MNP Report 550033001/2006

Environmental policy and modelling in evolutionary economics

A. Faber, K. Frenken[1], A.M. Idenburg[2] (editors)

1. Urban and Regional Research Centre Utrecht (URU), Utrecht University
2. DHV Consulting BV.

Contact:
Albert Faber
Netherlands Environmental Assessment Agency
albert.faber@mnp.nl

Netherlands Environmental Assessment Agency (MNP), P.O. Box 303, 3720 AH Bilthoven, the Netherlands;

Tel: +31-30-274 274 5; Fax: +31-30-274 4479; www.mnp.nl/en

Rapport in het kort

Milieubeleid en evolutionair economisch modelleren

Evolutionaire economie is een systeemaanpak, waarin activiteiten op microniveau in verband worden gelegd met effecten op macroniveau. In een door MNP georganiseerde workshop (Amsterdam, mei 2006) is door een groep Europese experts gediscussieerd over de mogelijkheden en onmogelijkheden voor de ontwikkeling van evolutionair economisch modelleren. Zulke tools zijn vooral handig voor de beoordeling van milieubeleid, waarin immers effecten op macroniveau worden nagestreefd, maar waarvoor de beleidsinstrumenten ingrijpen op activiteiten van actoren op een meer microniveau.

Evolutionaire economie is een theoretische methode die dynamische en meervoudige evenwichten centraal stelt, in plaats van een statisch, enkelvoudig evenwicht. Evolutionair economisch modelleren betekent vaak het maken van verkenningen met behulp van verklarende simulaties, meer dan het maken van (voorspellende) projecties naar de toekomst. Evolutionair-economische modellen kunnen generiek of specifiek van aard zijn. De laatste categorie modellen hebben een nadruk op een specifieke technologie, groep actoren (populatie) of mechanisme. In dit rapport wordt een aantal essays gepresenteerd, die al deze benaderingen verkennen. Elk essay neemt een geheel eigen uitgangspunt en perspectief en verkent van daaruit de mogelijkheden voor evolutionair modelleren.

Trefwoorden: evolutionaire economie, modelleren, milieubeleid

Preface

In 2005 the Netherlands Environmental Assessment Agency (MNP) published the report ‘Survival of the Greenest, evolutionary economics as an inspiration for energy and transition policies’ (see reference list in chapter 1). This report resulted from a joined research project with the Free University Amsterdam. The objective of that study was to investigate what insights evolutionary economics theory can provide for the design of an environmental policy that aims to stimulate innovations and a transition to a long term environmentally sustainable economy. The report offers an overview of the main literature on evolutionary economics and derives some core concepts from this theory. These concepts were subsequently used to assess and understand processes of change in economic structure, technological development and institutions, as well as to formulate guidelines for the role of government and the design of public policies.

The underlying report results from a complementary study to explore the possibilities of evolutionary economic *modelling* for policy design. To this end, four international experts on evolutionary economics and (environmental) innovation have been asked to write an essay: Paul Windrum (Maastricht University, Manchester Metropolitan University), Bart Verspagen (Eindhoven Technical University), Maïder Saint-Jean (Bordeaux University) and Wander Jager (Groningen University). These essays formed the background material for a one-day workshop held on 18 May 2006 in Amsterdam (see annex 1 and 2 to this report). These four essays are included in this MNP-report with an additional essay by Malte Schwoon (Hamburg University). Chapter 1 functions as a short introduction to the five essays and also provides a summary of the discussions during the workshop in Amsterdam held on 18 May 2006. It aims to formulate an advice on how to use evolutionary economic modelling for environmental research and policy.

We would like to thank the authors as well as the workshop participants for the lessons learned and for the clear insights they presented on evolutionary modelling for the environment.

This report, the workshop presentations and an account of the workshop are available at the MNP-website, www.mnp.nl.

The editors

Contents

Summary

In order to assess sustainability transitions, it is useful to make use of evolutionary economic modelling tools, since they include (more than other approaches) heterogeneity and bounded rationality of agents, which leads to a diversity in strategies and behaviour, as well as to path-dependency of incremental technological development. Multiple equilibria often act as attractors for diverse patterns of technological development. This approach offers the opportunity to study bottom-up emergence of system changes. Given these characteristics, evolutionary models can provide new understandings and new tools for the analysis and design of environmental innovation policy. In particular, following recent governmental initiatives in the Netherlands organized under the heading of *transition management*, evolutionary models may be helpful in assessing and developing policy strategies to trigger technological transitions from one technological system to a future technological system.

Modelling tools can be useful for policy making, as they enable to study and evaluate different policy measures ex ante. Evolutionary modelling often involves exploratory *simulations*, rather than making (predictive) projections of the future. A variety of models may be developed:

1. *Comprehensive evolutionary models* take into account all relevant aspects of technological development, but methodologically they have a high level of aggregation, and they may suffer from the many parameters needed;
2. *Specific evolutionary models* focus on understanding a specific aspect of technological innovation, thus greatly reducing the complexity of the modelling task at hand. The focus may be on *a)* a specific technology; *b)* a specific mechanism (e.g. a policy measure), or *c)* a specific population of agents; *d)* in a *modular approach* different parts of a technological system are modelled in separate sub-models, with clear interfaces among them.

In the essays in this report, these different approaches are all presented.

Paul Windrum's essay provides an overview of elements that could be included in a generic model based on a review of theory, models and case studies. The core thesis of the essay is that policy makers, firms and consumers are on the one hand the agents of path dependencies, supporting and maintaining current technological paradigms, while on the other hand they also are the agents of change. The replacement of old technology paradigms by new paradigms occurs when there are fundamental changes in the expectations, preferences, competences and policies of these agents. Windrum develops a co-evolutionary framework that captures the dynamics of technological successions and suggests conceptual building blocks for a comprehensive evolutionary model of technological transitions.

The essay of *Bart Verspagen* is a position paper on the possibilities of informing the (economic and environmental) policy debate by using quantitative evolutionary models. Verspagen argues that an evolutionary worldview implies that the existing quantitative modelling tools used for policy analysis are problematic. However, a number of main elements for an evolutionary-economic analysis can be distinguished and incorporated into quantitative models. This approach is followed by an energy transitions analysis with a study of micro co-generation technologies. This approach clearly follows a focus on a specific technology, strategy *a*.

Maïder Saint-Jean follows strategy *b*, by focusing on the impact of environmental standards on product and process innovation. Her essay aims at examining the impacts of particular policy options (emission standards and procurement policy) on clean trajectories by firms in an industry. This examination uses an evolutionary model of industrial dynamics to explore such impacts in the long range on the market structure and the innovation output of firms. Simulations with the model show that the rise in demand for environmental requirements, generated by tighter environmental standards, has different impacts according to the nature and timing of the standards. Regarding procurement policy, simulations show the existence of a critical mass of clients that value environmental characteristics of the product able to impulse a dynamics of innovation that induces a change in paradigm of the whole firms in the industry.

Wander Jager follows strategy *c*, discussing the behaviour of consumers as one of the relevant actors in the economy. He adopts the approach of multi-agent simulation, allowing for heterogeneity and social interaction. Jager develops a perspective on how behaviour could be formalised in environmental models, focussing on micro-level decision making of populations of agents, social interaction between agents, and intrinsic adaptation of decision-making and behaviour by the agents. The essay aims at describing a venue to increase this practical applicability of agent-based modelling, in particular with respect to testing policy measures in complex man-environment systems.

Finally, *Malte Schwoon* adopts a modular approach in his study on the technological system of fuel cells. Four types of agents are involved in modelling potential transition paths: car producers, consumers, fuel suppliers and the government. An evolutionary approach is adopted in order to deal with crucial infrastructure issues. The essay introduces an agent-based simulation model that puts together an existing producer competition model with a consumer model of adoption decisions in a modular way. It is applied to investigate the impacts of tax and infrastructure policies. Results suggest that consumers and individual producers are asymmetrically affected by taxes and public infrastructure investments, so that different types of resistance towards the policies can be anticipated. Moreover, there is evidence that large car producers might benefit from co-operation with fuel suppliers to generate a faster build-up of hydrogen infrastructure.

In May 2006 a workshop was organized in Amsterdam, where all the essays were presented and where a discussion on the use of evolutionary modelling took place. During the workshop consensus emerged that specific evolutionary modelling is more attractive than the first type of wide range theorizing, since it offers more perspective on practical modelling application, without having to make an unacceptably large number of assumptions beforehand. This aligns with theorizing on the intermediate range, in line with the characteristics of the problems at stake. Problems in the field of transitions to sustainability can be coined as intermediate range problems, allowing for a specific evolutionary modelling exercise.

Presently, the development of comprehensive, integrated evolutionary models for the study of technological systems is still regarded to be a bridge too far. The specific strategies to reduce the complexity of the modelling task at hand, then, apply in different contexts, dependent on the policy issues at stake.

Finally, it is essential to make sure that the empirical base in terms of mechanism and data is sound, in order to enhance calibration and validation of the model. More than in other economic modelling approaches, this step has often been neglected and it remains one of the biggest challenges for evolutionary modelling to enhance its empirical base.

1 Introduction: environmental policy and modelling in evolutionary economics

K. Frenken[1], A. Faber[2] and A.M. Idenburg[3]

1. Urban & Regional research centre Utrecht (URU), Faculty of Geosciences, Utrecht University, Utrecht (NL)
2. Netherlands Enivronmental Assessment Agency (MNP), Bilthoven (NL)
   Contact: albert.faber@mnp.nl
3. DHV Consulting BV, Amersfoort (NL)

This chapter functions as a short introduction to the five essays and also provides a summary of the discussions during the workshop in Amsterdam held on 18 May 2006. It does not aim to provide a literal account of these discussions. The objective is rather to formulate an advice on how to use evolutionary economic modelling for environmental research and policy. Section 1.1 will first explain shortly why environmental agencies should be interested in exploring the possibilities of evolutionary economic modelling. Section 1.2 summarizes the discussions on existing evolutionary economic models and the lessons they provide for new modelling activities. Finally, section 1.3 draws some conclusions.

1.1 Why evolutionary economic modelling

Where social science is increasingly used for sustainability assessment, a proper tool to include social issues is often lacking. Presently, this gap is often filled with a neoclassical economic approach. Neoclassical economics generally offers consistency as well as the availability of sound data, which makes it a strong basis for economic analysis. However, the neoclassical approach is not very fit to grasp system changes, radical innovations or socio-economic transitions, because it is not capable to deal with diversity of behaviour and imperfect rationality (see Van den Bergh et al. (2005) for a discussion on evolutionary versus neoclassical economics). An evolutionary economic approach could overcome some of these limitations of neoclassical models currently dominant (even if evolutionary models have limits of their own). In particular, evolutionary economics has a more realistic stance towards modelling economic agents. In contrast to neoclassical models, evolutionary models will have to take a number of defining characteristics into account (Nelson and Winter, 1982; chapter Verspagen in this report):

1. the bounded rationality of agents (firms, consumers, governments);
2. the heterogeneity of agents (firms, consumers, governments) and contexts (geographical, institutional);
3. the path-dependent nature of technological development creating irreversibility in the socio-economic system, and

4. multiple equilibria, which act as attractors but are rarely reached in time (rather than an instantaneously reached, single equilibrium as in neoclassical economic models).

Given these characteristics, evolutionary models can provide new understandings and new tools for the analysis and layout of environmental innovation policy. In particular, following recent governmental initiatives in the Netherlands organized under the heading of *transition management*, evolutionary models may be helpful in assessing and developing policy strategies to trigger technological transitions from one technological system to a future technological system (Grübler, 1998; Unruh, 2000, 2002; Hoogma et al., 2002; Frenken et al., 2004; Könnölä et al., 2006; Van den Bergh et al., 2005; Dosi and Grazzi, 2006; Carrillo-Hermosilla, 2006; Zhang et al., 2006). There is a growing consensus that societies should reduce their dependence on fossil fuels and shift to more sustainable alternatives (e.g., from a gasoline to a fuel cell car system). However, it is far from clear which of the alternative technologies should be selected and introduced at a large scale. Given the current uncertainties, the development of a large variety of technological options is to be preferred. On the concept of variety in evolutionary economics, see Saviotti (1996) and Stirling (1998, 2004).

The specific features and (therefore) added value of an evolutionary approach compared to traditional neoclassical approach in the context of technological transitions can be summarised in three aspects, which are outlined below.

1. *Technological trajectories and technological paradigms*

Neoclassical (endogenous) models of technology adoption focus solely on changes in factor prices that induce the level and the choice of technology. In the context of technological transitions, these models tend to favour policies that ‘correct’ prices through taxes and subsidies to reflect the true welfare consequences of various technologies. These price changes will then trigger firms to adopt – or start developing – cleaner technologies. Evolutionary economic models, on the other hand, focus on the idea that technological development follows certain *technological trajectories* of incremental change (due to learning, network externalities and increasing returns to scale) within the boundaries of a technological paradigm (Dosi, 1982; Frenken, 2006) or what has been termed a techno-institutional complex in the context of the carbon-based technologies (Unruh, 2002). Only occasionally, transitions occur between two technological paradigms, yet these transitions are triggered not only by higher prices of the old paradigm, but also by the exhaustion of technological opportunities in the old paradigm and new, or newly recognised, technological opportunities in alternative paradigms. The classic example is the limited effect of the oil crisis in the 1970s on the subsequent direction of technological development. Higher oil prices did not induce radically different alternatives but rather pushed for incremental innovations within the oil-based technologies, aimed at saving fuel or reducing pollution. By contrast, the current peak in oil prices seems to trigger R&D efforts of a more radical nature because profit returns along the current technological trajectory have decreased and institutional pressures towards sustainability have increased. The shift in dynamics between 1970s and today cannot be explained by a neoclassical model including price changes only, but is more in line with an evolutionary model that stresses the incremental innovation strategies of firms and governments. This is not to say, however, that economic incentives play no role at all in evolutionary modelling (e.g., essay Schwoon). Rather, an evolutionary approach stresses that price changes alone are insufficient to explain the rate and direction of technological development and that a broader perspective should be taken into account.

2. *Taking demand seriously*

A second fundamental difference is the treatment of demand. Rather than reducing human consumption to given preferences, which are articulated solely through prices, an evolutionary approach to consumption stresses that preferences are subject to change (Witt, 2001; chapter Jager in this report). Changing preferences may result in public action towards sustainability, for example in the form of imposing stricter regulatory standards, new labels, or even the banning of certain products or materials. In this context, it is crucial to have insight into the effects of standards on industrial dynamics (essay Saint-Jean) as well as into the political processes that lead societies to introduce particular standards (Könnölä et al., 2006). Changing preferences can also lead to user-led innovations as demanding 'niche users' are often crucial as the frontrunners to a technology's emergence and success (Von Hippel, 1988; chapter Windrum in this report). Thus, in an evolutionary world, the role of consumers lies not so much in reacting to price differentials but more in their ability to change their preferences, possibly triggering new out-of-paradigm niche markets.

3. *A co-evolutionary perspective on policy*

In evolutionary economics, institutions including government policy are seen as an integral part of technological paradigms rather than being independent from it. For example, part of the technological paradigm surrounding the current car system is a complex set of institutions including tax laws, environmental laws, fuel supply infrastructure, safety requirements, technological standards, traffic rules, trade treaties, consumer organisations, producer organisations, training and research institutions, brand names, consumer typologies, et cetera. These institutions strengthen the current paradigm and contribute to its economic efficiency and social acceptance. Environmental innovation policy thus not only requires specific policies to favour certain developments within the existing paradigm or the development of alternative paradigms; policies should also change the institutions inherited from the previous/current technological system in desirable directions. As such, evolutionary economics takes a meta-institutional perspective on technological development in which technological change, consumer demand and institutions co-evolve and mutually interact (chapter Windrum in this report).

1.2 Models

Evolutionary models can be used for policy making in a way similar to neoclassical models in that one can experiment with different policy measures and evaluate their effect ex ante (for example, in terms of welfare, in terms of CO_2 reduction, etc.). Another use of models is to specify the model as a game and to let stakeholders play with the model and discuss the results. Obviously, experimenting with different policy measures using models that simulate society is less costly and less risky than actual experiments with true policy. However, the insight gained through computer simulation crucially depends on the 'degree of correspondence' of the model to real-world society. Correspondence in this context does not necessarily mean that models that try to take into account more aspects of reality also achieve a higher degree of correspondence. What matters most is that models capture the most relevant mechanisms at work in specific social processes.

1.2.1 Existing evolutionary economic models

A wide range of existing models falls under the family name of 'evolutionary economics' (see especially the essay by Verspagen). In particular, there is a variety of approaches towards evolutionary economic modelling for studying technological development. Within the variety of evolutionary models one can distinguish broad comprehensive models from more specific models. Within the second type of models one can subsequently distinguish four different focusing strategies.

1. *Comprehensive evolutionary models*: models that take into account all relevant aspects of technological development. These models include different co-evolving populations (consumers, producers, governments) and different mechanisms (competition, innovation, learning, externalities, market segmentation) in a comprehensive and (more or less) integrated approach and a broad framework. Windrum's essay provides an overview of elements that could be included in such a model based on a review of theory, models and case studies. Comprehensive evolutionary models may serve to describe and understand the integrated *context* of large scale technological developments, without the disadvantages of including detailed bottom-up technical studies. Methodologically, these models have a high level of aggregation and they may suffer from the many parameters that are present in the model, due to the large number of (sometimes poorly understood) mechanisms of social processes included. This renders their validation significantly more difficult than simpler models.
2. *Specific evolutionary models*: these models focus on understanding a specific aspect of technological innovation, thus greatly reducing the complexity of the modelling task at hand. These specific models usually have less parameters and lead to more conclusive insight, but they run the risk of lacking a clear perspective on the broader framework in which technological development takes place. There are mainly four focusing strategies in specific evolutionary models:
    a. to focus on one specific technology;
    b. to concentrate on one specific mechanism;
    c. to focus on one specific population of agents;
    d. to follow a modular approach in which different parts of a technological system are modelled in separate models with clear interfaces among them.

All four strategies are exemplified in the essays. Verspagen follows strategy *a.* by focusing on micro co-generation technologies. Saint-Jean follows strategy *b.* by focusing on the impact of environmental standards on product and process innovation. Jager follows strategy *c.* discussing the behaviour of consumers as one of the relevant populations in the economy. Finally, Schwoon adopts strategy *d.* in his study on the technological system of fuel cells.

During the workshop consensus emerged that specific evolutionary modelling – which can also be termed middle range theorizing – is more attractive than the first type of wide range theorizing, since it offers more perspective on practical modelling application, without having to make an unacceptably large number of assumptions beforehand. Presently, the development of comprehensive, integrated evolutionary models for the study of technological systems is still regarded to be a bridge too far. The specific strategies to reduce the complexity of the modelling task at hand, then, apply in different contexts:

1. *Strategy a* is especially relevant when a policy maker is interested in assessing the potential of a specific new technology (including its sub-variants) as well as to investigate which type of policy is most helpful in stimulating the technology given specific technological characteristics and market conditions.
2. *Strategy b* is especially relevant when a policy maker wants to assess the effectiveness and efficiency of some general policy measure (subsidy, regulation, standards, etc.) in different types of markets c.q. technologies.
3. *Strategy c* is especially relevant when a policy maker is interested in evaluating different ways to influence particular types of agents, for example, consumers, truck drivers, firms, farmers, etc.
4. *Strategy d* is especially relevant to advance a theoretical understanding of the complete technological system. Furthermore, the different modules can be applied more specifically, using strategy *a*, *b* or *c*.

For each of these strategies the *system scale* and *time horizon* are very relevant parameters to take into account in defining the policy questions to address with the model. Verspagen points out in his essay that uncertainty is high in the study of large-scale systems, because many interconnected components are unpredictable. Because of the dependency between the components in the system at large, unpredictability multiplies at the system level. On the other hand, in micro-level studies there is a large amount of external factors as well as a large degree of heterogeneity. Therefore, evolutionary theory is regarded to be a theory of the intermediate range. Within this approach, it is useful to take into account the 'windows of opportunity' for policy makers, making use of larger scale political and technological dynamics (e.g. EU regulations) as well as of the interplay between short-term and long term policies. This momentum may greatly affect the diffusion of technologies and thus 'the right time' for political action aimed at stimulating environmental technologies (Sartorius and Zundel, 2006). This clearly drifts away from regular neoclassical approaches of instant reaction to e.g. a tax or other policy measures.

Most evolutionary economic models do not include policy options for influencing the outcome of the modelled processes. Only a limited number of examples of policy oriented evolutionary economic models (outside environment and ecology) is available. Although the concept of evolution has found its way in management theory, its applications in economics have been less frequent. Exceptions are the Swedish MOSES-model that is used for economic policy questions (Ballot and Taymaz, 1999), a recent model of technological transitions based on fitness landscapes (Schwoon et al., 2006), work on environmental policy in the essays by Saint-Jean and Schwoon, and the papers presented at a recent workshop on 'Agent-Based Models for Economic Policy Design', organized by Professor Herbert Dawid.[1]

---

[1] See: http://www.wiwi.uni-bielefeld.de/~dawid/acepol/program.htm

1.2.2 Recommendations for getting started and for a long-term research agenda

Based on the essays and workshop discussions (first working sessions) the following conclusions and recommendations can be made on evolutionary economic modelling to support policy studies:

- It is impossible to build a generic model for the economic development of our society at large.
- Focus on a technology (strategy *a*), a mechanism (strategy *b*) or a population of agents (strategy *c*), possibly using these strategies as modules for a more integrated approach to a technological system (strategy *d*).
- It is important to identify the relevant stakeholders, underlying mechanisms and dynamics, and the relevant technologies (foresights).
- Evolutionary model builders should therefore work in close association with experts in a particular field, as well as experts in different kinds of (technology) foresight studies.
- Statistical and empirical information on populations, psychological, sociological and economic theory is needed to provide an adequate input to the model. It is therefore advisable to start with a well-known and well-described (in terms of mechanisms and data availability) problem.
- An interesting approach might be to link the evolutionary model with other types of models: e.g. describe system options with an Input-Output approach using the DIMITRI-model (Idenburg and Wilting, 2004), and then simulate pathways to that system with an evolutionary economic model.
- For the longer term, it is considered important to have a number of different model (concepts) available dealing with similar phenomena or questions. Within such model diversity one can envisage some level of model competition, which could enhance learning and insight in modelling approaches.
- The research organisation could be structured as in the following scheme:

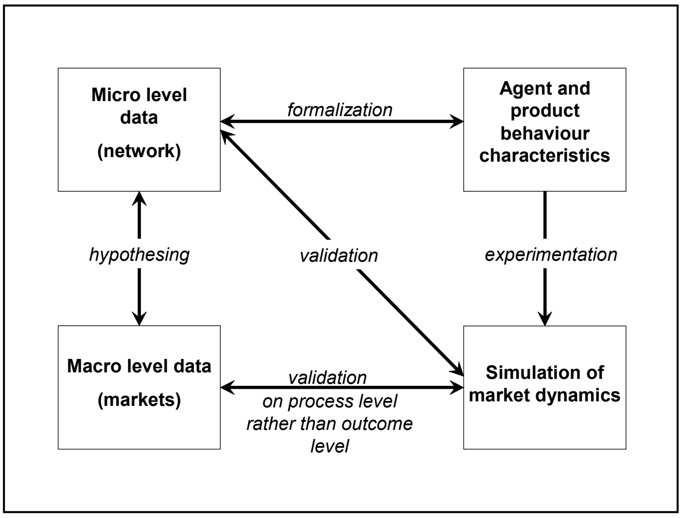

Figure 1.1 Scheme of future research organisation

1.3 Conclusion

Evolutionary economics provides a promising theory to assess and understand processes of change in economic structure, technological development and institutions, as well as to formulate guidelines for the role of government and the design of public policy in this context. Evolutionary economic models can in principle be used to experiment with different policy measures and to evaluate *ex ante* effects. However, although evolutionary economic modelling is promising in terms of describing mechanisms, there is no complete scientific consensus yet on a single modelling paradigm. A number of examples of evolutionary economic models on micro-level problems is available as a useful starting point.

Two essential recommendations stand out for further development of evolutionary modelling with respect to environmental policy analysis:

1. focus on a specific issue (a technology or a mechanism or a population), rather than taking into account all complexities;
2. make sure that the empirical base in terms of mechanism and data is sound.

It is important to realise that the development of evolutionary economic modelling is an evolutionary process itself. It may therefore be helpful to develop competing approaches and assess their relative performance rather than aiming at one modelling paradigm.

References

Ballot, G. and E. Taymaz (1999). Technological change, learning and macro-economic coordination: An evolutionary model. *Journal of Artificial Societies and Social Simulation* 2 (2). http://www.soc.surrey.ac.uk/JASSS/2/2/3.html

Bergh, J.C.J.M. van den, A. Faber, A.M. Idenburg, F.H. Oosterhuis (2005). *Survival of the Greenest, evolutionaire economie als inspiratie voor energie- en transitiebeleid*. MNP (report no. 550006002), Bilthoven (NL). http://www.rivm.nl/bibliotheek/rapporten/550006002.pdf

Carrillo-Hermosilla, J. (2006). A policy approach to the environmental impacts of technological lock-in. *Ecological Economics* 58: 717-742

Dosi, G. (1982). Technological paradigms and technological trajectories. A suggested interpretation of the determinants and directions of technical change. *Research Policy* 11: 147-162

Dosi, G. and M. Grazzi (2006). Energy, development, and the environment: an appraisal three decades after the ‘Limits to Growth’ debate. *LEM Working Paper* 2006-15. http://www.lem.sssup.it

Frenken, K. (2006). *Innovation, Evolution and Complexity Theory*. Cheltenham (UK), Edward Elgar

Frenken, K., M. Hekkert and P. Godfroij (2004). R&D portfolios in environmentally friendly automotive propulsion: variety, competition and policy implications, *Technological Forecasting and Social Change* 71(5): 485-507

Grübler, A. (1998). *Technology and Global Change*. Cambridge (UK), Cambridge University Press

Hoogma, R., R. Kemp and J. Schot, (2002). *Experimenting for Sustainable Transport. The Approach of Strategic Niche Management*, London: EF&N

Könnölä, T., Unruh, G.C., Carrillo-Hermosilla, J. (2006). Prospective voluntary agreements for escaping techno-institutional lock-in, *Ecological Economics* 57: 239-252

Idenburg, A.M., Wilting, H.C. (2004). DIMITRI: A Model for the Study of Policy Issues in Relation to the Economy, Technology, and the Environment. In: Van den Bergh, J.C.J.M., M.A. Janssen (eds.), *Economics of Industrial Ecology; Materials, Structural Change and Spatial Scales*, pp. 223-252. Cambridge (Mass.), MIT Press

Nelson, R.R. and S.G. Winter (1982). *An Evolutionary Theory of Economic Change*. Cambridge (Mass.), London, Belknap Press of Harvard University Press

Sartorius, C. and S. Zundel (eds) (2006). *Time strategies, innovation and environmental policy*. Cheltenham (UK), Edward Elgar

Saviotti, P.P. (1996). *Technological Evolution, variety and the economy*. Cheltenham (UK), Edward Elgar

Stirling, A. (1998). On the economics and measurement of diversity. *SPRU Electronic Working Paper* 28. SPRU, University of Sussex. http://www.sussex.ac.uk/spru

Stirling, A. (2004). Diverse designs, fostering technological diversity in innovation for sustainability. Paper presented at conference *Innovation, Sustainability and Policy*, Seeon (Germany), 23-25 May 2004

Schwoon, M., F. Alkemade, K. Frenken and M. P. Hekkert (2006), Flexible transition strategies towards future well-to-wheel chains: an evolutionary modelling approach, *FNU Working Paper* #114 http://www.uni-hamburg.de/Wiss/FB/15/Sustainability/WP-FNU-114_Flexible_transition_strategies.pdf

Unruh, G.C. (2000). Understanding carbon lock-in. *Energy Policy* 28: 817-830

Unruh, G.C. (2002). Escaping carbon lock-in. *Energy Policy* 30: 317-325

Von Hippel, E. (1988). The sources of innovation. New York, Oxford, Oxord University Press

Witt, U. (2001). Special issue: Economic growth - What happens on the demand side? *Journal of Evolutionary Economics* 11: 1-5

Zhang, J.T., Y. Fan, Y.M. Wei (2006). An empirical analysis for national energy R&D expenditures. *International Journal of Global Energy Issues* 25 (1-2): 141-159

2 The use of modelling tools for policy in evolutionary environments

B. Verspagen

Eindhoven Centre for Innovation Studies (ECIS), Technische Universiteit Eindhoven (NL).

Contact: b.verspagen@tm.tue.nl

ABSTRACT

This is a position paper on the possibilities of informing the (economic and environmental) policy debate by using quantitative evolutionary models. I argue that an evolutionary worldview implies that the existing quantitative modelling tools used for policy analysis are problematic. Then I summarize the main elements of an evolutionary way of analysis, and the way in which it can be incorporated into quantitative models. I conclude with an outline of a proposal for how to apply the ideas in the analysis of energy transitions.

2.1 Introduction

Evolutionary economics has been presented as a more relevant alternative to mainstream economics. It is rooted in the economic analysis of technological change and innovation, and argues that it can provide a more realistic theory of these phenomena. Since innovation is a societal process with wide-ranging impacts, evolutionary economics is very relevant for policy.

But at the same time, the direct policy implications of evolutionary theorizing are far from clear. For example, it is not clear if the policy implications from evolutionary economics differ from those of mainstream economics. Even if the foundations of the two theories differ, the policy implications may be similar, especially when formulated at a general level (‘stimulate technological innovation’).

Policy advice by economists has traditionally been based on quantitative simulation models that can be used to ‘predict’ the effects of policies, as if it were a laboratory setting. This has the advantage that the impact of policies can be assessed ex ante in a precise way (at least, if the model’s predictions are by-and-large correct).

This paper is concerned with the question whether such an approach is also possible using evolutionary economics (or evolutionary analysis in a broader sense). Is it possible to formulate quantitative evolutionary models that can be used to support policy? Given the tentative but affirmative answer to this question that I will give below, I will further ask whether the use of such evolutionary models differs from the use of the traditional economic policy models.

I will lay out my argument in the following way. In section 2.2, I will briefly summarize the foundations of the mainstream economics approach to quantitative policy modelling. In section 2.3, I will discuss the principle of evolutionary economic analysis, and define what I consider the most important elements of evolutionary thinking for the question formulated above. section 2.4 will discuss two particular approaches to modelling, i.e., the use of confidence intervals and scenario analysis, and their relevance for evolutionary policy modelling. Section 2.5 will present a list of more concrete guiding points for evolutionary policy models. Finally, this list will be used in section 2.6 to present the outlines of an example of an evolutionary policy model in the field of energy systems. More concrete, this model is aimed at modelling a potential transition to a hydrogen economy. Based on an existing model (Taanman, 2004), I will discuss how such an approach may be implemented, what kind of results we may expect, and how these results should be interpreted.

2.2 Economic policy models and the notion of equilibrium

Economics has a relatively strong influence on policy thinking through the use of large-scale econometric models that are used for simulations to support policy. In these models, as in general in economics science, the notion of equilibrium plays a large role. The usual definition of equilibrium that is used in economics points to a state of the economic system in which none of the economic agents (firms, consumers) has an incentive to change behaviour (e.g., charge higher prices, or buy more of a certain good). Without such an incentive, there is no factor (apart from random fluctuations) that may induce any change, hence the term equilibrium. In such a *static* equilibrium, nothing changes in the way the economy works. Economic policy models are based on a notion of *dynamic* equilibrium. In its basic form, dynamic equilibrium is a sequence of static equilibria.

Take, for example, the case of a simple model of supply and demand. The interaction between demand and supply will lead to an equilibrium that is characterized by a unique price and quantity sold/bought. As long as this equilibrium is not reached (i.e., the price is either too high or too low), buyers and suppliers have an incentive to change. If the price is too high, suppliers cannot sell all products they wish to sell (there is a supply surplus), or, in other words, buyers are not willing to buy everything the suppliers offer. Hence there is an incentive for the suppliers to change their behaviour, for example by offering their surpluses at lower prices. This continues until the market reaches a point where demand and supply are equal to each other, and none of the parties has an incentive to change behaviour. We have reached static equilibrium.

However, if some of the external (*exogenous* is the usual technical term) factors that determine the market outcome change, the nature of the static equilibrium changes. For example, if the supply curve in our market example shifts to the left (e.g., due to climatic circumstances), the equilibrium market price will go up and the equilibrium quantity will go down. The result is a *dynamic equilibrium path* in which the price goes up from period to another, and the quantity goes down.

The economic policy modelling tradition that starts with Tinbergen is based on this framework of a dynamic equilibrium path. More specifically, it assumes that a) the dynamic equilibrium path is unique and stable, b) that adjustment to static equilibrium is instantaneous, and c) we can calculate the equilibrium based on an empirical specification of the model that can be obtained by statistical procedures (econometrics).

These three assumptions, which I will discuss critically in section 2.3, enable the policymaker to compare a whole range of policy options by plugging them into the model, and interpret the outcomes as in terms of various variables that are of interest for the maximization of policy outcomes. On the basis of such a comparison, the most favourable policy outcome can be selected, and the respective policy can be implemented.[2]

Equilibrium is a cornerstone in this way of thinking, because it is an essential concept for the calculation of the effects of the policy variables. Changes in policy will change the equilibrium, and the measurable effect of policy is taken as the difference between those two equilibria.

The individual equations of the model that must be used to calculate the equilibrium are usually based on microeconomic theories of agent behaviour (this is the so-called micro-foundation of macroeconomics). For example, a supply curve will be based on a theory of producer behaviour (profit maximization) under restrictions set by market structure and production technology. In this step from micro to macro relations, the representative agent plays a large role. This is a notion that is used to aggregate outcomes of the microeconomic theory directly to the macroeconomic level, without the need to explicitly add up different behavioural patterns.

Uncertainty plays only a minor role in this approach. It enters the equations in the form of a random disturbance term (with very specific characteristics) that is added to each equation. Thus, it can be expected that the actual outcome that will be observed in the real economy differs slightly from the

---

[2] An additional problem is how the different variables (e.g., income growth, distribution, unemployment) should be weighted, but we will abstract from this here.

outcome predicted by the model, due to these random disturbances. But for policy analysis, the disturbances do not matter, since one may compare the different policy options on an ‘equal basis’ by always setting the disturbances to zero.

2.3 Evolution, equilibrium, policy, and modelling

Before I compare the above approach to policy modelling to a more evolutionary way of thinking, there is a need to define what is meant by such an evolutionary approach. Quite often, an evolutionary process is defined as one in which novelty and selection work hand-in-hand to produce change. Although this is obviously a correct, and often useful definition, I will not adopt it here. The reason is that it does not help us outline what the specific consequences of evolution for policy modelling are.

Instead, I define the following four crucial characteristics of a socio-economic evolutionary process. First, such a process is characterized by bounded rationality at the micro level, leading to significant variety of behavioural patterns. When faced with the same external environment, different agents (individual consumers, firms) may react in different ways, and show different behaviour.

Second, evolutionary processes are characterized by a certain degree of persistence of random events. In simple words, small random events may change the course of history. Rather than being additive to a deterministic equilibrium, small random events in evolutionary processes may accumulate into larger factors that may change the nature of the system and its history.

Third, if equilibrium plays any role in an evolutionary process, it certainly is in the form of multiple equilibria. A dynamic system that has a single, stable equilibrium, will, at least in the long run, always tend towards this single equilibrium. This makes prediction simpler (if, as I did in section 2, we assume that the equilibrium can be calculated). But in an evolutionary context, there generally are multiple equilibria, meaning that which particular equilibrium state is reached, depends on where the system starts (or, to take an advance on our discussion below, where it is pushed, for example by policy).

Fourth, in any evolutionary system, the speed with which equilibria are approached may vary over time (so-called punctuated equilibrium), but reaching equilibrium may take a long time. Moreover, and equilibria themselves are changing as a result of change in the system itself. As a result, equilibrium points in an evolutionary system are rarely actually reached. Instead, they serve as an attractor that pulls the system towards itself for a prolonged period, before giving way to a new attractor. The consequence of this is that we cannot take the equilibrium of an evolutionary model as a useful description of an actual future state of the world. Instead, we must model the path towards the equilibrium as an approximation of what the world may look like.

Note that each one of these four characteristics may be found in some specific economic modelling approaches, but that only in a truly evolutionary economic model, the four are found jointly. For example, Sargent (1994) uses the theory of bounded rationality and behavioural variety to model macroeconomic process, Krugman (1990) makes extensive use of the notion of multiple equilibria, the notion of persistence of random factors is central in the econometric debate about unit roots (Nelson & Plossner, 1982), and the debate on convergence in living standards (Barro and Sala-I-Martin, 1991) puts strong emphasis on transitory dynamics towards dynamic equilibrium.

The result of these four characteristics of evolutionary processes is that that evolution is very difficult to predict. Extending the argument from prediction in the (usual) time domain, it is also true that it is very hard to produce dependable 'counterfactuals' in an evolutionary model. In the biological/paleontological debate, this had led to the famous question (asked by Stephen Jay Gould) 'What would be conserved if the tape were run twice' (see also Fontana & Buss, 1994). This question refers to the thought experiment in which we would be able to run two parallel worlds, initially similar to our own, both in which evolution would take its course. After a significant amount of time had lapsed, would the two worlds look anything like each other, or like the one that we know now?

We can see how each of the four characteristics of evolution described above would contribute to producing widely diverging worlds. Bounded rationality and behavioural variety may lead individuals (once they had evolved) to go entirely different ways even when initial environments are similar, and this may in turn lead to different outcomes. The persistence of random events will lead to an accumulation of random events that is different from every realization of a stochastic process, again leading to completely different outcomes in the hypothetical parallel worlds. Multiple equilibria may equally fork the parallel worlds into completely different directions. Finally, when speed of the evolutionary change process differs between periods in each parallel world, this will induce again an element of difference between them.

One would thus tend to answer that 'not much' would be preserved if the tape were played twice. This implies that evolutionary processes are characterized by a high degree of strong uncertainty. If *n* (where *n* is fairly large) 'parallel worlds' that start out as being similar, may evolve to be quite different from each other after a while, this implies, first, that a large number of possible outcomes are thinkable, and, second, that it is impossible to predict which of these outcomes will actually prevail.

In such a situation, traditional methods of assessing risk may loose their relevance, since these are based on probability distributions. A probability distribution assumes both that the possible outcomes are known in advance, and that (an estimate of) a probability can be given for each. But when uncertainty is strong, the possible outcomes are unknown, and the probability distribution cannot be conceived.

Despite this strong level of uncertainty, there must be some bounds to evolutionary outcomes, if only because the laws of nature (which, at a higher level, may be subject to evolution themselves). Thus, evolution is a process in which the two factors of chance and necessity (Monod, 1970) are intermingled and determine the direction that a system takes. In evolutionary biology (see, e.g., the popular works of Dawkins and Gould), there seems to be some consensus that the chance side of this relationship is dominant, but I will argue below that the balance may be different in socio-economic evolutionary systems.

These characteristics of evolutionary processes largely invalidate the approach in building economic policy models that I discussed in section 2.2. Bounded rationality and the associated behavioural variety invalidate the idea of a representative agent, and hence makes the usual aggregation procedures impossible. Multiple equilibria invalidate the calculation of the single equilibrium that varies under policy variations, and introduces the need to consider starting conditions and define basins of attraction. The effects of stochastic processes and uncertainty invalidate the idea of a unique and calculable equilibrium. Finally, the importance of transitory dynamics detracts from the importance of the equilibrium notion itself.

Although it is obviously possible to discuss these issues at greater length, I will not do so here. Instead, I will focus the largest part of the essay on the positive implications of these four evolutionary principles for policy modelling.

2.4 Evolutionary analysis and existing modelling traditions

The main challenge to building evolutionary policy models is the fact that evolution is a process in which chance plays a significant role. The key feature of evolution is that small, random (and therefore unpredictable) events may have severe long-run consequences. This means that any simulation exercises performed with a policy model must be taken with extreme caution.

In this section, I ask the question whether any existing ways of dealing with uncertainty in quantitative models can help us deal with this feature of the evolutionary process. Two specific issues come to mind: first, sensitivity analysis and the augmentation of model simulations with confidence intervals and standard errors, and, second, scenario studies.

Initially, the outcomes of the policy models as described in section 2.2 were taken as point estimates, i.e., the specific dynamic equilibrium path that was produced by the model for a given set of policy parameters, was taken as the direct estimation of the impact of the proposed policy. This obviously does not consider the uncertainty that is embedded in these models. There are at least two sources of such uncertainty: potential parameter variations, and imperfect estimations of exogenous variables (including the variables related to the policy itself).

However, given that we have some information on the potential amount of (stochastic) variation in these two dimensions, we may actually produce not only the single dynamic equilibrium paths, but also produce an indication of how variable they are under reasonable stochastic variations. Hence, instead of using the parameter values obtained in econometric estimation, we may vary the parameters by using the standard errors of these estimations. Similarly, we can undertake sensitivity analysis of the model outcomes as a result of variation in exogenous (policy) variables. In this way, instead of a point estimate of the policy effect, we can obtain a confidence interval.

While confidence intervals are obviously a step forward compared to points estimates, they do not solve any issues related to the model structure itself. For example, a model that is based on the notion of a single equilibrium that is characterized by traditional economic reasoning, does not chance in nature by having it produce confidence intervals instead of point estimates. If the structure of the model and the main ideas underlying it is flawed, a more sophisticated sensitivity analysis will not rescue its predictive power.

Scenario analysis may be a more sophisticated tool of analysis that comes closer to the core evolutionary ideas. Scenario analysis is usually associated with the systems dynamics way of modelling (e.g., Hughes, 1999), but it is also used in more mainstream (economic) policy models such as those used by the Netherlands Bureau of Economic Policy Analysis. In scenario analysis, an existing policy model is used to generate a number of outlooks on the future. A scenario is specified as a combination of specific assumptions that can be associated with a broad narrative about potential ways in which the system that is being modeled will develop. It is not the intention of the scenario analysis to predict which scenario will take place, and this is a major difference with the mainstream policy models discussed in section 2.2.

Instead, the aim of scenario analysis is to explore the variety of potential outcomes under alternative assumptions. For example, in a model of the global (macro) economy, one may wish to investigate the general nature of different scenarios for the development of world trade. Then, one could specify one scenario in which world trade will stagnate (e.g., as a consequence of the outcome of international negotiations about liberalizing trade), and one scenario in which international trade will grow. One may then investigate how a range of variables (e.g., global income distribution, CO_2 emissions, etc.) will differ between the scenarios. In this way, an impression is obtained of how whether or not world trade will grow will change the world.

Scenario analysis is less pretentious in prescribing specific policies than the models we discussed in section 2.2. It gives insight into the available range of policies and the order of magnitude of their effect, rather than analyzing the exact impact of a specific policy. In this sense, it is closer to the principles of evolutionary systems as outlined above, because it recognizes the large degree of uncertainty present in the real world.

Although scenario analysis may certainly be useful, I maintain that, as a potential centerpiece of evolutionary model building, it is not very useful. As I will argue below, evolutionary models may well be used to conduct scenario analysis, and this is likely to add insights, but scenario analysis is not the saviour of evolutionary model builders. The reason for this is that at the heart of the models that are used for scenario analysis, we still have the same approach that is used to build the policy models I discussed in section 2.2. If the model itself is not built on evolutionary principles, using it for scenario analysis does not make it evolutionary.

2.5 Towards evolutionary policy models

Although, as argued above, we must be pessimistic about the possibility of existing risk-treatment techniques in quantitative policy models for dealing with ‘evolutionary uncertainty’, the prospects for using quantitative model tools in evolutionary policy analysis are not hopeless. This section will attempt to outline some possible ways of proceeding in this way. The key issue is about the mix between chance and necessity in the evolutionary processes that we wish to analyze for policy. What is the relative contribution of chance and necessity to evolutionary processes remains a matter open to debate. Arguably, the outcome of this debate will differ between pure biological and socio-economic evolutionary systems.

In biological evolution, the main source of novelty is random genetic mutation. Genetic mutation consists of errors in copying genetic information, and can be characterized as a truly blind process. Any specific genetic mutation that occurs in the history of a biological process may or may not lead to a ‘useful’ design change, but whether or not the change is ‘useful’ plays no role at all in generating the mutation itself. Hence Richard Dawkins’ metaphor of the blind watchmaker: mutations are not purposeful, although they may, ex post, prove to be ‘useful’.

In socio-economic systems, more complicated sources of novelty exist. An important source is behaviour of the micro-entities in the system (let’s say firms and consumers). This behaviour, although not fully rational in the sense of mainstream economics, certainly has a purpose (as conceived by the agent). Behavioural change is implemented for a reason, and in general terms we may say that this reason is to generate better performance of the agent who implements the change. In addition, while genetic mutations are memory-less (there is a positive probability that a copying error is reversed later on), socio-economic agents have the ability to learn on the basis of their previous experiences. This opens up the possibility of experimentation aimed at finding a ‘good’ strategy.

