SMITH SCHOOL OF ENTERPRISE AND THE ENVIRONMENT

UNIVERSITY OF OXFORD

Smith School of Enterprise and the Environment, University of Oxford

UBS –Smith School Essay Competition for Enterprise and the Environment

25 April 2014

Melting Innovation Policy Gridlock with F.I.R.E.S

By Dane Rook

Technological innovation is commonly offered up as both a 'solution' and counterargument on sustainability matters. Particularly with respect to climate change concerns, the supposed ingenuity of humankind has been put forth as a rationalisation for both radical policy that departs from convention, as well as adherence to the status quo. While it would be a misrepresentation to claim that proponents of the salvational nature of innovation strictly are concerned with substance over process (i.e. they think innovation will simply "happen" rather than being concerned with "how" it happens), it seems legitimate to say that not nearly enough scrutiny has been dealt out toward the particular contexts and mechanisms by which innovation is supposed to fruitfully transpire. One of the foundational bits of theory underlying this privileging of innovation outcomes over innovation processes has been the so-called "Porter Hypothesis" (Porter & Van der Linde 1995). The Porter Hypothesis (hereafter "PH") asserts that increasingly-stringent regulation and policy is advantageous for progressive innovation.

This essay takes the PH as a point of departure and argues that its characteristically atemporal nature (that is, the PH does not give enough consideration to multiple timescales) is more problematic than has previously been examined. The main inspiration for the alternative presented here comes from best-practice in modern entrepreneurship and venture capital – the genuine "gold standards" of systematic innovation. In seeking policy architectures that are conducive to innovation, it only seems reasonable that policymakers should look to those that are the demonstrable leaders in reproducible innovation. A particular hallmark of entrepreneurial success on path-breaking innovation is the 'iterated design' model in which innovators are inspired to "fail quickly and in a contained way".

Obviously this mentality seems curious for such a 'mission critical' issue as climate change policy. But in light of the extreme complexity and ambiguity that surrounds not only the physical science of climate change, but also the economic, social, and cultural impact from and responses to it, the entrepreneurial ethos of 'controlled, rapid experimentation' makes sense. So much time, cost, and political goodwill have gone into attempting to craft innovation policies that are "fail-safe", "robust", and sweepingly large-scale that achieving consensus on them seems almost essential. Yet, that consensus more often than not only entails and promotes a 'least common denominator' among policy makers and regulated parties – it does not often drive or construct settings that are amenable to the *best* innovation.

I therefore suggest that the best recipes for successful innovation policies should be realistic and pragmatic about failure; they should accept that small and quick failures that serve as platforms for iterative learning are better than 'grand plans' that are centrally-architected, have huge scales and long timeframes, require enormous political and social buy-in, and, if they fail, huge costs of not succeeding. As an alternative I propose here a new policy framework: the fast-interval regulatory exploration (FIRE). This mechanism takes inspiration from Porter's Hypothesis and shores it up by wedding it with proven best-practice from entrepreneurship. It requires new policy changes at frequent and set intervals that are designed and implemented on a rotational basis. Before describing FIREs in more detail, however, I turn to a more critical discussion of why they are so needed.

At its core, the PH contends that regulated parties innovate more when policy and regulation are stringent and progressively tightening than when they are lax. This makes intuitive sense, as research and development costs may be quickly re-cooped and innovation laggards punished. But the PH is overly-idyllic in some ways. Firstly, it simply proposes that policy will become ever more stringent. Presumably the relative standards of stringency are often pre-scheduled, as with planned emissions-reductions timetables. Yet, this long-term planning approach may be overly-straight-jacketing for nonlinear innovation. Such "smooth" paths to change as often appear in time-tabled policy schedules don't often take enough account for the fact that innovation is a cumulative process and occurs in 'bursts' after long periods of trial and experimentation that are characterised by series of small failures that lead to eventual breakthroughs that are hard to predict in terms of timing and size.

A more reasonable and innovation-sensitive policy construct should respect this nature of the innovation process rather than merely blithely assuming that, with sufficient pressure, innovation will just "happen" in a forecastable way. Second, the PH is prone to issues of mimetism or free-riding. That is, if those entities that put in sizeable efforts to innovate cannot enjoy sufficient competitive advantages from their research costs because their competitors can readily mimic or replicate their innovations rather than pay the full costs of innovation themselves, then the original innovators become motivated to "innovate less" than they might otherwise do. A more fair and practical system might have better alignment with game theoretic concepts that allow laggards to be punished appropriately and also allow leaders to fully enjoy (or at least more fully than otherwise) the fruits of their R&D efforts.

