fool’s errand: attribution of effects to an intervention is only one of the questions that policy makers want answered and it can be tackled in a variety of ways (Stern et al., 2012).

Revealingly, experimental methods are not being advocated for market-led interventions by their private-sector sponsors. Instead, the methods they use to assess social impact rely mostly on the vivid articulation of good intentions and on auditors' verification of goal achievement data unmoored from theory. Equally, big-data analytics have been used mostly as marketing instruments rather than for the design of public policies, while corporate social responsibility principles have remained voluntary and not subject to independent scrutiny.

In sum, recent prominent innovations in evaluation practice reflect the steady encroachment of the evaluation market by auditors and management consultants. In parallel, the organisational management literature, utilisation-focused evaluation textbooks, as well as the influential American Evaluation Association Guiding Principles urge evaluators to meet legitimate commissioners' needs whenever it is feasible and appropriate to do so. But giving pride of place to the utilisation of evaluation results has undermined evaluation independence in contexts where evaluation is treated as a private good instead of a public good. One unintended consequence has been a neglect of the social dimensions of policies and programs and the neglect of participatory and democratic evaluation methods.

Yet, protecting the public interest is the core remit of evaluation. To this end, evaluators should adapt their approaches and tools to the complexity of the operating context. Just as innovation in society should involve adaptation of approaches and methods tested in other contexts evaluators should make effective use of all the methods and tools available within the evaluation mainstream and, to achieve professional status, demonstrate that they have the knowledge, skills, and dispositions to use them judiciously.

References

- Azzarelli, S. (2009). Against intellectual monopoly. *Syracuse Science & Technology Law Reporter*, 21, 130–145.
- Bill & Melinda Gates Foundation. (2011). *A report by Bill Gates to G20 leaders*. Retrieved from <http://www.gatesfoundation.org/What-We-Do/Global-Policy/G20-Report>
- Chakravarthi, R. (1990). *Recolonization: GATT, the Uruguay round and the third world*. Penang, Malaysia: Third World Network.
- Chelimsky, E. (2012). Valuing, evaluation methods, and the politicization of the evaluation process. *New Directions for Evaluation*, (133), 77–83.
- Commission on Intellectual Property Rights. (2002). *Integrating intellectual property rights and development policy*. Retrieved from http://www.iprcommission.org/papers/text/final_report/reportwebfinal.htm
- Duflo, E., & Kremer, M. (2003, 15-16 July). *Use of randomization in the evaluation of development effectiveness*. Paper presented at the World Bank Operations Evaluation Department (OED) Conference on Evaluation and Development Effectiveness, Washington DC. http://faculty.las.illinois.edu/akresh/GBL298/Duflo-Kremer_Randomization.pdf
- Dutta, S., Lanvin, B., & Wunsch-Vincent, S. (Eds.). (2014). *The global innovation index 2014: The human factor in innovation*. Retrieved from <http://www.globalinnovationindex.org/content.aspx?page=gii-home>
- Ellerman, D. (2007). Best practice in education and training: Hype or hope. *NORRAG News*, 39, 16–18. Retrieved from <http://www.norrag.org/es/publications/boletin-norrag.html>
- European Commission. (2014, January 1). *Implementation, evaluation and performance reports* [online archive]. Retrieved from http://ec.europa.eu/cip/documents/implementation-reports/index_en.htm
- Fontana, R., Nuvolari, A., Shimizu, H., & Vezzulli, A. (2013). Reassessing patent propensity: Evidence from a dataset of R&D awards, 1977–2004. *Research Policy*, 42(10), 1780–1792. <https://doi.org/10.1016/j.respol.2012.05.014>