This has important consequences for the outcome of the evolutionary system. In the first place, the non-purposeful mutations in biology have a far greater potential range of impacts than the purposeful changes in socio-economic behaviour. Of all possible changes in behavioural patterns, the conscious economic agent will immediately rule out a number as non-sensible (even if they might make sense beyond the decision horizon of the individual agent). Biological evolution does not, at the level of the mutation itself, include any such selection. Thus, novelty in socio-economic evolutionary systems will be confined to a narrower (but possibly still rather broad) range than in biological systems.

Second, because agents in socio-economic systems can learn, as well as apply selection at their own micro-level, the speed at which evolution may take place will be much higher than in biological systems. In other words, the relevant time horizons in socio-economic systems are much shorter than those in biological systems. The emergence of mankind took millions of years, the emergence of the Industrial revolution several decades.

These two differences between biological and socio-economic evolution have consequences for the nature of the two evolutionary processes. In biological evolution, the potential for predicting which direction evolution will take is an impossibility. Carbon-based life on earth is a ‘magnificent accident’ indeed, and we should not expect something even broadly similar to emerge in a parallel world. But in socio-economic evolution, the range of directions that evolution may take may be smaller.

This does not imply that predictability of socio-economic systems is perfect, or even close to the level that is assumed by the policy models discussed in section 2.2 above. Socio-economic evolution remains a historical process in which contingencies play a role. It is different from a mechanistic process with perfect predictability. Predicting the motion of planets and other heavenly bodies using a Newtonian model remains a quite different affair from interpreting and analyzing evolutionary change in socio-economic systems. These latter systems are somewhere in between the clockwork world of Newton and the magnificent accident of Stephen Jay Gould.

Where exactly the systems that we are interested in are on this continuum, depends on the scope that we are taking, both in terms of time (how long do we want to look ahead?), and the range of phenomena we wish to look at. Contingencies and random factors are more likely to play a decisive role in making outcomes of evolutionary processes indeterminate when we look either at large scale systems of many interconnected components, or when we look at small-scale (micro) systems.

In the case of large-scale systems, indeterminacy is large because each of the interconnected components itself is unpredictable. Because of the dependency between the components in the system at large, unpredictability multiplies at the system level. The scope for building a precise quantitative evolutionary policy model for problems that require such large-scale systems analysis is thin.

At the micro level the problems are of a different nature. They stem from two sources. First, at the micro level, we have a large amount of external factors, each of which is the result of the large-scale system that we have discussed above. Second, behavioural patterns at the micro level are subject to a large degree of heterogeneity, and evolutionary theory as such does not have much to add about the way in which this heterogeneity can be analyzed. This is the domain of psychology, and possibly sociology or even (mainstream) microeconomics.

Evolutionary theory in the field of socio-economic processes, on the contrary, is a theory of the intermediate range (Merton, 1973). When and if we can formulate problems that can be analyzed in an evolutionary system in which not too many different domains of interaction are involved, the scope for using quantitative models for policy purposes are good.

What exactly an 'intermediate range problem' is, is hard to specify in more concrete terms. Probably the question of how Chinese economic growth will have an impact on the income distribution in the Netherlands in 2025 is an example of a too large-scale system to be analyzed in a precise quantitative way using an evolutionary policy model.

A sufficient but not necessary condition for an intermediate range problem can be formulated using the notion of multiple equilibria. If a specific policy problem is characterized by a small, but larger than one, number of equilibria, that can be clearly separated from each other, we may characterize this as a typical problem that can be modeled by evolutionary dynamics. Typically, problems in the field of transition analysis, e.g., environmental-friendly technological trajectories can be characterized in this way. I will therefore attempt to sketch the steps in modelling such transitions using evolutionary dynamics in the next section. Before doing so, however, I will formulate in the remainder of this section a number of general issues regarding the nature of evolutionary policy models.

In a pure technical sense, evolutionary models differ from more mainstream models in at least two ways that are important for policy analysis. The first one is the existence of multiple equilibria, and the second is the importance of variety in behavioural patterns.

Multiple equilibria provide a different perspective on policy than the one that is found in mainstream policy models. As summarized in section 2.2, the usual way of looking at policy analysis in quantitative models is to assume that policy may change the nature of the (single) equilibrium in the model (world). With multiple equilibria, this changes. In addition to policy changing the character of the equilibria, there is also an option to move the system out of the basin of attraction of one equilibrium, and into that of a different one.

This is a significant change of perspective in different ways. For example, it is not so clear that the 'Lucas-critique' is valid in the same way in the case of a world with multiple equilibria. Lucas (1976) argues that if economic agents have rational expectations, government policy may in many cases be inefficient, because agents calculate the effects of government policy, adjust their actions accordingly, and the effect of the policy may be counteracted by this. In a technical sense, the equilibrium of the model is the same whether or not government policy is affected. But if there are multiple equilibria, the response of the agents to government policies may leave the equilibria unchanged, but may still put the economy on a track towards a different equilibrium.

Also, if there are multiple equilibria, government policy has more options. If, for reasons of efficiency of policy instruments, some policies are not effective, other options may still be open. For example, it may be the case that the nature of each of the multiple equilibria depends on technology (e.g., the case of alternative energy systems), but government has insufficient information to select the agents that are best situated to advance a certain technology (this is the argument often used by those who oppose a government policy based on 'picking winners'). In this case, policy may be geared towards bringing the system in the basin of attraction of a different equilibrium, without having to pick winners (i.e., specific firms to subsidize) within or between alternative technologies. Instead, a general policy aimed at stimulating consumption may do the trick.

Thus, an evolutionary policy model must take the existence of multiple equilibria serious. But it is hardly to be expected that a generic model (i.e., set of equations that can be run on a computer) will tell us how many and which equilibria exist for a specific policy situation of interest. This is a task for exploratory analysis that must be performed before any particular model can be built.

This 'treatment' of multiple equilibria has two implications. First, it reinforces the argument about evolutionary policy model being theories of the intermediate range. We cannot build a generic model of the multiple equilibria that may attract the economic development of our society at large. We can only hope to build a model of the multiple equilibria of a problem in the intermediate range that we have carefully outlined by non-quantitative analysis before attempting to build a policy model.

Second, it implies that evolutionary model builders must work in close association with experts in a particular field, as well as experts in different kinds of (technology) foresight studies. This includes interacting with, for example, technical experts that work in a quantitative engineering tradition and who can help outlining the technology options, as well as using the heterogeneous 'art' of foresight studies in all its guises. The 'roadmaps' that foresight studies can produce should not be taken literal, but they can help in outlining in a general sense the various equilibria that serve as attractors in a socio-economic evolutionary system, as well as the factors that play a role in bringing the system towards one of these basins of attraction.

The second specific technical issue addressed by evolutionary (policy) models is behavioural heterogeneity. I have already argued that it is not the domain of evolutionary analysis to specify theories of individual behaviour. Instead, evolutionary theories take the population perspective, i.e., they describe the various types of agents that can be found in a population, and the way in which their behaviour may change under the pressure of selection and the generation of novelty.

There are two principal sources of behavioural variation in a population. The first is different characteristics between members of the population. Firms may differ in such dimensions as size, the products they produce, the technologies they use, their location, etc. Consumers may differ with regard to income, their preferences, their physical characteristics, etc. Such differences may induce differences in behaviour. The second source of behavioural variety lies in the notion of bounded rationality. Each individual agent may react differently to similar incentives, even in comparable circumstances. Exactly because individual behaviour is not completely rational (in the neo-classical economists' way), it is rather unpredictable, at least when analyzed from a population perspective.

In actual practice, these two sources of behavioural variety will interact, and it is difficult, if not impossible to separate them in terms of the empirical data that we have available. This is in strong contrast with the theoretical work in evolutionary economics, which has, traditionally since Nelson and Winter (1982), focused on the side of bounded rationality as a source of variety. This focus is at least partly the result of a desire of evolutionary economists to differentiate themselves from neo-classical economists. Critique of the assumption of strong rationality in mainstream neo-classical economics is obviously a cornerstone of evolutionary economic theory. Thus, the existing evolutionary economic models, without a single exception, put a lot of emphasis on variety between agents that results from agents using different rules of thumb, or other decision rules. Variety that is related to differences in agents' characteristics has attracted much less attention.

In my view, this is a tendency that, although it may have merits in a theoretical context, is not very useful for the type of evolutionary modelling perspective that I propose here. In the intermediate range empirical model that I propose, we must arrive at a single, or at most a few, aggregate behavioural patterns by aggregating variety at the micro level. In order to be able to aggregate, we need both detailed data on the differences in characteristics in the population, and information (or an assumption) about variety in behavioural patterns (bounded rationality). In this aggregation process, the idea of fully modelling bounded rationality at the micro level is not very useful, for at least two reasons. The first is that the question of what motivates and drives an individual agent is, in most cases, simply not relevant for the more aggregate population-level outcome. The second is that, under many circumstances, it will be impossible to specify bounded rationality in a different way than by *exogenously* specified varieties of 'rules of thumb'.

As a way out of this, I propose two potential solutions. The first is that we use micro-level (evolutionary) theories to specify a limited number of 'archetypal' patterns of bounded rationality, and link these to different population 'scenarios' in the overall model. As an example, one may derive from a detailed (psychological) theory of consumer behaviour a taxonomy of consumers into 'early and late adopters' (a real-world example would probably have a slightly more sophisticated classification), and link these to a specific fraction of the population to arrive at scenarios for overall population behaviour.

A second approach, however, may exist in simply using a single and rather straightforward assumption about actual bounded rationality in the population. This approach puts less emphasis on bounded rationality as a source of variety, and, instead, relies more on individual characteristics to generate the population diversity. When the single assumption on bounded rationality involves a (stylized) notion of optimizing, this strategy might appear as somewhat alien to the idea of evolutionary dynamics. Nevertheless, I argue that, if properly combined with variety in the characteristics of the population members, even such a simplified 'optimizing' approach can be useful at the level of intermediate range evolutionary models. Specifically, in the example of a modelling strategy that I will discuss in section 2.6 below, I will proceed along these lines, and use an explicit (although short-run) maximizing strategy for the population of adopters in the model.

Either way, these strategies depend to a large extent on the population variety that is generated by different characteristics in the population. Thus, there is, again, an important role for preliminary exploratory research. In this case, this must be aimed at describing, depending on the specific policy problem at hand, the user population, the way in which they may be affected by various factors in the model, and they way in which they may contribute towards moving the economy between equilibria. It can easily be seen that this requires different inputs than the type of foresight studies mentioned above. In this case, statistical information on user populations, as well as psychological, sociological and economic theory is needed to provide an adequate input to the model.

The emphasis on population dynamics in evolutionary models suggests a novel element in policy models in the form of game theory. Recently, so-called evolutionary game theory (Maynard-Smith, 1982) has asked the question under what circumstances novel 'strategies' (behavioural patterns) can 'invade' a population of existing behaviours. The concept of an *evolutionary stable strategy* (ESS) specifies a strategy that cannot be successfully invaded in this way. Possibly, evolutionary stable strategies are mixed strategies, i.e., a situation in which a part of the population plays one strategy, and another part plays a different strategy. Hence, behavioural heterogeneity plays a large role in evolutionary game theory.

The box on evolutionary game theory provides an example that illustrates the notion of an evolutionary stable strategy. The concept is important for the present analysis because it provides an analytical tool to analyze the potential switching between multiple equilibria in the case of heterogeneous populations. If an existing 'constellation' of behavioural patterns can be characterized as an ESS (say, a user population locked-into a particular technology), the prospects of for switching between equilibria may be much more thin than if the existing 'constellation' is not an ESS. In the latter case (not an ESS), a policy of attracting lead users may start a process of more or less automatic switching to a different equilibrium. In the former case (an ESS), policy may be more difficult and involve both coordinated user actions (persuading large amounts of consumers to switch at once, possibly by regulation) as well as specific policies aimed at changing the (relative) pay-offs of the two technologies (i.e., directly intervening at the level of technological innovation).

Box: Evolutionary Game Theory

We use the well-known example of the game of Doves and Hawks to illustrate the main idea of an evolutionary stable strategy (EES) in a repeated game. Imagine a context in which each individual in a population of players repeatedly meets a different individual to fight over a resource (e.g. food). In each meeting (fight), the player can play one of two strategies. The *Dove* strategy is not to fight, the *Hawk* strategy is to fight. If a Dove strategy meets a Hawk strategy, the Hawk takes control of the resource, and no fight takes place. If the two players both use the same strategy, both have a 50% chance of taking the resource, and they both face a cost. The cost is lower when both players play the Dove strategy than when both players play the Hawk strategy. The players do not have any knowledge of their opponent’s strategy before the meeting. A specific numeric example is in the following pay-off matrix.

Pay-off matrix for each meeting

	Opponent’s strategy	
Own Strategy	Hawk	Dove
Hawk	-30	50
Dove	0	5

Now suppose that the whole population is made up of Hawks. Clearly, all meetings will end with a pay-off of –30 for both players. What if, in this situation, one of the players considers switching to a Dove strategy? Obviously, this player will only meet Hawks, and therefore it will always get a zero pay-off (its opponent will have a pay-off equal to 50). Since 0 > -30, switching to a Dove strategy is beneficial for this individual player. What, on the contrary, if the whole population consists of Doves? Then, all meetings will end with a pay-off equal to 15 for both players. The player that considers switching to a Hawk strategy will be able to increase its pay-off to 50.

Clearly then, neither of the pure strategies Dove or Hawk is an evolutionary stable strategy. If the whole population consists of players with an identical strategy, it pays for the individual player to move to a different strategy. But now suppose that the players play so-called mixed strategies. This means that every time they play, they have a (fixed) probability of using a particular strategy.

As an example, let’s assume that all players have a 40% probability of playing Dove, and a 60% probability of playing Hawk. If a player plays Hawk, the expected pay-off is –30x0.6 +50x0.4 = 2, while if it plays Dove, the expected pay-off is 0x0.6+5x0.4 = 2. Thus, the overall expected pay-off is 0.6x2 (Playing Hawk) + 0.6x2 (Playing Dove) = 2. Now let’s consider whether changing this strategy is beneficial for an individual agent. If it is only this agent who changes strategy, it will face the same expected pay-off (since the frequency of strategies with neither its opponents nor the pay-off matrix has changed). Hence there is no incentive for an individual to change strategies.

What if the whole population is playing strategies at a different frequency than the 0.6/0.4? Let’s suppose we have a 0.5/0.5 probability for the two strategies. Then, the expected pay-off for Hawks is –30x0.5+50x0.5 = 10 and the expected pay-off for Doves is 0x0.5+5x0.5 = 2.5. Clearly, the higher pay-off for being a Hawk will induce players to play the Hawk strategy more often, i.e., to increase the frequency. This will remain beneficial until the Hawk frequency is 0.6, at which point the pay-offs of the two strategies even out.

We conclude that in any population that does not apply the 0.6 Hawk frequency, mutant strategies can successfully invade the population. Hence only the 0.6 Hawk frequency is an evolutionary stable strategy.

The application of applied evolutionary game theoretic models seems a promising avenue to investigate these issues for concrete policy situations. There is still a major challenge involved here, because evolutionary game theory is a highly abstract field, in which the theoretical models tend to make highly simplified assumptions about both the strategies that are open to agents (players), and about the degree to which (expected) pay-offs can be measured. Moreover, the default setting of evolutionary game theory, i.e., a context of repeated meetings with multiple other players, may not be very adequate for most problems found in the reality of environmental policy.

Although one of the first applications of evolutionary game theory has been to the 'Tragedy of the Commons' problem (e.g., Axelrod, 1984), which is certainly an environmental problem, it is not easy to see how the context of repeated interactions between multiple players extends to a broader class of environmental problems.

Thus, some of the same arguments that I raised against the use of mainstream economic models in policy analysis, may be valid against evolutionary game theory models. But it is also conceivable that specific (intermediate range) problems may be identified in which evolutionary game theory can be usefully applied. I will therefore recommend in the section below that applied evolutionary game theory becomes a standard element of evolutionary policy thinking. The purpose of applying evolutionary game theory in this way is to apply a wide range of potential behavioural patterns (including those that are hard to imagine for a present-day observer) in a stability test of existing behavioural patterns.

2.6 Some specific ideas on the modelling of transitions

In this section, I will reflect on how the ideas expressed above can be put into practice in terms of developing an actual evolutionary model aimed at supporting policy decisions in the field of environmental analysis. The case I will consider is that of a potential transition towards a hydrogen economy. How exactly this hydrogen economy is defined, and what current systems it will replace, will be defined below.

My argument will be based to a large extent on the work of Taanman (2004), which is summarized by Taanman et al. (2006). This contains a detailed diffusion model of alternative technological trajectories towards a hydrogen economy. The core of this model, so I will argue, may well serve an evolutionary policy model.

2.6.1 The hydrogen economy: policy issues

The hydrogen economy (Rifkin, 2002) is now a much hyped vision of the future of the world's energy system. Central in the hydrogen economy are the *fuel cell* and *hydrogen*. A fuel cell is a piece of machinery that generates electricity in an electrochemical way. It works by separating an electron from a hydrogen atom. Various types of fuel cells exist, and these can be classified both in terms of their technical characteristics (such as operating temperature, the material used for the electrodes, the type of fuel used, etc.), or in terms of their functional characteristics (such as mobile fuel cells, micro fuel cells, etc.). Thus, fuel cells may be used for a range of applications, such as in cars (instead of an internal combustion engine), in houses (for electric heaters, boilers, etc.), or in factories.

The fuel used in fuel cells is either pure hydrogen or some other fuel, such as methane, from which hydrogen is reformed inside the fuel cell itself. Hydrogen must be produced, and this can be done in various ways, some of which are sustainable, and some not. For example, one may produce hydrogen from water using solar power, which would be sustainable, or one may reform hydrogen from fossil fuels such as methane, which does not differ much from the existing methods of using fossil fuels in terms of sustainability.

With the range of different types of fuel cells and their varying applications, as well as the various ways of producing hydrogen, it is clear that hydrogen is a rather flexible way of providing energy. Hence the vision of a ‘hydrogen economy’, i.e., a complete system of production and consumption in which fuel cells and hydrogen are the sole carrier of energy.

I will set the task of formulating a realistic evolutionary policy model that can help us answer the question whether it is likely that the hydrogen economy can really replace the current fossil fuel economy. A secondary question is which particular policy measures can be envisaged to facilitate this transition.

The general way in which I will attempt to tackle this, keeping in mind the ideas expressed above, is to formulate the policy problem as one that typically fits the intermediate range for which evolutionary models can be used, and then to apply the principles of evolutionary analysis, such as a population approach and the (game theoretic) idea of mutant strategies. In general, this approach will imply that we collect and use a lot of specific information about the (future) hydrogen economy, rather than treating it as an abstract vision that can be characterized by a set of general equations.

2.6.2 Problem conceptualization

The set of factors that determine energy production and use can be characterized as a large-scale techno-economic system, with many complementarities. There are several factors that induce path-dependence in these systems. The first is that large-scale specific (infrastructural, but also non-material, e.g., in terms of knowledge) investments are necessary to support the system. A single actor is usually not able to finance these investments. Once in place, these investments represent a vested interest of the established players, which makes them less willing to switch to other technological trajectories. For a system that is challenging the vested interested (e.g., hydrogen), the large-scale investments represent a financial hurdle that is hard to overcome.

The second factor that induces path dependence is the fact that technological progress inside the system is strongly related to learning-by-doing and learning-by-using (i.e., dynamic increasing returns to scale). Hence new systems necessarily have to start at relatively low levels of productivity. Only by actually being implemented and used can productivity of the system grow. But with a more mature system in place, a new system may never reach levels of productivity that are competitive vis-à-vis the established system.

Theoretical work on competition between these technological systems has been presented by, among others, Arthur (1994) and David (1975). Models representing these processes have usually been formulated as dynamic models with multiple equilibria (e.g., Arthur et al., 1983). Each technological system is represented by one equilibrium (path). Depending on where the system starts, it locks-in to one of these equilibria. Once the lock-in has occurred, it is hard for the system to select a different equilibrium.

Thus, the specific problem area that I have chosen can be seen as one of multiple equilibria, lock-in and competition between (large-scale) technological systems. I have argued above that such a situation of multiple equilibria is a potentially good case of an intermediate range problem that can be successfully tackled by evolutionary models. I will consider the specific policy problem as one of potential transition from the current mode of energy production and use, towards one in which hydrogen and fuel cells play the central role. Obviously, this context is closely associated to current debates in environmental analysis and policy. Obviously, a large body of existing literature exists on this topic, much of which is applied to the specific Dutch policy context (e.g., Hoogma et al., 2002, Kemp and Loorbach 2005).

2.6.3 Preliminary field work

As I have argued above, I see an important task for technology foresight studies, as well as a general engineering understanding of a particular technology in the modelling process. The main purpose of this type of analysis is to outline the possible configurations of the equilibria in the process that is being modeled. This includes both the existing energy system (based on fossil fuels) and the system of which we wish to investigate the probability of transition, i.e., the hydrogen economy.

It is obviously beyond the scope of this paper to present a complete assessment of this type. I will therefore suffice by giving some general directions that this preliminary analysis should take.

For the existing energy system of fossil fuels, two major problems exist. The first is CO_2 emissions. This is increasingly seen as a large-scale problem by policymakers because of the greenhouse effect. However, since the greenhouse effect is politically disputed, commitment among policymakers to reducing CO_2 emissions is still not complete. Potential solutions for the CO_2 problem exist both within and outside the fossil fuel energy trajectory. Within, technological innovation may reduce emissions for a given amount of energy produced, or CO_2 may be captured and stored in a less harmful way.

Outside the fossil fuels trajectory, the hydrogen economy is a potential source of complete reduction of CO_2 emissions. The fuel cell itself does not produce any CO_2 or other harmful waste. Hydrogen can be produced both by using fossil fuels and without doing so. In the first case, CO_2 is produced, although it can potentially be captured and stored.

The second major challenge for the fossil fuel trajectory is the increasing scarcity of fossil fuels and the associated rise in energy prices. This cuts both at the supply side and the demand side. In terms of supply, it is true that the historical record shows a long history of discovery of ever-more amounts of oil and natural gas. But still, we know that these reserves are finite, and the day will that the reserves are so small that fossil fuels will become a too valuable resource to be the raw material for a global energy system. On the demand side, the industrialization and development of large countries such as China, India and Brazil is already putting pressure on oil prices, and this affects the traditionally developed countries.

Hence three major factors about which we need to form some kind of foresight in order to characterize the equilibrium development path of the current energy system based on fossil fuels, are the effects of scarcity of fossil fuels, the expected benefits (mainly in terms of CO_2 emissions) of technological innovation, and the societal attitude towards CO_2 emissions and the greenhouse effect. These foresights must be operationalized into three model variables/parameters: the future development of oil prices, the expected rate of reduction of CO_2 emissions in the use of fossil fuels, and the expected social pressure towards reducing CO_2 emissions.

For the hydrogen energy system, a detailed outline of technological possibilities and the technological efficiency that can be expected for each of them must be constructed using foresight techniques. Taanman (2004) focuses on the use of fuel cells for micro-cogeneration (*micro-warmtekrachtkoppeling*). This means that the fuel cell produces electricity and heat at the same time, and hence can be used to supply in the need for electricity and heating in buildings (both residential and non-residential). This means that the model in Taanman (2004) does not consider the complete hydrogen economy, but his modelling strategy can be applied to the more general case by replicating the model for other uses of fuel cells (e.g., automobiles).

Within the micro cogeneration application, based on an outline of foresight studies, Taanman distinguished different technological options. These differ in three dimensions. First, whether the electricity demand or the heating-demand is leading. If electricity demand is leading, the fuel cell is switched on when electricity is demanded (e.g., when the resident switches on the light). In this case, heat is produced as a by-product, stored and used when needed. In the case where heating demand is leading, the fuel cell is switched on when heating is demand and electricity is produced as a by-product. In this case, electricity can be supplied back to the electricity network, fetching a price paid by the electricity company.

A different dimension in which technologies differ is the way in which hydrogen is produced. The crux here is that the fuel cell is installed in the building itself, and hence the hydrogen needs to be available at the local level. This can either be produced centrally and transported to the locality by means of a new pipe system, or by mixing hydrogen with natural gas, for which an extensive transport infrastructure exists in the Netherlands. Taanman only considers the latter case, but his modelling strategy can also be applied to a completely new infrastructure. Hydrogen can also be produced locally (at the are level), in which case missing is not necessary. The costs of installing a local hydrogen production and distribution infrastructure differ between existing buildings and newly-built areas.

By combining the various options, a number of technological clusters can be formulated (e.g., heat-demand following fuel cells, decentrally produced hydrogen). Note that the options in the Taanman (2004) model are not exhaustive, because a number of the technological options has been fixed. Besides the choice for micro cogeneration, also a choice has been made to consider only the production of hydrogen from fossil fuels (natural gas). Again, such a choice has implications for the specific outcomes (sustainable hydrogen production is more expensive, certainly in the shorter run), but the general modelling strategy could easily be maintained even if the type of hydrogen production is considered as an additional dimension in the technological domain.

For each of the technological clusters, a number of parameters, such as technologically efficiency, specific infrastructural costs, etc. must be formulated. Obviously, since this is essentially a foresight analysis, the parameter sets must take into account variability of these expectations and investigate the sensitivity of the outcomes for this variability.

A different part of the model for which detailed data must be collected is the (potential) user population. Since Taanman focuses on micro cogeneration, his units of observation on the user side are buildings (and the people who inhabit or use them). He starts from a detailed description of existing buildings in the Netherlands, distinguishing different type of residential buildings (e.g., detached, semi-detached, corner, terraced, apartment), as well as different commercial buildings (e.g., shops, factories, agricultural). He then constructs three different typical ‘areas’, which consist of a specific mix of these types of buildings. These four types are urban, semi-urban and rural. Each of the types of areas can be existing or newly-built.

Using recent data from Statistics Netherlands, the user population (buildings) is described, and a set of projections is made for how each type of area will grow in the period until 2050. Obviously, projected population growth and planning policies are the main ingredients in these projections. Again, variability in the projections is important.

2.6.4 Model elements and some examples of results

The model constructed by Taanman (2004) consists of an aggregated set of adoptions decisions in the population at the level of the three areas (urban, semi-urban, rural). The adoption decisions are assumed to depend only on an economic problem: the costs of using hydrogen and fuel cells are compared to traditional ways of heating and electricity-generation. Hence, each member of the population (i.e., an ‘area’) will compare the costs of using the fuel cell with traditional ways of heating and buying electricity from a supplier in the market. When the costs of using hydrogen fall below those of the traditional system, the area is assumed to adopt the hydrogen system. Obviously, this decision will differ between new entrants in the population (i.e., newly-built areas) and existing areas, because of the differences in costs. Generally, the newly-built areas will adopt earlier than existing areas.

It is obvious that this particular way in which the model is formulated involves an assumption of optimizing behaviour at the micro-level, even if the optimization problem is specified in relatively simple terms (e.g., no aggregating of benefits over longer time horizons, or discounting of future benefits). This goes back to the discussion, in the previous section, about the sources of variety at the population level. As was proposed in the previous section, I am willing to accept this stylized description of consumer behaviour, as long as it is complemented by a sufficient level of variety generated by population characteristics. The Taanman model obviously includes this source of variety, although in a larger scale application one would probably want to elaborate this aspect more.

In this way, diffusion curves can be generated. Table 2.1 summarizes some of the results for the example of decentralized hydrogen production and electricity-demand-following fuel cells. The table documents the year in which the model produces a take-off of hydrogen use (i.e., the year in which the first user adopts), and the year in which the complete population has adopted. The results are produced for a standard set of parameters, in which two parameters are varied: the price of electricity (which is a competitor for the hydrogen system) and the price of natural gas (which is both a competitor, and an important input into the price of hydrogen itself).

The results, which are given for illustration of the general model outcomes only, clearly show that the two prices have a substantial impact on the results. For low electricity prices, the hydrogen economy does not take off before 2050, but for high electricity prices, the take-off is predicted to take place soon (5 years from now). The price of natural gas has less of an impact, but even here the variation from ½ to 1½ can make a difference of 22 years (in terms of the year in which complete diffusion is reached).

Table 2.1 Adoption years (take off – complete diffusion) produced by the model for different energy prices

Scale factor price of natural gas	Scaling factor price of electricity		
	½	1	1½
½	Hydrogen does not take off before 2050	2021 – 2027	2010 – 2012
1		2025 – 2034	2011 – 2015
1½		2032 - 2045	2013 - 2018

In line with the discussion above, we cannot take these results as predictions of what happens in the real-world evolutionary system in which the transition towards a hydrogen economy may (or may not) take place. We should take them as broad indications of the feasibility of a ‘hydrogen-equilibrium’ in the context of the multiple equilibria energy sub-system of the global (Dutch) economy. Rather than the end-result of a modelling analysis, they should be taken as the beginning or a more elaborate process.

2.6.5 Use of model results

The Taanman (2004) model is not an extremely realistic description of the actual evolutionary system of energy transitions. I prefer to call it an evolutionary model because it adheres to several major principles in the evolutionary theory of social and economic change, such as the existence of multiple equilibria, the modelling of the user side by means of a population approach in which heterogeneity plays a major role, and close interaction between the model and more qualitative foresight techniques (i.e., the use of detailed information about technological and other forecasts).

Before policy is actually evaluated using the model, I propose that the model results are qualified in a broader analysis of adoption dynamics using some of the principles of evolutionary game theory. This may seem odd, since the context of the model (the adoption, once and for all) of a fuel cell for micro cogeneration, does not resemble the context of multiple and repeated interactions between two players that we see in evolutionary game theory. Despite this, I think that the principles of evolutionary game theory can help us investigate the robustness of the behavioural patterns that the model predicts.

As was explained above, the model only considers economic decisions, based on user costs. In practice, costs are undoubtedly an important ingredient in adoption decisions, but they are far from the only ingredient. The existing literature on transitions (e.g., Hoogma et al., 2002) has outlined many of these factors, but without being able to provide a precise quantitative interpretation of how the different dimensions of the decision process interact and compare to each other.

Factors that cannot easily be put into a cost-benefit calculation (e.g., perceived safety, related to the explosive nature of hydrogen, or the desire to contribute to a cleaner environment) have been described in this literature, and one of the notions that has emerged is that of niche management. The idea, in a nutshell, is that users are heterogeneous (as they are in the Taanman model), and that some specific users are more willing to adopt than others because they have a specific reason to do so. A group of these users are called a niche. A niche may exist because a specific characteristic of the user group makes their benefits especially high (i.e., a group of ‘heavy users’), or because they have strong inter-group imitation dynamics, or even because they like to use a specific artifact because of different reasons than the actual beneficial impact that policymakers are interested in (e.g., care drivers that are interested in fuel cell-driven cars because of driving characteristics rather than environmental considerations.

Evolutionary game theory may help us model the behaviour of these niche groups in more details, and may be employed to answer the question under which circumstances adoption may actually take place within a niche, and under which circumstances the niche adoption may ‘spill over’ to the broader user population. Although this game, at least in the case of durable goods such as a fuel cell, is not played repeatedly, we may take the existence of a large population as a way of justifying the use of average pay-offs for each strategy that we can define at the individual level. Thus, instead of arguing that each individual player repeatedly plays the same game, and hence has the average pay-off associated with the frequency of strategies by its opponents (see the box on evolutionary game theory above), we can assume that the population pay-off is a weighted average of non-repeated decisions.

What would be needed for the use of evolutionary game theory in this way would be a specification of how behavioural patterns are mitigated through a user population. Such a ‘game theory plus’ approach may be usefully applied to investigate how sensitive the outcomes of the pure economic model above are for a more realistic set of behavioural alternatives.

With this information, of which I cannot here give any specific empirical indication since I have not developed the game theoretic tools to implement it, the model results could be related to various policy options. This would most usefully take the route of identifying policy objectives that would be necessary to achieve a certain goal (e.g., ‘how much increase in fuel cell efficiency do we need to reasonably achieve a take-off of the hydrogen economy in 2015?’), and then to find the policy instruments that can contribute to this goal. As long as the policymaker makes a realistic assessment of the potential of policy instruments, the whole traditional range of instruments can be applied towards this goal. Thus, both policies aimed at users (the selection environment) and at technology (trying to pick winners in the generation of evolutionary variation) are good candidates for an evolutionary policy. I see no reason at all why an evolutionary approach to policy would necessarily be restricted to either selection or variation generation. On the contrary, only a combination of various policy instruments is likely to achieve the necessary effects.

References

Arthur, W.B. (1994). *Increasing Returns and Path Dependence in the Economy.* University of Michigan Press, Ann Arbor

Arthur, W.B., Yu. M. Ermoliev and Yu. M. Kaniovski (1983). A Generalized Urn Problem and its Applications. *Kibernetika* 19 (1983): 49-56, and *Trans. in Cybernetics* 19: 61-71

Axelrod, R. (1984), *The Evolution of Co-operation*, Basic Books

Barro, R. J. and X. Sala-i-Martin (1991). Convergence across States and Regions. *Brookings Papers on Economic Activity*: 107-158

David, P.A. (1975). Technical Choice, Innovation and Economic Growth: Essays on American and British Experience in the Nineteenth Century, Cambridge: Cambridge University Press

Fontana, W. and L. W. Buss (1994). What would be conserved if ‘the tape were played twice’? *Proceedings of the National Academy of Sciences of the USA* 91: 757-761

Hoogma, R., Kemp, R. and J. Schot (2002). *Experimenting for Sustainable Transport, The Approach of Strategic Niche Management*, London: EF&N

Hughes, B.B. (1999). *International Futures: Choices in the Search for a New World Order*. Boulder (CO), Westview

Kemp, R. and D. Loorbach (2005). Dutch Policies to Manage the transition to Sustainable Energy. *Jahrbuch œkologische œkonomik 4 Innovationen und Nachhaltigkeit* (pp. 123-150). Marburg: Metropolis Verlag

Krugman, P. (1990). *Rethinking International Trade*. Cambridge MA, MIT Press

Lucas, R. (1976). Econometric Policy Evaluation: A Critique. *Carnegie-Rochester Conference Series on Public Policy* 1: 19-46

Maynard-Smith, J. (1982). *Evolution and the Theory of Games*, Cambridge: Cambridge University Press

Merton, R. M. (1973). *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago, University of Chicago Press

Monod, J. (1970). *Le Hasard et la Nécessité*. Paris, Seuil

Nelson, C. R. and C. I. Plosser (1982). Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications. *Journal of Monetary Economics* 10: 139-162

Nelson, R.R. and S.G. Winter (1982). *An Evolutionary Theory of Economic Change*, Harvard: The Belknap Press

Rifkin, J. (2002). *The Hydrogen Economy*. Penguin, Putnam

Sargent, T.J. (1994). *Bounded Rationality in Macroeconomics*, Clarendon Press

Taanman, M. (2004). *Diffusie van Waterstoftechnologie. Modellering van de inzet van waterstof voor micro-WKK in Nederland*, MSc thesis Technische Universiteit Eindhoven, Faculty of Technology Management

Taanman, M., De Groot, A., Kemp, R. & B. Verspagen, (2006). *Diffusion paths for Micro Cogeneration Using Hydrogen in the Netherlands*, UNU-MERIT Working paper 2006-025, Maastricht

3 Environmental innovation and policy: lessons from an evolutionary model of industrial dynamics

M. Saint-Jean

Institut Fédératif de Recherche sur les Dynamiques Économiques (IFREDE), Bordeaux IV University, Bordeaux, France.

Contact: saintjea@u-bordeaux4.fr

ABSTRACT

The paper is aimed at examining the impacts of particular policy options (emission standards and procurement policy) on clean trajectories developed by firms in an industry. An evolutionary model of industrial dynamics is used to explore such impacts in the long term on the market structure and the innovation output of firms. The peculiarity of the present model lies in the coevolution of two populations of agents (buyers and suppliers) so that the dynamics of the industry is governed by the coevolution of technology, user requirements and market structure. The first section introduces environmental innovations and the underlying evolutionary mechanisms in order to stress the specificity of environmental innovations and what are the main forces at stake to be introduced in a model examining clean technology development. Section 2 presents one specific model of industrial dynamics that we have developed to grasp such clean technology development. Section 3 examines the impact of different policy options on the firms' trajectories and the market structure. Computer simulations show that the rise in the environmental requirements of demand, generated by tighter environmental standards, has different impacts according to the nature and timing of the standards. Regarding procurement policy, simulations show the existence of a critical mass of clients that value environmental characteristics of the product able to impulse a dynamics of innovation that induces a change in paradigm of the whole firms in the industry. Lessons for energy policy are finally drawn.

3.1 Introduction

Environmental quality requirements as demanded by regulation and consumers represent new challenges for firms in their technology choice. Firms are led to innovate taking into account these new environmental performance criteria. As a dynamic and interactive process, innovation is determined by many factors like for example the intensity of competition, the prevailing costs and demand conditions, appropriability conditions, technological opportunities and the knowledge-base and absorptive capacity of firms. Thus an analysis of environmental innovations should not neglect the complexity of the determinants influencing innovative behaviour of firms. Given that environmental innovations result from learning and research activities of firms, they can be examined in the light of the evolutionary theory of technological change. Indeed, evolutionary works shed light on innovation and learning processes, with particular emphasis on diversity, selection and path-dependency.

From a policy perspective, it is also very important to take into account the context in which firms operate, the type of pressures they have to face and the way they interact for example in the supply chain. Given the number of factors potentially at stake, modelling is tool worth considering, in order to disentangle the effects of policy instruments on the dynamics of an industry.

Based on a model of industrial dynamics we examine different policy options like tighter emission standards and procurement policy. The model is used to exhibit the discriminating role of such policy options within buyer-supplier relations. The aim is to give guidance to the conditions of dynamic efficiency of policy instruments taking into account the coevolution of technology, user requirements and market structure.

This paper is organized as follows. Section 3.2 introduces environmental innovations and the underlying evolutionary mechanisms. Section 3.3 presents one specific model of industrial dynamics and section 3.4 examines the impact of different policy options on the firms' trajectories and the market structure.

3.2 Environmental innovations: their specificity and the underlying evolutionary mechanisms

This section aims at underlying what are the differences (if any) between environmental innovations and more 'traditional' innovations.

3.2.1 The double externality problem

Environmental innovations produce positive spillovers in both the innovation and the diffusion phase. This specificity is linked to the component of public good that characterises environmental goods. Positive spillovers in the diffusion phase are due to smaller external costs compared to competing goods on the market. This peculiarity, which is called the double externality problem, reduces the incentives for firms to invest in environmental innovations (cf. Jaffe et al., 2005). Therefore the double externality induces a second peculiarity which is the importance of the regulatory framework as a key determinant for environmental innovations (see for example Jaffe et al. (1995), Porter and van der Linde (1995), Brunnermeier and Cohen (2003), or Cole et al. (2005)).

Thus, the main difference with innovations in general is that the incentive for firms to develop, or to adopt, environmental innovations comes from the regulatory pressure. Thus in comparison with innovation in general -which is mainly driven by technology push and demand pull mechanisms- we have to add a third factor which is called by Rennings (2000) the 'regulatory push-pull effect'. Nevertheless environmental innovation is not a simple and direct response to regulation. As a matter of fact, there are many other factors that govern environmental innovations (and innovation in general), in particular the intensity of competition, the prevailing costs and demand conditions, appropriability conditions, technological opportunities and the knowledge-base and absorptive capacity of firms. Thus an analysis of environmental innovations should not neglect the complexity of the determinants influencing innovative behaviour of firms (cf. Kemp et al. (2000) or Jänicke et al. (2000)).

Moreover, even if the relationship between invention, innovation and diffusion is not different for environmental innovations than for other innovations, one peculiarity of environmental ones is that a large diffusion is always socially desirable. This feature is also due to the public-good character of environmental innovations. Consequently, environmental policy is also needed in the diffusion stage to support the adoption of environmental-friendly innovations.

Environmental policies rely largely on a regulatory normative approach by which public authorities impose some environmental objectives on the private actors. In general, regulation suffers from an inability to take into account the specificity of individual firms, and for this reason will generally not lead to the lowest cost solution. Modern regulation, however, such as in the context of the European acidification strategy or the national emissions-ceilings directive, often tries to take account of economic constraints such as investment cycles, available abatement technologies, and so on, in order to limit the burden for the regulated industries. Unlike market-based approaches, regulation does not give firms incentives to outperform whatever standard is set for them. Nevertheless regulation may be the preferred choice when it is necessary to avoid 'hot spots' of local pollution, or when it is imperative that a particular objective be met exactly. For a detailed discussion on the links between environmental policy instruments and innovation, see for example Fischer et al. (2003), Hansen et al. (2002) and Requate (2005).

Depending on the way regulation is designed, based for example on technology-forcing or performance standards, possibilities for adaptation will be different for firms. Indeed, technology-forcing standards set the technical means to reduce environmental impacts of some production activity whereas performance standards set environmental objectives to be reached within a given deadline, at the risk of incurring financial penalty, but giving enterprises the choice of how to achieve environmental targets. By setting up performance standards, public authorities 'force' firms to undertake significant effort so that the resulting innovation output could serve to formulate emission limit values, thus taking into account the progress made. Such standards play an important role in convincing companies that regulatory authorities will force laggards to react and consequently they are conducive in establishing a climate where firms consider innovation an integral part of their response to the global objectives set by the regulator.

Concepts from evolutionary economics enable to specify some features of innovation that are also particularly relevant for environmental questions. First, the change from one paradigm to another underlies a problem of transition that turns to be of particular importance when considering the shift towards new more sustainable techno-economic systems (transport, energy, agriculture, etc.). Second, path-dependency plays a major role in generating inertia and self-reinforcement in existing technological trajectories, that can result in a lock-in, i.e. the temporary dominance of an inferior technology. In this context, not only small events can have major effects on the technological development but also cumulative processes may lead to the persistence or emergence of suboptimal technologies and thus inhibits the transition towards another. Third, periods of stability are temporary and alternate with phases of turbulence where the transition to an alternative technological path is more easily accomplished, giving rise to windows of opportunity, i.e. the temporary existence of circumstances that allow novelty to get selected. Using a window of opportunity, innovative ‘entrepreneurs’ have a better chance of influencing the long-term direction of technological, economic or social development than during periods of stability outside the windows.

3.2.2 Policy dilemmas

The development of environmental innovations raises policy dilemmas in particular because they take place in a context of technological competition. Evolutionary theory emphasises that in such a context the market may select a suboptimal technology as increasing returns to adoption lead to a process of technology selection which is path-dependent, and to a large extent, irreversible and unpredictable. The works of Arthur (1989) and David (1985) put forward that technological substitution is not warranted, even when a superior technology is introduced in the market, because the incumbent technology benefits from increasing returns to adoption.

As underlined by Sartorius and Zundel (2005) regarding environmental innovations, two kinds of technological competition (old versus new and new versus new) deserve to be distinguished according to the ability of competing technologies to meet similar functions or rather different functions meeting different demands. Since the solution of an environmental problem generally defines a new function that matches the corresponding demand (set for example by a new regulation), several technologies executing this function happen to compete with each other on the level ‘new versus new’. But the new environmentally improved technologies have also to fulfil the genuine function of the established technology they are supposed to replace and this gives rise to a competition ‘old versus new’. Depending on the case, the circumstances behind the opening of windows of opportunity are different. In the first case (new versus new) the window is open in the early stage of competition between technologies developed for corresponding purposes. Increasing returns to adoption are typical determinants of this kind of window. In the second case (old versus new) the window typically tends to open if the investment cycle of an old technology comes to an end and new promising technologies are available at that time. However, the competition ‘new versus new’ cannot be analysed alone, but needs to be seen in close relation to the competition ‘old versus new’. Indeed, the window is the result of mutual interaction between the new technologies and their established counterpart. Whatever the case, the concept of window of opportunity leads to put into light time critical events and the importance of choosing the right point in time for government action. But it also raises the dilemma between the prolonged maintenance of technological diversity on the one hand and the realisation of economies of scale and scope on the other. All in all the coexistence between a strong process of exploitation of the existing paradigm and a process of exploration of various alternative technologies is the core of the issue on the relationship between environmental innovation and policy.

3.3 A model of industrial dynamics

The purpose of the model is to represent technological trajectories of firms that are guided by several dynamic forces such as path-dependency, market selection and supply-demand coevolution. Such model incorporates fundamental properties associated with the innovative process (Nelson and Winter, 1982; Dosi et al., 1988). First, the 'normal' patterns of technological change tend to follow 'trajectories' defined by specific sets of knowledge and expertise. Second, major discontinuities in the patterns of change are associated with changes in technological paradigms. Third, technical progress exhibits strong irreversibility features meaning that over time the dominant process of change implies improvements along the trajectories. So the model depicts technological trajectories that result from innovative efforts of firms whose investment choices may be dedicated to environmental performance. Innovation activities take place within the boundaries defined by paradigms with specific opportunities for improving environmental performance. For an empirical illustration, see Belis et al. (2004) or Oltra and Saint-Jean (2005). Questions to be answered are thus: is market selection able to favour the development of cleaner technologies? Under which conditions can cleaner technologies be developed and diffused?