Thirdly, the PH entails certain scale-circumscriptions and boundedness with respect to enforceability and commitment. It is not so good at smaller scales when regulated parties can simply pick up and leave for another jurisdiction (i.e. there is competitor/production mobility). Moreover, the PH becomes unwieldy at very large (especially global) scales as it requires increasingly enormous collective buy-in and policing. As international law demonstrated time and again (and the experiences under the UNFCCC and Kyoto Protocol bear out) this collective commitment and policing can be fiendishly difficult to achieve in practice. A replacement paradigm would do well to be functional across multiple scales simultaneously.

It seems plausible that FIREs might serve as an improved innovation framework relative to the Porter Hypothesis. The foundational basis for FIREs is that regulated parties pre-commit to a timetable of frequent policy and regulatory changes without complete definition of what those changes will be. It is now commonplace in many sustainability-oriented policy conventions and agreements which aim to promote innovation that a long planning and negotiation horizon is needed. The required time is spent "scenario building" and identifying contingencies and flaws in the proposed policy charge, and then haggling over language and targets (again, outcome-orientated but often procedurally insensitive).

FIREs intentionally reverses this approach. It urges regulated entities to pre-commit to a calendar of regular changes and amendments to policy that are spaced close together (thus "fast interval"). The standard, long-horizon approach demands a humungous number of scenarios to be forecasted, and, given the complexity involved, regularly ends up as myopic for missing some critical possibilities (i.e. it presumes near omniscience when in reality opacity and uncertainty are actually dominant on sustainability issues). Rather than aspiring to tame complexity, FIREs instead seeks to benefit from it. FIREs makes, in essence, the participating entities "anti-fragile" (see the book by Taleb (2013)) – it allows them to benefit from small shocks and setbacks in a collectively adaptive way. It thus gains from what the standard version of PH hopes to accomplish: certainty that some policy change will happen. Where it improves on PH, however, is in its short periodicity and adaptability, as well as its collectively-steered nature that promotes fairness and representation (to be discussed shortly).

The basic setup for FIREs is as follows. Parties looking to enter into a sustainability-improving policy arrangement that relies substantially on innovating technologically agree to a calendar for changes and sketch some general objectives. These general objectives are non-binding, but the calendar is binding. The key behind the calendar is that the important 'dates' are spaced at close intervals (e.g. one quarter or one month). There is no need for the parties to agree what policy changes specifically will come into force on those dates, only that *some* change will happen, and that the change will be range-bound in size: it will be above a certain significance threshold, but not so large as to expose the parties to excessive costs of failure.

The precedent for this setting of a minimum band for incremental policy already appears in many components of the UNFCCC, wherein parties undertaking projects throughout CDM or JI provisions must demonstrate positive and meaningful impact. The upper band, however, is new. It constrains the maximal costs of failure and promotes rapid uptake and creativity. To allow parties to adhere to the fast intervals of policy change under FIREs, it would be ideal for the calculations of these maximum and minimum bounds to be formulaic and predetermined so that they would be minimally onerous and maximally fair and transparent to all parties.

This combination of range-bounding impacts and fast-paced change is common in entrepreneurial and venture capital circles because it allows complexity to be navigated through learning iteratively; that is, it encourages participants to innovate by taking many “little bets”. A good many of these may not work, but because the costs of “failing small” are themselves small, the net value of learning outweighs these minor stumbles and leads, with much higher likelihood, to the “big breakthroughs” that are really the result of a process of small, cumulative changes. This process is ultimately evolutionary. In nature, evolution proceeds by small and progressive changes in genetic composition, not huge “innovations” (these innovations of large genetic scale are practically never viable). Natural selection permits “designs” that have worked in the past to make small “tweaks” – some of which fail, but some of which generate resounding successes. Under FIREs this principle is captured. If one round of change does not succeed, it can be quickly altered in the next (or other, later) rounds. This promotes safe experimentation and learning and allows adaptive change.