- Goldhammer, J., Mitchell, K., Parker, A., Anderson, B. & Joshi, S. (2014). *The craft of incentive prize design: Lessons from the public sector*. Westlake, TX: Deloitte University Press.
- Gordon, R.J. (2016). *The rise and fall of American growth*. Princeton, NJ: Princeton University Press. <https://doi.org/10.1515/9781400873302>
- Hu, A.G., & Png, I.P. (2009). *Patent rights and economic growth: Evidence from cross-country panels of manufacturing industries*. Retrieved from http://www.wipo.int/edocs/mdocs/mdocs/en/wipo_ip_econ_ge_5_10/wipo_ip_econ_ge_5_10_ref_huandpng.pdf
- Kogan, L., Papanikolaou, D. & Stoffman, N. (2012) *Technical innovation: Winners and losers*. Retrieved from <http://economics.mit.edu/files/8381>
- Kretschmer, T., Cardona, M., & Strobel, T. (2012). *ICT and productivity: A roadmap for empirical research*. University of Munich, Leibniz Institute for Economic Research. Retrieved from [http://www.lse.ac.uk/management/research/initiatives/lseTech/nef/events/ICT-and-Productivity-\(Kretschmer-Cardona-Strobel\).pdf](http://www.lse.ac.uk/management/research/initiatives/lseTech/nef/events/ICT-and-Productivity-(Kretschmer-Cardona-Strobel).pdf)
- Merton, R. K. (1964). Foreword. In J.Ellul, *The technological society*. New York, NY: Vintage Books.
- Organization for Economic Cooperation and Development. (1992). *Oslo manual: Proposed guidelines for collecting and interpreting technological innovation data* (1st ed.). Paris, France: OECD.
- Organization for Economic Cooperation and Development. (1997). *Oslo manual: Proposed guidelines for collecting and interpreting technological innovation data* (2nd ed.). Retrieved from <http://www.oecd.org/science/inno/2367580.pdf>
- Organization for Economic Cooperation and Development. (2010). *The OECD innovation strategy: Getting a head start on tomorrow*. Retrieved from <http://www.oecd.org/sti/45302349.pdf>
- Organization for Economic Cooperation and Development, & Eurostat. (2005). *Oslo manual: Proposed guidelines for collecting and interpreting technological innovation data* (3rd ed.). <https://doi.org/10.1787/9789264013100-en>

- Organization for Economic Cooperation and Development Centre. (2010, June 16). Economy: developing countries set to account for nearly 60% of world GDP by 2030, according to new estimates. *Perspectives on global development*. Retrieved from <http://www.oecd.org/dev/pgd/>
- Organization for Economic Cooperation and Development. (2012). *Innovation for development: A discussion of the issues and an overview of work of the OECD directorate for science, technology and industry*. Retrieved from <http://www.oecd.org/innovation/inno/50586251.pdf>
- ProGRess. (2014). *Promoting global responsible research and social and scientific innovation*. Retrieved from <http://www.progressproject.eu/news/1093-progress-brochure-d6-2-is-out/>
- Rockefeller Foundation & Global Business Network. (2010). *Scenarios for the future of technology and international development*. Retrieved from <http://www.rockefellerfoundation.org/uploads/files/bba493f7-cc97-4da3-add6-3deb007cc719.pdf>
- Romer, P.M. (1986). Increasing returns and long-run growth. *The Journal of Political Economy*, 95(5), 1002–1037. <https://doi.org/10.1086/261420>
- Romer, P.M. (1994). The origins of endogenous growth. *The Journal of Economic Perspectives*, 8(1), 3–22. <https://doi.org/10.1257/jep.8.1.3>
- Rothbard, M. (1970). *Power and market: Government and the economy*. Kansas City, KS: Sheed Andrews and Memeel.
- Schwab Foundation for Social Entrepreneurship (Producer). (2014). *The view from Davos: Social innovation in 2014* [video]. Retrieved from <http://www.schwabfound.org/content/view-davos-social-innovation-2014>
- Solow, R.M. (1987, July 12). We'd better watch out. *New York Times*, p. 36.
- Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R., & Betani, B. (2012). *Broadening the range of designs and methods for international development* (Department for International Development Working Paper 38). Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/67427/design-method-impact-eval.pdf

- Stufflebeam, D. L., & Webster, W. J. (1980). An analysis of alternative approaches to evaluation. *Educational Evaluation and Policy Analysis*. 2(3), 5–19. <https://doi.org/10.3102/01623737002003005>
- Ulku, H. (2004). *Innovation and economic growth: An empirical analysis* (IMF Research Department Working Paper WP/04/185). Retrieved from <https://www.imf.org/external/pubs/ft/wp/2004/wp04185.pdf>
- University of Manchester. (2016). *Evaluation of science and innovation policies*. Manchester Institute of Innovation Research. Retrieved from http://www.research.mbs.ac.uk/innovation/Portals/0/Documents/MIOIR_Evaluation_of_Science_and_Innovation_Policies.pdf
- World Health Organization (2002). *Global strategy for traditional medicine*. (Institute of Science in Society (ISIS) Report, 1 August 2002). WHO.

The author

Robert Picciotto is senior independent evaluation adviser, New Zealand Ministry of Foreign Affairs and Trade. He served as Director General, Independent Evaluation Group, at the World Bank from 1992 to 2002.

Email: r.picciotto@jeunot.net