3.3.1 Method and experimental settings

The model we have developed belongs to a wider family of models of industrial dynamics that integrate an evolutionary approach based on micro diversity of behaviours and the bounded rationality of decision processes. Such models also place central importance on a wide variety of devices for co-ordinating and shaping interaction between agents. The relevant evolutionary outcomes are emergent phenomena arising from processes of interaction. An emergent phenomenon is one that arises from the coordination of the activities of agents and creates a pattern of order not contained in the intrinsic properties of those agents. In such a framework, competition is not seen as a state of equilibrium induced by a particular market structure. Rather it is viewed as a dynamic process that depends on how the micro diversity of firms' behaviours results in changing market positions and on how competitive advantage is defined and leads to particular patterns of change.

Our modelling explicitly takes into account buyer-supplier relationships. The model deals with a population of rival suppliers in interaction with a population of industrial clients. On the one hand, suppliers modify the characteristics of their product thanks to R&D investments so as to adapt to demand pressures and to acquire competitive advantage. On the other hand, clients' requirements evolve so as to adapt to technological changes and modifications of industrial structures. Environmental pressures are synthesized with supply chain pressures (e.g. price, product quality). Figure 3.1 gives a schematic view of the basic interactions between supply and demand we have considered in an industry subject to environmental regulation.

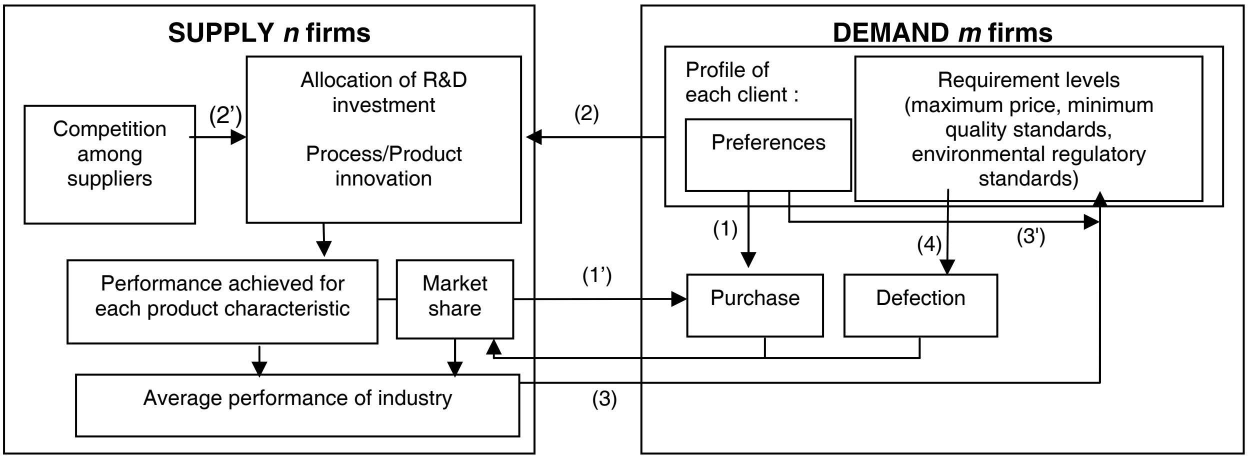

Figure 3.1 Supply-Demand Interactions:

Buyer preferences (arrow 1) and the overall product performance achieved by each supplier (arrow 1') determine the purchase decision. These elements are used by clients in their evaluation of suppliers. When suppliers are selected, they are given some information on the current profile of their buyers (arrow 2). This information will be used to direct R&D investments of suppliers. Additionally, competition among suppliers to get dominant position on the market will influence R&D allocation (arrow 2'). Innovative activities carried out by suppliers lead to improved average industry performance, which will modify requirement levels of clients (arrow 3), in proportion to the priority they assign to the considered characteristics (arrow 3'). Such requirement levels intervene in the decision to leave a supplier at the time of product replacement (arrow 4). In case of non-compliance with client requirements, there is defection and the client will turn to the other suppliers still active in the market. In the case of satisfaction, the client will renew the purchase among the same supplier.

R&D investments and innovation activities of suppliers

Each supplier devotes a part of its turnover to R&D activities. The R&D budget of firm *i* at time *t* is given by the following equation:

$$RD_{i,t} = \mu_i.P_{i,t}.B_{i,t} \qquad (1)$$

with $P_{i,t}$ the price of the product *i* at time *t* and $B_{i,t}$ is the install base of firm *i* at time *t*. μ_i stands for the fraction of turnover allocated to R&D. The install base of firm *i* represents the stock of clients that use the product *i*. By basing R&D expenditures upon the user stock, firms can get a rather stable R&D investment.

Let denote $X_{i,t}^h$ the performance level performed by firm *i* for the characteristic *h* at time *t*.[3] The amount of resources allocated to the characteristic *h* is given by:

$$RD_{i,t}^h = \delta_{i,t}^h.RD_{i,t} \qquad (2)$$

---

[3] In the following, X^1 corresponds to the productive efficiency, X^2 to the product performance, X^3 to the environmental quality at the process level and X^4 to the environmental quality at the product level.

where $\delta^h_{i,t}$ represents the rate of R&D investment dedicated to the improvement of characteristic *h*. For empirical evidence concerning environmental R&D, see Scott (2003). We assume that

$$\sum_{h=1}^{4} \delta^h_{i,t} = 1$$

The research level achieved for the characteristic *h* is:

$$R^h_{i,t} = \gamma . RD^h_{i,t} + (1-\gamma).R^h_{i,t-1} \tag{3}$$

This research level reflects the progressive contribution of the resources dedicated to R&D to the knowledge base of firms. It adaptively evolves to account for learning in knowledge production activities. γ is a parameter determining the speed to which the research level is adjusting to the current R&D budget dedicated to the characteristic.

Innovation is treated as a stochastic process and a two-step procedure is used to determine the innovation output. The first step determines if there is success or not. The second step consists in determining the increase in the new performance that results from the innovation. Thus, for each characteristic, the probability of the value improving depends on the R&D resources allocated to it:

$$\Pi^h_{i,t} = \pi_1 / (\pi_2 + \pi_3 . \exp(-\pi_4 . R^h_{i,t})) \tag{4}$$

π_1, π_2 and π_3 are the limiting parameters of the logistic function. The parameter π_4 determines the speed at which the maximal probability is approached. According to (4), the R&D returns are successively characterised by increasing and decreasing returns.

In the case of success, the innovation output is determined by a Cobb-Douglas-function that depends on the R&D budget invested on the characteristic ($R^h_{i,t}$), the cumulated experience on this characteristic ($E^h_{i,t}$) and the distance to the technological frontier that prevails for this characteristic ($X^{h\max}$):

$$\Delta X^h_{i,t} = \eta_0 . (R^h_{i,t})^{\eta_1} . (E^h_{i,t})^{\eta_2} . (X^{h\max} - X^h_{i,t-1})^{\eta_3} \tag{5}$$

η_0 is a scale parameter. Parameters η_1, η_2 and η_3 respectively reflect the intensity of R&D impact, of the experience and of the saturation of technological opportunities upon the magnitude of improvement of the characteristic. We assume that $\eta_1 + \eta_2 + \eta_3 = 1$. Hence innovation is a cumulative and firm-specific process.

The experience variable is itself subject to progressive adaptations according to the following equation:

$$E^h_{i,t} = \lambda . (MaxE \times R^h_{i,t}) + (1-\lambda) . E^h_{i,t-1} \tag{6}$$

According to this equation, experience depends on past experience and on the current research level achieved for the characteristic. However the accumulation of experience is limited by a maximum value (MaxE). λ represents the coefficient that weights the experience potential achieved thanks to the current research level on the characteristic.

The product price

The product price is determined by applying a mark-up rate over the production costs. Since our model does not consider explicit production factors, productivity gains that suppliers get by investing in R&D to improve their production process can be used as a proxy for the decrease in production costs. By doing this, we assume an inverse relation between the productive efficiency (the characteristic identified by X^1 in our formulation) and the price. The equation for the price is thus simply given by the following function:

$$P_{i,t} = (1+\theta_i).(1/X^1_{i,t-1}) \tag{7}$$

with θ_i the mark-up rate of firm *i*.

Technology space

Technological and environmental opportunities are represented by considering two paradigms (cf. Figure 3.2): paradigm 1 with low environmental potential and paradigm 2 with high environmental potential. The first potential represents conventional production practices with low opportunity along the environmental dimension whereas the second technological space offers higher opportunities since it is based on radically new production practices that reduce pollution at the source.

We assume that the switch carried out by a firm in the paradigm with high environmental potential leads to the following effects:

- A shift in the frontier achievable on the dimension ‘environmental quality of process’ ($X^{3\max2}$), the frontier on the dimension ‘productive efficiency’ remaining unchanged ($X^{1\max2}$). However, we also consider that a threshold exists in terms of productive efficiency ($X^{1\max1}$ in figure 2) gained by firms that evolve in the first paradigm. The experience variable (*E*) will decrease when the first limit ($X^{1\max1}$) is crossed and higher opportunities will be available ($X^{1\max2}$).
- A drop in the product performance (X^2).
- A decrease in the cumulated experience (*E*).

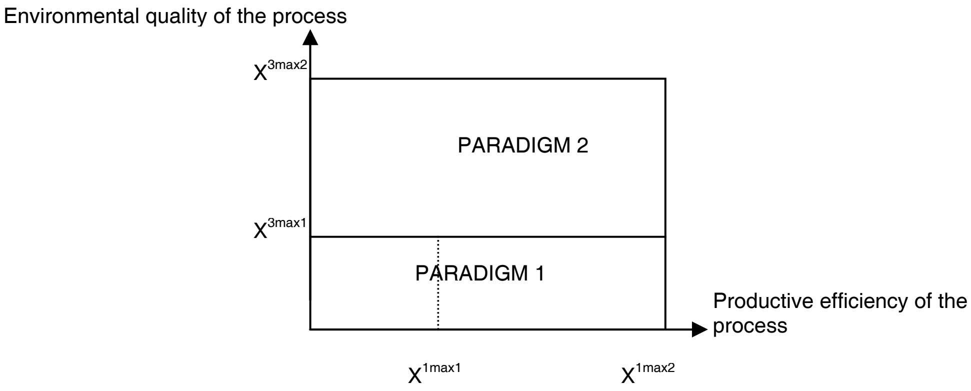

Figure 3.2 The space of technological and environmental opportunities in the model

Decision rules of clients

Three decisions concern the clients: the purchase of a product, the moment to replace the product and the decision to leave or keep the same supplier.

a) Purchase

Each client *j* is characterised by a probability of buying the product *i* as follows:

$$\Pr oba^{i}_{j,t} = (1/P_{i,t-1})^{x^1_j}.(X^2_{i,t-1})^{x^2_j}.(\widetilde{X}^3_{i,t-1})^{x^3_j}.(\widetilde{X}^4_{i,t-1})^{x^4_j}.(MS_{i,t-1})^{e} \tag{8}$$

with x^1_j, x^2_j, x^3_j and x^4_j the preference that client *j* attributes to the corresponding characteristics. These preferences reflect the positioning that client *j* adopts in the final good market. The features of the positioning are reflected in the weight assigned to each characteristic by the client. In order to limit the range of these parameters, we set them so that their sum equals to 1. Parameter e expresses the intensity of 'bandwagon effect' that a supplier with high market shares may exert on clients. We consider random draws for X^3 and X^4 when buyers have to read the score of these characteristics and choose which product to buy (the tilde refers to that). This is related to the fact that buyers have limited capabilities to perceive environmental quality of production activities or products. The perceived value comes from a draw in a normal law with average, the level achieved by the supplier for the considered characteristic, and with standard deviation (σ), a parameter calibrating the error degree of environmental quality evaluation.

So clients take into account the characteristics they perceive from the suppliers but also the suppliers' market shares. Market shares in the purchase decision exhibit two things. First, in a context of uncertainty and imperfect information, an important source of information comes from previous users. An agent who wishes to buy a new product will tend to refer to the choice made by the other users in the past. Market shares give such an indication. So, mimetic behaviour on the user side can help to choose among alternative suppliers. Second, links between suppliers and clients for product innovation require mechanisms allowing positive outcomes for both parts. The implementation of these mechanisms and the costs they incur may be so strong that they reduce transaction alternatives and entail high switching costs. So market shares in the purchase decision enable to account for switching costs that would prevent a client to choose alternative suppliers.

Market share of firm *i* at period *t* is determined with its install base, i.e. with its actual users' stock:

$$MS_{i,t} = B_{i,t} \Big/ \sum_{i=1}^{n} B_{i,t} \tag{9}$$

b) Product replacement

Each client replaces its product after *T* periods, with *T* settled randomly between 1 and 10. Under this hypothesis, clients need a certain period of use before buying a new product, which is generally the case with intermediary or equipment goods. By doing this, a distinction is made between the client stock (the install base) and the current sale flow of a supplier since we assume that clients have different purchase rhythms and thus do not renew the product at each time period. In this case, each supplier has to manage the user stock of its own product even if, for example, its current sale flow is zero.

c) Defection

In order to account for the defection/voice decision, the following procedure is formalised: if all the requirement levels of client *j* are satisfied, then there is continuation of the relationship with the current supplier *i*; else, defection occurs and a new supplier is chosen among those remaining on the market (cf. purchase procedure a).

At the time of product replacement, the client compares the performance achieved by its current supplier for each characteristic and its own requirement levels. If all the requirement criteria are fulfilled, then the client keeps the same supplier and renews its purchase among it. In this case, the corresponding supplier records a sale ($N_{i,t} = +1$) while its install base remains unchanged. On the contrary, if the client leaves its supplier, this is because the client is not satisfied with the supplier's performance on at least one of the four characteristics with regard to its minimum requirements. In this case, the corresponding supplier loses one client ($B_{i,t} = -1$).

Inter-firm interactions

Inter-firm interactions involve three mechanisms: the transfer of information from demand to supply, the updating of R&D investment allocation and the evolution of requirement levels.

a) Transfer of information

Two types of data are used to guide supplier's R&D allocation:

- The product characteristics that are both a priority for the clients and represent a source of technological lead for the supplier. Formally, priority characteristics are those endowed with the highest weight. Characteristics with a weight greater or equal to 95% of the maximum weight are considered to be priorities. To be one of the leaders, we assume that a firm has to reach a performance level greater or equal to 95% of the highest performance on the considered characteristic. In total, a positive score ($Z_{i,t}^h = +1$) is given to characteristics that meet these two conditions.
- The product characteristics for which supplier performance is inferior to that required as a minimum by clients and which are likely to cause defections at the end of the product use period. Formally, a negative score ($W_{i,t}^h = +1$) is registered for the characteristic with a performance level below the one required by the client.

b) Evolution of the allocation of R&D investment of suppliers

On the base of the information gathered by the supplier *i* during each purchase cycle, the allocation of R&D among the characteristics is updated. Let $RDIndex_{i,t}^h$ denote the R&D index for the characteristic *h* that is used in the evolution of R&D rates. It is subject to progressive adjustments according to the following equation:

$$RDIndex_{i,t}^{h} = (1-\alpha).RDIndex_{i,t-1}^{h} + \alpha\left((\beta.Z_{i,t}^{h} + (1-\beta).W_{i,t}^{h}) \Big/ (\beta.\sum_{h} Z_{i,t}^{h} + (1-\beta).\sum_{h} W_{i,t}^{h}) \right) \quad (10)$$

where $Z_{i,t}^{h}$ represents the positive score assigned by the supplier *i* to the characteristic *h* when this characteristic is both a priority for its current clients and a source of technological lead compared to its direct competitors. $W_{i,t}^{h}$ represents the negative score assigned by the supplier i to the characteristic h when the performance achieved on this characteristic is above the requirement level of its current clients. α is the speed to which R&D index adjusts to information raising from the sales performed by each supplier. β stands for the relative importance attributed to the positive indicators compared to the negative ones. R&D indices are then normalised such that:

$$\delta_{i,t}^{h} = RDIndex_{i,t}^{h} \Big/ \sum_{h} RDIndex_{i,t}^{h} \quad (11)$$

c) Evolution of minimum requirements of clients under the influence of technological advances in the industry

The requirement levels evolve through time according to average performance of industry and to the importance attributed by the clients to the considered characteristic. The following equation gives the dynamics of requirement level assigned by the client *j* to the characteristic *h*: We only explicit here the formal equation of the dynamics of requirement levels for the quality characteristics (X^2, X^3 and X^4) which are expected to increase. The equation for the maximum price is slightly different and adapted so as to take into account the decreasing updating of its level.

$$levelX_{j,t}^{h} = levelX_{j,t-1}^{h} + \varepsilon.(\max(0, x_{j}^{h}.(\overline{X}_{t}^{h} - levelX_{j,t-1}^{h}))) \quad (12)$$

with $\overline{X}_{t}^{h}$ the average performance of the industry for the characteristic *h* weighted by suppliers market shares, i.e.

$$\overline{X}_{t}^{h} = \sum_{i=1}^{n} MS_{i,t}.X_{i,t}^{h}.$$

Parameter ε represents the difference share between average performance and requirement level that is passed on the evolution of aspiration levels, x_{j}^{h} the preference (weight) attributed by the client *j* to the characteristic *h*.

Exit process

Exit of suppliers occurs both when the install base is equal to 0, i.e. the user stock of the supplier is exhausted, and when the sales are equal to zero for a minimum of four periods. In this case, the current turnover does not permit any R&D investments and innovation. In the model, there is no entry of new suppliers or new clients.

3.3.2 Results and limits

The model incorporates mechanisms that give the industry the capacity to self-organise, i.e. to cause a global structure to emerge that did not exist previously and that may appear as the outcome of various interactions between agent populations. A self-organisation process such as this implies that an economic system can, by starting from the same initial situation, evolve differently depending on the uncertainties of its history. This leads us to emphasise which possible scenarios of industrial dynamics are likely to be generated under the set of assumptions we have considered. Such scenarios are used to explore the impact of tighter standards upon the trajectories of firms and upon the market structures.

The structure of the model is such that no analytical solution exists and only simulation trials can enable us to infer the properties of the model. The inductive analysis of the properties of industrial dynamics aims to explain the patterns of interaction between firms and the characteristics of evolution of the industry. Some precautions have to be taken however in order to stress emergent properties and regularities of the industrial dynamics and thus to guarantee some robustness of the results.

The reference configuration

a) Initialisation and simulation trials

Initial values of parameters and variables are presented in detail in the appendix (tables 1 and 2). We have considered a population of 12 suppliers interacting with a population of 200 clients. Each client makes a purchase during the first period and then renews the purchase of the product after *T* periods with *T* settled randomly between 1 and 10. Each simulation run comprises 500 iterations. Time series analyses conduced over more than 500 periods showed that the industrial structure converges on an asymptotic state characterised by a high level of market concentration after 500 periods.

Suppliers are initially identical so that the differences likely to emerge from the dynamics of the system result from the competition process and from their specific interactions with the set of clients. Whatever the characteristics, the initial technological level is set to 1. Each characteristic is endowed with an upper limit that represents technological constraints. We assume the coexistence of two paradigms differentiated in terms of potential for improving environmental quality of the process (cf. Figure 3.2).

Two distinct groups of clients are considered. Differences across both groups come from different sets of preferences and different willingness to pay. The first group of clients (G1) strongly weight environmental characteristics whereas economic characteristics such as price and product performance are weakly weighted. The inverse is assumed for the second group of clients (G2). This group is dominant in the market and initially represents 80% of demand.

Group 1 is supposed to have a relatively high maximum price, consistent with the assumption that this type of clients is willing to pay a higher price for products meeting environmental criteria. Group 2, which pays great attention to price and product efficiency, has a relatively low maximum price. The point is that a client that wants to adopt an environmental positioning on the final good market, will consequently make stronger demands on environmental characteristics than on price to its suppliers since it will allocate a greater budget in order to buy such a product. This is why weightings and maximum price are set in consistency with the strategic positioning of each type of client.

The requirements in terms of minimum performance of product are assumed to be identical for both groups of clients. As to the requirements in terms of minimum environmental performance, these are the same for both groups since they are enforced by regulatory authority and apply homogeneously to the industry. However, given the dynamics we have considered, the various requirement levels evolve differently across the client groups in accordance with their sets of preferences and with the average progress of the industry.

Given the stochastic character of the dynamics, one simulation alone does not prove the existence of an emergent property since the system is characterised by several random processes. Furthermore properties are obtained under specific initial conditions and parameter values. Therefore, in order to guarantee some robustness of the results and to stress the regularities of the industrial dynamics, a high number of simulation runs and a sensibility analysis of parameters need to be carried out.

As we are focusing on the impacts of environmental standards on industrial dynamics in this paper, we have limited the analysis to one set of parameters, the so-called reference configuration. The results come from 50 series of simulation, of 500 periods each. But in order to better observe the forces behind the system dynamics, we have also examined individual series of simulation that characterise the most likely scenarios and summarise the behaviour of the simulation model. These individual series will be used to study the impact of standards. Finally, to implement the model, we have used the programming system LSD (Laboratory Simulation Development) developed by M. Valente at the IIASA[4].

b) Two emergent market structures

Two alternative types of industrial structure emerge in the long term:

- *A concentrated structure constituted by firms with a specialisation on price and product performance.* The firms that survive after 500 periods are characterised by a high R&D investment in the economic performance of processes (productive efficiency) and products (product performance). These firms are labelled PROD firms. They succeed in complying with environmental regulation. However these firms do not change paradigm. This scenario emphasises the emergence of a monopolistic or oligopolistic structure, with firms oriented toward the improvement of price competitiveness and product performance and evolving within the paradigm with low environmental potential. Over a battery of 50 simulation runs, 26 series correspond to this scenario. We call it *scenario HO* for homogeneous oligopoly.
- *A concentrated structure where a 'green' market niche coexists with a low number of dominating firms characterised by a technological lead on price and product performance.* The 'green' market niche results from the survival of a firm characterised by a high environmental R&D investment that enables it to change paradigm, but also by a high price and a low product performance. This scenario emphasises market segmentation with the emergence of a green market niche dominated by a firm that mainly orients its R&D activities

---

[4] A complete report of the model, including equations and computer programming of the model, can be provided on request.

toward the improvement of environmental quality – which we call ENVI firm – and that evolves in the paradigm with high environmental potential. This niche coexists with (at least) one PROD-type firm specialised in price and product performance which dominates the market. Over a battery of 50 simulation runs, 20 series corresponding to this case have been observed. Over a battery of 50 simulation runs, we have observed four specific cases which were characterised in the long run by a domination of firms specialised in economic performances and that had changed paradigm. The low frequency of apparition of this type of situation led us to focus on the analysis of the two other more frequent cases. We call it *scenario MS* for market segmentation.

Scenario HO exhibits 'design dominant' features. Indeed, if we refer to the Abernathy-Utterback model of the innovation life cycle (Abernathy and Utterback, 1975), the first phase is characterised by a market convergence to a single design or 'dominant design'. An industry shake-out occurs and a phase of process innovation can start. The final phase is characterised by market stability with the leading firms maintaining their position through incremental innovation. Scenario HO corresponds rather well with this sequence of events which results in unassailable market position. On the contrary, scenario MS is characterised by a bipolar structure which makes the dominant design compatible with a market niche. This configuration is consistent with the analysis made by Windrum and Birchenhall (1998).

Although we do not develop this section further, in order to concentrate on the standard impacts, a detailed examination of both structures, based on individual series of simulation, has highlighted important forces behind the dynamics of the system (Saint-Jean, 2005). In summary, the intensity of competition that prevails across the leader firms on the most demanded characteristics, the capture of a leader group of green users and the early change of paradigm tend to condition the emergence and survival of the green market niche.

Limits

Three types of limits need to be underscored: methodological problems related to simulations in general; lacks in the specific model presented in the paper; gap between evolutionary modelling and environmental innovations. These limits may pave the way for further lines of research.

Main methodological problems[5] related to simulations are:

- The stochastic characteristic of the dynamics. This involves that only one simulated history of the system does not enable to validate the existence of an emergent property. Moreover the properties result from particular initial conditions and specific parameter values.
- The high number of parameters. This makes it fastidious to explore the whole space of initial conditions. Sensitive analysis may thus focus on a subset of relevant parameters for the phenomena under consideration.
- The empirical calibration of the model. Simulation then enables to test the validity of a theoretical model on the basis of a calibration of parameters upon empirical data. Such

---

[5] For a detailed discussion on methodological issues raised by simulation, see e.g. Lane (1993), Ostrom (1988), Valente (2005).

utilisation of simulation requires a great amount of consistent data for the phenomena under study.

The model we have developed does not grasp a certain number of economic phenomena that are yet important to apprehend innovation processes, like for example:

- no sectoral differences are taken into account;
- there's no real price strategies of firms;
- effective financial constraints do not apply;
- the role of final consumers is not explicitly incorporated;
- no new innovative entrants are considered.

Regarding environmental innovations, our model fails to properly grasp the following aspects:

- the anticipation of environmental regulation by firms and its impact on firm's innovation strategy;
- the issue of 'transition management' consisting of 'a deliberate attempt to bring about structural change in a stepwise manner' (Kemp and Rotmans, 2001) stresses an orientation towards system innovation. Such system innovations are at the core of many sustainability issues such as the one affecting energy, transport, chemistry or agriculture.

Bearing these limits in mind, we use simulation as an explorative tool to follow step by step the complex dynamics that result from many interactions between entities of the model. The heuristic dimension, introduced by exploratory simulation, enables to focus on the interactions of forces that are assumed to drive the phenomena. The model we have developed needs to be resituated in such exploratory perspective.

3.4 Policy implications

The model is used to study the impact of emission standards and the role of procurement policy on the industrial dynamics. Lessons will then be drawn for energy policy.

3.4.1 The role of stricter pollution standards

Initialisation

We consider four different cases depending on the nature of the standard, i.e. depending whether its application concerns the process or the product, and on the timing of intervention. Two periods (100 and 200) have been considered so as to take into account the differentiated stability situation of the industrial structure at these dates. At time 100, whether the scenario HO or MS is being considered, the industrial structure is characterised by a relatively low degree of concentration and the firm's specialisation is not yet established. Thus, the use of tighter standards at this date can be considered as occurring relatively early. On the contrary, at time 200, the industrial structure appears to have stabilised in both scenarios HO and MS. We examine the impact of tighter standards that are twice as much as those initially enforced. This means a level of 3 for process standards and a level of 4 for product standards.

In the model we present below, requirements of environmental performance enforced by regulation and transmitted through demand on suppliers of intermediary goods are represented by two variables: *PerfMinIp* and *PerfMinIm*. We assume that initially these pressures are exerted homogeneously. But they are likely to evolve heterogeneously under the differentiated impulse of clients' demand. The variables *PerfMinIp* and *PerfMinIm* which both characterise demand of environmental quality in our model play a role at three different levels:

- They justify the client's decision to leave a supplier when they replace the intermediary good. The number of defections is proportional to the requirement levels of environmental quality since clients will only choose stable relationships with suppliers when they have achieved regulatory compliance.
- They orient R&D activities of suppliers toward environmental R&D. Indeed, suppliers record which characteristics are underdeveloped regarding the minimum requirements of their buyers. This is taken into account in the R&D activities of suppliers.
- They contribute to raise the levels of environmental quality that are required by the group of clients with high sensitivity toward environmental protection criteria. Indeed, the variables *PerfMinIp* and *PerfMinIm* change according to the average environmental performance of industry weighted by market shares and according to clients' preferences. The effect of environmental quality demand on the adjustment function of requirements depends on the firm's activities of environmental innovation and on the sensitivity of clients toward environmental characteristics.

In the reference configuration, we have considered the following initial values for *PerfMinIp* and *PerfMinIm*: 1.5 and 2 respectively, similar for the whole set of clients.

The impact of emission standards at the product level

a) HO scenario

In the HO case, the introduction of a tighter environmental standard at the product level leads to an increase in industrial concentration in the long term (cf. Figure 3.3). The introduction of the standard leads, through demand requirements, to a transitory increase in the average rate of R&D investment dedicated to environmental quality of the product. An increase such as this leads to improve the environmental performances of the industry (cf. Figure 3.4 and Figure 3.5) proportional to the rise in environmental quality requirement levels of the whole clients. Though transitory, the reallocation of R&D budget toward the environmental characteristic of the product gives an advantage to those firms with competence cumulated in this field. They can reach the new requirement levels faster. The tightening of regulation mainly contributes to increase market opportunities for firms with strong environmental competencies. These firms can perform some product differentiation on the environment and so they can keep away from strong price competition. The introduction of the standard also operates a selection among the leader firms on price, that target clients with a high sensitivity on price but a low one on the environment.

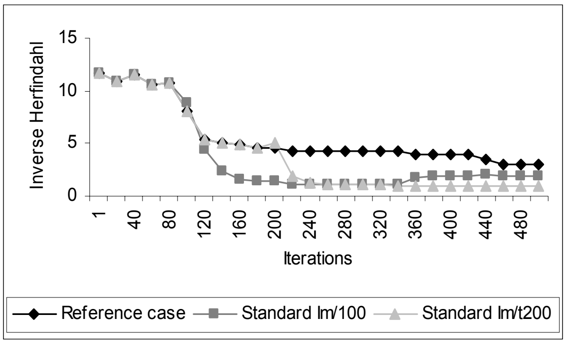

Figure 3.3 Impact of tighter product standards for scenario HO (inverse Herfindahl index)

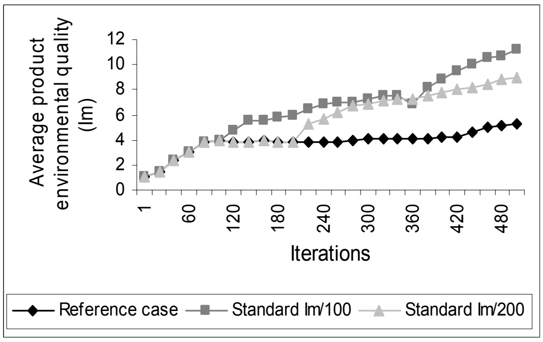

Figure 3.4 Impact of tighter product standards for scenario HO (average product environmental quality)

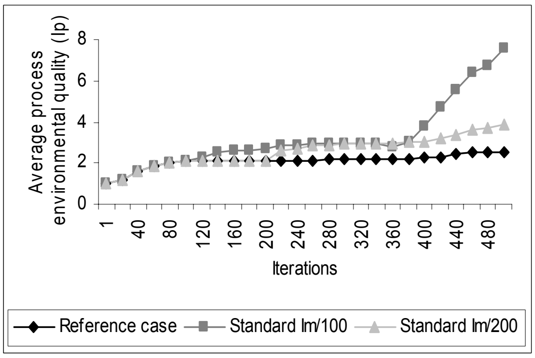

Figure 3.5 Impact of tighter product standards for scenario HO (average process environmental quality)

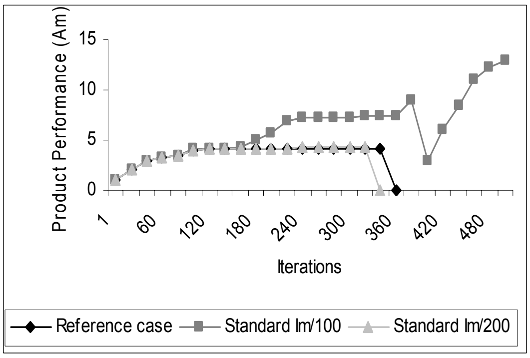

Figure 3.6 Impact of tighter product standards for scenario HO (product performance)

We note that the product standard favours paradigm change of firms since they succeed in improving the environmental quality of the process along with the product one. In this case, the product standard can represent a lever of action sufficient to allow not only an improvement of the product environmental quality but also a significant improvement of the process environmental quality.

Simulation results show that the introduction of standard at time 100 makes it possible for an ENVI-type firm to survive in the long run. On the contrary, the application of standard at time 200 is too late to allow such market differentiation. Figure 3.6 illustrates the impact of a tighter product standard upon trajectories of product performance (Am) of an ENVI-type firm. The standard enforced at time 100 leads the ENVI-firm to improve the product environmental quality as well as the product performance, which means innovation offsets. Such innovation offsets result from an efficient combination between economic and environmental characteristics of the product.

In the HO case, we conclude that the early introduction of a product standard gives the possibility of acting before the emergence of a dominant design which is driven by PROD-type firms, i.e. firms betting on price competitiveness and product performance, and which would represent a lock-in into technological trajectories characterised by low environmental content.

b) MS scenario

In the MS case, the introduction of a tighter environmental standard at the product level leads to a relative decrease in the degree of industrial concentration in the long run (cf. Figure 3.7). This is due to the survival of imitating firms which follow the leader firms specialised on environment and price. The decrease in concentration allows heterogeneous environmental performances to develop. Figure 3.8 represents the variation coefficient for the global (product and process) environmental performance of firms over time.

We note that the increase in the diversity of environmental performances is higher for a standard at time 100 than for a standard at time 200. Such increase results from the rise in clients' requirements regarding product environmental quality, which implies a reallocation of R&D investment toward environmental characteristics. Improvements in product environmental quality (Im) depend on the firm's competences that have already been accumulated at the time of standard intervention. In the long run, the increase in diversity leads to higher product environmental quality (cf. Figure 3.9) in comparison to the reference case and also to higher process environmental quality (cf. Figure 3.10), in particular when the standard is applied at time 100 rather than at time 200. However, in the long run the level of diversity is lower since firms reach the limits of technological potential faster.

The main beneficiaries of a tighter product standard are the followers, i.e. those firms able to achieve a particular global performance high enough to attract and keep clients with higher requirement levels. The presence of imitators is associated with a decrease in the market shares of the leader firms. Thus the situation of local monopoly achieved by leader firms thanks to their specialisation tends to be questioned. However the leader firms continue to benefit from a first-mover advantage.

We note that ENVI-type firms are the only ones to change paradigm. In other words, whatever the timing of intervention, a tighter product standard has no effect on the paradigm change of PROD-type firms.

In the MS case, we conclude that the introduction of a tighter product standard allows the survival of follower firms that lie behind the specialised leaders and is associated with an increase in the diversity of environmental performances. However, the resulting increase in product environmental quality does not simultaneously lead to an improvement in the process environmental quality that could be high enough for PROD-firms to experience a paradigm change.

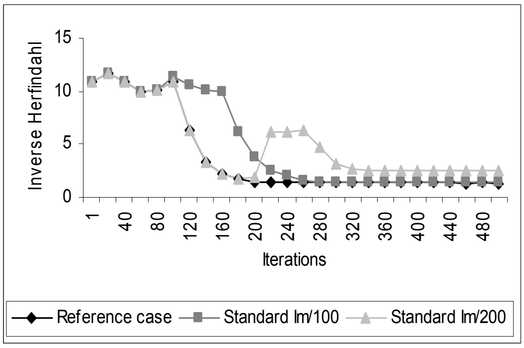

Figure 3.7 Impact of tighter product standards for scenario MS (inverse Herfindahl index

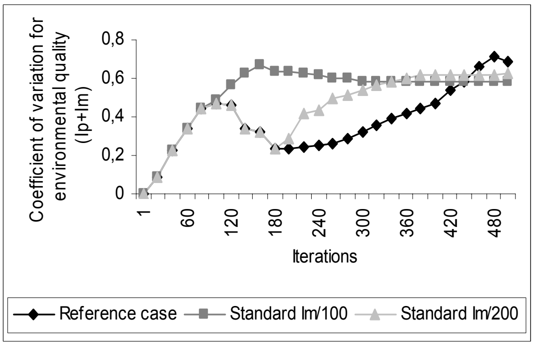

Figure 3.8 Impact of tighter product standards for scenario MS (coefficient of variation for environmental quality)

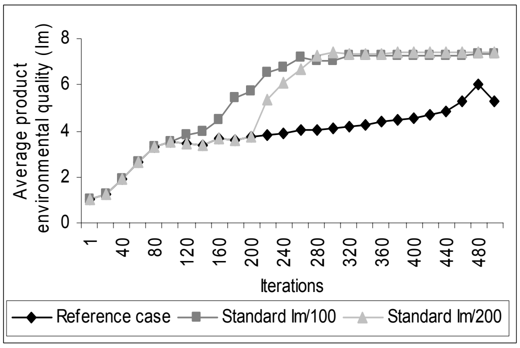

Figure 3.9 Impact of tighter product standards for scenario MS (average product environmental quality

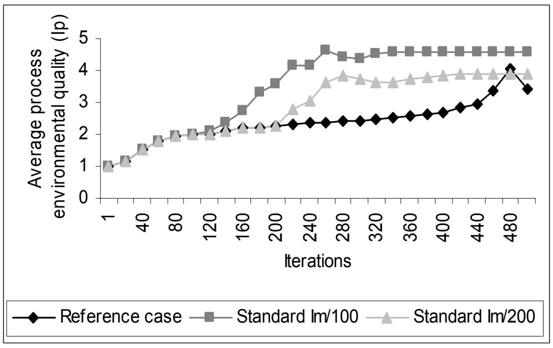

Figure 3.10 Impact of tighter product standards for scenario MS (average process environmental performance)

The impact of emission standards at the process level

a) HO scenario

The introduction of a tighter environmental standard at the process level leads to an increase in the market shares of ENVI-type firms in the long run. Indeed the cumulated market shares of ENVI-type firms during the last period is 4% when the standard is applied at time 100 and 23.5% when the standard is introduced at time 200 whereas this share is zero in the reference case.

The introduction of the standard has few effects on the degree of concentration. However, it significantly affects the characteristics of clean technologies. The introduction of the standard at time 100 leads all firms, especially the PROD-type ones, to change paradigm, which is associated with an improvement of the average environmental quality of the process (Ip) (cf. Figure 3.11). Competition thus takes place between firms with different specialisation but all progressing in the green paradigm. The late application of the standard at time 200 strongly punishes the firms that are below the new regulatory requirements and that do not possess the sufficient capabilities to close the gap. This induces a progressive decline of their market shares and leads them to exit the market. However, the late introduction of a process standard does not enable all firms to change paradigm. Indeed, some PROD-type firms locked in the first paradigm are able to survive in the long run even if their market shares tend to decline in the last periods.

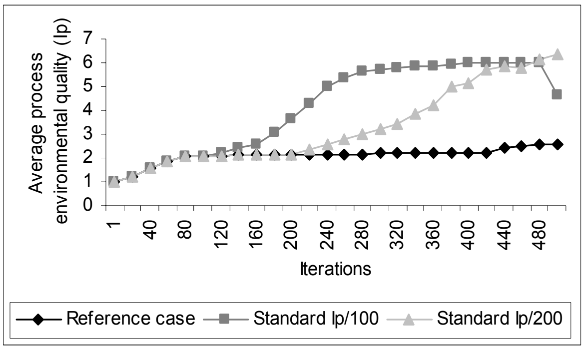

Figure 3.11 Impact of tighter process standards for scenario HO (average process environmental performance

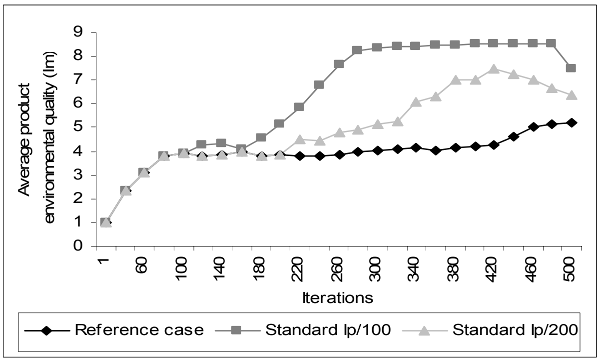

Figure 3.12 Impact of tighter process standards for scenario HO (average product environmental quality)

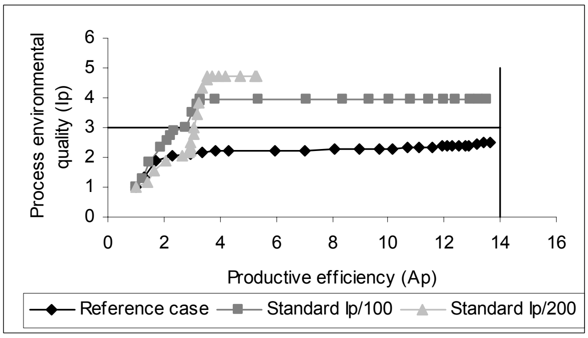

Figure 3.13 Impact of tighter process standards for scenario HO (process environmental performance in relation to productive efficiency)

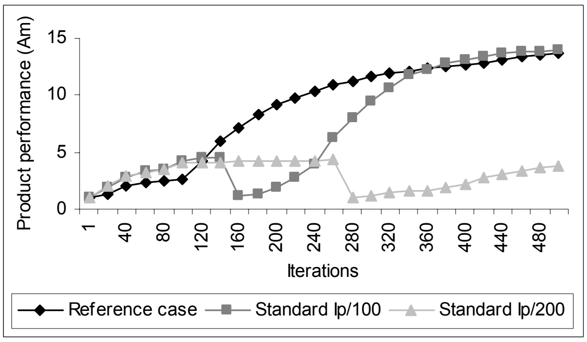

Figure 3.14 Impact of tighter process standards for scenario HO (product performance)

From Figure 3.12, we note that the tightening of the process standard has a positive effect on the environmental quality of the product, in particular if the standard is introduced early. Figure 3.13 hows the impact of the process standard on the evolution of the process characteristics of a PROD-type firm. We note that the process standard leads the firm to change paradigm. When the standard is applied at time 100, the productivity level of the firm is close to the upper limit set in the green paradigm. On the contrary, when the process standard is introduced at time 200, the productivity level is lower in the long run but it is associated with a higher process environmental quality. Figure 3.14 shows that the process standard at time 100 also leads the PROD-firm to reach a very high level of product performance. Thus, the late introduction of a process standard leads to a decrease in the economic performance of PROD-firms since they are forced to invest in environmental R&D to meet the user's needs.

In the HO case, the impact study of a tighter process standard raises the following conclusions:

- The early application of the standard leads to the paradigm change of all firms and contributes to the emergence of a market niche. However the survival of the market niche is weakened by the price competition of rival firms in the green paradigm. The standard also leads to a strong improvement in the average environmental quality of the product in the industry.
- The late introduction of the standard occurs when the firm's specialisation is well established and competition is strong between PROD-type firms. The application of the standard jeopardises the PROD-type firms that have not accumulated strong environmental competencies. This explains their difficulty to quickly reallocate their research activities

towards the improvement of process environmental quality and their resultant lag in complying with the new standard.

b) MS scenario

Simulation results show that the late introduction of a tighter process standard leads to an increase in the average process environmental quality (cf. Figure 3.15) that leads to firms shifting towards the green paradigm.

Figur 3.16 shows that there is a higher impact on the environmental quality of the product if there is a late introduction of the standard compared to introduction during a non stabilised stage of the industrial structure.

The introduction of the standard at time 200 contributes to increased market opportunities for firms achieving intermediate economic performance (price and product quality) compared to the specialised leaders (cf. Figure 3.17).

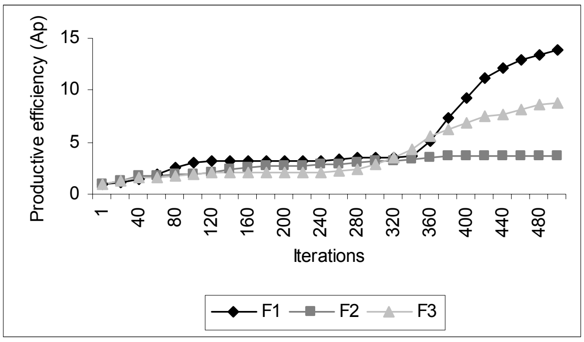

Figure 3.15 Impact of tighter process standards for scenario MS (productive efficiency)

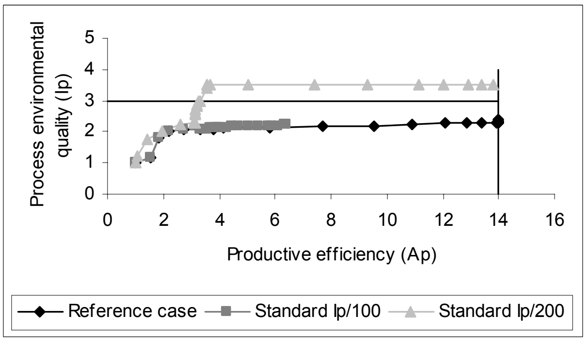

Figure 3.16 Impact of tighter process standards for scenario MS (process environmental quality)

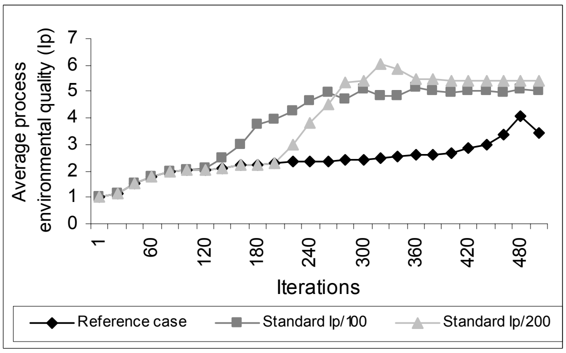

Figure 3.17 Impact of tighter process standards for scenario MS (average process environmental quality)

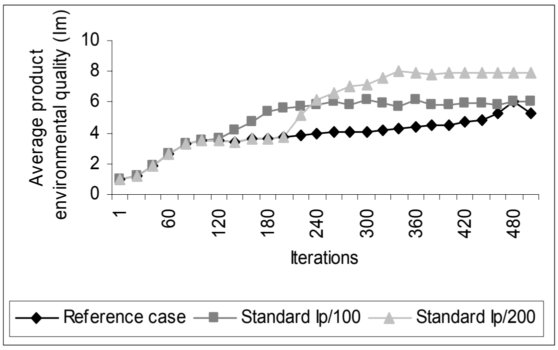

Figure 3.18 Impact of tighter process standards for scenario MS (average product environmental quality

On the contrary, the process standard at time 100 initiates strong instability in the supplier-user relationships due to a low differentiation in the supply of firms. This prevents buyers from selecting appropriately the suppliers and prevents suppliers from exiting the market. In the simulation, only one ENVI-type firm succeeded in differentiating and achieving a relatively stable market share of 25%.