The equitable nature of FIREs stems from its rotational nature. Changes to policy under FIREs should be driven (and generally funded) by the regulated participating entities. Give the short timescales between changes, it thus might become burdensome for every entity to have to propose a policy change at each date on the policy calendar. But, if there are two dozen entities and changes occur each month, then any given entity would only have to propose a change once every two years. The participant-driven nature is crucial. It may require some outside authority to enforce, but the rotational nature of FIREs is what mitigates the free-rider impacts in innovation.

An example implementation of FIREs might proceed as follows. Two dozen US states (roughly half of the total) agree to introduce policy changes to their collective sustainable energy regulations (ignore for now whether or not doing so would step on the toes of the associated federal regulators – this is merely a stylised example) every month. On rotation, a new state proposes two changes each month. It and the other states must vote on which of the two changes to implement. They can implement two, but they must implement at least one unless one of the voting states vetoes. If this happens then the vetoing state loses its next turn in the rotation to propose. But, the originally-proposing state is incentivised to propose sound changes, as it can be 'punished' by the proposals in later rounds.

This setup does not resolve Arrow's (1950, 1951) "voting paradox" but it does make the system more fair by giving asymmetric power to one state (the proposer) in any one round, but symmetric power to all over all rounds. Drawing from ideas on game theory (see Myerson 2001) this setup should quickly achieve "fair" solutions. In particular, the pace of changes can be governed in advance by the lower threshold for the 'size' of the proposed changes. Likewise, much of the riskiness of failures can be controlled through the upper bound on the size of the change. In many setups, these threshold parameters could be set by an external regulator or governing body (in the US example, perhaps the environmental protection agency – EPA). But, importantly, much of the content of innovation is steered by the regulated parties. Free-riding is reduced because the specific *content* of innovation leaders' proposed changes can be used to punish free-riders and a balance is freely created.

Also of note, the FIREs framework is flexible in scale. It could conceivable apply to companies, regions, municipalities, nations, or even people as individuals or families. It clearly should work best when entities are similarly-sized and subject to the same laws or jurisdictional rules. This flexibility may allow parties keen to try it to not have to await "higher levels" of approval (e.g. counties in the UK wanting to enter into it could voluntarily or even bindingly do so without awaiting national-scale approval; and not all counties in the UK would need to partake for benefits to arise for it for participants).

The exploratory nature of FIREs which is premised on learning may feel uncomfortable for some, but in essence it captures the same mechanisms that market-based approaches to sustainability (like carbon markets) have promised but so often failed to deliver. In essence, those other approaches are typically over-architected; they have too little inbuilt flexibility and are prone to "crashes" because of unforeseen (i.e. "unplanned") contingencies, which are the norm rather than the exception in today's complex political economy. Forcing incremental change at a bounded but secure rate makes FIREs adaptable and anti-fragile because it does not need to "see around every corner" to succeed. It is nimble and makes itself perpetually effective and never redundant because changes are always being forced. If participants are at all sensible, those changes proposed should always be the most pressing (highest-priority) at the time because the proposer will only have one shot in that cycle to be in the leadership role. This juxtaposition of relatively short ("fast") intervals between changes but comparatively long ("slow") intervals between any one entity's ability to exercise authority should make the overall system more responsible because it effectively makes authority a relatively scarce resource. One could even "tweak" voting weights to make the system more or less fair as necessary.

In closing, I wish to suggest why the time may now by ripe for implementing some variant of the FIREs system for innovation policy on sustainability. It seems that the vast majority of plans for sustainability policy nowadays invoke a need or role for ongoing technological innovation. A prevailing philosophy has been that the innovations themselves should be relatively unrestricted, but that the outcomes from innovation should not be. This logic is relatively convoluted and inverted as the gridlock over such policies typically stalls over the target outcomes (and timing thereof) simply because the feasible set of outcomes *depends upon* the plausible set of innovations that can take place in the anointed timeframe. While the desire to not restrict innovation is admirable, the whole edifice is out of touch with entrepreneurial best-practice on promoting innovation from being shackled to a set of insufficiently flexible plans that make the embedded costs of failure larger than they need to be. Very few would now contest the assertion that, across many social, environmental, and economic domains, action rather than inaction is necessary. By pre-committing to a set schedule of adaptive/open-ended actions, change can be guaranteed in a way that is "contextually fit" and derived from competitive trial-and-error (learning) rather than relying on grand plans and compromising negotiations. I hope that FIREs or some similar framework might therefore fulfil more of the immense promise of the Porter Hypothesis than has been achieved to date.