The remaining market is shared by the other firms. In spite of the high requirement level of process environmental quality enforced at time 100, only two firms succeed in changing paradigm in the long run. In fact, the increase in client demand for process environmental quality – resulting from the new standard – contradicts the predominant pressure of clients on price. From period 100, competition tends to be centred on product performance (Am) thanks to strong demand and to low minimum requirements of all clients. This gives rise to an increase in the average rate of R&D investment dedicated to this attribute.

The higher level of competition on the product performance increases client requirements on this attribute, which limits all the more the possibilities for reallocating R&D towards other characteristics. In this case, the R&D level allocated to the environmental characteristic of the process is insufficiently high to innovate on this dimension, which thus jeopardises the possibilities to progress on this dimension.

Figur 3.18 represents the trajectory of a leader PROD-type firm in the three cases considered. It illustrates that progress on productive efficiency is limited by the introduction of the standard at time 100. In this case, the maximum level reached at the last period is 6.41 for a process environmental quality of 2.22. On the contrary, when the standard is applied at time 200, the leader PROD-type firm is not only close to the upper limit prevailing for this dimension (14) with a score of 13.81 but it also achieves a level of process environmental quality of 3.51, i.e. it succeeds in changing paradigm.

From an environmental point of view, the increase in process environmental quality requirements resulting from the late application of the standard is more efficient as it enables higher levels of environmental and economic performance to be reached. In other words, in the MS case, the late enforcement of a process standard, i.e. in the stabilisation stage of the industrial structure, leads to innovation offsets.

Discussion

From this set of results, we draw the conclusion that the rise in the environmental requirements of clients, generated by tighter environmental standards, has different impacts according to the nature and timing of the standards:

- A tighter product standard enables a greater increase in the average environmental quality of the product if it is enforced early rather than late. The product standard has also a positive side-effect on the process environmental quality. In particular, if an exclusive dominant design emerges on the market because of strong competition between PROD-type firms, the early application of the product standard leads to a shift in paradigm for firms.
- A tighter process standard enables an increase not only in the average process environmental quality but also in the average product one. The early application of the standard tends to be more efficient in the case of an homogeneous oligopoly dominated by PROD-type firms.
- On the contrary, in the case of a market segmentation characterised by the emergence of a green market niche the late application of a tighter process standard allows higher levels of environmental and economic performance to be reached.

Finally, the results exhibit that in the scenario of an exclusive dominant design, independent of the type of standards, it is important to act relatively early before the specialisation of leader PROD-type firms has stabilised, which allows firms to take action before the lock-in into a technological path with low environmental content. In the scenario of coexistence of a dominant design and a green market niche, it is important for the product standard to be implemented prior to the process standard in order to enable the followers to survive and to encourage innovation offsets for firms. In such cases, emission standards may prevent both a situation of lock-in on the supplier side and a situation of behavioural inertia on the user side. Standards may thus enable a preservation of certain forms of technological and behavioural diversity.

3.4.2 The effect of procurement policy

The model is used to analyse the impact of an increase in the number of ‘green’ clients. The aim is to determine whether a threshold exists beyond which all the firms in the industry are driven to change paradigm. Such a threshold would justify for example the role of procurement policy to generate an impulse effect toward the development of environmental innovations.

Considering different initial proportions of green clients, Figure 3.19 represents the average cumulated market shares -over a battery of 50 simulation runs- of firms that have changed paradigm at the last period (t=500). Standard deviation for each series is also reported.

We observe that, for an initial proportion of 25% of green clients, the market is characterised in the long term by a significant increase in the cumulated market shares of firms that have changed paradigm. On average, cumulated market shares amount to 48%. However, the dispersion as measured by the standard deviation is relatively high compared to other configurations. For initial proportions of green clients comprised between 25% and 70%, cumulated market shares of ‘green firms’ increase but less than proportionally and with a decreasing dispersion with the initial percentage of green clients. Beyond an initial proportion of 70% of green clients, we note that the totality of surviving firms in the long run have succeeded in changing paradigm.

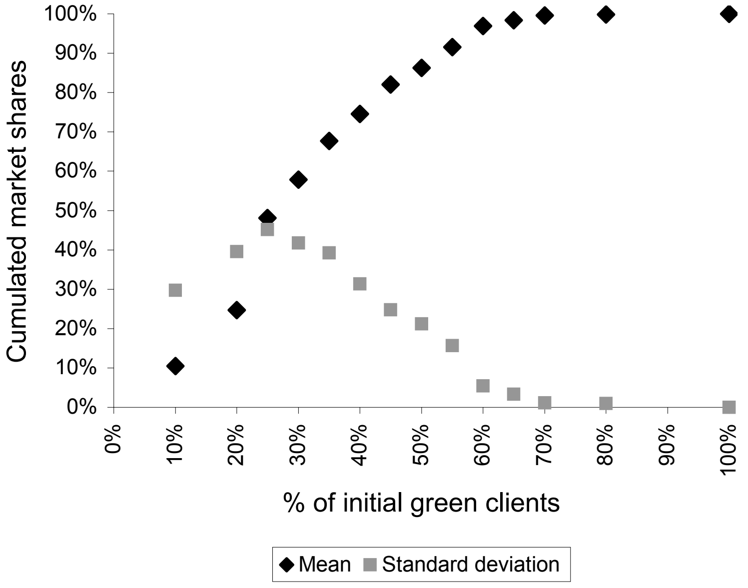

Figure 3.19 Cumulated market shares at the last period (t=500) of firms that have changed paradigm for different initial proportions of green clients

Such result suggests that the orientation of technological trajectories of firms toward the green paradigm needs a strong increase in the clients sensitive to environmental quality. If environmental requirements from clients tend to generalise then firms will be forced to re-orient their R&D investments toward environmental quality improvements. Such environmental research activities will lead to increase environmental performances at the industry level and thus to increase environmental requirements of the green clients beyond initial levels imposed by regulation. Such process urges the competing firms to invest in environmental R&D, which contributes to knowledge and experience accumulation in the field and to increase the firm’s probabilities of environmental innovations. All in all, firms succeed in achieving environmental performances at the process level above the superior limit set in the first paradigm via the increase in environmental requirements of the green clients. Finally, a critical mass of 70% of clients that value environmental characteristics of the product is able to impulse a dynamics of innovation that induces a change in paradigm of the whole firms in the industry. Such result suggests that demand of environmental quality needs to be very strong to reach such purpose, which tends to raise the limits of public action with this only tool.

3.4.3 Lessons for energy policy

One main specificity of environmental innovations is to be ‘regulatory push-pulled’. So policy actions are essential to drive environmental innovations along particular performance directions. For specific issues related to energy technologies and policy, see for example Menanteau (2000) and Norberg-Bohm (2000). Several aspects can be outlined here:

- It is important for firms to be able to anticipate the implementation of new regulation and to adapt more or less proactively depending on various factors such as the firm's bet and strategy, the stage in the investment cycle, the technological and market opportunities, the nature of the knowledge base. This conjunction of factors stresses the interactions between the determinants of industrial dynamics and the ability to comply with and/or to anticipate the environmental policy. Thus what does matter is not only the date of application of the policy but also the time lag given to (or bargained by) the actors before the policy measure comes really into force i.e. the regulatory threat.

- Public authorities can orient R&D directly (for example by financing innovative projects on new environmental technologies or cooperative R&D agreements; technology roadmaps) and indirectly (by making public purchase conditional to environmental improvements) and thus modify – quantitatively and qualitatively – the selection environment by imposing environmental quality as performance criteria for public markets but also by playing the role of a large customer. Rather than being designed in isolation one with each other, a system of instruments related to environmental regulation on the one hand and to innovation and diffusion support on the other hand should be designed. The stake is to implement policy that supports environmental innovation as a dynamic process and that takes into account the different stages from innovation to diffusion.

- As already underlined by several authors and as illustrated by the particular evolutionary model of industrial dynamics developed in this paper, time really matters and dynamic processes seem to be very important for environmental policy. The time of implementation is of great importance and the impact of policy instruments are very likely to depend on the stage of industrial structure and the stabilisation of technological change. This echoes a result from Nill (2004) according which the dosage, design and timing of instruments proves to be crucial.

- In the model, the survival of a green market niche results from the adjustment between an environmental leader firm and a group of green clients characterised by high environmental requirement levels and high willingness to pay. Moreover an early change in paradigm experienced by environmental pioneering firms turns to open a window of opportunity for the development of a green market niche. So qualitative coordination that prevails between vertically related firms can be an appropriate channel through which regulation can be transmitted. Public support can be used to facilitate adaptation of suppliers, in particular small and medium enterprises, that have to experiment a new learning phase. Technology transfer policy and technology demonstration and deployment are of primary importance in such a context.

3.5 Conclusion

Further lines of research may be suggested in order to better grasp the interactions between environmental innovation and policy through evolutionary modelling. Three directions may be sketched:

- Environmental innovations involve many changes at different levels, in particular in infrastructure able to receive the new technology. Environmental innovations are thus said to be part of system innovations. The differentiated development of each sub-system can create bottlenecks that can hinder technological development and diffusion. Evolutionary modelling would be useful to examine such issue (see for example Schwoon, 2005).
- It would be helpful to better explicit supply-demand coevolution in the supply-chain (‘filiere’) and to examine environmental innovations along the whole product life cycle. This would allow to assess how the development of green products is constrained by the development of series of innovations (input, process or intermediary goods) along the supply chain.
- Efforts should be made to analyse technological competition in association with technological complementarities that develop between established technologies and the new ones on the one hand and between the various alternatives on the other hand. Such analysis would enable to grasp the process of ‘hybridisation’ of technologies, as part of the firm’s strategies and so endogenous to the interactions between techno-economic change and environmental regulation.

References

Abernathy, W.J. and J.M. Utterback (1975). A dynamic model of product and process innovation. *Omega* 3: 639-656

Arthur W.B. (1989). Competing technologies, increasing returns, and lock-in by historical events. *Economic Journal* 99:116-131

Belis-Bergouignan M.C., V. Oltra and M. Saint-Jean (2004). Trajectories towards clean technology: example of volatile organic compound emission reductions. *Ecological Economics* 48: 201-220

Brunnermeier S.B. and M.A. Cohen (2003). Determinants of environmental innovation in US manufacturing industries. *Journal of Environmental Economics and Management* 45: 278-293

Cole M.A., R. Elliot R. and K. Shimamoto (2005). Industrial characteristics, environmental regulations and air pollution: an analysis of the UK manufacturing sector. *Journal of Environmental Economics and Management* 50 (1): 121-143

David P.A. (1985). The economics of QWERTY. *American Economic Review* 75: 332-337

Dosi G., C. Freeman, R. Nelson, G. Silverberg, L. Soete (eds) (1988). *Technical Change and Economic Theory*. Pinter (London and New York), Columbia University Press

Fischer C., I. Parry and W. Pizer (2003). Instrument choice for environmental protection when technological innovation is endogenous. *Journal of Environmental Economics and Management* 45: 523-545

Hansen O.E., B. Sondergard and S. Meredith (2002). Environmental policy instruments and their effects on innovation. *European Planning Studies* 1 (1): 37-56

Jaffe A..B., R.G. Newell and R.N. Stavins (2005). A tale of two market failures: technology and environmental policy. *Ecological Economics* 54: 164-174

Jaffe A..B., S. Peterson, P. Portney and R. Stavins (1995). Environmental regulation and the competitiveness of U.S. Manufacturing: what does the evidence tell us? *Journal of Economic Literature* XXXIII: 132-163

Jänicke M., J. Blazejczack, D. Edler and J. Hemmelskamp (2000). Environmental policy and innovation: an international comparison of policy frameworks and innovation effects. in J. Hemmelskamp et al., *Innovation-oriented environmental regulation, theoretical approaches and empirical analysis*, 125-152

Kemp R. and J. Rotmans (2001). *The management of the co-evolution of technical, environmental and social systems*. Paper for international conference Towards Environemental Innovation Systems. 27-29 Sept 2001, Garmisch-Partenkirchen (D)

Kemp R., K. Smith and G. Becher (2000). How should we study the relationship between environmental regulation and innovation? In Hemmelskamp J., Rennings K. and F. Leone (eds). *Innovation-oriented environmental regulation.* ZEW Economic Studies, Physica Verlag, Heidelberg, New York

Lane D. (1993). Artificial worlds and economics, parts I and II. *Journal of Evolutionary Economics* 3: 89-107 et 177-197

Menanteau P. (2000). Learning from variety and competition between technological options for generating photovoltaic electricity. *Technological Forecasting and Social Change* 63: 63-80

Nelson R., S. Winter (1982). *An evolutionary theory of economic change.* The Belknap Press of Harvard University Press, Cambridge (Mass.)

Nill J. (2004). Instrumentation of Time Strategies for an Ecological Innovation Policy: a New Role for Subsidies? The Cases of Innovation in Iron Production and Housing in Comparison. Discussion paper IÖW 61/04.

Norberg-Bohm V. (2000). Creating incentives for environmentally enhancing technological change: lessons from 30 years of US energy technology policy. *Technological forecasting and social change* 65: pp.125-148

Oltra V., M. Saint-Jean (2005). The dynamics of environmental innovations: three stylised trajectories of clean technology. *Economics of Innovation and New Technology* 14 (3): 189-212

Ostrom T.M. (1988). Computer simulation, the third symbol system. *Journal of Experimental Social Psychology* 24: 381-392

Porter M. and C. van der Linde (1995). Toward a new conception of the environment-competitiveness relationship. *Journal of economic perspectives* 9 (4): 97-118

Rennings K. (2000). Redefining innovation – eco-innovation research and the contribution from ecological economics. *Ecological Economics* 32: 319-332

Requate T. (2005). Dynamic incentives by environmental policy instruments-a survey. *Ecological Economics* 54: 175-195

Saint-Jean M. (2005). Analysis of the coevolution of suppliers and users through an evolutionary modelling – the case of environmental innovations. *European Journal of Economic and Social Systems* 18(2): 255-284

Sartorius C., Zundel S. (eds) (2005). *Time strategies, innovation and environmental policy*. Edward Elgar, Cheltenham (UK)

Schwoon M. (2005). *Simulating the Adoption of Fuel Cell Vehicles.* Working Paper FNU-59

Scott J.T. (2003). *Environmental Research and Development,* Edward Elgar, Cheltenham (UK)

Valente M. (2005). *Qualitative simulation modelling*, Paper presented to the 4th European Meeting on Applied Evolutionary Economics (EMAEE), 19-21 May, Utrecht, The Netherlands

Windrum P. and C.R. Birchenhall (1998). Is Life Cycle Theory a Special Case?: Dominant Designs and the Emergence of Market Niches through Co-evolutionary Learning. *Journal of Structural Change and Economic Dynamics* 9: 109-134

Appendix 3.1: initial values of parameters in the reference configuration

Table 3.A1: Parameter values

Parameter	Value
Initial number of suppliers	n=12
Initial number of clients	m=200
Requirement level parameter	$\varepsilon = 0.01$
Perception degree of environmental quality	$\sigma = 0.25$
Sales effect on R&D index	$\alpha = 0.01$
Relative importance of positive score	$\beta = 0.09$
Characteristic-specific R&D rate	$\delta^h = 0$
Experience level	$E^h = 1$
R&D rate on turnover	$\mu = 1$
Research level	$R^h = 1$
Initial performance level	$X^h = 1$
Initial price	$P = 3$
Initial R&D index	$RDIndex^h = 0$
Speed parameter for experience	$\lambda = 0.01$
Speed parameter for research level	$\gamma = 0.1$
Mark-up rate	$\theta = 200\%$
Initial market shares	MS = 1/12
Upper product performance bound	$X^{2max} = 14$
Lower product performance bound	$X^{2min} = 1$
Productive efficiency threshold in paradigm 1	$X^{1max1} = 4$
Upper productive efficiency bound	$X^{1max2} = 14$
Upper bound of process environmental quality in paradigm 1	$X^{3max1} = 3$
Upper bound of process environmental quality in paradigm 2	$X^{3max2} = 13$
Upper bound of product environmental quality	$X^{4max} = 14$
Scale parameter for innovation output	$\eta_0 = 0.01$
Innovation elasticity of research level	$\eta_1 = 0.45$
Innovation elasticity of experience	$\eta_2 = 0.1$
Innovation elasticity of distance to the upper bound	$\eta_3 = 0.45$
Maximum experience level	$MaxE = 3$
Minimum sales	$N^{min} = 1$
Parameter of the innovation probability	$\pi_1 = 0.035$
Parameter of the innovation probability	$\pi_2 = 0.05$
Parameter of the innovation probability	$\pi_3 = 0.65$
Parameter of the innovation probability	$\pi_4 = 0.4$
Bandwagon effect	e = 0.1

Table 3.A 2: Typology of user groups in the reference configuration

Profile	**User group**	
	G1	**G2**
Weight assigned to price	$x^1 = 0.05$	$x^1 = 0.45$
Weight assigned to product performance	$x^2 = 0.05$	$x^2 = 0.45$
Weight assigned to process environmental quality	$x^3 = 0.45$	$x^3 = 0.05$
Weight assigned to product environmental quality	$x^4 = 0.45$	$x^4 = 0.05$
Maximum price	$level\ X^1 = 4$	$level\ X^1 = 1$
Minimum requirement level for product performance	$level\ X^2 = 2$	$level\ X^2 = 2$
Minimum requirement level for process environmental quality	$level\ X^3 = 1.5$	$level\ X^3 = 1.5$
Minimum requirement level for product environmental quality	$level\ X^4 = 2$	$level\ X^4 = 2$
Initial proportion	20%	80%

4 Technology successions and policies for promoting more environmental friendly technologies

P. Windrum

UNU-MERIT, University of Maastricht (Netherlands) and Centre for International Business and Innovation (CIBI), Manchester Metropolitan University Business School (UK).

Contact: p.windrum@mmu.ac.uk

ABSTRACT

The core thesis of this paper is that policy makers, firms and consumers are the agents of path dependencies that support and maintain technological paradigms. They are also the agents of change. The replacement of old technology paradigms by new paradigms occurs when there are fundamental changes in the expectations, preferences, competences and policies of these agents. The first part of this paper (sections 2 to 4) reviews the research on path-dependency in sequential technology competitions that has been conducted since 1999. The second part of the paper (sections 5 and 6) develops a co-evolutionary framework that captures the dynamics of successions. This can be used by those interested in promoting more environmentally friendly technologies as well as as conceptual building blocks for a comprehensive evolutionary model of technological transitions.

4.1 Introduction

This paper is a contribution to the Netherlands Environmental Assessment Agency project ‘Environmental policy and modelling in evolutionary economics’. The core thesis of this paper is that policy makers, firms and consumers are the agents of path dependencies that main support and maintain technological paradigms. They are also the agents of change. The replacement of old technology paradigms by new paradigms occurs when there are fundamental changes in the expectations, preferences, competences and policies of these agents.

In Windrum (1999), I made the point that past research on the diffusion of new technologies ignores the existence of powerful path dependencies that build up around established technologies and lock-out new technologies. At that time, a body of research existed on path-dependency in contemporaneous competitions between rival variants of a new technology. However path-dependency was not being considered in relation to sequential competitions between old and new technologies. Implicitly, it was assumed that path dependencies did not matter. The notable exception was David’s 1985 paper on the lock-out of the DVORAK keyboard by the QWERTY keyboard. It highlighted the need for research into this issue.

My 1999 paper began to explore factors of path-dependency that affect the probability of new technology adoptions. I started out by making a link with evolutionary ecology. Traditional diffusion models are best suited to discussions of ‘early colonists’, i.e. the very first technologies that establish themselves in virgin territory and, hence, are not competing with pre-established technologies. Subsequent technologies, however, must directly compete with, and displace, established technologies. This is the basis of a ‘technology succession’. A succession is difficult because path-dependency means the selection environment – made up of firms, consumers and policy makers – has a predisposition in favour of the old technology. In other words, the selection environment is less open to radical new technologies than it was in earlier stages. The new technology must do something *novel* and *different*. This was the starting point for my subsequent research in this area.

The first part of this paper (sections 4.2 to 4.4) reviews the research on path-dependency in sequential technology competitions that has been conducted since 1999. The second part of the paper (sections 4.5 and 4.6) develops a co-evolutionary framework that captures the dynamics of successions. This can be used by those interested in promoting more environmentally friendly technologies. Section 4.2 reviews existing research on the sources of market path-dependency by firms and consumers. I extend the discussion to consider the key role played by the path dependencies of policy makers and how this helps to establish and maintain a technological trajectory. There is an important shift in perspective here. Rather than viewing the policy maker as an independent rational planner, we need to view the policy maker as an interconnected agent within a *complex selection environment* that comprises firms, consumers and policy makers.

Section 4.3 discusses a set of conditions for a technology succession - i.e. the conditions that are necessary for a new technology to displace an established technology. This is derived from theoretical papers by Shy (1996), Malerba et al. (1999), Saviotti and Pyka (2004), and Windrum and Birchenhall (2005), and from empirical research conducted by Islas (1997), Yamamura et al. (2005), and Windrum (2001, 2004, 2005). Section 4.4 identifies a set of strategies that have been successfully used by firms to break the path dependencies supporting established technologies. Understanding these strategies, and the basis of their success, is essential for policy makers interested in promoting new, more environmentally friendly technologies.

Section 4.5 introduces a co-evolutionary framework that comprises policy makers, consumers and firms. It begins by outlining the general approach of interacting agents operating within a complex selection environment. Section 4.5.1 discusses two case studies, the evolution of the car-based transport system and refrigerants. The cases studies provide us with a set of *stylised facts* that are to be captured in the formal framework. These concern the behaviour and actions of agents, and the interactions between agents. Section 4.5.2 proceeds to formalise these behaviours, actions and interactions for each of the three key agents: policy makers, consumers, and firms. A number of core key concepts and ideas are introduced at this stage: trans-trajectory or *'deep' path-dependency*, *conceptual innovation*, *new consumption possibilities* facilitated by new technologies, and the *co-evolutionary learning* of adaptive agents in dynamic, changing environments. Section 6 concludes by identifying a set of policy lessons that can be gleaned from this approach, and the set of instruments that can be used by policy makers to promote technology successions.

It is important to clearly define the *scope* of this research. The focus here is sequential technology competitions and, more specifically, the conditions under which new technologies can displace established technologies. In a sense, this research topic harks back to an old research question, one that dates back to Schumpeter (1912, 1939) and was restated and reinvigorated in the work of Freeman (Freeman, 1982; Freeman et al., 1982). This older research question concerns the periodic introduction of key new technologies that set into train long-run economic cycles, known as 'Kondratiev Waves'. As stated earlier, the focus here is not on the periodicity of new technologies but on the conditions under which new technologies are able to displace old, established technologies in sequential competitions. This is a key issue for policy makers and all others interested in the promotion of new, more environmentally friendly technologies.

What are the essential features of technology competitions? Not all competitions between rival products are included in the definition of 'technology competitions'. Technologies competitions are a particular subset of product competition. Technology competitions have two particular features. First, there is a technical differentiation (non-compatibility) between the competing technologies. For instance a DVD disk is incompatible with a VCR player. iPods do not play CDs. Second, there tend to be significant switching costs for adopters. New technology goods tend to be expensive, and there are invariably non-pecuniary set-up costs associated with learning how to use a particular technology product. It is due to both technical incompatibilities and the existence of switching costs between rival technology products that technology competitions are zero sum games, i.e. they are winner-takes-all competitions.

Technology competitions are important for technical, commercial and policy reasons. Technically, standards are essential for the integration and development of technological systems. The internet, for example, is a complex technology that comprises numerous interacting components. In order for the internet to work, content (media and services), hardware (cables, routers, servers, PCs), software (operating systems, browsers, and e-mail), communication protocols (WWW and TCP/IP), and design conventions (that provide website ergonomics and functionality to the user) must all interface and work together. In terms of commercial advantage, the proprietary ownership and control of a standard technology is a key strategic device through which a firm can influence and control an industry. For instance, the term 'WINTEL PC' indicates the two key corporations that control the PC industry by virtue of having proprietary control of the underpinning operating system and chip set technologies. Finally, technology competitions have significant implications for policy making. Policy makers must operate within an envelope of possible options that are set by technologies. Technologies are one important factor that frames and gives direction to policy. Alternative technologies tend to be associated with different envelopes, and so policy makers are deeply interested in the different possibilities associated with new and/or alternative technologies.

It is also important to clarify the *phase* of technology development that is being considered. Here we are specifically concerned with the *diffusion* of a new technology. This is the third, and final, phase of technology development discussed by Schumpeter (1912, 1939). The first phase, invention, is usually related to some empirical or scientific discovery. In itself, an invention has no economic or social significance, and typically offers no hints about possible applications. The second phase, innovation, is the point at which the invention is actually applied for the first time, whether this takes the form of a product or a process. The first applications of a technology are invariably crude and inefficient. Not only is their performance usually poor compared to existing (alternative) technologies, but the (fixed) production costs are likely to be very high. Hence, innovations are not automatically capable of diffusing. They are like swan's eggs, requiring a (possibly long) gestation period in which further basic research and development is needed to develop them[6]. If this is not possible, they will perish. Rosenberg (1982) observes that survival of the new technology requires the establishment of a protected space in which further development can be achieved. This can take the form of distinct niche or sub-niche in the market, which may be complementary to the established technology, or else take the form of public sector support, where users are often also contributors to the R&D process (Hoogma et al., 2002; Geels, 2005).

Our focus is on the third phase of a new technology, its diffusion. Diffusion involves the widespread assimilation of a technology within a politico-socio-economic setting. To carry on the analogy of the swan's egg, an innovation is an attractive duckling at the outset of this diffusion stage. This highlights an important difference between the innovation phase and diffusion phase of a technology life cycle. In the innovation phase, a technology will survive provided it shows sufficient promise or potential to a key group of supporters – even if it is initially inferior in many respects to the old technology. In the diffusion phase, new technology goods must directly compete on quality and price with old technology goods in the mass market. While the mass market offerings may not be fully fledged swans, they must at the very least be attractive ducklings!

Having identified the phase that we are studying, we next need to make an important distinction between a *technology succession* and *technology substitution*. In a technological substitution, a new technology is used in the same way as the old technology. It is adopted because it offers a superior quality/price performance in the same basic use. The new technology has a superior performance in one or more service characteristics that are common to both it and the old technology. For example, the compact disk (CD) replaced the vinyl LP in domestic music systems in the 1980s. Consumers adopted the CD as a storage medium for music because it is more convenient (i.e. is smaller in size), requires less maintenance (no need to clean disks regularly), individual tracks can be accurately and easily selected (using a remote control unit), and is far less prone to degradation (i.e. scratches) than vinyl. Hardware manufacturers saw an opportunity to increase profits through sales of new CD players, while record companies realised that significant profits could be generated if, in addition to new material, consumers could be convinced to repurchase previously owned material, this time in the CD format.

---

[6] Less poetically, Mokyr (1990) calls them 'hopeful monstrosities'. 'Hopeful' because they have particular features that are of interest, and 'monstrous' because of their initial crudeness and inefficiency.

In contrast to a technological substitution, a technological succession – the focus of this paper - opens up new consumption possibilities. These new consumption possibilities are not provided by the old technology. The new possibilities may not have been evident in the invention or innovation stages, but they certainly become apparent in the diffusion stage, as producers and consumers experiment with the new technology, possibly combining it in novel ways with other technologies that appear in this phase. One of the case studies we shall look at is the car. The car initially competed with alternative forms of urban transport (predominantly horse-drawn vehicles such as trams). Subsequently, new markets emerged as a consequence of wider social, economic and political changes. The development of suburban living after WWII saw the emergence of new type of user - the suburban car commuter.

The other characteristic of a technological succession is the entrance of new firms into the market. The shift from LPs to CDs did not change the way in which music was made, recorded, packaged or distributed. Indeed, the 'big four' industry labels – SonyBMG, EMI, Warner and Universal - actively promoted CD. CD did not threaten their control of the music publishing industry. Indeed, it enabled them to significantly increase their sales base and profits as consumers bought new material at higher prices on the CD format and repurchased existing titles as they replaced their LP collections with new CD versions. This contrasts strongly with the internet. The big labels still control 70% of the world's music market, and mainstream radio and TV airplay, but the internet has opened up new ways of recording, packaging and distributing music. Downloads, peer-to-peer file sharing, pod casting, and on-line radio stations challenge their traditional business model and their market power. March 2006 saw the first MP3-only No.1 single in the UK, bypassing the traditional distribution network completely. There has been a rejuvenation of independent labels and enabled a new generation of music entrepreneurs to enter the market, as well as the more familiar names of Napster, Rio and Apple (an established company that has used the new technology to enter this lucrative industry). Bands are also exploiting the new technology to their advantage. Arctic Monkeys and Editors released self-recorded demo versions of their songs over the internet in order to build fan base in the UK and Europe. This gave them a very strong position when it came to negotiating a contract with (independent) record labels.

4.2 Path-dependency

Our core thesis is that successions require fundamental changes in the beliefs and actions of policy makers, firms and consumers because these are the agents of path dependencies that maintain existing technology paradigms. A first step in understanding successions is a clear appreciation of the path-dependency that maintains established paradigms. Here we consider the various sources of path-dependency associated with firms (4.2.1), consumers (4.2.2) and policy makers (4.2.3). The discussion makes clear just how powerful these factors are and hence, why successions occur so infrequently. These factors directly affect the timing and frequency of new technology adoptions.

4.2.1 Supply side factors

The earliest discussions of path-dependency highlighted the importance of firms in determining both the rate and direction of technological innovation. Key contributions were Atkinson and Stiglitz (1969) on localised technological improvement, the historical studies of Rosenberg (1969) and David

(1975), and the work of Vernon (1966) and Abernathy and Utterback (1975) on product lifecycles. Path-dependency in R&D arises because firms' knowledge and technological expectations are built up cumulatively over time, frequently as a consequence of trial-and-error learning. In this way, firms focus on just a few potentially fruitful avenues for R&D, to the exclusion of other possibilities. The discussion was later elaborated into the idea of trajectories by Nelson and Winter (1977), Dosi (1982), and Freeman (1982). This seeks to capture the cumulative nature of the search process of firms. To a greater or lesser degree, these translated Kuhn's (1962) theories of scientific development into the technology realm. For instance, Dosi's concepts of 'technological paradigms' and 'technological trajectories' and 'heuristics' are related in the following manner:

'(T)echnologies develop along relatively ordered paths shaped by the technical properties, the problem-solving heuristics and the cumulative expertise embodied in *technological paradigms*... A *technological trajectory* is the activity of technological progress along the economic and technological trade-offs defined by a paradigm' (Dosi and Orsenigo, 1988, p.16; italics in original).

Nelson and Winter (1977) suggest that mechanisation and scaling are two key heuristics that guide firms in their search for new innovations. Through increased mechanisation, firms can reduce (expensive) labour inputs and thereby lower production costs. Scaling involves the improvement of product performance through increasing/decreasing the size of the product. For example, larger engines are more efficient than smaller engines. By contrast, the speed of microchips improves as the signal paths on circuit boards are reduced. A particularly important empirical study of trajectories is that of Sahal (1985). Using long-run data for the aircraft, tractor, and computer industries, he sought to identify stable relationships between clusters of key product features that had been scaled over time.

The nature of the artefact itself may give further direction to innovative search, leading to well-defined trajectories of incremental innovation. The work of Vernon (1966), and Abernathy and Utterback (1975) on the emergence of dominant designs is of particular importance in this respect. Contemporaneous technology battles between competing variants of a new technology product can lead to the emergence of a single, 'dominant' design. Through competition, market consensus is established regarding the core set of product characteristics to be produced and with which other technologies must interact. Over time, incremental innovations improve the overall quality of the dominant design, and the range of applications to which it can be applied, further enhancing the diffusion of the technology.

A far more radical version of the dominant design is the technology architecture thesis (Clark, 1985; Anderson and Tushman, 1990; Henderson and Clark, 1990; Vincenti, 1990). It is observed that technological products typically comprise a set of elements which must work together effectively. This interoperability is controlled by an 'architecture'. An oft cited example is a computer operating system. It is argued that alternative technology variants are defined by their architectures and that these are incompatible with one another. Hence, market selection of a dominant design is in fact the selection of a particular architecture. Using this concept, Henderson and Clark (1990) derived a set of very different classification of radical and incremental innovation. According to their classification, 'incremental innovation' involves improvement in the performance of one or more elements of the design (e.g. an improved version of a chipset). A 'modular' innovation involves replacing an old technology component with another or the addition of a new technology component to improve performance (e.g. the introduction of car airbags to improve passenger safety). Importantly, a 'pure' modular innovation leaves the architecture unchanged. Conversely, an 'architectural innovation' takes the same basic set of elements (modules) and recombines them in a novel manner in order to improve performance. Finally, ‘radical’ innovation involves both architectural and modular change. Obvious examples in the computer industry, for instance, were the IBM 360 mainframe computer, the PC, and the internet. It is argued that the technology architecture thesis helps explain the highly path-dependent nature of innovative search. Radical innovations are competence-destroying and so affect the underpinning knowledge of the firm (Anderson and Tushman, 1990). They also have serious organisational consequences because new ways must be found in which to assemble and optimise the production of new architecture and modules (Henderson and Clark, 1990).

4.2.2 Demand side factors

Following the work of Arthur and David, there has been much interest in demand side factors affecting path dependence. Arthur’s work on ‘lock-in’ focuses on contemporaneous technology competitions amongst competing variants of a new technology. By contrast, David’s seminal empirical study of the QWERTY keyboard (David, 1985) discusses the dynamics of sequential technology competitions – in this particular instance there was a market ‘lock-out’, with the DVORAK keyboard failing to displace the older, established QWERTY design.

At the core of this discussion is the interaction of individual choices. Path-dependency arises because new adopters’ take into account the choices already made by previous adopters. Hence, an adopter’s preference (utility) function contains an autonomous individual component and an inter-personal component, such that the payoff (Π) associated with each technology variant at time t is

$$\Pi_t = X_{ij\,t} + r\,(n_{j\,t-1}) \tag{1}$$

where X_{ij} is individual i’s personal preference for technology j, and r is a term that captures the increasing returns to adoption.

Strong path-dependency arises when the population of adopters is relatively homogeneous with regards to the autonomous component (X_{ij}). If this is the case then decisions are frequency-dependent and the process reduces to a Polya urn model in which decisions depend on the relative market shares of competing technology products (Bassanini and Dosi, 1998). The properties of the Poya urn model are of interest. We know with certainty that the market will lock in to one of the competing technology variants (a winner-takes-all competition), resulting in a monopoly. However, because selection is frequency-dependent it is impossible to predict ex ante which particular variant will emerge as the winner[7].

The discussion complements Leibenstein’s famous 1950 paper on consumer demand. Leibenstein discussed four factors influencing the demand for a product; two of these are price factors and two are quantity factors. With regards to price, there is the conventional ‘law of demand’ with the quantity

---

[7] While it does not concern us directly - the focus of this paper is sequential technology competitions - there has been an important debate regarding whether markets can lock-in to an inferior quality variant in contemporaneous standards competitions. David and Arthur suggest that it is possible. This has been challenged, most notably in a series of papers by Liebowitz and Margolis (1990, 1996, 1998). They argue it cannot in the short-run or the long-run. Liebowitz and Margolis criticise the empirical data (short-run) and argue that side payments will ensure that better technology variants win out in the long-run. Hence, they argue, it is not possible to lock-in to an inferior variant.

demanded inversely related to market price. By contrast, a second price effect is the ‘Veblen effect’. Veblen’s ‘Theory of the Leisure Class’ was the first to focus on the interactions between humans and artefacts in an institutional context. Artefacts are a means of communicating group membership (social status in Veblen’s discussion). The ‘Veblen effect’ manifests itself as a positive relationship between price and the quantity demanded. The effect is opposite and symmetrical to the law of demand. As with Adam Smith’s famous water-diamond paradox, the high price of diamonds makes this commodity exclusive for those that can afford it and which to distinguish themselves from those that cannot afford them. Although water is essential for life, it is plentiful and cheap and, hence has no such distinction value.

On the quantity side, Leibenstein’s discussion of the ‘bandwagon effect’ foreshadows much of the work conducted in recent years. Not only did he identify the positive relationship between the attraction of a good and the number of previous adopters (*nj*), but he also discussed the importance of bounded returns to adoption for competition. Leibenstein identifies the ‘snob effect’ as an opposite and symmetrical effect to the bandwagon effect. He posited that the desirability of a product, as a snob good, falls the more widely it is adopted because it looses its exclusivity. As we shall see later, the discussion of upper bounds on increasing returns to adoption and the drive for exclusivity and distinction are important factors in sequential technology competitions.

4.2.3 Policy factors

Economists focus on markets with firms and consumers, and so pay less attention to the nature and extent of path-dependency amongst policy makers. In political science it is commonly assumed that the policy maker is a ‘rational planner’ who is free to consider all feasible actions, and to take whatever course is optimal unbound by past decisions. I will argue that policy makers need to adopt a very different position; one that explicitly recognises the path-dependency of their own actions, and the role which policy itself plays in promoting or locking out new, alternative technologies.

The existence of path-dependency in policy-making is fairly self-evident and (hopefully) not too controversial. Policy is the sum total of laws and regulations regarding a particular set of issues. Policy is highly cumulative. Current policy is in part the result of a long stream of decisions taken over time. For instance, tax policy is not made anew each year. Policy makers lack the capability to enact whole new policies every year. Further, genuinely new policies (as opposed to incremental shifts in old ones) present radical dislocations. The social order can accommodate occasional radical dislocations, it cannot accommodate them every year. It is therefore for good reason that policy is highly path-dependent.

Path-dependency is evident in general trends that are shared by nations. It also helps to account for persistent variation between nations. Take, for example, post-war political attitudes in western Europe and the USA towards the role of the state. In the immediate aftermath of WWII, the new Keynesian consensus saw both a necessity and a role for macroeconomic intervention by the state. It also encouraged state intervention and public ownership of key sectors, such as transportation, education and health, where the quantity and/or quality of private sector provision was perceived to be inadequate. The Keynesian consensus was called into question in the 1970s and in the 1980s was overturned by a new pro-market rhetoric. Policy makers fell in love with the private sector. The upshot has been dramatic and wide ranging. First, there was the advocacy of a much reduced role for the state in macro management. In Europe this has led to the transfer of monetary control to an independent European Central Bank, something unimaginable in the Keynesian era. Second, there were calls for a smaller state. This prompted the privatisation of publicly owned firms and the outsourcing/competitive tendering of remaining basic services. Later, there was the adoption of private sector management practices (e.g. the ‘New Public Management’ movement), and more recently the promotion of public-private sector funding of large investment projects such as new hospitals, schools and other projects that would previously have been funded by public monies alone.

Within this broad and shared pattern there are national differences. These differences highlight localised path dependencies. Take health sector reforms as an example. The UK public (and the vast majority of politicians) views the publicly owned National Health Service (NHS) as a national treasure. It is, in effect, sacrosanct. Hence, no political party will consider policy measures that encourage greater private sector involvement in health service delivery. In the Netherlands and in Germany, policy reforms will tinker with, but not overhaul, their systems of insurance funding. In the USA, the Clinton administration failed to pass legislation to revamp its medical system, by doing away with private insurance (except as a supplement) and having a federally funded system to cover all citizens regardless of employment status, income and age.

With regards to technological path dependence, political institutions themselves often play an important role in supporting and, hence, mainlining lock-in to established technologies. The car provides a clear example. The efficient running of a car-based transport system requires the co-ordination of traffic flow and parking spaces. The former includes support services such as road lighting, road maintenance, traffic signals and signs, repair garages, and break-down services. These are provided by a mix of private and public sector providers. Indeed a complicated regulation environment is present in nearly all countries, with a combination of national and local government regulatory bodies responsible for the formulation and delivery of urban and environmental planning programmes covering road construction, urban development and traffic control. Public sector institutions are additionally involved in the provision of safety-related functions such as proficiency tests for drivers, regular mechanical tests for car safety, road police, and accident and emergency services. Finally, a system of taxation operates to levy car users for these publicly provided services.

The car was actively supported by successive governments in the US and in Europe throughout the post-war era. The development of the car, rather than collective modes of transport, was viewed as the most effective means of increasing mobility (Flink, 1988). The car was a symbol of modernity, associated with notions of freedom and democracy in Europe (Mom et al., 1997) and the USA (McShane, 1994). As an icon of modernity, the car was perceived as part of a wider socio-economic change that included, amongst other things, the rise of suburban living and the relocation of branch plant manufacturing and light industry from cities to new out-of-town industrial estates. Somewhat ironically, road construction subsequently became a symbol of urban regeneration once these industries had vacated the traditional industrial districts. As the car became woven ever more finely into the fabric of society, it changed from a luxury good to a necessity.

Despite our understanding of the environmental impact of car based transport systems, path-dependency makes it exceedingly difficult to change policy. On the one hand, there are real constraints on policy options. Alternatives modes of passenger transport may not readily exist in many areas. In Europe and the US, urban and suburban tram and rail networks, and inter-city rail infrastructures were destroyed in order to facilitate the development of the car. Ironically, where the privatisation of rail and bus services has occurred, there has been a further ‘rationalisation’ of the alternatives, most noticeably in rural areas but also in urban areas, leading to increased car use. Like King Canute[8], policy makers are well aware of the limitations of their power, and of the political dangers inherent in setting unattainable goals. On the other hand, national and local governments are enmeshed in the effective running of a car-based transport system. This often leads to well-known ‘silos problem’. Rather than working towards a common agreed goal, some government departments (usually those charged with responsibility for car transport) will act to improve the efficacy of car transport, thereby maintaining lock-in, while others act to limit car transport and champion the alternatives.

4.3 Necessary conditions for a technology succession

The innovation literature discussed in section 2 contributes enormously to our appreciation of path-dependency, and helps us understand why successions occur so infrequently. Nevertheless, successions do occur. Drawing on recent theoretical and empirical research, this section identifies the conditions under which a succession can occur. This further enhanced our understanding of the factors determining the timing and frequency of new technology adoptions.

As noted in the introduction, economists have traditionally ignored the presence of old paradigm path dependencies in their models. Hence, models of sequential technology competitions in the presence of path-dependency are thin on the ground. Four important exceptions are the models of Shy (1996), Malerba et al. (1999), Saviotti and Pyka (2004), and Windrum and Birchenhall (2005). In addition, there are some empirical papers by Islas (1997), Yamamura et al.(2005), and myself (Windrum 2001, 2004, 2005)[9].

Pulling this theoretical and empirical research together, we can identify a set of necessary conditions for a technology succession:

1. The functional equivalence of new and old technology products (Shy 1996; Windrum and Birchenhall, 2005). For example, the car, bus, train, tram, motorbike and bicycle are alternative types of mechanical passenger transport, each with its own particular merits in terms of journey times, cost per km, flexibility in the timing of journeys, and the pollution generated per km. Yet they all perform the same basic function – they transport a person from one geographical place to another – and so they are competing alternatives.
2. Novelty. The new technology products must offer users new consumption possibilities, previously unavailable using the old technology products (Windrum, 2005; Windrum and Birchenhall, 2005).

---

[8] To demonstrate the limits of his power to his subjects, the English king had his throne set on a beach. King Canute sat on the throne as the tide was coming in and famously ordered it to stop.

[9] It is more useful to report the findings of this research rather than enter into an involved discussion of the details of the models and the empirical research.

3. New consumer types. These new consumer types are willing and have the finances competences to experiment with, and champion, the new technology products (Malerba et al., 1999; Windrum and Birchenhall, 2005).
4. These new user groups must be willing to trade-off the benefits of an established technology, e.g. those associated with a large installed base of old technology users, against the novel consumption possibilities of the new technology products (Shy, 1996; Malerba et al., 1999; Windrum and Birchenhall, 2005).
5. New firms entering the market (Malerba et al., 1999; Windrum and Birchenhall, 2005). New entrants go hand-in-hand with new consumer types in established a new technological paradigm.
6. New market entrants bring with them new conceptualisations of what the market is, and what it can become (Windrum, 2005; Windrum and Birchenhall, 2005). This is fundamentally important in the development of an alternative technology paradigm.
7. Financial capital. The availability of venture capital is a key factor affecting new industry start-ups and market entry by firms operating on other markets (Malerba et al., 1999, Saviotti and Pyka, 2004).
8. The R&D response of established technology firms is a key determining factor. One response is for old technology firms to step up their R&D programmes, engage in product and process innovation, and thereby improve the quality/price performance of the old technology products. This is known as the ‘sail ship effect’[10]. If old technology firms innovate more successfully than new technology firms, improvements in quality/price performance may be sufficient to see off the challenge posed by new technology entrants (Windrum and Birchenhall, 2005). An alternative response is for old technology producers to switch camps and set up production of new technology goods (Malerba et al., 1999). This may be an attractive proposition for firms with relatively small market shares in the old technology industry. Like new start-up firms, they may view the new technology as an opportunity to be become a major industry player. The ability to successfully switch strongly depends on the transferability of knowledge and competences from the old to the new technology (Gort and Klepper, 1982; Anderson and Tushman, 1990; Malerba et al., 1999). Finally, it may be a dominant old technology firm that develops and launches the new technology. While this is less common, the notable example was IBM’s championing of the personal computer as a serious business machine. The development and launch of the IBM PC gave it credibility amongst the business community, and became the dominant design.
9. New policy models. The development of a new technology paradigm requires the development of a ‘new policy model’. At the core of the new policy model is an alternative ‘mentality’ of policy practice (Foucault, 1972). This translates into a relatively coherent, explicit cluster of policy positions and practical measures. Old models and mentalities are embedded in institutional structures and arrangements, and can persist long after the

---

[10] The term was first coined by Gilfillan (1935) when referring to the rapid spurt of technical improvement in sailing ships that followed the introduction of steamships in the in the 1860s.

technologies which they originally supported have disappeared. Hence, a new policy model must displace the old policy model if a new technology paradigm is to develop.

10. The most visible (and readily quantifiable) aspects of a succession are new product designs and new process technologies (Shy, 1996; Islas, 1997; Malerba et al., 1999; Saviotti and Pyka, 2004; Windrum and Birchenhall, 2005; Yamamura et al., 2005; Windrum, 2005).
11. Timing. The probability of a succession occurring may differ at different moments in time, with distinct 'windows of opportunity' arising.

Let us investigate in more detail the time related conditions of technology successions. This has not been addressed in the literature on sequential technology competitions, and is important for effective policy formulation.

Investment and innovation cycles

The probability of a succession occurring may be affected by investment cycles, economic cycles, and patent cycles. Some industries have distinct investment cycles tied to the scrapping of existing product vintages. For instance, the investment cycle in the computer industry has a periodicity of around 3 years, offering new entrants a distinct window of opportunity to launching alternative, new technologies. Discussion of the link between macroeconomic cycles and the introduction of new technologies dates back to work of Mensch (1979) and Freeman et al. (1982). Empirical evidence indicates that initial investments in new technologies are more likely to occur at the top of the economic cycle, when the investment climate is favourable and venture capital more readily available, that at the bottom of the economic cycle. Patents play an important role in the R&D strategies of pharmaceutical firms and in certain manufacturing industries. Patents not only cover existing product technologies but also prospective, alternative technologies. It is not uncommon for dominant (old technology) firms to engage in defensive patenting. Dominant firms take out patents on core aspects of new technologies, not in order to produce these themselves but to deny access to potential new rivals, thereby protecting their established technology products and the profits generated by these old technologies. The R&D activities of new technology entrants are unrestricted once such patents have elapsed.

Generation-based cycles

Distinct windows of opportunity exist on the demand side that can give rise to regular cycles. These may be linked to generations and be very frequent, as in the fashion and music industries. Each generation of teenagers seeks to distinguish itself from its parents and, just as importantly, the previous teenage cohort (older bothers and sisters). This gives rise to well-documented cycles of teenage clothing and music fashions, which occur every 3 to 4 years. Other demand side cycles, with longer periodicity, are linked to the product lifecycle. As Leibenstein (1950) observed long ago, highly successful products eventually saturate markets and upper limits exist on the potential network externalities. What is more, when a product becomes widely diffused, it loses its exclusivity – it is no longer 'hip' and loses its appeal (i.e. it has a zero snob effect). As marketers are well aware, consumers are potentially very interested in the next new product to come along. This is because the new product has a strong snob effect - kudos is attached to the new technology product because ownership differentiates and sets apart early adopters from the rest (who are still using the old technology product). In a very real sense, the success of an established technology breeds its own destruction.

Supply side cycles

There are additional supply side aspects of the product lifecycle that can give rise to windows of opportunity and, hence, cycles. Technology trajectories eventually run into decreasing returns as technological opportunities are exhausted. Notably, there are upper limits to scaling and mechanisation. ‘Wolff’s Law’ states that physical limits impose boundaries on the gains that can be achieved through scaling (Mensch, 1979). For instance, quantum mechanics imposes a physical limit on the ability of microchip manufacturers to increase chip speed through miniaturisation. Sahal (1985) additionally observes that limits to scaling arise from the non-linearities that exist between the interdependent components that make up a product. A design can work effectively within a range of scale values but beyond a certain threshold, further changes in size require changes to be made to both the form and structure of a product. Sahal illustrates his argument using the piston propeller airplane engine. Scaling of the piston engine was limited by increasing vibration and by the tips of propellers, which became increasingly inefficient as one approached the speed of sound. It was this physical limit that led aeronautical engineers to consider R&D into new, alternative engine designs.

In addition to limits on performance improvements achievable through scaling, there are limits to gains in production costs available through increasing scale. Upper bounds on scale economies are well discussed in economics, and are associated with physical limitations and with the loss of managerial control of the production process as scale increases. This opens a window of opportunity for the adoption of technologies that improve the organisation and managerial control of production. There may also be longer-run cycles associated with fundamental shifts from old to new production paradigms. Hölzl et al. (2006) discuss the switch from Fordist to Post-Fordist technologies. Fordist mass production technologies enabled firms to modularise production activities, and at the same time centralise management and R&D in order to increase control over production and product development processes. Post-Fordist technologies make a decentralised and externally modularised architecture possible. Where Fordist paradigm focused on reducing the costs of internal coordination, the Post-Fordist paradigm focuses on reducing costs of external coordination, facilitating greater internetworking of firms along the supply chain.

The core concepts of Post-Fordism are lean production and just-in-time delivery. Herein lies the importance of new internet technologies. They enable an effective flow through of information. Through a modularisation and reintegration of their activities, producers can accommodate ongoing improvements in component design by other firms without the need for changes elsewhere. This opens up the potential for product innovation along the supply chain, while simultaneously enabling firms to offer a broader range of designs to the end consumer. Greater customisation is possible, as is the ability to respond to changes in demand for different features. Finally, there are huge savings in component inventories. The net result is large efficiency gains through a reorganisation of external relationships. This new paradigm, supported by internet-based technologies that facilitate reconfigurations in firms’ organisational architectures, enables the integration and monitoring of production processes that are external to the firm.

Conditions 1-11 involve fundamental discontinuities amongst the three sets of agents: consumers, firms, and policy makers. These must *all* occur in order for a technology succession to be possible.

4.4 Commercial strategies for successions

We have discussed the necessary conditions for a succession that have been identified by the literature. Let us now examine the strategies that private sector firms have used in order to win sequential technology competitions. As we shall see, new technology entrants cleverly set up the necessary conditions for a succession, while simultaneously exploiting opportunities associated with weaknesses of old technology firms and products. It is important for policy makers to understand these strategies and when they are most effectively played.

4.4.1 Demand-side strategies

Case studies indicate that new market entrants can successfully employ a number of strategies to overcome the large installed user base enjoyed by established firms.

- To start with, the new entrant could technically differentiate its product so that it is incompatible with the old technology (for example, CD’s not being compatible with LP’s). This clearly sets up a winner-takes-all competition in the minds of consumers. Alternatively, the new entrant can design its technology as a complement rather than a full substitute to the dominant technology as to profit from the already large installed base (for example, nineteenth century steam ships used for inland shipping being complementary to sailing ships used at sea).
- Having done this, success lies in understanding the nature of consumer demand and formulating appropriate strategies. Consumer demand is invariably heterogeneous and subject to change, providing late technology entrants the opportunity to overturn established firms with large installed user bases. This is true even when there is a well-established dominant design and a market is dominated by a few large firms.
- Successful strategy depends on identifying and developing new user types interested in pursuing the alternative consumption possibilities that are facilitated by the new technology. This is an important point. Users are interested in the new consumption possibilities, not in the technical features of the technology per se. The strategy is therefore to identify and support new user types who wish to differentiate themselves through the consumption of new products. As noted previously, there may be a new generation of users or else an existing set of users may wish to set themselves apart through the development of new consumption patterns (the snob effect).
- Heterogeneity may mean there is latent demand amongst dissatisfied users of a dominant design. A good example is provided by the camera industry (Windrum 2005). This comprises two distinct types of amateur user: the occasional user, and the serious hobbyist. Both were being sold the same dominant design, the viewfinder camera, in the 1950s. The introduction of the two radically new camera designs, the single lens reflex and the 126 enabled new entrants to take over the market, which split into two clear segments. The basis for their success was the recognition of latent, unsatisfied demand and the identification of designs with more attractive consumption possibilities for two distinct types of consumer.
- Another strategy is to build up a core following in a trend-setting group which other groups aspire to, rather than trying to win over all consumer types. This strategy is frequently linked

to a branding strategy that targets leading magazines and other media to quickly build brand awareness amongst the wider public.

- A well-established marketing strategy is the short-term price offer / ‘give away’. Here users familiar with an established product are encouraged to try out the new alternative. This can be an effective way of quickly building an installed user base, and is popular with companies seeking to gain a rapid internet presence.
- Finally, the cross-leveraging of installed user bases is a strategy that can be played by existing industry firms with a large installed user base in a related industry market (Windrum 2004). This late entry strategy that has been very successfully used by Microsoft. By linking the browser market (where it was weak) and the operating system market (where it was strong) it was able to leverage its installed base of Windows and Office across to the browser market and thereby gain control of this market as well. More recently, it successfully used the same strategy in the media player market.

4.4.2 Supply-side strategies

Case studies have also identified a number of supply-side strategies that have been used by new market entrants to overcome the large installed user base of an established firm.

- New entrants can enter mature markets by engaging in radical product innovation. Through radical innovation, late Japanese entrants successfully entered a series mature manufacturing industries in the 1960s and 1970s, such as cameras, hifi and motorbikes (Windrum, 2005; Yamamura et al., 2005).
- Radical process innovation is another means of entering a mature industry. For instance, Pilkington’s invention of the float glass process is an example of a radical process innovation that enabled a late firm entrant to dominate large scale glass production (Uusitalo, 1995; 1997). A key element in the success of late Japanese entrants in the car, motorbike and electronic industries was the development of lean production and just-in-time delivery (Windrum and Birchenhall, 2005).
- A strategic factor discussed by Porter (1985) is better/improved access to key local/national resources, i.e. wages and other input cost advantages.
- A new entrant may have new organisational structures that more effectively manage internal and external resources. Organisational innovation is the means by which firms can reorganise their hierarchial structures, internal procedures, and external relationships along the supply chain (Hölzl et al., 2006)
- Building alternative/superior distribution channels is an important strategic objective. A key aspect of Microsoft’s victory in the browser war was its exploiting its strength, and Netscape’s weakness, in the traditional and new distribution channels for browsers, i.e. PC manufacturers and internet service providers (ISPs) (Windrum, 2005).
- The formation of strategic alliances is another key strategy (Pyka and Windrum, 2003). This can take the form of open licence agreements between firms that the produce the same product. Here R&D efforts are shared, through the purchase of a licence, by all those firms wishing to produce the technology. The key potential advantage of this strategy is the ability

to quickly build a critical mass of producers, all of whom are committed to backing and developing the technology. This was the strategy used in the development of the IBM PC.

- An alternative basis for a strategic alliance is the production of complementary goods. In contrast to the open licence strategy between producers of the same product, here the alliance is between producers of goods that complement one another. This not only requires the purposeful design of interoperability between the set of complementary goods, but also the creation and management of linkages between their respective competences and knowledge bases. This entails an understanding of what partners can reasonably expect of each other (their relative strengths and weaknesses) and how partners' competences can be synthesised together to create an innovative product (Pyka and Windrum, 2003).

4.5 A co-evolutionary policy framework

Having identified necessary conditions for successions and strategies used by firms to bring them about, let us move on to consider the role of policy makers and policy options that encourage the diffusion of more environmentally friendly technologies.

To start with, we need to flesh out and develop the notion of the policy maker being one of the key agents that make up a *complex selection environment*. In the framework that I shall put forward there are three types of interacting agent: firms, consumers and political policy makers. The technology beliefs and actions of each type of agent are shaped over time by the beliefs and actions of the other agents. This interaction establishes what I have elsewhere called 'co-evolutionary learning'. In Windrum (1999) and Windrum and Birchenhall (2005), market interactions were assumed to take place within a given regulatory and societal environment, enabling the discussion to be limited to consumers and firms. Here that simplifying assumption is relaxed and the political policy maker is endogenised within the framework.

Co-evolutionary learning has two aspects. First, technology products are the objects via which different sets of agents communicate their expectations, mentalities, desires and competences. Technological change is the consequence of inter-agent learning. Rather than being an independent causal factor, a product is a mediation device. One cannot understand the emergence and use of a technology without an explicit examination of the way in which the (possibly conflicting) interests of different agents interact through a product. It is this inter-agent mediation that leads to technological change. The features of technologies change and substitutions occur as consumers, firms and policy makers interactively learn about the new possibilities associated with the production, consumption and environmental impacts of new technologies.

The second aspect of co-evolutionary learning is that feedbacks between agents also change the beliefs and actions of firms, consumers and policy makers over time. As discussed in section 3, successions require radical changes to occur in the mentalities, behaviours, and actions of agents. A succession involves far more than the substitution of one set of technology products with another. It involves the displacement of existing consumer preferences by new consumer classes with alternative preference sets, the displacement of established market firms and production structures by new firms with new production structures, and old policy regimes with new policy regimes. In sum, it is a gestalt shift. It is the replacement of one paradigm of mentalities, behaviours and actions with a new paradigm (Figure 4.1). This radical shift is what Schumpeter describes in his ‘gales of creative destruction’. It is widespread in reach, and deep in impact.

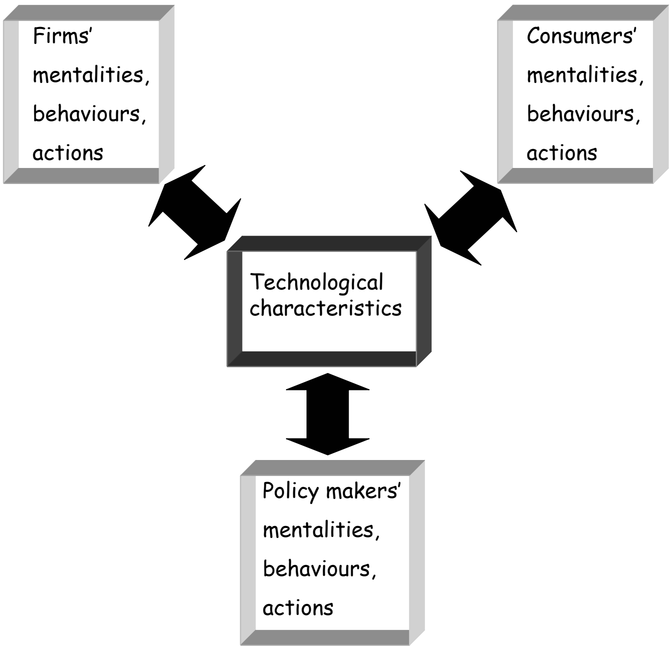

Figure 4.1 Co-evolutionary learning mediated by changing technologies

This co-evolutionary framework captures three spheres that shape environmental policy: the business sphere dominated by market relationships between customers and suppliers, the regulatory sphere of national and local government, and the wider societal sphere in which public opinion about issues are influenced by consumer organisations, environmental groups, the media, public opinion leaders and independent scientists.

4.5.1 Case studies

In order to develop the framework, section 4.5.1 considers two case studies: the evolution of car based transport and the evolution of refrigerants. The case studies provide a set of initial stylised facts that we wish to capture in the framework, and a set of outcomes we would expect the framework to reproduce (as an initial validation check). Section 4.5.2 develops the co-evolutionary framework by formalising the behaviours, actions, and interactions of our three agents: policy makers, consumers, and firms.

Case study: the evolution of car based transport and the evolution of refrigerants

The actions of policy makers with respect to the car are particularly interesting. The modern industrial city was a 19th century phenomenon. Steam power enabled the relocation of manufacturing towards cities, where most consumers lived and where labour could be found. Relocation in turn made it even more attractive for the population to relocate from the countryside and there was a dramatic growth of cities Europe and the USA. This growth was accompanied by health and hygiene problems. Swelling urban populations produced unprecedented densities of horse and human excrement, much of which found its way on to public streets. Since the middle ages in Europe, epidemics and diseases such as cholera, malaria, and tuberculosis were linked to 'miasmas' – a poisonous vapour created by decaying organic material. In the 19th century, health theorists morphed the notion. Using their knowledge of the respiratory process, they 'placed the blame on exhaled carbon dioxide in unventilated rooms and sewer gas, an often colorless, odoroless gas given off by inadequately flushed plumbing or poorly cleaned privies' (McShane 1994, p. 24). The solution was proper ventilation to remove these gases. This gave rise a 'public hygiene movement'. The movement led to a seismic shift in public opinion, to fundamental changes in urban living, and impacted on the nature of politics itself. Modern local (civic) politics was born in this era. This started with an expansion of street cleaning departments and the construction of publicly funded sewers and water systems for clean drinking water. This required the levying of local taxes to pay for this infrastructure and the development of municipal departments to construct and maintain it. The development of new municipal organisations necessitated an amendment of the liberal political tradition of upholding the rights of the individual. In the UK, for example, intervention was legitimated by the new Benthamite utilitarian philosophy of 'bringing the greatest benefit to the greatest number'.

Local government also began to take responsibility of the urban street away from individual citizens. Residents had been individually responsible for the area of road in front of their houses. This responsibility was taken over by local government. Raised pavements were constructed. This separated pedestrians from horse traffic and their excrement, leading to improvements in road safety and hygiene. City planning departments were created, charged with the responsibility of changing the urban city itself in order to reduce disease and improve health. In Europe and the USA, the target was densely packed, ill ventilated row houses that were inhabited by the city working classes. Other social groups viewed these as the source of a moral and social ill health, as well as physical ill health. The new city planners removed these tenements and worked with private sector firms to open up spaces and let in 'healthy' light and air in the rebuild. The astounding transformation of Paris provided the template: a sprawling mass of dense and poorly constructed housing being replaced with grand avenues and public parks (the lungs of the city).

It was in this new urban environment, at the constellation of social, economic and political attitudes, that the car was born. The public hygiene movement had led to the removal of human excrement from the street. The one remaining source of organic pollution was the horse. As discussed previously, the pollution associated with horses was well understood by the end of the 19th century, the pollution associated with cars was not. There was another factor favouring the car. This was the desire for the new concept of 'suburban living'. The concept was a child of the public health movement. Suburban living was both the moral and physical healthy alternative to the city.

In the medieval European city, suburbs (which literally mean homes under defensive city walls) had been prohibited or at least heavily discouraged because they could give shelter to an attacking army. With urban land within the city walls at a premium, there was little physical space for the rich to segregate themselves from the poor. Interestingly, the pattern of segregation that did exist within the city walls was the opposite of the modern city. Because transportation was primitive and most people walked to work, city centre residences commanded higher prices. Relatively poorer people tended to reside on the more remote outskirts. This pattern is still recognisable in the 17th and 18th century buildings of Amsterdam. Areas close to the old city wall, such as Mokum, were the dwelling of the city poor while the rich lived around along the central portions of the grand canals, well within the city walls.

The pattern had started to change in the 18th century as horse pulled carriages and wagons improved. The important change came with the introduction of the railroads and the electric tram (streetcar). The railways enabled the upper classes to relocate to the countryside and city to commute to work on a daily basis. The electric tram enabled the middle classes to relocate to newly built suburbs from the 1890s. Travel by tram was much faster than by horse and, unlike train travel, daily tram travel was affordable for the middle classes. The maximum speed of an electric tram was 12 mph compared to 4 mph or a horse drawn omnibus. The increase in speed translated into a significant expansion of the radius of land accessible for settlement, from 12.26 square miles with the horse drawn omnibus to 113.86 miles with the electric tram (Bass Warner Jr., 1973). In practice, the desire for suburban living remained an unsatisfied ‘new taste’ until the advent of the car. By the 1930s the car had become the dominant form of urban travel in the USA, while in Western Europe the transition occurred shortly after WWII. There was the emergence of new type of commuter: the suburban car commuter.

This highlights two important issues in technology successions. The first is the role played by *intermediate technologies*. The pollution and congestion created by horse based transportation provided a ‘window of opportunity’ for new transportation technologies. It was the electric tram, not the car, which started to displace the horse as a means of mass urban transport. While the electric tram would later be displaced by the car (and in this sense came to represent an ‘intermediate’ technology), it played an essential role in a wider transformation process of the city (Nye, 1990).

First, it was the diffusion of the electric tram which resulted in the disappearance of the horse drawn omnibus. Horses remained in certain niches - private taxis, freight transport, and in the countryside where electric power line were too expensive to erect given low population densities. Second, as noted already, it was the electric tram which enabled suburban living to be explored for the first time. Due to relatively low fares and its greater speed, the electric tram made it possible for the middle classes to live in suburban communities, far from the city centre. Third, the electric tram assisted the transformation of the city centre itself. The upper and middle classes may have vacated the city centre as a place of residency but they continued to use it as a centre for business and entertainment activity. The city centre transformed into the ‘central business district’ we know today, with a concentration of business and entertainment activities such as department stores, theatres, museums and cinemas. Fourth, electric trams were a key factor in the reconfiguring of the street. In the medieval city, the street was a place for social interactions, for public gatherings, and for trade. This had started to change in the horse era. Parks became the setting for public gatherings and trade was removed from the street and conducted in high street and corner shops. Yet, prior to the 1890s, many streets still fulfilled their function as a social meeting place. Now, as the upper and middle classes started to relocate, the street was increasingly defined as a transport artery which needed to be kept clear in order to ensure the free flow of traffic (McShane, 1994). Fifth, the electric tram further enhanced the idea of high speed transport. This had first been introduced by the advent of train travel between cities. Now the notion of high speed was being applied to travel within the city (Nye, 1990; McShane, 1994).

This leads us to another key concept, that of *deep path-dependency* across paradigms. Successions involve new technologies that in some way do something different to the old technologies that they displace. At the same time, the new technologies share certain features of the old technologies. Successions involve a sequences of technologies that unfold, one form another. Each succession contains elements of continuity and change. This trans-paradigm or ‘deep’ path-dependency can manifest itself in numerous ways. First, there is *conceptual innovation*. The concept of the street – its meaning and its uses – was in part reorganized and reinvented by the electric tram. This was subsequently further reorganized and reinvented by the car.

With regards to conceptual understandings about what new technologies are and what they can do, these are often framed by the preceding technologies. For instance, the terms used with technologies often refer to a prior technology because these were a convenient means of conveying the meaning of the new technology to the first generation of new adopters. Take the car. The concept of private (individual) automated passenger transport was radically new. Yet the concept of what a car 'is' was conveyed in its original name of the 'horseless carriage'. What is more, the power of a car engine was (and still is) measured in 'horse power'. The PC was another radical innovation. It broke with the previous technology of a central mainframe processor, operated by specialist staff, connected to users who operated dumb display terminals. Yet the very term 'personal computer', also known as the 'desktop', conveyed the idea that this is a small, stand alone version of the mainframe designed to fit on one's desk. The aeroplane displaced the ship as the primary means of mass passenger ocean transport. It is a radically different technology and means of physical travel. However, consumers of flight travel could readily understand that the 'aircraft' is 'piloted' by a 'captain' and his 'cabin crew'. As 'passengers', they 'stow' their luggage when they arrive 'onboard' and are served in-flight meals by air 'stewards' and 'stewardesses'. Commercial aeroplane firms deliberately adopted terms and concepts used in luxury ocean liner travel in order to make this radically new technology more familiar. The transference of terms and ideas assist users in the cognitive transition from one technology to another.

One also finds this combination of change and continuity in the technical features of new technologies, i.e. as well as containing the novel technical elements which define them they retain elements of the previous technology. Thus the electric tram was a motorised version of the horse drawn omnibus. It was still a public transport vehicle, but it facilitated the movement of the middle classes to the suburbs, where began to engage in private technology consumption. This paved the way for the car. The earliest cars were indeed carriages (often built by stagecoach firms) powered by an engine rather than a horse. But they were a *private* transport vehicle; one that would enable the middle classes to really engage in and develop a distinct, suburban lifestyle.

Here, then, is the core understanding of technology successions. They involve the unfolding of new technological trajectories from old trajectories as sequences of new products alter technology space, thereby facilitating the evolution of different consumption opportunities and lifestyles over time. It was not the technical features of the car *per se* that led to its rapid adoption. It was the demand for suburban living, a concept first explored in the era of the electric tram, that drove its rapid diffusion after WWII. Cars enabled the middle class to truly explore the new set of consumption opportunities of suburban living. What is more, the car was championed and popularised by the media and by policy makers. The media, particularly Hollywood and US consumer product advertisers forcefully shaped the identity of 'modern suburban living', perhaps most distinctively in the new consumption-based lifestyle imagery of the 1950s 'American Dream'. This was reinforced by the active support of successive governments in the US and Europe throughout the post-war era. The development of the car, rather than collective (public) modes of transport, was viewed as the most effective means of increasing mobility. The car was a powerful symbol of 'modernity', associated with notions of freedom and democracy, to be championed in the Cold War. As the car-based suburban living became woven ever more finely into the fabric of western society, so car ownership became a basic necessity.

Policy makers and political institutions at all levels (local, regional and national), became enmeshed in the effective running of the mass car transport system. The efficient running of a car-based transport system requires the co-ordination of traffic flow and parking spaces. The former includes support services such as road lighting, road maintenance, traffic signals and signs, repair garages, and break-down services. These are provided by a mix of private and public sector providers. Indeed a complicated regulation system is present in nearly all countries, with a combination of national and local government regulatory bodies responsible for the formulation and delivery of urban and environmental planning programmes covering road construction, urban development and traffic control. Public bodies are additionally involved in the provision of safety-related functions such as proficiency tests for drivers, regular mechanical tests for car safety, road police, and accident and emergency services. Finally, a system of taxation operates to levy car users for these publicly provided goods.

In Europe and the USA, the image of the car changed as the negative environmental impact of intensive car use became apparent in the 1970s and 1980s. But rather than supporting moves to jettison the car, public opinion in these countries views the car as a ‘necessary evil’. Rather than championing alternative technologies, policy makers are currently supporting established vehicle manufacturers and oil companies in their attempts to find ‘technology fixes’ for the worst excesses of the car, thereby ‘greening’ the existing car-based transport paradigm. A series of incremental innovations have been made to the car engine since the 1970s, such as catalytic converters, lead-free petrol, electronic engine monitoring systems to improve emissions from petrol engines, and most recently the launch of hybrid petrol-electric battery cars (e.g. the Toyota Prius).

Case study: the evolution of car based transport and the evolution of refrigerants

In contrast to the car, policy makers have actively targeted refrigerants as a means of reducing greenhouse gas emissions. In part, this no doubt reflects the fact that it is easier to implement policy in this area. There are a handful of chemical firms that produce refrigerants and the key adopters are supermarkets and manufacturers of refrigerators rather than individual households. The business-to-business refrigerant market is thus divided into two parts: low temperature applications (freezers) and medium temperature applications (fridges). In terms of volume, supermarkets (low temperature applications) are by far the largest users. The initial draft of the first Montreal Protocol agreement in 1987 did not call for a full phasing out of chlorofluorocarbons (CFCs) but this changed in light of new scientific evidence. EU legislation was already one step ahead of the Montreal Protocol in this respect, and since 1987 has consistently introduced tighter phase out schedules (Landis Gabel, 1995; Glynn, 2002).

In this technically complicated area, opinion has been shaped by scientists, specialist parts of the media, and by activist groups such as Greenpeace rather than by the general public. Policy makers have been quick to respond to shifts in opinion – and there have been a few! First, there was the shift against CFCs. The problem was that there was not an immediate alternative to hand. This was reflected in the 1990 Protocol, which did not legislate in favour of an alternative but simply established dates for the phase out of CFCs. It was left to chemical manufacturers to identify an alternative. Their first response was hydrochlorofluorocarbons (HCFCs). HCFCs were yet another example of a quick ‘technology fix’ intended to maintain an existing technological trajectory. These had actually been produced since the 1930s and so chemical manufacturers did not need to engage in radical R&D effort. The key selling point of HCFCs to supermarkets and refrigerator manufacturers was that it was a direct substitute that did not require alterations to existing machinery. This meant that early adopters could be seen to be showing concern and making a positive environmental statement. Unfortunately, while HCFCs are less damaging than CFCs, they also contain chlorine, the active ozone depleting agent. This was highlighted by scientists, the specialist media, and by Greenpeace who actively targeted the largest supermarkets with a campaign to stop the use of HCFCs. In 1992 the second Montreal Protocol agreement was amended to include HCFCs as well as CFCs, setting phase out dates for HCFCs of 99.5% by 2020 and 100% by 2030.

The next option put forward by chemical manufacturers was hydrofluorocarbons (HFCs). As the name indicates, HFCs do not contain chlorine and so are not ozone depleting. Unlike HCFC, HFC was a new chemical with no previous history of use. The first HFC was first launched by ICI, who began production of HFC-134a in late 1990. This could be used in medium temperature applications only. Later, Du Pont introduced HFC-404A which could be used in both low and medium temperature applications (Glynn, 2002).

A key advantage of HFCs, as far as supermarkets and refrigerator manufacturers were concerned, was that they did not require the replacement of existing refrigeration systems or of existing practices. Hence, following the definition given in section 2, HFCs were an incremental innovation. The one problem that new HFCs posed was their incompatibility with traditional lubricants. Chemical firms developed a new set of lubricants, though there were some initial teething problems. Still, a technology fix seemed to have been found and there was a rapid take up of HFCs. There was a real expectation on the part of chemical manufactures that, having been set a clear task by the new legislation, a solution to the CFC problem had been found. Unfortunately for them, opinion (and policy) shifted yet again in the mid-1990s. The scientific community and activist organisations successfully pushed for a change in focus. There was a shift from a narrow focus on ozone depleting chemicals to a wider consideration of global warming due to total CO_2 emissions. A key event was the introduction of the new concept of 'total equivalent warming impact' (TEWI). This replaced the previous standard indicator, the ozone depleting potential (ODP) of chemical CFC substitute. As noted, HFCs have a zero ODP because they do not contain chlorine.

TEWI came out of two studies that were jointly funded by the US Department of Environment and the Alternative fluorocarbons Environmental Acceptability Study (AFEAS). TEWI comprises two parts. One is a measure of the 'direct' global warming potential (GWP) of a chemical. This is determined by the extent to which molecules scatter infra red radiation, thereby affecting the ability of the ozone layer to repel harmful rays. The other part is a measure of the 'indirect' contribution of energy consumption to CO_2 emissions. The adoption of TEWI was by policy makers marked a policy gestalt shift. From now on the debate was no longer just about chemical refrigerants but the energy consumption of refrigerator systems. The clear message of TEWI is that the main contributor to global warming is the indirect effect associated with energy consumption. The shift was enacted in the 1997 Kyoto Protocol that was signed by 171 nations. Despite the opposition of the chemical industry, the Kyoto Protocol included HFCs along with other greenhouse gases. Prior to this, the Danish government has already announced it would unilaterally phase out HFCs.

Glynn (2002) discusses the challenge faced established chemical manufacturers by a strategic alliance of new market entrants at the turn of the century. Whereas the established chemical firms continued to back HFCs, the new entrants championed hydrocarbons (HCs). This directly challenged the established chemical oligopoly. This new strategic alliance comprised Greenpeace, DKK Scharfenstein, and a number of manufactures of hydrocarbons (HCs), such as Calor in the UK. These were not new start-up firms but an existing set of organisations with strategic strengths in HC technology, marketing and distribution, and the political lobbying of governments and supermarkets. Greenpeace worked with DKK Scharfenstein (Germany) to develop a new refrigerator technology called ‘Greenfreeze’. It was successful in getting a number of prominent UK supermarkets, such as Tesco and Iceland, to purchase HC refrigerators for a number of new supermarket sites. It also persuaded the UK Department of Environment to purchase a number of its refrigerators for research purposes in 1993. The key technical problem was to obtain sufficient supplies of HC refrigerants, given the chemical industry was opposed to HCs. This led Greenpeace to collaborate with firms from other industries. In the UK, for instance it established a relationship with Calor. Calor is the largest supplier of LPG in the UK and already had significant expertise in HC technology, which it had already used to help the aerosol industry move away from CFCs. Further, Calor had international marketing expertise and a global distribution network.

The chemical manufacturers and this new strategic alliance made a series of claims and counter claims regarding the relative merits of HFCs and HCs. Greenpeace and Calor highlighted the fact that HCs have a zero GWP, because they break down into their natural components before reaching the ozone layer, while HFCs have a high GWP. HCs can be used with traditional lubricants and, more importantly, require a significantly smaller refrigerant charge. This means they are in principle more energy saving than HFCs, with a lower TEWI and greater cost savings for adopters (ENDS Report 248, 1995). Yet, while Greenpeace and Calor made inroads with the supermarkets, they had problems selling to refrigerator manufacturers. According to Calor, it is unable to break an alliance between the chemical companies, compressor manufacturers and the refrigerator manufactures that has built up over 50 years and which is led by the chemical companies (Glynn, 2002).

The chemical and refrigerator manufacturers responded to the HC challenge in two ways. First, they highlighted the safety concerns associated with HC refrigerants. HCs are highly flammable and care in needed in their handling. Calor acted to counteract this tactic. In the UK, Calor successfully pushed for the national certification of service engineers handling HCs, and developed the training programmes. It did not succeed in the US, however. Fears of litigation prevented HC technology from establishing itself in the US. The second tactic was the introduction of new HCFC refrigerator designs that were not only more energy efficient but which significantly reduced leakage rates. Given that leakage rates of old systems were around 30%, the impact on TEWI and on users’ energy bills was significant. This enabled the HFC manufacturers to claim that, despite having a high GWP, HFC refrigerators are more energy efficient than HC refrigerators and so have overall have a lower TEWI. This is disputed by Greenpeace and Calor.

At the time of writing it seems that the established industry players are seeing off the hydrocarbon challenge. Part of this success lies in the new R&D effort along the fluorocarbon trajectory. Indeed, it is an excellent example of the sail ship effect discussed in section 3. The threat posed by the combination of new legislation and a new set of market entrants promoting an alternative technology, prompted new R&D by the established chemical and refrigerator manufacturers along the existing technology trajectory. Part of the success lies in the continuing control of the distribution channels by the chemical and refrigerator manufacturers, and the close relationships they have built up with supermarkets over the previous 50 years. As discussed in section 3, the reaction of the established players is important. If a number of these were to switch to support hydrocarbon technology then the situation would change. As it stands, they are maintaining the old the fluorocarbon trajectory. Finally, there is the timing of the HC challenge. The supermarket investment cycle in refrigeration is long – around 15 years. Hence most investments are associated with the opening of new supermarket sites. Glynn (2002) suggests that the HC challenge was 3 to 4 years too late. With the writing on the wall for CFCs, supermarkets had already started to invest in the fluorocarbon replacements – first HFCFs and then HFCs – into their new sites before the HC option became available. Following the discussion of investment cycles in section 3, there will not be another window of opportunity for at least another decade unless something dramatic changes.

4.5.2 Formalising the co-evolutionary framework

Policy maker

It is traditional for political scientists to model the policy maker as a rational and strategic actor who wishes to be elected and subsequently re-elected in the future. Policy is the means of securing (re)election. For each policy there is assumed to be a distribution of voters' opinions on a particular issue. The policy maker therefore seeks to identify the opinion of the median voter, as this maximises the probability of (re)election. In order to capture the dynamics of technological change, this model must be altered. Like firms and consumers, policy makers are boundedly rational agents that are engaged in the open-ended search of dynamically changing environments. This is due to two factors. First, the interaction between agents means policy makers not only face problems in collecting and processing information, they must also deal with the algorithmic complexity of the non-linear interactions and their ability to define preferences over expected actions, events, and outcomes. The policy maker is not initially endowed with an understanding of the underlying structure of the environment in which (s)he operates but must develop, through experience, a representation of the underlying structure. Second, radical innovation involves the introduction of new technological objects into the environment that alters the underlying structure and, hence, the payoffs associated with alternative policy actions. Agents operate in the presence of Knightian uncertainty: they cannot know, *ex ante,* the outcomes of a particular course of action (Knight, 1921). This is why successions require policy makers to develop fundamentally new policy mentalities. Radically new behaviour and action on the part of firms and consumers means old policies will no longer work. Policy makers are required to develop new mental models.

This appears to lead us to a problem. On the one hand, it is suggested that radical technological change leads to a change in the mental models of boundedly rational policy makers. On the other, it has been found that the development of new mental models by policy agents is a prerequisite for a succession occurring (section 3 above). So how is this chicken and egg circularity to be broken? There are a number of different possible avenues. Let us here consider a boundedly rational policy maker whose objective is seeking future re-election ($elect_{t+1}$). The likelihood of re-election is a function of the outcome of past elections ($elect_{t-1}$) and current policy (pol_t).

$$elect_{t+1} = f(elect_{t-1}, pol_t) \tag{2}$$

The outcomes of past elections are given, so the control variable that maximises this objective function is current policy (pol_t). Current policy itself comprises two components, previous policy decisions (pol_{t-1}) and current public opinion about key issues ($opin_t$).

$$pol_t = f(pol_{t-1}, opin_t) \tag{3}$$

Linking up with the discussion in section 2.3, the variable pol_{t-1} introduces a strong element of path-dependency in current policy making and available options. There may be good reasons for inertia and path-dependency. First, policy changes impose real administrative and technology costs on the policy maker. Second, changes impose social costs on individuals, and the rational policy is aware of this and will take this into account.

The second component $opin_t$, is a factor for change. Politicians are sensitive to large swings in public opinion on key issues. Public opinion is a very hard phenomenon to measure, let alone model. In terms of environmental policy, an important link does exist between technology use and public concern regarding its environmental impact. We have seen governments make rapid policy changes in response to concerns about pollution (McShane, 1994; Grübler, 1998; Glynn, 2002; Flannery, 2006).

Threshold effects often exist. Initially, the pollution generated by a particular technology tends to go unrecognised, only making its presence felt in the later stages of diffusion when a large number of adopters are using the technology. For instance, threshold effects existed for the burning of coal fires and the health problems associated with city smogs, and for the use of petrol engine cars and their associated health and environmental problems. Indeed, in the early days of the car it was perceived as a healthy alternative to horse transport. Street pollution due to horse urination and droppings was a major health issue in the early 1900s. In New York, for instance, horses daily produced 2.5 million pounds of manure and 60,000 gallons of urine. This accounted for twothirds of all street filth. Roads were frequently clogged by dead horse carcasses - some 15,000 dead animals being removed from the streets each year. (Flink, 1988). Infectious diseases such as typhoid, tetanus and tuberculosis were known to be harboured and carried via dried excreta. Taking the form of airborne dust, it passed through nasal passages to infect the lungs (Flink, 1988; McShane, 1994). At that time people could neither know nor guess at the health problems associated with mass car use. Hence, public opinion viewed horses negatively and the car positively.

Along with the car, a major source of greenhouse emissions is refrigeration. The physics of mechanical refrigeration are simple. A liquid refrigerant evaporates as it moves through pipes, sucking heat from an inner compartment and dissipating it through external coils. An electrically powered compressor then turns the gas into a liquid, and the cycle begins anew. The first refrigerants were sulphur dioxide and ammonia. These were known toxic agents. When introduced in the 1930s, chlorofluorocarbons (CFCs) were hailed as a new, safe alternative. In 1973 it was first realised that CFCs could reach the upper layers of the atmosphere and destroy ozone, and it was not until the 1980s that a sufficient body of empirical evidence was collected to indicate that this was actually happening in practice (Glynn, 2002).

The existence of pollution thresholds provides a key dynamic for policy change. At some point the negative externalities (in the form of pollution) of adoption become noticeable and continue to increase as the technology j continues to diffuse,

$$\hat{p}_j = \max\{0, p_j - p_j \min\} \tag{4}$$

where: $\hat{p}_j$ is the observed level of pollution that is generated by n users of technology j,

p_j is the real level of pollution generated by n users of technology j,

p_j min is the threshold.

As agents learn about and better understand the causes of the pollution, so opinion about the technology changes. Environmental opinion takes into account the observed pollution associated with different sets of technology products. Where there is a discrete choice between an old technology ($\hat{p}_0$) and a new technology ($\hat{p}_1$), environmental opinion is given by:

$$\text{environmental opinion} = \hat{p}_0 - \hat{p}_1 \tag{5}$$

Substituting (5) into (3) we derive

$$pol_t = f(pol_{t-1}, \hat{p}_0 - \hat{p}_1) \tag{6}$$

Changes in public opinion are important because they affect the probability of future election. This puts pressure on policy makers to revise policy, possibly even leading to new world views that champion the emergence of new alternatives that are more environmentally benign. In this event, there is a major break with past policy. Of course, the alternatives that emerge may themselves have negative environmental impacts, which can only be identified as they diffuse. This can in turn lead to another revision of mental models and policies in the future.

Placing this discussion within the co-evolutionary framework, the changing views of policy makers are linked to, and interact with, the changing views of consumers and entrepreneurs. Consumers and entrepreneurs are, after all, the majority of voters and like policy makers they are influenced by the activities of environmental lobby groups, the media and others who to a large extent shape public opinion. This has important consequences for the frequency and timing of successions. A new technology will quickly displace an established technology if policy makers, consumers and existing producers develop a new set of mental models around a new technology.

Whether or not a shift in policy actually occurs depends on how the forces for change play out against the countervailing and forces for path-dependency. These are captured in equation 3 where, on the one hand, a change in opinion is a force for radical change while, on the other, past policy decisions and election voting are forces for continuing path-dependency along the old technology trajectory. Empirically, we observe that radical upheavals are not the usual case and, as already stated, there are good reasons why humans do not engage in constant social upheavals. Rather than jettisoning an established technology in favour of a new alternative, a common initial response is to try to find technology fixes for the worst aspects of the established technology.

Let us next consider the drivers of consumer and firm behaviour and action within the co-evolutionary policy framework.

Consumers

There is a population of individual consumers. This is assumed to be fixed in size. Each individual consumer evaluates, and chooses between, a set of alternative *consumption possibilities* in each time period. These distinct consumption possibilities are associated with different group lifestyles or *consumer types* (as we shall call them). A consumer type is tied to the use of a particular technology. The introduction of a new technology facilitates the development of a new consumer type. A new consumer type will grow if it is supplied by firms with good quality technology products at affordable prices, and there is a supportive legislative policy (or, at least there is not a discriminatory policy). If a new consumer type continues to attract individual consumers, it will grow and eventually displace the old, established consumer types (a related approach has been developed by Aversi et al., 1999).

Note how this approach differs to the Arthur model (equation 1). Here individual tastes and preferences are not assumed to be innate and fixed from the outset. Tastes and preferences evolve over time as individuals have new experiences with radically new technology products that they come into contact with. This is done through individuals joining and leaving consumer types to which other individual consumers belong. These other consumers will also be learning about new preferences over time. This approach is very much within the spirit of Becker and his work on social economics (see Becker, 1996).

In each period, an individual consumer *i* evaluates as set of existing consumer types (T_1, T_2, $\ldots T_t$). The payoff Π_j associated with a consumer type using technology *j* at time *t* is

$$\prod_{jt} = c_{jt} + r_{jt} - \hat{p}_{jt} \tag{7}$$

where $c_{jt} = \frac{C_j}{p_j}^* - \frac{C_{jt}}{p_{jt}}$

$r_{jt} = an(b-n)$

$\frac{C_j}{p_j}^*$ is the optimum quality/price combination for consumer type *j*,

$\frac{C_{jt}}{p_{jt}}$ is the current quality/price combination currently offered by firms to consumer type *j* in period *t*,

r_{jt} is the returns to adoption associated with consumer type *j* in period *t*,

$\hat{p}_j$ is the observed level of pollution that is generated by *n* users of technology *j*.

The first term (c_{jt}) on the right-hand side of (7) captures the private good aspect of consumption while the second and third terms (r_{jt} and $\hat{p}_{jt}$) capture the public good aspect of consumption. r_{jt} is a quadratic function of the number of consumers that have previously joined this consumer type. This captures both the positive and negative externalities discussed in section 2. Initially, the positive network utility discussed by Arthur dominate, However, as increasing numbers of individual consumers join type *j* so the negative externality starts to dominate. In terms of the car, the most obvious negative externality effect is *traffic congestion* caused by other car users simultaneously commuting to and from work, and to and from holiday destinations. There may also be a snob effect of the type discussed by Liebenstein. In the presence of negative externalities, there is an upper limit on *r*. What is more, there may eventually be a decline in the value of *r* as increasing numbers of users continue to join this type. If this is the case, then there is a direct incentive for individual consumers to search for alternative consumer types. Here lies the potential for a group of consumers being willing to experiment with, and adopt, a new technological alternative. Ironically, through its very success, an old technology regime lays the seeds of its own destruction.

An individual consumer will decide to join a new technology consumer type T$_1$ or an old technology consumer type T$_0$

$$\text{if} \qquad \begin{cases} c_1 + r_1 + \hat{p}_1 \geq c_0 + r_0 + \hat{p}_0 & \text{join } \mathrm{T}_1 \\ c_1 + r_1 + \hat{p}_1 \leq c_0 + r_0 + \hat{p}_0 & \text{join } \mathrm{T}_0 \end{cases} \tag{8}$$

Alternative service characteristics and new consumer types

We need to be more precise about the meaning of the term ‘quality’. Saviotti-Pyka (2004) and Windrum-Birchenhall (2005) use Lancaster’s characteristics approach (1971). Lancaster observed that a product is not demanded for itself but because of the stream of services that it provides the users over its lifetime. This gives us a precise meaning of C_j. Windrum (1999) and Windrum-Birchenhall (2005) suggest that what distinguishes alternative technologies is the distinct sets of service characteristics they offer, i.e. each technology offers something that the other technology cannot. Thus, when comparing old and new technology products, users are comparing the *different sets* of service characteristics offered by each. Tying this observation to the earlier discussion of consumption possibilities, I suggest there is a relationship between the service characteristics of technology C_j and the consumer type T_j it facilitates.

$$C_j \Leftrightarrow T_j$$

The double arrow indicates that they are not independent and that there is a correspondence between the two sets. Here lies the significance of new technologies: they *open up new consumption possibilities*. When there are different competing technologies, individual consumers are able to choose between different consumption possibilities (lifestyles) associated with alternative consumer types.

The discussion of the car provided a clear example of this. There are a number of service characteristics that distinguish the car from other (public) modes of urban transport. First, it offers the user a flexible, single source method of travelling between any two points. Second, it is an explicitly individual, rather than collective, form of mobility. These service characteristics facilitated the development of suburban living, turning what had hitherto been little more than been an unfulfilled aspiration into a reality. This new consumer type – the suburbanite – was adopted by the aspiring middle classes. They wished to imitate the upper social groups and move away from the inner cities (which were now left to the working classes). Here social differentiation combined powerfully with a snob effect. The middle classes were abandoning one consumption type for another: *urban living* for *suburban living*. The adoption of this new consumption type promised a move away from ‘city pollution’ to ‘healthy suburbs’, and the middle classes were actively supported in their ambitions by political elites and the media. As highlighted, a technological succession is associated with the emergence of new consumer classes with new preference sets. This contrasts sharply with technological substitutions, where users adopt a new technology because it better fulfils the same role, increasing consumer utility over an unchanging set of preferences.

Let us consider the distinction made between the private and public consumption components of (7). The service characteristics offered to users by an individual design may include its individual pollution performance. For instance, the Prius hybrid engine car offers consumers a 'greener' option. But the aggregate level of car pollution depends on all the different types of car design that comprise the current stock of cars in use, how many trips (and their distance) are made using the current car stock, and how many trips are made using alternative modes of transport. This distinction between individual and collective (aggregate) is the basis of a well-known paradox in environmental economics. Namely, improvement in the pollution performance of an individual product design may actually lead to an increase in environmental pollution if it leads to a significant increase in the use of pollution contributing products. Further, as noted, environmental impact of technology use is often non-linear, captured by the threshold specification of $\hat{p}_{jt}$.

Firms

Firms are heterogeneous with respect to the set of service characteristics *C* that make up their product designs, and the consumer type *T* that they target. Here we shall assume that firms do not switch target consumer types. This is a stylised fact gleaned from the case studies.

In each period, every firm has a current design, a productive capacity (setting an upper limit on output), and a non-negative inventory of stock carried over from the previous period. The price of its design is determined by a fixed mark-up on the unit cost of production (i.e. prices do not adjust to clear the market)[11]. This means that coordination of market supply and demand occurs through quantity adjustments. Firms adjust output and capacity in light of past demand.

Firms compete by offering a combination of service characteristics, with a consequent price, they believe will be more attractive than those offered by their rivals. In this way, a firm effectively offers consumers a distinct point in a multi-dimensional service characteristic/price space. Product innovation is the means by which firms search this multi-dimensional space. Unit cost is the sum of an average fixed cost (a common fixed cost Φ that includes a fixed cost for innovation, divided by the firm's level of production y) and an average variable cost that is a function of the good's design (the vector of service characteristics offered by the design)[12]. Average variable costs of the design are taken to be independent of the level of production. The average total cost $\overline{TC}$ is given by

$$\overline{TC} = (\Phi/y) + (\Sigma_k \gamma_k c_k(x_k)) \tag{9}$$

---

[11] Fixed mark-up pricing is a common feature of a number of evolutionary models. Probably the best-known piece of research in this area is Hall and Hitch (1939). Their study of 38 businesses found that the most common pricing procedure was average cost with a 'normal' mark-up. The same finding has appeared in more recent studies in the US and UK. More recently, more than half the 72 US firms (with annual revenues of $10+ million) interviewed by Blinder (1991) reported that cost-based pricing was a moderate or very important factor in explaining price adjustment, while 37% of respondents in the Hall, Walsh, and Yates (1997) study of 654 UK companies use a cost-based pricing rule.

[12] In order to simply, this average variable cost function (mapping designs on to unit variable cost) is assumed to be a fixed convex function that is common to all firms. The marginal cost of each service characteristic k is positive and increasing. The partials of the average cost function are positive, and the diagonals of the Hessian are positive.

where γ_k are constants and the c_k are monotonically increasing, convex functions of the *k*th service characteristic. Firms set prices according to a simple mark up rule:

$$p_{jt} = (1+\eta_{jt}) \times \overline{TC}_{jt} \tag{10}$$

where $\overline{TC}_{jt}$ is the *j*th firm's average total cost in period *t* and η_{jt} is the *j*th firm's mark up in period *t*. To simplify, let us assume there is a common and constant mark up, so that $\eta_{it} = \eta$.

At the beginning of every period, each firm offers a quantity (a stock q_{jt} plus current production y_{jt}) of design x_{jt} at a price p_{jt} that reflects both the variable cost of producing the current design and an average fixed cost. Given sales s_{jt} and the level of production y_{jt} a firm's net revenue σ is

$$\sigma_{jt} = p_{jt}s_{jt} - \overline{TC}_{jt}\, y_{jt} \tag{11}$$

This profit is added to its monetary wealth M_{jt}, which changes in each period according to $M_{jt+1} = M_{jt} + \sigma_{jt}$.

Successful firms, with high levels of sales and production, gain a direct advantage from their lower average fixed costs and (in turn) lower prices, making their goods more attractive to consumers. Where the growth of productive capacity is financed from initial wealth or profits, so a firm with relatively high levels of sales, and thus relatively high profits, will be able to finance a higher growth of capacity. Loss making firms, by contrast, will initially use up their monetary wealth and, once exhausted, will finance itself by reducing (i.e. selling) capacity[13]. Once capacity is exhausted, the firm is bankrupt and exits the market.

Each firm is randomly assigned a target consumer class. Its design strategy is to maximise the utility function of this target class. As noted, we will assume that firms do not switch between consumer classes. Consequently, success depends on a firm's ability to innovative. Product innovation involves the creation and evaluation of new designs in each period. New designs are created through a combination of imitation (of the service characteristics of successful rivals) and through the firm's own R&D activities. These are modelled using a modified genetic algorithm (see Windrum and Birchenhall, 2005). As a consequence of performing R&D, there is a random mutation in one or more service characteristics. In the evaluation process, the firm uses its knowledge of the utility function of its target consumer class to determine whether the proposed design should be put into production or else the existing design should be retained. In other words, a firm will only implement the proposed design if this raises the utility to the target class.

It is worth emphasising, once again, that a 'design' is a particular point in the service characteristic space and *not* a point in an engineer's technical space. Windrum and Birchenhall (2005) simplify by assuming that each firm knows the utility function of their target consumer classes but does not know how to implement an optimal design. The technical problem facing the firm is the construction of an optimal design that maximises the utility of the target consumer type *given* a set of production and

---

[13] In the models of Malerba et al. (1999), and Saviotti and Pyka (2004), this process is tempered by venture capitalists and other financiers. These may bankroll a firm for a sufficient time in order for it to identify a more competitive design.

innovation costs[14]. Since firms can only alter their current designs through innovative search – the filtered process of imitation and mutation – in design space, there is no guarantee they will produce designs with characteristic / price combinations that are optimal for their target consumer type.

Key variables

The framework contains a number of key variables that affect the probability of a succession occurring.

On the *consumer* side, a necessary precondition for a succession is a set individual consumers who are willing to switch away from an existing consumer type and to experiment with a new type. This situation arises when there are negative externalities to belonging to an established consumption type. In the framework there are two types of negative externalities:

1. negative externalities associated with the number of previous adopters, such as physical congestion and snob effects. These place an upper limit on r and can even lead to decreasing returns to r.
2. pollution p generated through the use of artefacts.

The timing of a new technology is therefore important. There may be new entrants with new technology designs, and policy making setting a new policy environment, but if individual consumers are not willing to experiment with these new consumption possibilities, then a succession will not occur.

With regards to *firms*, timing is also important. A set of interested consumers and an appropriate policy environment is of no value without a set of innovative firms that are willing and capable of developing radical new designs. Factors affecting the probability of a succession in our framework are

3. the market entry conditions, notably barriers to entry such as high set up costs due to capital intensity of production, i.e. high fixed costs Φ (Windrum and Birchenhall, 2005). Malerba et al. (1999) and Saviotti and Pyka (2004) highlight the availability of venture capital is a key factor affecting new start ups (though not established firms).
4. the quality /price of the initial set of designs offered by new technology entrants. Windrum and Birchenhall (2005) found a trade off exists between the direct utility of the characteristics offered by a product and the indirect utility of product price. Consequently, a new technology, offering superior characteristics, will not necessarily displace an old technology if the price differentials are large.
5. the subsequent R&D performance of old and new technology firms is an essential factor. New entrants stimulates R&D by old technology firms (the sail ship effect). If new firms are more effective innovators then the probability of a succession is high. However, if the old technology firms are more effective innovators, then a lock out is likely to occur (Windrum and Birchenhall 2005).

---

[14] This assumes that the whole of the service characteristic space is technically feasible, and that this optimal design is the one that maximises the intrinsic utility of the target consumer class. Recall that improving a service characteristic increases direct utility but also increases cost and price, thereby reducing indirect utility. In the model intrinsic utility is a strictly concave function of the design vector and so the optima will be the unique stationary point of this function.

6. time is a key variable. In Windrum and Birchenhall (2005) a succession is more likely to occur (a) the shorter the time old technology firms have (i.e. prior to new technology firms entering the market) to innovate and develop designs that closely match the preferences of their target consumers, and (b) the longer new firms have to innovate and turn their initial set of designs into a set of designs that are optimal for their target consumer type.

Turning to key variables associated with the *policy maker*, timing is once again important. A policy change that creates a legislative environment which favours a new technology is a fundamental prerequisite for a succession. The framework captures a number of factors that affecting the probability of this occurring. In equation 3 current environment policy depends on previous environment policy decisions (pol_{t-1}) and current public opinion about key environmental issues ($opin_t$). Given this, the probability of a policy change occurring depends on

7. the strength of path-dependency pol_{t-1} on current policy;
8. the extent to which the policy maker is willing to discount path-dependency pol_{t-1} in favour of changes in current opinion $opin_t$;
9. the extent and speed to $opin_t$ changes as a consequence of observed environmental pollution $\hat{p}$ generated through technology use;
10. speed of adjustment from a change in policy view to the implementation of the new policy. As we saw in the case studies, this was relatively quick in the case of CFCs but has been extremely slow in the case of the car.

4.6 Conclusions

Let us conclude drawing together the policy lessons that can be gleaned from this co-evolutionary approach, and discussing the set of policy instruments that are available to policy makers for promoting technology successions.

To start with, alternative technologies compete with a complex selection environment. This environment contains policy makers, firms, consumers and other agents. A policy maker may be a very important agent within this complex selection environment, but (s)he is still only one agent. Technology successions involve the unfolding of a sequence of technologies which alter the space of service characteristics, thereby facilitating a change in consumption possibilities and lifestyles over time. When successions occur there is the displacement of existing consumer preferences by new consumer classes with alternative preference sets, the displacement of established market firms and production structures by new firms with alternative production structures, and displacement of old policy regimes with new policy regimes. It is, in every sense of the word, a gestalt shift - the replacement of one paradigm of mentalities, behaviours and actions with another.

Policy makers who wish to engage in the promotion of technology successions need to understand the dynamics and these processes and must, as a consequence, engage in a different vision of policy. In traditional political science the policy maker is a rational decision maker who knows with certainty all the available options and the payoffs to each option. This is rather like the farmer who understands perfectly the seasons, knows the optimal time to plant and harvest, which crops to sow, how to control for pests and so on. Here, by contrast, we accept that policy maker is not a rational planner with perfect information but a boundedly rational agent that operates in a world of Knightian uncertainty. The policy maker is an agent engaged in the search for new possibilities – hopefully ones that are more environmentally friendly that the current options, although this can never be truly known *ex ante*. The more appropriate image for this policy maker is the explorative hunter of the unknown. It is a process of continual, ongoing search in an environment that is dynamically changing over time.

Within this conceptual framework, what are limitations and opportunities of policy? The case studies on refrigerants and the car provide a number of policy insights.

- There is a web of interrelated political, commercial and social interests that co-evolve around an established technology. These generate the path dependencies that support an existing technological trajectory.
- Support for a technology fix is invariably the first reaction of policy makers to the pollution generated by an established technology. If a technology fix cannot be found the legislators will consider more radical solutions.
- Firms do not see the wider picture. Indeed, one should not expect them to see the wider picture. In the case of CFC legislation, for instance, better understanding of the consequences of refrigerants on the ozone layer has been driven by scientific discoveries, by environmental pressure groups, and by policy makers.
- The focus of firms is narrow because R&D and industry position are highly path-dependent. As we see for both car and chemical companies, their success has been built on a particular set of knowledge and skills. Further, their industry position rests upon a set of established relationships with consumers, policy makers and other firms along the supply chain. R&D and industry position are invariably threatened by new, alternative technologies and the entry of new firms. This is why chemical manufacturers continue to champion fluorocarbons and oppose hydrocarbons, and why car manufacturers champion hybrid engines.
- Policy makers try to shape public opinion, but they are also highly sensitive to changes in opinion. In environmental policy, changes in opinion have had a significant impact on policy. Here changes are strongly influenced by scientific discoveries, by interest groups and the media as well as by firms and consumer groups.
- Having said this, the case studies also indicate that the speed of policy changes will be affected by a number of factors. First, there is the time horizon. Changes need to have a fairly immediate impact, or at least be seen to be having an impact. This is the case in the case of setting phase outs for refrigerants (CFCs then HCFs and now HCFCs). This short-termism is driven by the electoral cycle. Since policy makers are looking for re-selection at the next election, policy changes made today must having an impact on the electorate before the next election takes place. Second, the scope of policy change is limited by the availability of alternative technologies. This involves factors beyond the policy maker's control. Not only must more environmentally benign alternatives be available, but they must be actively championed by (new) firms and consumers if they are to replace existing technologies. Hence, policy makers could legislate for the phasing out of CFCs but were not in a position to positively legislate for an alternative. They are unable to legislate against mass car transport until an alternative is in place and being championed by a key set of firms and consumer types.

- Policy makers support new alternatives on the basis that they 'appear' to be more environmentally benign. The actual long-term environmental impact of a new alternative *cannot* be known *ex ante*. The pollution impact may not necessarily have been evident in the invention or innovation stages, but becomes apparent in the diffusion stage as producers and consumers continue to experiment with the new technology, possibly combining it in novel ways with other technologies that appear in the diffusion phase. It is ironic that the public health movement, which began in the late 19th century as a reaction to the pollution and disease of overcrowded cities, gave rise to the concept of suburban living which in turn helped to promote the car and the pollution impact that it has had. There is an important policy message here. Shifts in policy, driven by shifts in opinion, may not always turn out to be ideal. Clearly, it would have been easier if one could have known the downsides of mass car transport *ex ante* but, as stated, they operate in conditions of Knightian uncertainty and so this is simply not possible.

Building on past research, the paper has identified a set of necessary conditions for a succession. Pulling these together:

1. There must be a functional equivalence between the new and the old technology products, e.g. alternative modes of transport and alternative means of refrigerating perishable food.
2. Novelty. The new technology products must offer users new consumption possibilities, based on service characteristics that were not available in the old technology products.
3. The emergence of new consumer types, with individual consumers willing to trade-off the novel consumption possibilities of the new technology against the benefits of an established technology.
4. Windows of opportunity. The probability of a succession occurring varies at different moments in time. These can be related to economic and investment cycles, and to different stages of the life cycle.
5. Consumer dissatisfaction with an established technology will arise if there are negative network externalities, or an upper limit to network externalities has been reached.
6. The pollution generated by an existing set of technology goods is a further factor that can cause consumers to consider a new alternative.
7. New technology firms that enter the market develop new technical competences and new conceptualisations of what the market is, and what it can become.
8. The availability of start-up capital is essential for new firms.
9. Market entry is easier when there are low initial set-up costs (e.g. due to low capital intensity).
10. The quality /price of the initial set of designs offered by new technology entrants must be sufficiently competitive to immediately attract consumers.
11. The relative R&D performance of old and new technology firms. New technology firms must be more successful innovators than old firms. If this is not the case, then improvements in quality/price performance of old technology products will lock out the new technology. This means that new firms must more successfully engage in both *product* and *process* innovation.

12. New entrants may enjoy better access to key local/national resources, i.e. wages and other input cost advantages.
13. New entrants may have superior organisational structures that more effectively manage internal and external resources.
14. Superior distribution channels are another key source of competitive advantage.
15. Late market entrants can gain advantages through the formation of strategic alliances. Producers of the same product group can use open licence agreements, while producers of complementary goods can establish interoperability between their designs.
16. Time itself is a key variable. A succession is more likely to occur the shorter is the time old technology firms have to develop an effective set of product designs prior to new technology entry, and the longer new technology firms have to develop a set of designs that are optimal for their target consumer type.
17. New policy models. The development of a new technology paradigm requires the development of a new policy model leading to a shift in policy.
18. The probability of a policy shift depends on the strength of policy path-dependency on current policy and the extent to which the policy maker is willing to discount path-dependency in favour of changes in environmental opinion (based on new scientific understanding, the actions of interest groups, and the media).
19. There are factors that affect the speed with which changes in environmental opinion translate into policy change.
20. Finally, there are factors affecting the speed of adjustment from a change in policy view to the actual implementation of a new policy view.

Finally, let us consider the set of *policies* and *policy instruments* that support technology successions within our co-evolutionary framework. We have discussed the sources of path dependencies and webs of supporting interests for an old technology which, once in place and having stabilised over long periods of time, are exceedingly difficult to break. Policy needs to identify and act on these path dependencies. At the same time, policy needs to promote the emergence of new consumer types and new technology firms.

The starting point is policy itself. Promoting new conceptual understandings amongst other agents requires policy makers engage in prior conceptual innovation, in policy and an assessment of policy on firms and consumers. This is an ongoing process of learning. Policy makers cannot know the final outcomes of technology successions *ex ante* but learn, through interactions with other agents, about their benefits and costs as new technological trajectories unfold. The co-evolutionary learning means policy makers must expect disappointment with regards to initial expectations about a new technology's environmental impact. These can only be truly identified when technologies actually diffuse (and even then, usually with a time lag). It is important that policy makers adopt a long term view in the formulation of environmental policy. Successions involve unfolding sequences of technological displacements. It is therefore important to evaluate 'intermediate technologies'. These technologies may not ultimately deliver on their initial environmental promise. However, they may play an important role as a stepping stone to a subsequent technology that is more environmentally benign.

Policy makers need to consider the timing of policy changes. If distinct windows of opportunity exist then there will be an optimal timing and frequency of new technology adoptions. To do this, policy makers must take into account data on successions. This requires the development of new policy indicators, such as the network externalities of consumers using established and new technologies, and rates of entry and exit amongst old and new technology firms.

Given that government institutions are themselves invariably enmeshed in supporting an established technology, policy change needs to identify and address areas in which government institutions provide complementary goods, or where legislation supports the established technology and locks out alternatives. In the example of the car, we observed that governments are deeply embedded in the provision of both complementary public goods and legislation that maintain car based transport systems. Of course, in advocating change, policy makers must have a viable alternative to champion. Where strong action has been taken successfully, such as road charging in London (the 'congestion charge'), a viable and readily available alternative was in place. Where this was not the case, such as in the Netherlands, the policy was highly unpopular and failed.

Demand side policies should encourage the development of new consumer types and discourage continuing adherence to old consumer types. The government can lead by example and be purchase of new technologies (Freeman et al., 1982). In some cases, the government may itself be a major customer for an industry. In others it may be an influential purchaser, even if it is not a major customer, and so can give a lead. This was case for the UK Department of Environment when it purchased a number of greenfreeze refrigerators for research purposes in 1993.

More traditional policy instruments are taxes and legislation. By raising taxes on established technologies and subsidising new technologies, policy makers can change the relative prices of old and new technology products and thereby alter their relative quality/price performance. Through environmental legislation, policy makers can ban an established technology outright (as in CFCs) or else specifically target one or more features of the old technology that give rise to pollution. If these taxation and legislation changes are persistent and strong enough, then the policy can lead to a switch in the consumer population away from an old consumer type to a new consumer type.

Government policy may seek to assist the demand side strategies used by new entrants to overcome the network externalities enjoyed by old technology firms. For instance, changes may be made to competition law. The introductory price offer or 'give away' (e.g. two for the price of one) is a well-established marketing strategy, used by UK and the US firms, to encourage consumers to try out a new product. This is practice is currently illegal in Germany. Major changes would need to be made to German competition law if its national government were to support the use of this late entrant strategy. Competition law would also need to be changed in order to facilitate the cross-leveraging installed user bases. This was one of the contentious issues in the US Department of Justice and EU cases against Microsoft.

Government policy may seek to alter the length of the investment cycle by legally specifying the maximum period in which consumers to repurchase a particular technology product. This has been used by the Japanese government to speed up the reinvestment cycle for cars. If older vintages are more polluting than newer vintages, changing the scrapping rate will have a significant impact on total pollution. Suppose this policy were applied to an established technology but not to the new technology. It would change the investment horizon, and hence the net present value of services, in favour of the new technology and against the old technology.

Turning to the supply side, policy should encourage new technology producers and discourage old technology producers. There are various policy instruments available. Some are quite traditional. For instance, taxes and subsides on final goods have already been discussed[15]. Another traditional policy instrument is preferential subsidies to R&D performed by new technology firms.

Less traditional are instruments that are designed to encourage new firm entry. Venture capital is important for new start up firms, though not for existing firms that move into anew market. If there is a shortage of private sector venture capital then government may set up its own venture capital funding. Government may also need to pull out of the public provision of complementary infrastructure and services to the old technology and start developing infrastructure and services that are complementary to the new technology. This may approve highly contentious. For example, each time national governments slow down road building programmes or seek savings in traffic policing, there are outbursts in the media and strong pressure is applied by lobby groups.

To summarise, effective policy-making requires more that a set of strategy recipes. It requires a policy framework in which one can identify the optimal timing and likely impacts of alternative policy strategies. To this end, the paper began the process of outlining such as framework. It is a framework that captures the co-evolutionary learning of interacting agents – policy makers, consumers and firms – as they explore successive sequences of new technologies over time.

The framework captures the twin forces of path-dependency and change that characterise technology successions. Understanding these twin forces is central to successful policy formulation. This entails as appreciation of the role policy itself plays in maintaining path-dependency to an established technology or in promoting change to a new alternative. Herein lies a fundamental shift in perspective. One must break away from the traditional view of a policy maker as an independent rational planner. First, the policy maker is not independent but is one agent, amongst a number of interconnected agents, that makes up a complex selection environment. Second, the policy maker does not have perfect information, does not know all of the different options that may become available and does not know the final payoffs associated with different policy choices. Instead, the policy maker is a boundedly rational agent who operates in a world of Knightian uncertainty. Policy makers, like other agents, learn by engaging in the search for new possibilities. Hopefully, these are more environmentally friendly that the current options. Policies promote technologies which appear to be more environmentally friendly *ex ante*, but this can never be truly known until they have diffused *ex post*. An analogy was drawn between this type of policy maker and the hunter who explores the unknown. The hunter policy maker in engaged in an ongoing search of an ever changing environment – change that is due to the emergence of new technological trajectories.

Understanding the dynamics of technology successions is, of course, essential for policy. In addition, effective policy making requires the identification of windows of opportunity, and the exploitation of deep (trans-trajectory) path dependencies in order to change the behaviour, actions and beliefs of firms, consumers, and government institutions. These have been discussed in the paper, as has the development of new policy indicators. The optimal timing and design of policy requires novel policy indicators, such increasing and decreasing network returns and firm entry/exit, together with insight into that factors that affect these variables. Finally, the paper has identified a set of policy instruments

---

[15] In practice, it is more efficient to collect taxes and pay subsidies to firms rather than to consumers.

that can be used by policy makers to promote successions to new, more environmentally friendly technologies.

References

Abernathy, W.J. and J.M. Utterback (1975). A dynamic model of product and process innovation. *Omega* 3: 639-656

Anderson, P. and M.L. Tushman (1990). Technological discontinuities and dominant designs: a cyclical model of technological change. *Administrative Science Quarterly* 35: 604-633

Arthur, W.B. (1988). Competing technologies: an overview. In G. Dosi, C. Freeman, R. Nelson, G. Silverberg and L. Soete (eds.). *Technical Change and Economic Theory*. London: Pinter

Arthur, W.B. (1989). Competing technologies, increasing returns and lock-in by historical events. *Economic Journal* 99: 116-131

Arthur, W.B., Y.M. Ermoliev and Y.M. Kaniovski (1987). Path-dependent processes and the emergence of macrostructure. *European Journal of Operational Research* 30: 294-303

Atkinson, A.B. and J.E. Stiglitz (1969). A new view of technological change. *Economic Journal* 79: 573-578

Aversi, R., G. Dosi, G. Fagiolo, M. Meacci and C. Olivetti (1999). Demand dynamics with socially evolving preferences. *Industrial and Corporate Change* 8: 353-408

Bass Warner Jr., S. (1973). *Streetcar suburbs: the process of growth in Boston 1870-1900*. New York: Atheneum

Bassanini, A. and G. Dosi (1998). *Competing technologies, international diffusion and the rate of convergence to a stable market structure*. IIASA Interim Report (IR-98-012), Laxenburg, IIASA

Becker, J.S. (1996). *Accounting for Tastes*. Cambridge, Mass.: Harvard University Press

Blinder, A.S. (1991). Why are prices sticky? Preliminary results from an interview study. *American Economic Review* 81(2): 89-96

Clark, K.B. (1985). The interaction of design hierarchies and market concepts in technological evolution. *Research Policy* 14: 235-251

David, P.A. (1975). *Technical choice, innovation and economic growth. Essays on American and British experience in the nineteenth century*. Cambridge: Cambridge University Press

David, P.A. (1985). Clio and the economics of QWERTY. *American Economic Review* 75: 332-336

Dosi, G. (1982). Technological paradigms and technological trajectories. *Research Policy* 11: 147-62

Dosi, G. and L. Orsenigo (1988). Co-ordination and transformation: an overview of structures, behaviours and change in evolutionary environments. in G. Dosi, C. Freeman, R. Nelson, G. Silverberg and L. Soete (ed.), *Technical Change and Economic Theory*. Pinter: London

ENDS (1995). *Calor gas makes strides in refrigerant battle*. ENDS Report 248: 26

Flannery, T. (2006). *How man is changing the climate and what it means for life on Earth*. London: Allen Lane

Flink, J.J. (1988). *The Automobile Age*. Cambridge, Mass.: MIT Press

Foucault, M. (1972). *The archaeology of knowledge and the discourse on language*. New York: Pantheon

Freeman, C. (1982). *The economics of industrial innovation* (2^{nd} ed.). London: Pinter

Freeman, C., J. Clark, and L. Soete (1982). *Unemployment and technological innovation: a study of long waves and economic development*. London: Pinter

Geels, F. (2005). *Technological transitions and system innovations: a co-evolutionary and socio-technical analysis*. Cheltenham: Edward Elgar

Glynn, S. (2002). Constructing a selection environment: competing expectations for CFC alternatives. *Research Policy* 31: 935-946

Gort, M. and S. Klepper (1982). Time-paths in the diffusion of product innovations. *Economic Journal* 92: 630-653

Glynn, S. (2002). Constructing a selection environment: competing expectations for CFC alternatives. *Research Policy* 31: 935-946

Grübler, A. (1998). *Technology and global change*. Cambridge: Cambridge University Press

Hall R.L., and C.J. Hitch (1939). Price theory and business behaviour. *Oxford Economic Papers* 2: 12-45. Reprinted in T. Wilson and P. W. S. Andrews (eds.) (1951). Oxford Studies in the Price Mechanism, Oxford: Clarendon

Hall, S., M. Walsh and A. Yates (1997). How do UK companies set prices? *Bank of England Working Paper* 67. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=114948.

Henderson, R. and K. Clark (1990). Architectural innovation. *Administrative Science Quarterly* 35: 9-30

Hölzl, W., A. Reinstaller and P. Windrum (2006). Organisational innovation, information technology, and outsourcing to business services. in L. Rubalcaba and H. Kox (eds) *Business Services in European Economic Growth*. London: Palgrave MacMillan. Forthcoming

Hoogma, R., R. Kemp and J. Schot (2002). *Experimenting for sustainable transport. The approach of strategic niche management*, London: EF&N

Islas, J. (1997). Getting around the lock-in in electricity generating systems: the example of the gas turbine. *Research Policy* 26: 49-66

Knight, F.H. (1921). *Risk, Uncertainty, and Profits*. Chicago: Chicago University Press

Landis Gabel, H., (1995). Environmental management as a competitive strategy: the case of CFCs. In H. Folmer et al. (eds). *Principles of Environmental and Resource Economics: A Guide for Students an Decision Makers*, Cheltenham: Edward Elgar

Leibenstein, H. (1950). Bandwagon, snob and Veblen effects in the theory of consumer demand. *Quarterly Journal of Economics* 64 (2): 183-207

Liebowitz, S.J. and S.E. Margolis (1990). Fable of the keys. *Journal of Law and Economics* 33: 1-25

Liebowitz, S.J. and S.E. Margolis (1996). Should technology choice be a concern of anti-trust policy? *Harvard Journal of Law and Technology* 9: 286-318

Liebowitz, S.J. and S.E. Margolis (1998). Dismal science fictions: network effects, Microsoft and anti-trust speculation. *Policy Analysis*: 324

Malerba F., R.R Nelson, L. Orsenigo and S.G. Winter (1999). History friendly models of industry evolution: the computer industry. *Industrial and Corporate Change* 8 (1): 3-41

McShane, C. (1994). *Down the asphalt road: the automobile and the American city*. New York: Columbia University Press

Mensch, G. (1979). *Stalemate in technology: innovations overcome depressions*. Ballinger, Cambridge

Mokyr, J. (1990). *The Lever of Riches*. Oxford: Oxford University Press

Mom, G., C. Werff and A. Bos (1997). *De Auto: van avonturenmachine naar gebruikscoorwerp*. Deventer: Kluwer

Nelson, R.R. and S.G. Winter (1977). In search of useful theory of innovation. *Research Policy* 6: 36-76

Nelson, R.R. and S.G. Winter (1982). *An Evolutionary Theory of Economic Change*, Cambridge (Mass.), Harvard University Press

Nye, D. (1990). *Electrifying America: social meanings of a new technology*. Cambridge, Mass.: MIT Press

Porter, M.E. (1985). *Competitive advantage of nations*. New York: Free Press

Pyka, A. and P. Windrum (2003). The self-organisation of strategic alliances. *Economics of Innovation and New Technology* 12 (3): 245-268

Rosenberg, N. (1969). The direction of technical change: inducement mechanisms and focussing devices. *Economic Development and Cultural Change* 18: 1-24. Reprinted in Rosenberg, N. (1976). *Perspectives on Technology*. Cambridge: Cambridge University Press

Sahal, D. (1985). Technological guideposts and innovation avenues. *Research Policy* 14: 61-82.

Saviotti P.P. and A. Pyka (2004). Economic development by the creation of new sectors. *Journal of Evolutionary Economics* 14 (1): 1-36

Schumpeter, J.A. (1912). *The theory of economic development: an inquiry into profits, capital, credit, interest, and the business cycle*. Oxford: Oxford University Press

Schumpeter, J.A. (1939). *Business cycles: a theoretical, historical and statistical analysis of the capitalist process*, New York: McGraw-Hill

Shy, O. (1996). Technology revolutions in the presence of network externalities. *International Journal of Industrial Organisation* 14: 785-800

Uusitalo, O. (1995). *The flat glass industry: the effect of float glass on industry structure*. Helsinki: Helsinki School of Economics and Business Administration

Uusitalo, O. (1997). Development of the flat glass industry in Scandinavia 1910-1990: the impact of technological change. *Scandinavian Economic History Review* 3: 276-295

Veblen. T. (1953). *Theory of the leisure class: an economic study of institutions*. (Revised edition). New York: New American Library

Vernon, R. (1966). International investment and international trade in the product life cycle. *Quarterly Journal of Economics* 80: 190-207

Vincenti, W.G. (1990). *What engineers know and how they know it. Analytical studies from aeronautical history.* Baltimore: John Hopkins University Press

Windrum, P. (1999). *Unlocking a lock-in: towards a model of technological succession*. MERIT (Merit Research Memoranda 1999-010), University of Maastricht

Windrum, P. (2001). Late entrant strategies in technological ecologies: Microsoft's use of standards in the browser wars. *International Studies of Management and Organisation,* special issue on 'Innovation management in the New Economy. 31 (1): 87-105

Windrum, P. (2004). Leveraging technological externalities in complex technologies: Microsoft's exploitation of standards in the browser wars. *Research Policy* 33: 385-394

Windrum, P. (2005). Heterogeneous preferences and new innovation cycles in mature industries: the amateur camera industry 1955-1974. *Industrial and Corporate Change* 14 (6): 1043-1074

Windrum, P. and C. Birchenhall (2005). Structural change in the presence of network externalities: a co-evolutionary model of technological successions. *Journal of Evolutionary Economics* 15 (2): 123-148

Yamamura, E., T. Sonobe and K. Otsuka (2005). Time path in innovation, imitation, and growth: the case of the motorcycle industry in postwar Japan. *Journal of Evolutionary Economics* 15 (2): 169-186

5 Simulating consumer behaviour: a perspective

W. Jager

Faculty of Economics (dept. Marketing), Rijksuniversiteit Groningen.

Contact: w.jager@bdk.rug.nl

ABSTRACT

This paper is aimed at providing a perspective on how human behaviour can be formalised in environmental models. First of all, an argument is provided why multi-agent simulation is the appropriate tool to study behavioural dynamics. Next it is being argued that a generic model of human behaviour is needed, because environmental relevant behaviour is not a specific category of behaviour. Such a generic model would include human needs and decision-making as elementary parts. Because the practical use of models including human behaviour is largely determined by the possibility to conduct policy experiments, the formalisation of rules is being structured along the dimensions of product, price, place and promotion, these elements constituting the basic marketing strategies to influence behaviour. The paper concludes with a discussion on validation and prediction of model outcomes, and a number of issues to be addressed in the development of applicable models.

5.1 Introduction

Multi-agent simulation is increasingly being advocated – and used - as a new tool to explore the dynamics in various kinds of systems where human behaviour plays a critical role. Agent-based modelling implies that a population of artificial people is constructed, thus allowing to model two critical elements of many systems in which people behave: people are different (heterogeneity), and affect each other (social interaction). As such, there is an increase in publications using this methodology in fields such as market dynamics, crowd and riot control, demographic developments, traffic behaviour, organisational performance, and man-environment relations. In this paper we aim to focus on the application of multi-agent simulation in understanding man-environment relations. In recent years there has been an increasing number of publications on agent-based simulation in environmental management (see e.g. Hare and Deadman (2004) for an overview). Hare and Deadman (2004) define six requirements to classify agent-based simulation models in environmental modelling. These requirements are: (1) coupling social and environmental models, (2) micro-level decision making of populations of agents, (3) social interaction between agents, (4) intrinsic adaptation of decision-making and behaviour by the agents, (5) population level adaptation (birth- and death-rates), and (6) multiple-scale level decision-making.

Our purpose in this paper is to draw a perspective on how behaviour could be formalised in environmental models, focusing on the above requirements 2, 3 and 4. For the purpose of policy-making experiments, we postulate that the availability of agent-based simulation models that validly describe behavioural dynamic processes are a prerequisite. Critical issues to be discussed relate to the added value of using multi-agent simulation, how to formalise human behaviour in agent-based models, how to experiment with policy measures using agent-based models and how to validate models. Whereas the multi-agent methodology is increasingly being recognised as a suitable tool to explore the dynamics of human behaviour in various contexts, the methodology is still in its infancy concerning practical applicability, and discussions are running on how social simulation may contribute to a better understanding of real-world dynamics (e.g., Deffuant et al., 2006). This paper aims at describing a venue to increase this practical applicability, in particular with respect to testing policy measures in complex man-environment systems.

5.2 Why multi agent simulation?

The experimental tradition within the social sciences has resulted in an abundance of laboratory studies revealing how various personal and contextual factors influence people's behaviour. Results obtained in fields focusing on, e.g., attitudes, interpersonal processes and social dilemmas, are often relevant in understanding how people interact with their natural environment. However, one of the main problems of translating experimental findings to real-world situations is the complex nature of reality. Where labstudies usually focus on isolated or a very confined set of variables in a well-controlled environment, in real life human behaviour is determined by a multitude of constantly changing and interacting variables. For example, the type of car we drive has a significant impact on emissions of carbon-dioxide, yet only a few consumers – if any – base their preference and eventual selection of a car on its environmental consequences. Rather, what car we drive is determined by a large number of factors. First of all, the mere fact that most of us own a car relates to the decision of many consumers before us to use a car, which indicates that a whole infrastructure has emerged supporting car-use. This also affects our life in terms of where we settle, where we work, and what kind of social and leisure-time activities we employ. Having identified such factors as a basic driver of why most of us own a car, the specific type of car we drive is subject to many, sometimes conflicting factors such as our need for space, the budget we have available, the locations which we visit, operating costs, our environmental concern, the importance of safety, brand and color preferences and many more. In deciding amongst all possible options, we usually use a limited set of factors, and friends often play a critical role in defining a 'socially acceptable' set of alternatives. One level of complexity can be related to the fact that consumers are heterogeneous concerning their weighting of the factors. Where some focus primarily on the affordability of a car, others also value the environmental performance, or prefer a sports car or even an all-road car as an expression of their identity. A next level of complexity is related to the continuous change of these preferences. We experience changes such as becoming parents, moving to another house, changing oil prices and taxes, the introduction of new car-models, discussing with friends on what cars are hot or not and the like. Moreover, information on the effects of emissions on local air quality and global warming is often complex, unclear, and under debate. As a consequence our preferences are moving, and the next car we buy may be very different from our current one.

Complexities as such are typical for environmental dilemmas in general, and can be related to the level of the individual, the social system (group) and the environmental system. On a personal level, factors such as knowledge, attitudes, goals (stakeholders), power, personality and the like determine the individuals' disposition towards performing certain behaviour. Whereas the effects of single (or a confined set) of relevant factors is often known, for combinations of such factors the effects are unclear, as the effects interact in a complex manner, and cannot be simply added or subtracted. Next, at the level of the social system, complexities arise from large populations of heterogeneous people, interacting through networks that are subject to change. The resulting group processes are very complex and often unpredictable, such as the opinions of a population, diffusion of new behaviours and fashions. Complexity of the environment arises from multiple interacting environmental processes (e.g., global warming affects the climate, which affects chances on floods and storms), the multiple environmental outcomes of some behaviours (e.g., compared to gas, diesel fuel combines lower emissions of carbondioxide with higher emissions of particles) the different time-scales involved, the large time lags of effects, and several feedback and feed-forward mechanisms.

Whereas systematic experimentation is possible by changing a limited number of factors in experimental settings, e.g., the effect of environmental information on valuation of fuel efficiency of cars, laboratory experiments are not a suitable tool to study human behaviour in complex environmental settings because effects of interventions may depend strongly on the context in which they are implemented, including other policy measures. For example, people may respond very differently to (combinations of) policy measures, and may respond to what other people are doing, causing effects to be sometimes rather unpredictable. Case-studies offer an alternative tool for studying behaviour in complex environments. Whereas these may provide valuable information on the complexities in given domains of consumer behaviour, due to the lack of experimental control case studies do not provide the researcher with causal behavioural mechanisms. Rather, a methodology is needed that allows for experimenting with behavioural processes within actors, social processes between actors and interactions between actors and the environment. Agent-based simulation is a tool that offers a perspective on simulating human behaviour in complex environments.

First, agent-based simulation allows for experimenting with the complexities at individual, social and environmental levels by formalizing populations of artificial humans, called ‘agents’, in an artificial world. Many factors can be included in these formalizations, and using computer simulations one can conduct thousands of experiments in a short time, thus allowing for exploring the effects of many combinations of factors. Such experiments also reveal to what extend certain combinations of factors result in fairly robust outcomes, or, on the contrary, result in outcomes that are very susceptible for minor changes in the factors.

Second, agent-based simulation allows for the modelling of interactions between individuals. Assuming that social interaction causes information and norms to spread, the accumulation of these interactions can be studied on a population scale. The number of agents to be included in a computer simulation depends only on the power of the computer, but even with an ordinary PC it is possible to simulate populations from 10,000 to 1 million. This allows for showing how population phenomena, e.g., opinions on the energy issues, may emerge from interactions at the local level.

Finally, agent-based simulation allows for experimenting with policy measures without harming people and the environment. Via simulations the mid-term and long-term effects of policy measures can be studied in scenarios. Different scenarios including differing forecasts of economic development and environmental quality could be used to test the efficacy of policy-measures under different conditions. Experiments can be repeated under the same starting conditions with different policy measures (and assumptions on their effect on individuals) as many times as we want. Therefore it is possible to simulate different policy strategies to examine which policy may be optimal in the specific situation.

5.3 What behaviour to simulate?

Asking ‘how to simulate behaviour’ first requires answering the question ‘what behaviour to simulate’. Whereas we indicated to have an interest in behaviour with environmental relevant consequences, we have to realise that ‘environmental relevant behaviour’ is not a really fruitful concept for two reasons. First, in principle all behaviour has environmental consequences, so all behaviour is environmentally relevant. Whereas it can be stated that some behaviours have a larger (negative) impact on the environment, focussing exclusively on these behaviours neglects behavioural alternatives that have a lesser impact. Just studying environmental harmful behaviour may cause that one overlooks viable strategies increasing the appeal of environmental benign behaviour.

Second, people usually do not perform behaviour primarily because of its environmental consequences. On the contrary, most people are not - or partly - aware of how their behaviour affects their environment. This is because people do not behave primarily to generate certain environmental effects, but rather effects pertaining to their own living conditions. They buy a particular car to serve their transportation needs given their financial capabilities. Moreover, they appreciate issues such as styling, performance and the like, and issues like emissions are most likely only to be mentioned as ‘nice’ if they happen to be positive. Negative information on emissions – if made available – is more likely to be filtered away according to cognitive dissonance processes. As a consequence, one should realize that we should study human behaviour in general rather than considering ‘environmental relevant behaviour’ as a category on its own. Henceforth we are interested in identifying generic strategies for behavioural change, which may be used to stimulate more environmentally benign behaviour. Essential from a policy-making point of view is that this implies that people may be stimulated to perform more environmentally benign behaviour for other reasons than the environment: one may decide to buy a Toyota Prius (a low-emission hybrid car) because of status reasons – several Hollywood stars already adopted such a car - rather than for environmental reasons (alone). Hence, the conclusion is that if we are interested in simulating behaviour of people in relation to environmental outcomes, we do not formalise ‘environmental behaviour’, but rather need a more generic model of human behaviour.

5.4 Why not to use too simple rules

Concluding that we should study human behaviour in general, the next issue is how to formalise generic human behaviour in agent-based models. Many agent-based simulations have been developed using very simple rules, where the agents try to optimise their outcomes. Especially in ‘econophysics’ researchers replicate macro-level phenomena using simple agent rules. This approach is not fruitful, because replicating macro-level phenomena does not automatically imply that the relevant micro-level processes are captured in the model. An example is the simulation of flocking by Reynolds (1987). His work on the flocking boids has become a key example how simple local rules lead to complex macro behaviour. Reynolds used three rules for each agent: avoid collisions with nearby flockmates, attempt to match velocity with nearby flockmates and attempt to stay close to nearby flockmates. Due to measurement errors by the flockmates, an impressive flocking like behaviour comes out of this model. Reynolds mentions that ‘success and validity of these simulations is difficult to measure objectively. They do seem to agree well with certain criteria and some statistical properties of natural flocks and schools which have been reported by the zoological and behavioural sciences. Perhaps more significantly, many people who view these animated flocks immediately recognize them as a representation of a natural flock, and find them similarly delightful to watch’ (Reynolds, 1987). One might derive the impression that we have a better understanding of flocking behaviour. However, research on schooling of fish illustrate that we lack a good understanding of the micro-behaviour of fish in relation to schooling. Indeed, information about the behaviour of nearby neighbours is found to be a crucial factor in empirical studies, but which behavioural rules are in use is a puzzle and so far computational models fail to reproduce observed behaviour in detail (Camazine et al., 2001). The same applies to many agent-based models of human behaviour. Whereas macro-level phenomena are often replicated, the agent-rules at the micro level are not founded in empirical observation of behavioural theory. The main problem is that if policy measures are tested in such simulations, the macro-level outcomes are determined by how these policy measures affect the processes at the micro-level. If this representation is not correct, the processes will be unrealistic, and the resulting macro-level policy effects may be far from realistic.

From the perspective of modelling human behaviour, several researchers have shown the impacts of using different formalisations. One of the early agent-based models that showed how different assumptions on human behaviour affected man-environment interactions was developed by Bousquet, Cambier, Mullon, Morand and Quensiere (1994) in their simulation model of a society of fishermen in Niger. Two main categories of objects in this simulation are concerned with the ecological dynamics (including fish) and the dynamics of fishing (including households). The model is aimed at simulating the decisions made by households (the micro-level) in a fluctuating environment. The ecological dynamics refer to the reproduction, growth, mortality, migration and competition of fish in four different biotopes (river, channels, ponds and flooded areas). In the simulation, various characteristics of the fish population may change (e.g., evolution of biomass, weight and age structure) depending on the fishermen's decisions and fishing strategies.

Their first simulations started with economically rational households. Simulation of a two-year period with two flood periods showed that all fishermen reacted on the same moment in the same way, depending on the state of the ecosystem. That is, they all started fishing in either the river, the channels (if filled), the ponds (if not dry), the plains (if flooded) or they started working on the land (agriculture), all depending on the water level. A second simulation tried to include environmental and human variability by means of variability of catch. Also in this simulation the fishermen tended to react in the same way on the same moment, however, the changes were less immediate (there was more diversity) than in the previous simulation. As such, during some moments in time some fishermen were fishing in the river whilst others are working on the land. A third simulation introduced a variable risk perception. The optimistic fishermen selected their behaviour by reference to their best achievement in the previous two weeks. The pessimistic fishermen referred to the worst achievement in the last two weeks, and the risk-neutral fishermen referred to their average achievement. Equipping the fishermen with different risk-perceptions (1/3 optimistic, 1/3 risk-neutral and 1/3 pessimistic) further increased the diversity in behaviour. Thus far, the households did not communicate directly with each other. To introduce communication, a next simulation allowed households to have two or five social relations. Bousquet et al. (1994) concluded that the size of the communication network might change the use of space, although the results are not clearly interpretable. A next simulation used anthropological data on where certain tribes are fishing. Not surprisingly the simulation shows that fishermen belonging to different tribes ('Bozo' vs. 'Somono') are fishing at different locations. The various simulations summarised here also revealed differences with respect to the resource dynamics, that is, the number and biomass of fish during the two-year period in the simulation.

Bousquet et al. (1994) demonstrate that the simulated man-environment interactions are sensitive to the precise formalisations of human behaviour. Moreover, the more complex rules, inspired by empirical observation, generated behaviour that demonstrated more 'face validity'. However, the development of agent rules tool place on a rather ad-hoc basis, and behavioural theory was not being used in a structural way. This translates in the formalisation of only one need (food), whereas it is obvious that more needs are involved in human behaviour. The question that rises then is how to formalise human behaviour, and what theories to implement.

5.5 What to include in agent rules?

The earlier example on car-use shows a number of prototypical elements of human behaviour that have to be captured in agent rules. Different attributes play a role when buying a car, such as space, economy, safety and design. Basic needs behind these attributes such as subsistence (travelling to job), safety and social status may be distinguished. Moreover, when actually buying a car is seems obvious that we do not behave like an optimising homo-economicus. Comparing all possible cars on all attributes is virtually impossible. Rather, we select from a small subset of cars that we like, for example because they fit with our transportation needs, their good brand image and the positive evaluation by friends. Moreover, learning and cognition also plays a critical role, as our preferences and attitudes towards certain (brands of) cars have been shaped by previous direct experiences, experiences from other people and information we obtained earlier. Hence, many theories on generic human behaviour and underlying determinants and processes can guide the development of agent rules to simulate manenvironment interactions. Janssen and Jager (2003) proposed that theories on needs, decision-making processes and processes of (social) learning constitute key issues to be modeled because they respectively describe the motivation to perform behaviour, choice processes in selecting behaviour, and storage of positive and negative experiences after performing behaviour.

5.5.1 Human needs

Basic human needs are considered to be the basic drivers of human behaviour. From a policy point-of-view one may state that in stimulating environmentally benign behaviour, one should try to develop policy measures that increase the need-satisfying capabilities of these behaviours, and at the same time decrease these capabilities for environmentally damaging behaviours. The basic question arising here is: what needs do people have, and how to they govern our daily behaviour? Whereas traditional economics often state that the concept of needs is obsolete, and consumers are assumed to have wants that are insatiable in principle, the fields of consumer research, economic psychology, marketing studies and motivation research have all provided a rather rich foundation for producers, retailers, marketers and advertisers wanting to know how to design and sell products that consumers will buy. These attempts to develop an understanding of consumer motivations have drawn quite specifically from the needs-theoretic framework that formal economics has rejected (see e.g., Jackson et al, 2004). Whereas different meanings of needs can be found in the literature, we adhere to the definition of needs as underlying internal forces that drive or guide our actions. For example, a need for safety might refer to the underlying drive that people have to protect themselves and the motivation that this provides them with to build houses, buy clothes, enact punitive legislation against criminals and so on. This definition distinguishes needs from wants because firstly, needs are considered non-negotiable; and secondly, the failure to satisfy a need has a detrimental effect on the overall health of the individual.

Within this definition of needs several inventories have been presented describing universal human needs, the hierarchical ordering of Maslow (1954) undoubtly being the most renown one. From the bottom to the top of his needs-pyramid, Maslow (1954) distinguishes physiological and safety needs, needs to belong and be loved, and then ‘higher’ cognitive, aesthetic and moral needs. The lower-order needs in this hierarchy, Maslow called material needs; the middle-order needs were referred to as social needs; and the higher-order needs, Maslow called growth or ‘self-actualisation’ needs. Maslow (1954) argued that needs low in the hierarchy must be at least partially satisfied before needs higher in the hierarchy may become important sources of motivation. This hierarchical approach to human needs has also drawn criticism because it (1) appears to deny access to the satisfaction of higher needs in less developed country populations, (2) legitimises a distribution of power in favour of those who specialise in so-called ‘higher’ needs – such as intellectuals and ascetics (Galtung 1990) – in developed country populations, (3) over-emphasises the individualistic nature of needssatisfaction, and (4) understates the importance of society, culture and the natural environment, by treating these as secondary in importance to individual motivation. Max-Neef (1991, 1992) developed a taxonomy of human needs avoiding the disadvantages of a hierarchical ordering. Max-Neef makes a distinction between nine ‘axiological’ needs – subsistence, protection, affection, understanding, participation, identify, idleness, creation, and freedom – against four ‘existential’ categories: being, doing, having and interacting. Without going in too much detail, the main point is that human actions – addressed by Max Neef as satisfiers - may have effects on single or several needs simultaneously. Max-Neef (1992) makes a distinction between the following types of satisfiers:

- destroyers or violators occupy the paradoxical position of failing completely to satisfy the need towards which they are directed;
- pseudo-satisfiers generate a false sense of satisfaction of the need;
- inhibiting satisfiers satisfy one need to which they are directed but tend to inhibit the satisfaction of other needs;
- singular satisfiers manage to satisfy a single category of need without affecting;
- satisfaction elsewhere; and
- synergistic satisfiers manage simultaneously to satisfy several different kinds of needs.

Within the context of human behaviour it seems obvious that often conflicts emerge between needs at different levels. For example, whereas the need for safety may be better satisfied buying a large heavy car, the need for identity, e.g. focusing on being an environmentally responsible person, may be satisfied more buying a smaller energy efficient car. Hence many products that are being bought can be understood as being inhibiting satisfiers.

5.5.2 Human decision making

Related to the issue of needs is how people make a decision in performing a particular behaviour, and what information is being used in making a decision. The neoclassical consumer theory assumes consumers to be rational utility maximisers, having fixed preferences that are complete and transitive, and always wanting to have more. Assuming multiple needs, consumers are assumed to consider the satisfaction of their needs as a multi-attribute optimization problem. However, many economists and psychologists challenge the assumptions of this paradigm because of several arguments (e.g., Hayakawa, 2000: p. 2; Bowles, 1998). First, the time available for information gathering and cognitive processing is limited and has to be distributed amongst the many decisions people make in their daily lives. Hence, consumers often construct preferences on the spot when confronted with a (new) decision problem (e.g., Bettman, 1979; Payne et al., 1992; Slovic, 1995). Second, people live in a social environment that affects the preferences they have. Currently an increasing number of economists discuss the relevance of preference change for economic research, and incorporate these ideas in their studies (Pollak, 1978; Witt 1991; Güth and Yaari, 1992; Bowles, 1998). Third, people base their decisions frequently on incorrect information due to their limited (and diverse) cognitive abilities. For example, the amount of attentional resources that people can invest is affected by factors such as age, arousal, and emotional state (Hasher and Zacks, 1979). Also psychological biases, such as are apparent in framing effects (e.g., Tversky and Kahneman, 1979) affect decision-making. And fourth, often the environment in which people make their decisions is uncertain by nature, making risk taking an inherent trait of the decision process.

The seminal work of Simon (e.g. 1976) on bounded rationality offers a perspective on why using simple decision strategies such as habits and complying with a norm may be a rational thing to do.

The essential argument is that humans optimise the full process of decision-making (procedural rationality), not just the outcomes (substantive rationality, Simon, 1976). This holds that consumers may decide that a certain choice problem is not worth investing a lot of cognitive effort (e.g., buying groceries), whereas another choice problem requires more cognitive attention (e.g., buying a car). The less important a decision problem is the less cognitive energy one is willing to invest in the decision. This implies that the less involved people are in making a decision, the more likely it is that they will use a simple heuristic in making a choice. Rather than using a multi-attribute approach where behavioural options are being judged on all their need satisfying capabilities, in such low-involvement conditions people are more likely to consider a single – or limited set – of needs in their decision process. Obviously, the needs that are most likely to be in focus are expected to dominate the decision process. Hence it can be expected that in low-involvement conditions environmental relevant outcomes are likely to be undervalued in the decision making process.

Two dimensions in human decision-making

The strategies – or heuristics – people employ in the decision-making process can be organized along the dimensions of cognitive effort and social v.s. individual orientation. This has important implications, because the use of social heuristics such as imitation and normative behaviour has a strong effect on how new behaviours diffuse through society.

Cognitive effort

Three basic factors have been found determining the cognitive effort people invest in a decision task, namely (1) consumer involvement, (2) complexity of the decision and (3) constraints. Involvement refers to the importance of the consequences of the decision. Tversky (1969; 1972) demonstrated that when a decision is less important (in terms of consequences), decision-makers are more likely to use a simpler heuristic instead of using all information available. The importance of a decision can be related to the underlying needs that are associated with the decision. The attention of a consumer for a decision problem will be higher the more important the goals or needs involved. This attention may be voluntary, in case new opportunities are being discovered, but also involuntary when the satisfaction of need is being jeopardised (e.g., Kahneman, 1973).

Complexity of the decision is related to the number of opportunities one may choose from, the number of attributes that these opportunities possess and the degree to which trade-offs between different needs elicit emotional distress. Many decision tasks confront people with a multitude of relative simple opportunities, such as clothing, sunglasses and furniture, offer an enormous range of opportunities to choose from. Other decisions are complex because of the multitude of attributes that are involved in the alternatives. Sometimes decision problems are characterised by both a multitude of opportunities, each of which possesses many attributes. Buying a car is such a decision problem, where many models and types exist, and attributes such as comfort, power, style, safety and fuel consumption have to be taken into consideration. Constraints relate to the time available to make a decision and the decision-makers cognitive capacity and knowledge level which determine the degree of cognitive processing.

Social versus individual decision-making

The next dimension addresses the social versus individual orientation of the decisionmaking process. In some instances people make up their mind individually, whereas in other situations the behaviour or advice of other people is being used. This process is also addressed and word-of-mouth (WOM) Social influences and WOM play a pivotal role in many consumer behaviours. These social influences may range from exchanging detailed information on various products between consumers to simply imitating the behaviour of another. Three critical issues are (1) when to compare, (2) who to compare with, and (3) how to compare.

When to compare

A first important factor here is the availability of clear information on the opportunities from which to choose. The less information available, or the more complex and contradictory this information is – as often the case with environmental and social outcomes –, the more likely one will look at others people to get an indication of the best outcome. Especially behaviour having social consequences is subject to such complexities and hence will elicit much social processing. Uncertainty is a key concept in understanding this use of social information. The less information a person has on the opportunity characteristics, the more uncertain he/she may be regarding what course of action to take. Festinger (1954) already indicated that social processing is more likely to occur under uncertainty. When behaviour is selected following the use of a social heuristic, this uncertainty will be resolved. Especially when the decision process is automated, people may hardly experience the uncertainty leading towards the use of the social heuristic. Some people have a lower tolerance for uncertainty than other people, and hence are more likely to look at the behaviour of others, and are more sensitive to social information.

A second factor that determines a person tendency to use social information is the visibility of other people's behaviour. This is more a constraint on the possibility of using social information. When the behaviour and opportunity use of other people is less visible, it is less likely that social information can be used in the decision making process.

The use of social information is especially of importance in understanding the diffusion of new behaviours. When large numbers of people strongly rely on the behaviour of others, new innovative behaviour hardly diffuse through society (e.g., Delre, Jager and Janssen, 2004). However, when a critical mass already adopted the new behaviour, this social susceptibility will propagate the further diffusion of the new behaviour. Hence the use of social information may both inhibit and stimulate the diffusion of new behaviour, depending on the number of people that already adopted. Here the connectivity between people plays a critical role. The more scale-free a social network connecting people is – indicating that some people are very well connected (hubs) and many others are more loosely connected – the faster a diffusion process may be. In the context of environmental relevant behaviour it is important to realise that social processes may propagate the diffusion of environmentally benign – or endangering – behaviour without environmental consequences of the behaviour playing a dominant role. Henceforth the recycling of bottles and paper may normative driven behaviour rather that an expression of environmentally conscious behaviour, and the buying of SUV vehicles may be strongly influenced by what other – successful - people in one's environment do.

Who to compare with

Social Comparison Theory (Festinger, 1954) states that people are motivated to consciously compare their opinions and abilities with those of other people. Comparison processes occur both at the individual and group level (Faucheux and Moscovici, 1972). These comparisons follow dimensions such as the possession of material goods, financial means, status, principles, attitudes and skills. With respect to opinions, people have a drive to roughly conform to others. With respect to abilities, people have a drive to be (somewhat) superior to others. A critical point is that similarity is a multidimensional concept. As such a consumer may compare with different people depending on the similarity dimension that is most relevant in a given comparison context. For example, when comparing on sports drinks, consumers are most likely to compare with people engaging in the same type of sport, where other dimensions such as social status or political beliefs are less influential. This implies that comparison processes may be directed at different people, depending on the issue at stake. This also relates to the underlying structure of the social network connecting consumers. It is assumed that much comparison takes place within dense networks where people are (spatially) clustered among many similarity dimensions, such as income, religion, culture, social status and the like. However, the less a particular type of consumption is clustered among these dimensions, the more likely it is that more distant links – if related to the comparison dimension that is relevant in this context – will be used. Network studies demonstrated that the type of network that is being used in a given context is critical in the diffusion of new behaviours (e.g., Janssen and Jager, 2003; Delre et al, in press). Hence in studying social comparison processes in consumer behaviour it is critical to identify the similarity dimensions that are relevant for a certain type of consumer behaviour.

How to compare

Concerning how to compare a distinction can be made on the type of information that is being exchanged in the comparison process. Basically a distinction can be made between more informative and normative conformity behaviour. Informative conformity implies accepting information from others as evidence about the reality (Deutsch and Gerard, 1955). Informative conformity can be demonstrated in two ways. First, individuals may search of obtain information from knowledgeable others in their social environment and use this information for their decision-making. This information may be specific on product characteristics (nutritional value of a brand of sport drink), but also more generic on product categories (the advantages of sport drinks in general). Second, individuals may make inferences about reality based on the observed behaviour of others (Park and Lessig, 1977). This implies imitating the behaviour of another person, assuming that the other person is eligible.

Normative conformity influence relates to the individual’s desires to comply with the positive expectations of others (Deutsch and Gerard, 1955). As such, normative conformity behaviour is driven by a desire to obtain social approval from others and to achieve a sense of belonging (Cialdini and Goldstein, 2004). This influence can result in decisions that conform to the dominant opinion or decision in one’s social network. This is also addressed as a simplifying strategy.

5.6 Formalising agent rules

In the previous section a quite extensive description has been described of the determinants and processes that appear to be relevant in consumer behaviour and that could be captured in agent rules. In formalising agent rules we have to capture essential behavioural principles in a comprehensive set of simple rules. Besides the rules describing the behavioural drivers and processes of humans, we also have to formalise the actual behaviours the agents can perform. A promising framework for this is provided by the so-called four P's of marketing (McCarthy, 1960): product, price, place and promotion. Combining these four factors constitute the so-called marketing mix. Concerning the product several characteristics determine the final utility (or satisfaction) of consumers with a product. Aspects that are mentioned often are brand name, functionality, styling, quality, safety, packaging, repairs and support, warranty and accessories and services. Whereas 'product' is often defined in a rather narrow sense, only addressing the physical properties of a product, it can be stated that behaviour in a more general sense can be captured under this heading. This would imply that also aspects that play a critical role during the performance of behaviour – i.e. the use of a product – can be incorporated. In the case of cars this would imply that also aspects related to e.g., different driving styles can be described in terms of utilities, such as operating costs, safety, comfort and the like. Price is also a key factor, and in particular with respect to policy making it is clear that pricing, in terms of taxes, fines and subsidies, is a frequently implemented strategy to change behaviour. Within marketing often mentioned aspects of pricing are pricing strategy (skimming, penetration) retail price, volume discounts and wholesale pricing, cash and early payment discounts, seasonal pricing, bundling, price flexibility and price discrimination. However, also different tax regimes and the like can be captured under pricing. Concerning place the aspects relate to the use of distribution channels, market coverage, specific channel members, inventory management, warehousing, distribution centres, order processing, transportation and reverse logistics. Finally, promotion refers to promotional strategies (push & pull), advertising, personal selling and sales force, sales promotions, public relations and publicity and the marketing communications budget.

Whereas it seems obvious that this full complexity of the marketing mix cannot be translated in a transparent – read simple - social simulation model, the distinction between the main marketing-mix components of product, price, place and promotion would allow for the development of simulations that also allow for the exploration of human behaviour in different conditions, and the effects of policy strategies. In particular in an environmental context the testing of (governmental) policy measures should also allow for the possibility to enter new products (or behavioural options) in the system. This would allow for testing policy measures in different conditions of market dynamics, where both consumers and producers are affected by these measures and respond accordingly.

In the following a formalisation will be provided of some main constituents of agent rules following the 4 P's framework. This formalization is not conclusive, but is aimed at providing an applicable framework.

5.6.1 Product

In an economical sense a product is a physical object or service that satisfies a markets need. A product is the complete bundle of benefits or satisfactions that buyers experience when purchasing, using and disposing a product, and thus constitutes the sum of all physical, psychological, symbolic, and service attributes. In this sense product can be defined in a broad sense, including a wide range of behaviours. Reverting to the example of cars, the product here not only involves to the physical item (metal and plastics), but also relates to how it is being used (frequency, driving style), its psychological meaning (freedom, safety), its symbolic meanings (e.g., identity and status) and service aspects (available infrastructure). Hence when the term product is being used the reader must bear in mind that a broad definition, including product related behaviour, is meant.

A basic assumption is that agents strive towards being satisfied with a product – or the behaviour associated with it. This implies that products and associated behaviours may differ concerning the need satisfying capacities – or utility - they have. Hence it is proposed that utility is being defined as the needs that are being satisfied by products and associated behaviours. Obviously, in line with the discussion on needs in section 5.1, products and behaviours differ concerning the number of utilities that are coupled to a product, and often trade-offs exists where alternative products score better on one utility, but worse on another. The utility of a product is first dependent on the degree to which a product or behaviour matches the individual preferences of an agent. This individual preference can be understood as a weighted multi-attribute composite of product characteristics such as brand, functionality, styling and quality as described earlier (e.g., Lancaster, 1966).

Individual preferences

Two different kinds of individual preferences may be distinguished. This distinction dates back to the work of Thurstone (1931). First, an individual preference may take the shape of ‘the more, the better’. This is usually being addressed as a ‘vector model of preferences’. This kind of preference relates to factors such as quality, service, flexibility and reliability (e.g., as measured with SERVQUAL: Parasuraman et al., 1985; 1988). The basic formulation for this type of product preferences is:

$$U_{inj} = A_{jn}$$

With: U_{inj} = Utility of consumer i on attribute n for product j
A_{jn} = Score of product j for attribute n

Formalising Ajn as a value between 0 and 1 results in a utility score between 0 and 1.

The second type of individual preference relates to a relative position on a scale, which implies an ‘ideal point’ type of preference. This kind of preference relates to e.g., design, colour, taste and the like which may have an optimum utility for a consumer on a more arbitrary position of the scale. This can be formalized as follows:

$$U_{inj} = 1 - |A_{jn} - P_{in}|$$

With: U_{inj} = Utility of consumer i on attribute n for product j,
A_{jn} = Score of product j for attribute n,
P_{in} = Preference of consumer i for attribute n.

Formalizing both the product attribute score as the consumer’s preference as a value between 0 and 1 results in a utility score between 0 and 1. Including heterogeneity of the consumers for P_{in} allows for modelling markets where consumers have different preferences (taste). Within the fields of marketing and psychology researchers often use Multi Dimensional Scaling techniques to measure preferences and construct perceptual maps of consumers’ preferences (e.g., Sheppard, 1962a; 1962b; Kruskal, 1964). An important assumption we make here is that preferences are not fixed, but can be changed due to new information and social influence, as will be discussed in the section on promotion. An increasing number of economists discuss the relevance of preference change for economic research, and incorporate these ideas in their studies (Bowles and Gintis, 2000; Akerlof, 1984; Kahneman et al, 1986; Tversky and Kahneman, 1974; Pollak, 1978; Witt, 1991; Güth and Yaari, 1992; Bowles, 1998). Also within the field of marketing the concept of preference change is accepted by many scholars.

Social preferences

Besides individual preferences, consumers also express social preferences for products, as the utility of a product often depends on what other agents use the product. Here the Veblen effect, referring to conspicuous consumption (Veblen, 1899) – or keeping up with the Joneses (Duesenberry (1949) – becomes manifest, expressing that consumers weight their individual preference and social preference for a product in deciding what they aspire (see e.g., Chao and Schor, 1998). For matters of simplicity we propose that this social utility is considered as one of the product attributes. However, this social utility does not depend on the product attributes, but on the number of other people in ones environment that are using the product. This can be formalized as follows:

$$U_{\mathrm{inj}} = N_{\mathrm{j}}/N$$

With: U_{inj} = Utility of consumer i on attribute n (here the social attribute) for product j
N_{j} = number of neighbours consuming product j.
N = number of neighbours

The calculation of the social attribute is very susceptible for the type of network being formalized. In earlier work (Janssen and Jager, 2003; Delre et al, in press) it has been demonstrated that different formalizations, such as a regular lattice, small world network or scale free network have major impacts on both the speed as the degree to which a new product gains market share.

Weighting of preferences

Different attributes will be weighted differently, both on a market level as well as on individual consumer level. Some markets display a high sensitivity of consumers for what other consumers are doing (e.g., clothes, furniture and other fashionable goods), whereas other markets are less dominated by social preferences (e.g., many groceries). Also in some markets the quality of a product will dominate consumers preferences (e.g., books, music), whereas in other markets service level may play a more prominent role (e.g., computers, software, (used) cars). Also, in markets where consumers are highly involved, consumers are likely to take more attributes into consideration than in a market typified by low consumer involvement. For example, when buying a house, a typical consumer will rate the options on more aspects than when buying food. The amount of cognitive effort involved in the consumer decision-making process thus can be represented by the weights of the respective attributes. To model markets that differ with respect to the importance of attributes (including markets that differ concerning the social relevance of products) and consumer decision-making processes, a weighting function can be attached to each U_{inj} capturing the utilities for separate attributes:

$$U_{ij} = \frac{\sum_1^n \left(\beta_n * U_{ijn} \right)}{n}$$

With: U_{ij} = Utility of consumer i for product j, ranging from 0 to 1,
β_n = Weighting of attribute n, ranging from 0 to 1,
U_{ijn} = Utility of consumer i for product j for attribute n.

The values of β can be set according to the type of market/product and the attributes taken into consideration by consumers. This allows for a simple formalisation of decision-processes along the dimension of cognitive effort as discussed in section 5.5.2 on decision-making. Basically, a higher involvement can be formalised as more attributes taken into consideration when calculating the utility of a product. Concerning the second dimension – social versus individual orientation in the decision-making process – it has to be realised that this is not captured by the social utility of behaviour. Rather, both the exchange of norms – as captured by the social utility - and information on other nonsocial product attributes play a role in social processing. Hence the social dimension of decision-making will be captured in section 5.4 on promotion, in particular where word-of-mouth effects are being discussed.

Empirical data, preferably derived from regression weights related to various product attributes, can be used to assign values of β in simulating a particular type of market. However, individual data, e.g. originating from surveys, have to be used to estimate the heterogeneity amongst consumers as regarding the weights they attach to different attributes, as typically regression functions do not account for individual differences. In particular volatile markets will display heterogeneity of consumers concerning the attribute scores and weights. Besides heterogeneity between consumers, also within consumers these scores and weights may fluctuate over time. In addressing heterogeneity at the consumer level, the ß values can be formalized on an individual level:

$$U_{ij} = \frac{\sum_1^n \left(\beta_{in} * U_{ijn} \right)}{n}$$

With: U_{ij} = Utility of consumer i for product j, ranging from 0 to 1,
β_{in} = Weighting of attribute n for consumer i,
U_{ijn} = Utility of consumer i for product j for attribute n.

Including heterogeneity in the formulation of consumer utility allows for composing a market where consumers attach different weights to the attributes, and engage in decision processes that differ with respect to the cognitive effort (weighting of attributes) and social susceptibility (β for the social attribute).

Substitutable versus non substitutable attributes

Whereas for substitutable attributes the above mentioned formalization suffices, for nonsubstitutable attributes we can formalize consumer thresholds, indicating beyond what score (‘more is better’) or deviation (‘taste’) a product is not taken into consideration for consumption. This can be used in modelling the selection of an evoked set of products. This can be modelled according to the Elimination-by-aspects model of Tversky (1972):

For the ‘more is better’ type of preferences holds:

If $T_{\mathrm{in}} > A_{\mathrm{jn}}$ then product is not considered
With: T_{in} = Threshold of consumer *i* for attribute *n*,
A_{jn} = Score of product *j* for attribute *n*.

Setting T_{in} at high levels (close to 1) implies that only products with a very high score on attribute *n* will be accepted for that attribute.

For ‘taste’ preferences holds:

$T_{\mathrm{in}} > 1 - |A_{\mathrm{jn}} - P_{\mathrm{in}}|$
With: T_{in} = Threshold of consumer *i* for attribute *n*,
A_{jn} = Score of product *j* for attribute *n*,
P_{in} = Preference of consumer *i* for attribute *n*.

Setting T_{in} at high levels (close to 1) implies that only small deviations of product *j* on attribute *n* will lead to an acceptance of the product for that attribute.

These thresholds also allow for a more elaborated formalisation of involvement, as highly-involved agents can be formalised as having high thresholds, accepting only products that really match their preferences. Consumers having a lower involvement are more accepting towards products that do not match their preferences. In combination with the number and weighting of attributes (βs) involvement can be formalised as the number of attributes being used in the decision process, their relative weight, and the maximum deviations from the own preferences that are accepted. This allows for discriminating between agents that carefully scrutinise a market in finding a product that matches all their preferences (e.g., a house, a car) versus agents that are easy to satisfy with a product that is acceptable on one or two of the most important attributes (e.g., many groceries).

In simple experiments the thresholds can be set equally amongst simulated consumers to model markets that differ with respect to generic consumer preferences. However, it seems obvious that consumers display heterogeneity on their thresholds, and hence it is also possible to formalize individual thresholds.

Product development

For reasons of simplicity it is recommended to strive towards a simple (aggregate) formalization of an agent’s individual preference for a product. Assuming heterogeneity in preferences, producers may change the attributes of a product as to target certain groups of agents having particular preferences (segments). Also policy measures can be formalised as changing product attributes. Reverting to the car example, changes in infrastructure, improving alternative modes of transportation, closing city centres for cars, imposing minimal standards on the performance of cars (e.g., emission levels) have an impact on the experienced utility of a range of competing products, i.e. modes of transportation. In this sense policy makers can also often be understood as producers. Product development focuses exclusively on the attribute values of the product (A_{jn}), which can be designed by the producer. Obviously, the social attribute cannot be designed by the producer, although there are promotional strategies available in trying to change this attribute. This will be discussed in the section on promotion.

5.6.2 Price

In valuing a product, consumers typically make a trade-off between what they receive, and what they have to give up to acquire and use a product or service (e.g., Woodruff, 1997). Focusing on price we may use the concept of 'value-for-money' in a model (e.g., Sirohi et al., 1998; Sweeny et al., 1997), indicating that consumers may use a trade-off between the price and the utility of a product. This weighting of the price will depend on the budget of consumers, expressing that higher income consumers will be less sensitive to price than low-income consumers. This however is mainly of importance in markets where the products are more expensive. In many markets – e.g., many groceries - the available budget hardly affects the decisionmaking process. Formalized:

$$V_{ij} = U_{ij} * B_i * (1 - P_j)$$

With: V_{ij} = Value for money of product *j* for consumer *i*,
U_{ij} = Utility of consumer *i* for product *j*,
P_j = Price of product *j*, ranging from 0 (free) to 1 (expensive),
B_i = Budget of consumer *I*, ranging from 0 (no budget) to 1 (unlimited budget).

According to this formalization, the value for money V_{ij} will be closer to the utility of the product U_{ij} the lower its price P_j and the higher the consumers budget B_i.

Assuming that the value for money indicates the consumers' likeliness to buy and use a product, linking V_{ij} with P_j provides an indication of the price elasticity for product *j* given a particular market with many competing products. The more products provide about the same U_{ij}, the more sensitive consumers will be for price differences. Moreover, elasticities will differ for consumers having different budgets, indicating that the lower the budget of a consumer gets, the higher the price elasticities will be. Also for some products the costs are clearer in advance and stable than for other products. For example, where the purchase price of a car is quite clear, the operational costs are more complex. Here fixed costs such as taxes and insurances come at play, and partly controllable flexible costs such as fuel economy and non-controllable costs such as fuel prices play a role. It seems obvious that the more complex and uncontrollable costs are, the worde the perception gets of value-for-money.

Available information on the price elasticities for various product categories provide a point of application for assigning values to price and budget for different markets. Within a social simulation context it would be obvious to include heterogeneity in the agents as regards their price elasticities, thus expressing differences in available budget. This also relates to segmentation of a market, where the marketing mix as regards price and tax-policies may be used to address specific types of consumer – in this context based on income and price perception. Hence, products may be assigned a price, and agents may differ concerning the willingness to pay.

Simulation experiments may reveal how different segments of consumers – differing on attribute preferences and price sensitivity – respond to pricing and taxing strategies. Experiments may also reveal if and how the existence of different price segments affects the dynamics of the overall market.

5.6.3 Place

Concerning place a critical aspect would be the availability of the product. This can be translated in terms of effort required to buy a particular product. When two about equally good products are available, a consumer generally prefers to buy the product at a nearby shop instead of at a more distant located shop. This implies that a spatial dimension should be included in the simulation model. This would require that products may differ concerning the average effort required to obtain the product. This may be a critical variable. For example, when consumers have a slight preference for ecological food, but this is not sold in the supermarket that is visited on a regular basis, it is likely that the effort involved in visiting a shop that does sell this food causes consumers to accept non-ecologically grown foods.

Assuming heterogeneity in agents – here concerning their location with respect to the outlet of different products - agents differ concerning the evaluation of the place. Basically, a denser distribution network could be represented as agents having a more equal –and on the average lower – effort in buying the product. Products having fewer outlets would cause a larger heterogeneity amongst agents concerning the effort, also resulting in a higher average.

A simple formalization of this would be including distance as one additional attribute in the model calculating the overall utility for the consumer (U_{ij}). This attribute may have different values for different consumers, thus expressing the distance to the place in a simple manner. Weighting this distance attribute (with a β) would allow for weighting the importance for this distance attribute. This might allow for distinguishing between markets where distance is more important (e.g., for groceries people prefer a nearby shop) versus markets where distance hardly plays a role (e.g., products sold on the Internet).

Obviously many more factors can be captured under the heading of place. Besides the distance, also aspects such as quality and image of the shops vending the product, time-slots available for buying the product (e.g., compare 24/7 Internet shopping with the often limited opening hours of service providers such as banks and brokers), and the other products that can be bought at the same location (creating synergy). Factors such as these could all be captured in a series of attributes designated to formalize place-factors in the overall utility function.

In more complex spatial models it would also be possible to include a spatial density of the location of the agents, thus allowing for distinguishing between more rural and urban areas. This would translate into a more sophisticated approach of the availability factor. However, it is expected that such sophisticated models are more appropriate to address issues such as shop location planning, rather than modelling typical product markets.

5.6.4 Promotion

Finally, promotion relates to how the agents are being informed about a (new) product. Promotion can take the shape of a planned action to inform (groups of) consumers about a product, but also relates to consumer informing one another about (new) products. The latter process is addressed as word-of-mouth (WOM).

Promotion by producer

The promotion of the producer relates to informing consumers about products. The broad definition of producer and products also implies that for example governmental informational campaigns can be captured under this heading. In the context of the proposed formalization we may distinguish between addressing the relative weight of attributes, as well as the attribute scores. Hence promotion strategies could either focus on convincing consumers to attach more weight to a product attribute on which the product scores well (increasing the $ß$), or on convincing the consumers that their utility for attribute n would be higher than they currently believe (increase U_{inj}). An additional strategy might be to inform consumers about other consumers (famous role models) that already use a product, thus affecting the social attribute.

A big issue here is who to address. The classic Bass model (Bass, 1969) makes a distinction between mass media influences and word of mouth effects. However, this distinction seems to be obsolete, as the increase in media channels has changed the mass-media more into a multitude of channels targeting specific segments, and viral-marketing strategies (recently being addressed as 'buzz-marketing') utilize the processes of word-of-mouth. Hence, in a simulation context it would be practical to formalize (and test) promotional strategies that differ concerning the number of people addressed, and the degree to which a particular type of consumers (segment) is addressed. This would allow for a more sophisticated study of the efficacy of promotional strategies. This would fit with the current practice of marketers spending much effort in addressing specific segments of the markets, thus aiming their promotional activities to specific types of consumers. These segments may be based on a combination of attributes as discussed above and income groups. As such the formalization of the product and price factors allow for selecting groups of consumers to target with promotional strategies. Simulation experiments may be used to reveal what kinds of strategies are most effective to persuade particular segments of consumers. For example, some consumers are more sensitive to communications on product features (attributes) whereas others are more likely to respond on communication on price-cuts.

The response of consumers to promotional activities may range from completely accepting the information to rejecting the information and even developing a more negative perception of the product. People may spend more or less cognitive effort in elaborating the information promoted. This is captured in the *Elaboration Likelihood Model* (ELM; Petty and Cacioppo, 1986), which discerns a *central* and a *peripheral* route to attitude change. The central route pertains to the elaboration of pure arguments in a persuasive message and/or new information. Here people are motivated and capable of processing the arguments of the message, whereas peripheral processing is more likely when people's motivation to elaborate is low, and or their cognitive processing ability is limited (i.e., complex issues). The peripheral route is concerned with the elaboration of form aspects or cues of a message such as the number of arguments, the credibility and the attractiveness of the source. Social Judgment Theory (Sherif and Hovland, 1961) offers points of application for modelling central processing (see e.g., Jager and Amblard, 2004). Here consumers may move towards the position advocated by the producer (e.g., attribute value A_{jn} or weight of an attribute i for product j $ß_{ij}$) if this position is relatively close to the own position. On the contrary, if the advocated position is very remote from the own position, the consumer may display a reactance effect moving further away from the advocated position. Finally, for mediocre differences the consumer may not change at all. Besides the direction of the change, also the degree of change may be modelled. This allows for modelling the susceptibility of consumers to change their opinions after receiving a promotional message. The change of consumers opinion, e.g. their perception of a product attribute, is thus a function of the difference between own position and promoted position, weighted by their susceptibility for change.

For peripheral processing often source effects are being reported. This implies that if one likes the source, one is more willing to accept the position promoted by this source. Such effects have been modelled in the context of opinion dynamics by Jager and Amblard (2005). Also processes like these, relevant in the context of product endorsement by famous people, can be formalized quite simply in market simulations. In this application the effects of a so-called meta-actor are included, thus representing the effects of mass-medial campaigns – or public statements of e.g. politicians – on opinionshifts in large populations.

Because marketers have a limited budget for promotional activities, it is important to consider what consumers to address. Consumers are linked in social networks, and promotional campaigns directed at specific consumers (or segments) may generate word-of-mouth effects that further contribute to the efficacy of promotional campaigns. Hence it is important to include a perspective on WOM effects in promotional campaigns.

Word-of-mouth

Word-of-mouth (WOM) relates to consumers informing one another about a new product. Whereas the first dimension of consumer decision-making, as discussed in section 5.2, focuses on cognitive effort, the second dimension distinguishes between individual versus social decision-making. WOM here indicates the socially oriented decision strategies, where individual strategies exclude this type of exchange. When engaging in social processing, people exchange information and norms with people they know. This implies that information and norms are moving over a network of people. Earlier work demonstrated that the size and shape of networks is crucial with respect to the speed and degree to which information is being spread and new products are being adopted (e.g., Janssen and Jager, 2003; Delre, Janssen and Jager, 2004). Essential elements that can be captured in an agent-based model are the average number of contacts between the agents, the heterogeneity concerning the number of contacts, and the distribution between local versus distant contacts between agents. These elements allow for the formalisation of various types of networks, ranging from regular lattice, small-world networks, random networks and scale-free networks. Based on empirical data –from surveys – it is possible to get indications of both the type as use of social networks in particular markets. Obviously, for products having more social relevance and which are more complex by nature (many attributes) it is more likely that consumers use others to come to a decision.

As discussed in section 5.2, different types of information may be exchanged through social networks. First, consumers may exchange information concerning the product utilities (U_{inj}). This social informative strategy implies the communication of the attributes of a specific product, e.g., the fuel consumption of a particular brand of car. Next, consumers may also on a more generic level discuss the importance of certain attributes, such as e.g. the importance of safety of a car. Here consumers exchange information on their weighting of the attributes. Finally, a normative strategy just considers the number of neighbours consuming a particular product without considering further information. This social normative process is captured by the social attribute as defined in the section on product.

Also here central and peripheral processes can be formalized in modelling in what direction and to what degree consumers are likely to change their attributes, attribute weights following social interaction.

Recent work (Delre et al., 2006) shows that the efficacy of promotional strategies strongly depends on these WOM effects. First simulations indicated that an optimal effect can be derived when random promotion and specific group directed promotions are carefully balanced. Also the timing of additional promotional strategies is crucial in boosting a diffusion process. Whereas there results are still preliminary, it offers a modelling perspective on exploring the efficacy of promotional strategies in relation to which consumers to address, when to address them, and how to address them (type of information).

5.7 Validation and prediction

Obviously, the validation of models is a big issue. We want to have trust in our model behaviour so that we can use them to test policy strategies. However, if we are dealing with complex behavioural domains, we have to realize that the developmental path – as captured in empirical data – might have developed in a complete different direction. For example, the major propulsion of cars could have been steam based (external combustion) rather than gas based (internal combustion, having resulted in cars being about twice as efficient with fuel. Whereas steam and gasoline engines were both facing technological difficulties a century ago, the internal combustion engine gained market share and locked-in the market as the dominant propulsion system. Realising that the current state of affairs could have been completely different if, e.g., early consumers had had less (unrealistic) fears for exploding steam engines, this exemplifies that calibration of a computer simulation model against an empirical data set describing a single event at the macro-level is a risky business, unless the macro-effect can be observed in many conditions and represents a kind of stylized fact. The latter would imply that data are available on a larger set of comparable macro events, and that all these data show a (qualitatively) comparable trajectory of developments. One example would be the distribution of market shares of fast-moving consumer goods, which typically show larger market shares for first entrants (e.g., Robinson and Fornell, 1985). However, counterexamples are available. Another example is the take-off of new products, which typically differ for product categories and countries (e.g., Tellis et al., 2003).

Replicating (or mimicking) the macro results with a model is not contributing to a better understanding of the mechanisms behind the empirical phenomenon, as the same mechanism might have yielded different outcomes. Using simulation models to develop point-predictions (including uncertainty margins) of future system states, no matter how attractive these may seem, seems to be an unrealistic option for policy making in complex systems. Whereas macro-level data are useful to identify complex behaviour of e.g. a market, as expressed by volatility of market shares, they provide no means for validating the processes that lead towards these outcomes. Hence, empirical validation should here focus at the level of processes rather than outcomes. This implies that data should be collected concerning the actual behaviour and decision making process of individual people, and the driving factors behind these decisions. Data on behaviour can be obtained in many ways, e.g. sales data can be obtained on an individual consumer level by using loyalty-card data and panel data. Concerning the decision-making process it is necessary to approach people with questions directed at the related variables (e.g., dominating needs, cognitive effort, social orientation). As heterogeneity of the population is an essential attribute in complex systems, such a validation should focus on the distributions of the critical factors in the population rather than the averages.

Whereas a better understanding of the dynamics in such complex systems does not help in developing (reliable) point estimations of the effects of policy, the contribution resides in a better understanding of how to dynamically manage such systems. Simulation experiments may reveal the different developments that may emerge from the same underlying behavioural determinants and processes. Policy experiments could focus on managing such simulated systems such as to stimulate developments in a pre-selected direction. For complex behaving systems, this would imply that – given a almost identical starting situation – in one simulation run less policy-measures may be necessary to achieve a desired outcome than in other runs. This signifies the importance of dynamical policy making. Rather that developing a worked out plan for policy measures, the policy maker should focus on a goal to reach, have a toolbox of policy measures available, and respond quickly to developments in the system that are unwanted. Simulation studies may contribute to exploring how different combinations of policy-measures affect the system as a whole in different dynamical conditions. This would contribute to the development of a more dynamical policy-making style, combining experience with managing complex systems with a close monitoring and policy responses to system behaviour. The main challenge here would be the development of valid simulation models that offer a tool for policy makers to experiment with policy measures in a dynamical context. This would also facilitate discussions on what policy strategies to implement given certain (unforeseen) developments in systems. For real-world systems this implies that simulation tools contribute to exploring possible developments in the system, and viable policy strategies given particular developments. Metaphorically speaking, policy makers should thus behave more like sailors. Sailors master the complexities of navigating their ships through various weather and sea conditions, which are often unpredictable by nature. Yet, their profound understanding of these dynamics allows them to respond efficiently in keeping their course and arriving at their port of destination. In contrast, many (social) scientists and policy makers endeavour to predict the destination of a voyage given a ship's initial position. Sailors would burst out laughing understanding these attempts. Rather, learning from how sailors deal with a dynamical system, policy makers should focus on understanding the principles of the underlying dynamics, and learn to respond dynamically on changes in the system, whilst keeping an eye on the goals they have.

5.8 Conclusions

In this paper it has been argued that the application of social simulation models in general, and in the specific case of studying man - environment relations, should be based on agent rules that are based on empirical and theoretical connotations on generic human behaviour. Next, following the structure as provided by the four P's – product, pricing, placement and promotion, a structure has been provided to formalise behaviour in rules for agents. The rationale behind this approach is that for stakeholders interested in changing consumer behaviour, such as product developers, marketers and policy makers, as well as scholars studying market dynamics, it is important to include a perspective on strategies to influence behaviour and aggregated markets. Obviously, including all formalizations as presented in this paper in a single simulation model would result in a very complex model, being not suitable to systematically explore the effects of various factors. To maintain transparency in simulation experiments, we therefore consider the full set of formalizations as a conceptual framework guiding the process of developing an experimental design. Hence it is advocated here to start building relatively simple models, and consecutively extend the model along the lines as expressed by the conceptual model. Following such a procedure it is possible to develop more complex models that are suitable for testing policy measures in more complex settings.

In conducting such a series of experiments it is recommended to use micro-level data to initialize the parameters of the model, and to have macro-level data to validate the simulation results against. Acknowledging the fact that in complex systems the empirical data reflect only one possible outcome of many, a valid simulation model should in principle be capable of reproducing the empirical outcomes – and underlying processes – as one of the possible outcomes. Hence in a simulation of the development of private car-use, simulation outcomes might result in various scenarios, such as dominance of steam engines, but should at least in a number of cases demonstrate how the current state of affairs emerged. Such an approach allows for making causal inferences concerning how the macro behaviour – i.e. dynamics – is related to specific behavioural determinants and processes in a particular system.

Another issue deserving attention is the competition between stakeholders such as producers and policy makers. Obviously, marketing strategies implemented by one producer will affect the market shares of competing producers. As such, in the beginning of the car-market producers of cars adverted the risks of using other fuel systems, and thus tried to rise fear in the public, e.g. for explosions of steam engines, or breaking limbs when starting a gas-fuelled car. Hence a typical product market involves a number of producers competing for market shares, implementing market strategies aimed at maintaining and increasing market share. On top of that stakeholders such as policymakers are setting the stage concerning tax-regimes, minimal product requirements and the like. Such interdependencies between different stakeholders involves strategic decision-making by producers and policy makers alike, e.g., in developing products for specific consumer segments (preferences and attribute weights), addressing specific groups of consumers – e.g. those using the product of a competitor - with promotional strategies, and responding to new developments on markets. In understanding complexities in man environment systems from a consumer behaviour perspective it would be necessary to include a perspective on interdependencies between producers and policy makers. In formalizing such a perspective in a simulation model we emphasize the importance of using a framework on consumer behaviour as sketched in this paper. This would allow for the formalization of strategies for behavioural change as signified by the four P's, and allow for experimenting with competition between different stakeholders along the dimensions of product characteristics, pricing, placement and promotion.

This paper provides a perspective on what behaviour to formalise in agent-based modelling to explore man-environment dynamics, provided a perspective on how to formalise this and drew a perspective on how to perform policy experiments. However, developing simulation models that are valid and applicable for testing policy measures remains a big challenge, and will require significant investments from scientists from various disciplines and policy makes as well. The author hopes that this paper contributes to the development of such simulation tools, and hopes that this will materialise in the use of simulation models in the policy development process, thus contributing to the development of effective and acceptable policy measures aimed at managing man-environment relations in an effective way.

References

Akerlof, G.A. (1984). *An economic theorist's book of tales*. Cambridge (UK): Cambridge University Press

Bass, F. M. (1969). A new product growth for model consumer durables. *Management Science* 15: 215-227

Bettman, J.R. (1979). *An Information Processing Theory of Consumer Choice*. Reading (MA): Addison-Wesley

Bousquet, F., C. Cambier, C. Mullon, P. Morand and J. Quensiere (1994). Simulating fishermen's society. In: N. Gilbert and J. Doran (Eds.) (1994). *Simulating societies: The computer simulation of social phenomena*. London: UCL Press

Bowles, S. (1998). Endogenous Preferences: The Cultural Consequences of Markets and other Economic Institutions. *Journal of Economic Literature* 36: 75-111

Bowles, S. and H. Gintis (2000). Walrasian economics in retrospective. *The Quarterly Journal of Economics* 115: 1411-1439

Camazine, S. J.-L. Deneubourg, N.R. Franks, J. Sneyd, G. Theraulaz and E. Bonabeau (2001). *Self-Organization in Biological Systems*. Princeton University Press

Chao, A, and J.B. Schor (1998). Empirical tests of status consumption: Evidence from women's cosmetics. *Journal of Economic Psychology* 19 (1): 107-131

Cialdini, R. B. and N.J. Goldstein (2004). Social Influence: Compliance and Conformity. *Annual Review of Psychology* 55: 591-621

Deffuant, G., W. Jager and S. Moss (2006). Dialogues concerning a (possibly) new science. *Journal of Artificial Societies and Social Simulation*. 9 (1). Avalaible at http://jasss.soc.surrey.ac.uk/9/1/1.html

Delre, S.A., W. Jager and M.A. Janssen (in press). *Diffusion dynamics in small-world networks with heterogeneous consumers*. Computational & Mathematical Organization Theory

Delre, S.A., W. Jager and M.A. Janssen (2004). *Percolation and innovation diffusion models compared: do network structures and social preferences matter?* Paper presented at the M2M2 workshop of the ESSA2 conference, Valladolid, Spain, September 2004

Delre, S.A., W. Jager, M.A. Janssen and T. Bijmolt (2006). *Promotion strategies for innovation diffusions in networks of consumers*. Paper presented at the workshop on Agent-based models of market dynamics and consumer behaviour, 17-18 January, 2006, University of Surrey, Guildford, (UK)

Deutsch, M. and H.B. Gerard (1955). A study of normative and informative social influences upon individual judgment. *Journal of Abnormal and Social Psychology* 51: 629-636

Duesenberry, J.S. (1949). *Income, Saving, and the Theory of Consumer Behaviour*. Harvard University Press, Cambridge

Faucheux, S. and C. Moscovici (1972). Social influence, conformity, bias and the study of active minorities. In: L. Berkowitz (Ed.). *Advances in experimental social psychology* 6: 149-202

Festinger, L. (1954). A theory of social comparison processes. *Human Relations* 7: 117-140

Galtung, J. (1990). International development in human perspective. In: J. Burton, (Ed.). *Conflict: human needs theory*. New York: St Martin's Press

Güth, W. and M.E. Yaari (1992). Explaining Reciprocal Behaviour in Simple Strategic Games: An Evolutionary Approach. In: U. Witt (Ed.). *Explaining Process and Change*. Ann Arbor: University of Michigan Press: 23-34

Hare, M. and P. Deadman (2004). Further towards a taxonomy of agent-based simulation models in environmental management. *Mathematics and Computers in Simulation* 64 (1): 25-40

Hasher, L. and R.T. Zacks (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology*: General, 108: 356-388

Hayakawa, H. (2000). Bounded rationality, social and cultural norms, and interdependence via reference groups. *Journal of Economic Behaviour and Organisation* 43: 1-34

Jackson, T., W. Jager and S. Stagl (2004). Beyond Insatiability: Needs theory, consumption and sustainability. In: Reisch, L. & I. Røpke (Eds.). *Consumption: Perspectives from ecological economics*. Cheltenham: Edward Elgar

Jager, W. and F. Amblard (2004). Uniformity, Bipolarization and Pluriformity Captured as Generic Stylized Behaviour with an Agent-Based Simulation Model of Attitude Change. *Computational & Mathematical Organization Theory* 10 (4): 295-303 (9)

Jager, W., and F. Amblard (2005). *Multiple attitude dynamics in large populations*. Paper presented at the Agent 2005 Conference on: Generative Social Processes, Models, and Mechanisms, Argonne National Laboratory, The University of Chicago, October 13-15, 2005

Janssen, M.A. and W. Jager (2003). Self Organisation of Market Dynamics: Consumer Psychology and Social Networks: *Artificial Life* 9 (4): 343-356 (fall 2003)

Kahneman, D. (1973). *Attention and Effort*. Englewood Cliffs, NJ: Prentice-Hall

Kahneman, D., J.L. Knetch and R.H. Thaler (1986). Fairness and the assumptions of economics. *Journal of Business* (59, 4-part 2): 285-300

Kahneman, D. and A. Tversky (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica* 263-291

Kruskal, J.B. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* 29: 1-27

Lancaster, K. (1966). A New Approach to Consumer Theory. *Journal of Political Economy*: 132-57

Maslow, A.H. (1954). *Motivation and Personality*. New York, USA: Harper and Row

Max-Neef, M. (1991). *Human-Scale Development - conception, application and further reflection*. London: Apex Press

Max-Neef, M. (1992). Development and human needs. In Ekins, P. and M. Max-Neef (Eds.). *Real-life economics: Understanding wealth creation*. Routledge, London

McCarthy, E. J. (1960).*Basic Marketing: A Managerial Approach*. Homewood (Ill.): Irwin

Parasuraman, A., V.A. Zeithaml and L.L. Berry (1995). A conceptual model of service quality and its implications for further research. Journal of Marketing, 49 (4), pp. 41-50

Parasuraman, A., V.A. Zeithaml and L.L. Berry (1998). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing* 64 (1): 12-40

Park, C.W. and P.V. Lessig (1977). Students and Housewives: Differences in Susceptibility to Reference Group Influence. *Journal of Consumer Research* 4: 102-110

Payne, J.W., JR. Bettman and E.J. Johnson (1992). Behavioural Decision Research: A Constructive Processing Perspective. *Annual Review of Psychology* 43: 87-131

Petty, R.E. and J.T. Cacioppo (1986). *Communication and persuasion: central and peripheral routes to attitude change*. New York: Springer Verlag

Pollak, R.A. (1978). Endogenous Tastes in Demand and Welfare Analysis. *American Economic Review* 68 (2): 374-379

Reynolds, C.W. (1987). Flocks, Herds, and Schools: A Distributed Behavioural Model. *Computer Graphics* 21 (4): 25-34

Robinson, W.T. and C. Fornell (1985). Sources of market pioneer advantages in consumer goods industries. *Journal of Marketing Research* 22: 305-317

Shepard, R.N. (1962a). The analysis of proximities: Multidimensional scaling with an unknown distance function. I. *Psychometrika* 27: 125-140

Shepard, R.N. (1962b). The analysis of proximities: Multidimensional scaling with an unknown distance function. II. *Psychometrika* 27: 219-246

Sherif, M. and C.I. Hovland (1961). *Social Judgment*.Yale University Press, New Haven (CT)

Simon, H.A. (1976). *Administrative behaviour: a study of decision making processes in administrative organizations*. New York: Harper

Slovic, P. (1995). The Construction of Preference. *American Psychologist* 50: 364-371

Tellis, G.J., S. Stremersch and E. Yin (2003). The international take-off of new products: the role of economics, culture, and country innovativeness. *Marketing Science* 22 (2): 188-208

Thurstone, L.L. (1931). Measurement of social attitudes. *Journal of Abnormal and Social Psychology* 26: 249-269

Tversky, A., (1969). Intransitivity of preferences. *Psychological Review* 76: 31-48

Tversky, A., (1972). Elimination by aspects: a theory of choice. *Psychological Review* 79: 281-299

Tversky, A., and D. Kahneman (1974). Judgement under uncertainty: heuristics and biases. *Science* 185: 1124-1134

Veblen, T., (1899). *The Theory of the Leisure Class*. New York: Macmillan

Witt, U. (1991). Economics, Sociobiology, and Behavioural Psychology on Preferences. *Journal of Economic Psychology* 12: 557-573

6 Managing the transition to fuel cell vehicles – policy implications of an agent-based approach

M. Schwoon

International Max Planck Research School on Earth System Modelling

Research Unit Sustainability and Global Change, Centre for Marine and Atmospheric Sciences

Hamburg University

Hamburg, Germany.

Contact: malte.schwoon@zmaw.de

ABSTRACT

Supply security and environmental impacts are two major concerns related to current gasoline and diesel-based individual transport. A promising way out is the transition to fuel cell vehicles (FCVs) running on hydrogen. To date, scenario studies of hydrogen infrastructure build-up and sales of FCVs often ignore that there is a ‘chicken-and-egg-problem’: car producers are not willing to set up large scale FCV production lines that are necessary for cost reductions as long as there are no hydrogen filling stations. On the other hand, fuel suppliers will not put up a hydrogen refuelling infrastructure unless there is sufficient demand from FCVs on the road. To overcome this problem, public infrastructure programs and (tax) incentives for consumers to buy FCVs are indicated. There are actually four types of agents involved in the transition problem: car producers, consumers, fuel suppliers and the government. In this paper, I argue that neoclassic technology adoption models with homogeneous, fully rational decision makers are not suitable to describe potential transition paths in the specific situation of infrastructure-dependent FCVs. Thus, I discuss an agent-based simulation model that puts together an existing producer-competition model with a consumer model of adoption decisions in a modular way. It is applied to investigate the impacts of tax and infrastructure policies that are pronounced enough to overcome the chicken-and-egg-problem. Results suggest that consumers and individual producers are asymmetrically affected by taxes and public infrastructure investments, so that resistance towards the policies can be anticipated. Moreover, there is evidence that large car producers might benefit from co-operation with fuel suppliers to generate a faster build-up of hydrogen infrastructure.

6.1 Introduction

The current crude oil-based individual transport system is not sustainable for economic and environmental reasons. Oil is a non-renewable resource and even though in the past discoveries of new oil fields and especially improved exhaustion methods have repeatedly extended the statistical reach of oil, there is extensive evidence that we get close to the oil peak (Bentley, 2002). Once the oil peak is reached, resource depletion will be accelerated; and given current demand, prices are likely to increase substantially. An additional (socio-) economic problem is the imbalanced distribution of the world's oil reserves, which are concentrated in the politically instable region of the Middle East. From an environmental point of view, internal combustion engine vehicles (ICEVs) running on gasoline or diesel are major contributors to greenhouse gas (GHG) emissions. They account for more than 20% of total GHG emissions in the US (EPA, 2006) and for about 16% in the EU (EEA, 2006). ICEVs also cause significant local air pollution. Advancements of end-of-pipe technologies (3-way-catalytic-converter, diesel particulate filter) have been substantial, but in the past, technological progress has often been compensated at least partly by an increase in the number of cars and/or car use (Friedrich and Bickel, 2001).

Already the oil crises of the 1970s, together with ever increasing emission standards later on, initiated research programs of car manufacturers towards alternative fuel/vehicle concepts. Since the 1990s research focuses on fuel cell vehicles (FCVs) running on hydrogen, with R&D investments in the range of billions of dollars (van den Hoed, 2005).[16] Today, every major car manufacturer has a small fleet of FCVs being tested in daily life situations. Also, some fleet tests of buses and taxis have started and technological problems with respect to onboard hydrogen storage for a sufficient range, reliability of the fuel cell or cold start are basically solved. Together with a positive public perception of the technology as being ‘compact, silent, efficient, and emission-free’ (Farrell et al., 2003: p. 1357), FCVs are now a technological option rather than a vision.

A large scale introduction of FCVs would directly solve problems related to local emissions, because FCVs only emit water vapor. Using hydrogen as the main energy carrier in individual transport implies the option to diversify energy sources and, thus, lower the dependency on crude oil. Early hydrogen demand in the introductory phase would probably be met using natural gas as the hydrogen source. This would shift GHG emissions (mainly CO_2) from the vehicle to the fuel production side, with only minor overall emission reductions (at relatively high costs per ton of CO_2 abated; EC-JRC, 2006). But later on, substantial GHG emission reductions could be achieved, if hydrogen generation from fossil fuels is combined with carbon capture and sequestration - once these technologies have proven reliable on a large scale and are accepted by the public - or if hydrogen is directly produced from renewable energy sources like wind power or biomass.[17]

---

[16] Before, battery/electric vehicles have received most attention, but technological problems with respect to capacity, recharge speed and weight of the battery seemed unsolvable.

[17] With carbon capture and sequestration even coal might become a reasonable source for hydrogen production, which otherwise would actually imply drastic increases of CO_2 emissions compared to the current system (EC-JRC, 2006). The advantage of coal would not only the much higher amount of resources compared to other fossil fuels, but also its more even distribution over world regions.

In this paper, I assume that reduction of local emissions, diversification of energy supply and the long-term potential to reduce GHGs are sufficient to let governments consider policies to encourage a significant penetration of FCVs in the car market.[18] These policies must be pronounced enough to overcome the problem of the missing hydrogen refuelling infrastructure, often referred to as the chicken-and-egg problem of hydrogen and fuel cells. Fuel cells are extremely expensive and significant cost reductions are only feasible, if they are produced on a large scale. But car manufacturers are not willing to make substantial investments in product lines as long as missing refuelling opportunities prevent consumers from buying. On the other hand, oil companies, as the major filling station operators, will not set up a hydrogen production/distribution network and hydrogen outlets at the stations without demand generated from FCVs on the road.

There exist a wide range of studies that develop scenarios of the introduction of FCVs for different regions; see e.g. Thomas et al. (1998), Moore and Raman (1998), Ogden (1999, 2002), Stromberger (2003), Mercuri et al. (2002), Sørensen et al. (2004), Oi and Wada (2004), Hart (2005). In the majority of studies, costs of building up a hydrogen infrastructure are estimated given certain scenarios of the development of the number of FCVs, starting in certain (local, commercial) niche markets before entering the large market of private consumers. The implied assumption is that the government must set up the necessary infrastructure and then there will be a smooth and successful diffusion in the market. But a substantial governmental commitment in setting up refuelling infrastructure would be unprecedented and unlikely given budget constraints of public authorities. Policies on emission reductions of the transport sector in the past were either direct regulation, taxes on polluting (old) technology or tax exemptions for the new technology. Yet, there must be a direct incentive for consumers to buy the (more costly) FCVs, as the pure existence of infrastructure is not sufficient. The willingness to pay for ‘environmental friendliness’ of a car is way below the expected additional costs for the fuel cell (Steinberger-Wilckens, 2003). Moreover, within car buying decisions, the environmental impact is just one feature in addition to other characteristic like size, acceleration and also psychological motivations like status. Thus, joint tax and infrastructure policies are indicated to promote the introduction of FCVs.

The diffusion problem can be considered extremely complex. There are four (types of) agents involved: car producers, consumers, fuel suppliers and the government. There are dynamic interactions between the different types of agents, but producers also affect each other as do consumers. However, the policy problem of introducing the new technologies remains of ‘intermediate range’ as defined by Verspagen (this report) and can therefore be approached with evolutionary methods. In this paper an agent-based simulation model based on Schwoon (2006a, 2006b) is discussed that puts together an existing producer competition model (Kwasnicki, 1996) with a product adoption model (Janssen and Jager, 2002) in a modular system. The links between the modules are established in a ‘natural’ way: consumers decide between the cars offered by the producers and the producers try to meet consumers’ preferences and estimate the demand. Fuel suppliers are represented by a (non-linear) increase of the number of filling stations with a hydrogen outlet as a reaction to increases in the number of FCVs on the road. The advantage of modularity and rather simple connections between the modules are that modelling improvements, as, e.g., a more

---

[18] The model presented later is calibrated to the German (compact) car market, but is easily transferable, e.g., to Japan, or the state of California, which have previously introduced strong emission control policies, if necessary without (inter-)national coordination.

complex decision model of fuel suppliers can be implemented ‘locally’, without changing the overall structure of the model.[19]

The purpose of the model is to investigate the impacts of different policies on the different agents and their success in promoting the new technology. Thus, the government is only represented as the exogenous driver of changing behaviour. This allows for normative interpretations.[20] The simulation results presented in this paper suggest that tax/infrastructure policies have strong (asymmetric) impacts on certain groups of agents and lead to concentration in the market. These impacts are so far neglected by scenario studies on infrastructure and fuel cell technology costs.

In this paper I will first argue in paragraph 2 that for the specific problem agent-based computational modelling seems to provide additional insights compared to traditional technology adoption models with externalities. In paragraph 3, I will discuss the main features of the model. A selection of results is presented in paragraph 4; and in the concluding paragraph 5, some shortcomings of the specific model and agent-based simulations in general are discussed.

6.2 Traditional analysis of adoption externalities

Katz and Shapiro (1985) state that ‘positive consumption externalities arise for a durable good when quality and availability of postpurchase service for the good depend on the experience and size of the service network, which may in turn vary with the number of units of the good that have been sold’ (p.424). FCVs are exactly such durable goods. The more FCVs have been sold (i.e., the higher the so called userbase), the more hydrogen filling stations and maintenance facilities will be set up making a FCV more valuable for later adopters.

Katz and Shapiro (1985, 1986) and Farrell and Saloner (1985, 1986) introduce a general theoretical framework to analyze welfare and strategy implications in the presence of adoption externalities. It is applied to show that usually two equilibria exist: an adoption and a non-adoption one. The non-adoption equilibrium can also be interpreted as a lock-in situation, with persistence of the old technology. The adoption equilibrium can only be reached if consumers expect a high enough future userbase, so that they then benefit from being part of that userbase, and it is assumed that firms have some influence on these expectations. But this framework cannot be applied to the case of FCVs, in which consumers, who make buying decisions, consider the compatibility with the current refuelling system and not with the future one. Thus, non-adoption would be the only reasonable equilibrium and the result is basically another description of the earlier mentioned chicken-and-egg problem.

The descriptive character of the standard neoclassical framework developed by Katz and Shapiro (1985, 1986) and Farrell and Saloner (1985, 1986) is the main reason why it appears to be inappropriate for FCV diffusion analysis. It lacks normative implications in order to compare different

---

[19] The same holds for changes within the modules. For example, implementing cost reductions due to learning by doing only affect producers decisions. Changes in the consumer module, e.g., a different representation of the buying decision, would require adjustments also in the demand estimation procedure of the producers, but does not influence the general structure.

[20] A fully specified approach would also include dynamic feedbacks, since the government would evaluate policy success and, if necessary, make adjustments. In other words, a government module would be required.

diffusion policies, given that the diffusion of the new technology is preferred, e.g., to reduce environmental impacts. Moreover, it offers a static description of the existence and characteristics of equilibria that does not allow, e.g., predicting the number of users during the transition from the non-adoption to the adoption equilibrium. Yet, a description of the transition process is crucial, as for car technologies it might take decades between introduction and full penetration. Another drawback of neoclassical models holds that consumers, producers and also products are assumed to be homogenous. In reality, consumers are heterogeneous not only with respect to preferences for a wide range of car characteristics, but also with respect to refuelling needs, i.e., their need for compatibility varies.[21] Car manufacturers are different with respect to their size, profitability and research success. And their products might be similar in a broad sense of functionality, but are certainly not perceived as homogeneous. All these shortcomings of the traditional framework seem to be substantial in the context of a new car technology. Thus, taking them together justifies a departure from an analytically tractable neoclassical framework in favour of an agent-based computational model that allows for heterogeneity of agents and normative insights into transition dynamics.[22]

6.3 An agent-based FCV diffusion model

The downsides of the traditional framework imply the advantages of the agent-based modelling approach. Agents can be modeled to represent heterogeneity in characteristics and behaviour. The development of ‘macro’ variables (e.g. the penetration rate of the new technology) emerges from dynamic interactions and decision-making of agents on the ‘micro’ level. It is assumed that agents are myopic (ruling out strategic long term behaviour) due to the complexity of the decision problems. In Schwoon (2006a) there are four different types of agents: car producers, consumers, filling station owners and the government. Car producers follow heuristic decision rules, because they are uncertain about hydrogen infrastructure development and own research success. In addition, technology choices and price decisions of the competitors are unknown. These uncertainties together cannot be described by probability distributions, ruling out traditional intertemporal expected profit maximization methods. As stated above, consumers behave myopically in buying the car that maximizes their current utility, because disutility of wrong usebase/infrastructure predictions would be immense.

---

[21] Already Katz and Shapiro (1985) identify the missing representation of consumer heterogeneity as a limitation of their approach.

[22] Note that this conclusion is not general but refers to the specific FCV diffusion problem. Within the set of assumptions, the traditional framework seems to be appropriate (and has widely applied to) diffusion processes in information technologies (examples are Brynjolfsson and Kemerer; 1992; Gandal 1994; Economides and Himmelberg, 1995). For many these technologies and particularly for software products, switches from the non-adoption to the adoption equilibrium are fast. These products often have a non-zero direct use value without compatibility (that remains if one bets on the wrong horse). This direct utility is usually higher for the new technology. Moreover, the loss associated with choosing the wrong technology with respect to the user base is small compared to the car case.

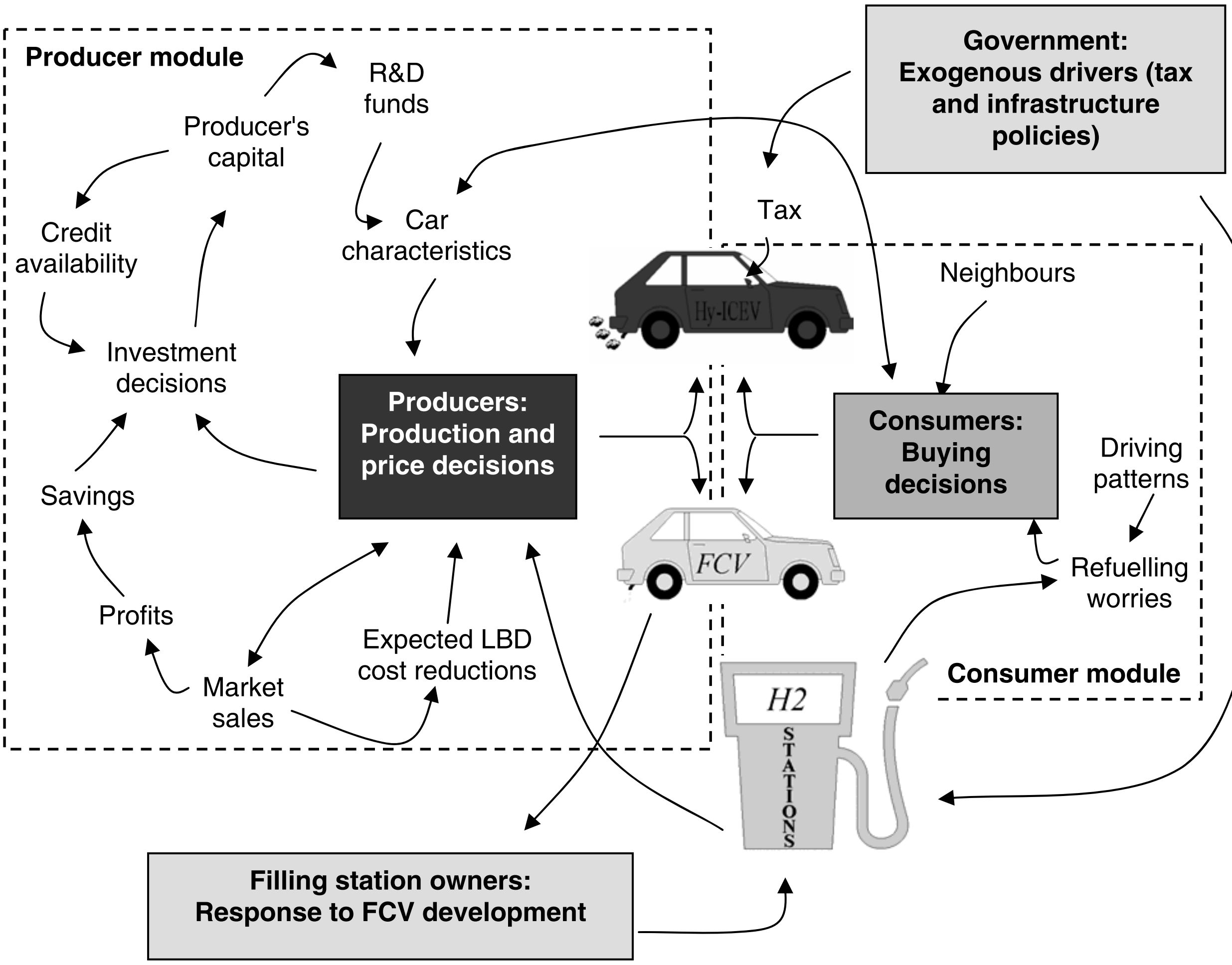

Figure 6.1 Scheme of the model, highlighting the producer and consumer modules

Figure 6.1 shows a scheme of the model. For a full description of the model and its calibration see Schwoon (2006a; 2006b). An arrow from variable A to variable B should be interpreted as ‘A is a major determinant of B’. In reality, the government can actively influence technological choice by announcing and implementing policies. Their success is then evaluated and, if necessary, adjustments are made. However, to compare the implications of different (long term) policies, dynamic interactions of the government are not included. Policies simply represent exogenous drivers that follow certain scenarios. The policies investigated include differently scheduled taxes on newly bought conventional cars and investments that increase the share of filling stations with a hydrogen outlet. Policies are introduced around 2010. Therefore, conventional technology refers to already advanced hybrid electric internal combustion engine vehicles (Hy-ICEVs), which perform much better than average current ICEVs with respect to fuel efficiency and local emissions. The model is restricted to the compact car segment, which is a likely segment for the introduction of FCVs on a large scale, because almost all the cars applied in today's small fleet demonstration projects are based on conventional cars of that segment. Within the segment, cars are similar but slightly differentiated between producers according to size, design, chassis or special equipment. If a producer switches from the production of a Hy-ICEV to a FCV all these features remain the same, including characteristics like acceleration, noise etc. Thus, conventional cars and FCVs of the same producer differ only in production costs and fuel availability.

Filling station operators are assumed to react collectively towards changes in the share of FCVs within newly registered cars. If the share reaches a new peak, hydrogen outlets are added, but existing ones are not deconstructed. So the behaviour of filling station operators is basically simplified to a positive feedback loop between hydrogen filling stations and the number of FCVs, representing the main source of adoption externalities. Thus, filling station owners are not represented as optimizing ('software'-) agents like consumers and car producers, whose decision making is sketched in the next two subsections.

6.3.1 The consumer module

The decision-making of the consumer is based on the consumat model introduced by Jager (2000) and applied in the context of environmentally friendly products in Janssen and Jager (2002). The original consumat approach allows for four cognitive strategies of the consumers (repetition, deliberation, imitation, and social comparison). A deliberating consumer compares all the cars from the different producers and takes the one that maximizes his utility function. With respect to cars, this is likely to be the predominant strategy. Repetition and imitation are not included, as they seem to be more appropriate for products that are frequently bought and less expensive. Social comparison refers to a strategy, where consumers evaluate only the utility of the car most of their neighbours drive and compare it with the utility they would get from buying the brand again that they are currently driving. This strategy arises from uncertainty about product characteristics and reduces the decision space to the directly perceivable environment. The impact of consumers doing social comparison is shown in section 6.4.3.

A consumer evaluates a car according to the utility function[23]:

$$U_{k,t}^{tot}(c_{i,t}) = \frac{\beta_k U_{k,t}(c_{i,t}) + (1-\beta_k) SN_{k,t}(FCV_i)}{\left(p(c_{i,t})(1+tax_t(1-FCV_i))\right)^{|\varepsilon|}} RFE_{k,t}(FCV_i) . \tag{1}$$

$U_{k,t}^{tot}(c_{i,t})$ is the total utility consumer k receives from buying car $c_{i,t}$ (from producer i) at time t. FCV_i is a binary variable that is 1 if the car under consideration is a FCV and 0 for a Hy-ICEV. The first component is direct utility, which measures the difference between car characteristics ($z_{i,j,t}$) and individual preferences ($pref_{k,j,t}$) for the characteristics according to

$$U_{k,t}(c_{i,t}) = 1 - \frac{1}{n_j} \sum_{j=1}^{n_j} \left| z_{i,j,t} - pref_{k,j,t} \right| , \tag{2}$$

where n_j is the number of characteristics being evaluated. Characteristics and preferences are randomly initialized to vary between 0 and 1. Correspondingly, direct utility also varies between 0 and 1. Producers can change characteristics via R&D. Therefore, the consumer's direct utility can be

---

[23] The equations included here should provide a general notion of how the components of the model shown in Figure 6.1 translate into a computational framework. For the full system of equations see Schwoon (2006a).

1 at the maximum if all characteristics exactly meet his preferences and is limited to zero in the opposite case.

The decision to buy a new technology like a FCV might be affected by the neighbours. The term ‘neighbours’ refers to the social environment (members of the family, colleagues, friends, etc.) that is relevant for the buying decision. If the car under consideration is a FCV, then the satisfaction of social needs $SN_{k,t}(1)$, equals the share of neighbours, who already have a FCV (otherwise, $SN_{k,t}(0)$ is computed as the share of neighbours, who still drive conventional cars). But this pressure towards conformity varies from consumer to consumer according to the weight β_k (which is positive and does not exceed 1). People with a β_k close to 1 decide rather independent of their neighbours (and are, therefore, likely to be early adopters of the new technology), while others follow the technology choice of their environment.

Utility is normalized by (after-tax) price and the magnitude of the responsiveness of utility towards price changes is represented by the elasticity ε. Note that taxes are zero in the case of a FCV. The unusual inclusion of price in the utility function is basically a short cut for not having an explicit budget constraint. Moreover, it allows for a direct trade off between price and fuel availability that has been measured in empirical studies. The impact of fuel availability on utility is represented by a ‘refuelling effect’ $RFE_{k,t}(FCV_i)$, which is defined as

$$RFE_{k,t}(0) = 1 \qquad \text{(for Hy-ICEV)} \qquad (3a)$$

and $$RFE_{k,t}(1) = 1 - DP_k \cdot \exp(-\gamma s_{H2,t}) \qquad \text{(for FCVs).} \qquad (3b)$$

Refuelling matters only for FCVs and depends on individual driving patterns of the consumer (DP_k) and the share of filling stations that have a hydrogen outlet ($s_{H2,t}$). The driving pattern represents the individual refuelling needs. For example, the decision might differ for buying a second car. Then, long distant trips with a high ‘refuelling uncertainty’ might be done with the conventional first car and the FCV, as the second, can be regularly fueled at a familiar filling station, e.g., on a weekly shopping trip.[24] Driving pattern and fuel availability combined are constructed to be in the range from 0 to 1, so that overall utility of a FCV can actually be 0. Initialization of driving patterns and the parameters ε and γ are jointly calibrated in order to get a price/fuel availability trade-off that is in line with the estimates of Bunch et al. (1993) and Greene (1998).

6.3.2 Heterogeneity of consumers

The description of the consumer module showed three sources of heterogeneity that determine different product choices of consumers. Firstly, consumers have different preferences for certain car characteristics (via $pref_k$). This is particularly decisive if cars with the same technology are compared. Secondly, consumers are differently influenced by their neighbours on the buying decision (via β_k), and thirdly they differ in their driving pattern and, therefore, in their refuelling needs (via DP_k). The

---

[24] The term driving pattern is used to indicate that it is the actual driving behaviour of a consumer that determines his individual refuelling needs. However, there might also be a psychological effect that perceived refuelling needs are higher than actual refuelling needs, but this is not addressed separately.

latter two sources of heterogeneity determine which consumers are most likely to be early adopters of FCVs, namely those with low refuelling needs, who decide independently of their neighbours.

6.3.3 The producer module

The supply side of the model is based on Kwasnicki's (1996) behavioural model of producers competing in a market of slightly differentiated products. Producers are price setters with limited market power depending on their market share. In each period, the individual producer sets the price that maximizes the following objective function:

$$\max Obj_{i,t} = (1-W_{i,t})\frac{INC^{e}_{i,t}}{\sum_{i=1}^{n_i} INC_{i,t-1}} + W_{i,t}\frac{q^{e}(c_{i,t})}{\sum_{i=1}^{n_i} q(c_{i,t-1})}, \tag{4}$$

$$\text{with} \qquad W_{i,t} = \exp\left(-\eta\frac{q^{e}(c_{i,FCV,t})}{\sum_{i=1}^{n_i} q(c_{i,FCV,t-1})}\right). \tag{5}$$

The objective is a weighted average of expected income $INC^{e}_{i,t}$ relative to total income of all producers in the previous period ('expected income share') and expected number of cars sold $q^{e}(c_{i,t})$ relative to the total number of cars sold in the car market in the previous period ('expected market share'). The parameter η calibrates the weight $W_{i,t}$, which is constructed in a way that large producers, i.e., producers with an expected high market share, have a higher preference for income, whereas small producers put more emphasis on market share. The latter can be interpreted as a survival strategy.[25]

The maximization is subject to capital constraints. For each price there is a certain expected quantity q^{e} and income INC^{e} implied. The values are derived following a certain sequence of computations. The sequence is not meant to mimic the order of an actual decision process, but reflects how available information is used to make an optimal decision under uncertainty. The sequence can be broken down into five parts, sketched below.

1. The utility of a car for a consumer depends on his preferences for certain car characteristics, driving patterns, social needs, and the (after-tax) price. It is assumed that the producer can estimate averages of these values at least of his customers (who bought in the previous period), e.g., from after sale questionnaires or maintenance reports. Thus, the producer estimates the expected (average) utility of his car, which is called expected competitiveness. R&D success improves the expected competitiveness, but is subject to some randomness. However, the producer can directly influence the expected competitiveness by setting the price (as the decision variable!).

---

[25] This is not directly implied by equation (5), but follows, because expected market share is deduced from current market share.

2. The producer observes the competitiveness of the cars of the competitors and extrapolates a trend from previous development. If the expected competitiveness of the own product exceeds the estimated average competitiveness of all the cars in the market, the producer expects his market share to increase and vice versa.
3. Expected total demand is estimated given expectations on the average price level of the market; and the own price decision affects the expected average price level depending on the market share of the producer. Expected total demand and the expected market share derived in the previous step allow the computation of the expected quantity.
4. The expected quantity can only be produced if sufficient capital is available. Capital depreciates over time, but can be increased using retained earnings from previous periods or from lending at the capital market. The individual credit line depends on the amount of existing capital (as collateral), and it is assumed that producing FCVs is more capital intensive than producing Hy-ICEVs.
5. From the expected quantity and the price, the expected income INC^e is computed as revenue minus variable costs and the value of the objective function can be obtained. As long as the producer has not switched to FCVs, the optimization is done twice, once for Hy-ICEVs and once for FCVs. The switch is made if producing FCVs leads to a higher value of the objective function.

In the case of the production of FCVs, the variable costs in the last step are assumed to decline due to learning by doing (LBD) as in Schwoon (2006b). Learning effects have been observed for a wide range of energy related technologies (Neij, 1997; Mackay and Probert, 1998; Wene, 2000; McDonald and Schrattenholzer, 2001; Neij et al. 2003; Junginger et al., 2005). They are expected to occur also for fuel cell and hydrogen related technologies (Rogner, 1998; Lipman and Sperling, 1999; Tsuchiya and Kobayashi, 2004). LBD leads to a negative relationship between cumulative output and production costs. In the model, LBD is restricted to the fuel cell and hydrogen tanks. Other components and also internal combustion engines show learning effects. But due to the fact that cumulative production already reached billions, they are assumed to be negligibly small.

Implementing LBD in the model requires some refinements of the optimization process, because current production levels affect future costs. Thus, expected (relative share of) income in equation (4) is replaced by its expected net present value computed over a certain decision horizon (with constant quantities and declining costs for FCVs). This generates an inconsistency in that producers create expectations beyond the next period. But this changed optimization only affects the switching decision and not the price decision. It is necessary, because with LBD the likelihood of switching would otherwise increase with the length of the model time step. A longer time step leads to higher production quantities and, therefore, higher cost reductions. The length of the model time step (three months) is determined by how often prices and production quantities can be adjusted. The decision to switch, though, should incorporate projections over several years.

6.3.4 Heterogeneity of producers

Producers cannot do optimal pricing based on intertemporal expected profit maximization, because the behaviour of competitors, R&D success, infrastructure build-up, etc. are uncertain and (altogether) do not follow probability distributions. Therefore, a wide range of potential heuristics, i.e., alternative objective functions exists. Kwasnicki and Kwasnicka (1992) show that in a similar setting the above objective function outperforms the majority of alternatives with respect to long term profits. Here, all producers use the same objective function as being the ‘best one available’. However, two sources of heterogeneity between producers remain. The first one is that their products initially have different characteristics and changes due to (individual) R&D success underlie some randomness. The second, more important one is size (in terms of market share). Size determines not only the weight in the objective function,[26] but also market power, credit availability and R&D expenditures (which are positively correlated with R&D success).

6.3.5 Connecting modules

The main connection between producers and consumers is the selling process. After the producers set prices and adjusted their production capacity, the actual total demand in the market (i.e., the number of potential consumers who buy a car) is derived from the actual average price of the cars offered. Buying consumers are chosen, depending on how long they already have their old car. This approximates the behaviour that in times of generally increasing prices, consumers tend to drive their old cars longer and in times of decreasing prices more new cars are bought. Consumers evaluate the cars as described in section 6.3.1 and make their orders. Producers construct only as many cars as consumers order, up to their capacity limit. So, there is no excess supply (inventories are omitted). This implies that producers, which overestimated the demand for their products, are penalized by their overinvestment in capacity, but not by high variable costs. In the case of excess demand, not all consumers can be satisfied, because a period is not long enough for capacity extensions or price increases. If a consumer cannot get his favourite product, because it is sold out, he will choose a less preferred product and he can actually end up with nothing and has to wait for the next period.

There are also indirect interactions between producers and consumers. As stated above, producers gather information about the preferences of their customers and target R&D activities accordingly. Consumers’ preferences, on the other hand, are influenced by average car characteristics in the market, representing a marketing effect similar to Valente (1999). Due to the simplistic representation of the fuel suppliers and the government, there is no real connection to the other modules, but rather a direct feedback or respectively an exogenous influence.

6.3.6 Calibration and scenarios

The model is implemented in the Laboratory for Simulation Development (LSD).[27] Its calibration aims at mimicking some of the main features of the German compact car segment. There are 12 important producers in the segment of compact cars in Germany with market shares exceeding 2% and a dominating producer (*Volkswagen*) with a market share of about 1/3. To simulate the asymmetric situation, initial market shares are drawn randomly from a normal distribution with mean

---

[26] Note that expected market share is correlated with current market share.

[27] LSD is an open source environment for C++ programming. Its main features are discussed in Valente and Andersen (2002). It is available at http://www.business.aau.dk/lsd/lsd.html.

12/100% and a standard deviation of 10%.[28] For computational reasons 6400 different consumers are modelled. In the control run without any policy about 125 consumers buy each period, i.e., if we assume that each consumer represents about 2,000 similarly behaving ones, we end up at one million sales per year, which corresponds to the size of the compact car segment.

It is assumed that by 2010, 400 fuel stations will offer hydrogen (i.e., approximately 3% of the filing stations in Germany).[29] Results are shown only for tax policies or additional 'major H_2 program' scenarios, which represent a public infrastructure program that provides 160 additional hydrogen outlets at existing filling stations each year.[30] Two different tax scenarios are implemented and eventually combined with the infrastructure program. One is a 'shock tax' with an instantaneous 40% tax in the year 2010 on (newly bought) conventional cars. The tax hits the market, so as to directly push FCVs into it. Alternatively a 'gradual tax' is used with a quarterly increase of 1% ending up also at 40% in 2020, where agents can smoothly adjust to the new circumstances. The scenarios represent extreme cases for demonstrative purpose. The 40% tax level represents not only purchase taxes, but also the net present value of total lifecycle taxes (on ownership, insurance, fuel etc.). Compared to present car taxes in Europe (as listed in Burnham, 2001), 40% is at the low end of current rates.

In the central case, a learning rate of 15% for fuel cell related technologies is assumed.[31] This means that costs decrease by 15% for a doubling of cumulative output. For the simulations here, learning is fully appropriated by the producers, i.e., learning spillovers are neglected (contrary to Schwoon, 2006b). Learning takes place only on the national market, i.e., global learning effects of international producers introducing FCVs in several markets at the same time are ignored. Thus, the results are relevant for a situation, in which a national government decides to push in a solo attempt the introduction of the new technology in a market of comparable size to the German market. An example for such a policy in the history of pollution regulation of cars is the independent introduction of unleaded fuels and the support of 3-way catalytic converters in Germany, preceding most other countries in Western Europe (Westheide, 1987).

6.4 Results

6.4.1 Diffusion projections

As a benchmark to compare the success of different policies Figure 6.2 shows the share of FCVs within all newly registered cars. All figures presented refer to averages of 100 simulation runs using different random initializations. Diffusion comes along with an increase in the share of filling stations with hydrogen outlets (see Figure 6.3). The shock tax directly forces at least one producer to switch to the production of FCVs. Public infrastructure speeds up diffusion at the beginning, but later on the

---

[28] The minimum market share is 2% and the sum of all market shares is scaled to sum up to 100%.

[29] 2010 is chosen arbitrarily as the starting point of the policies that should move FCVs out of niche applications.

[30] This is equal to the 'high exogenous H_2'-scenarios in Schwoon (2006a).

[31] Dutton and Thomas (1984) present data of 100 estimates of learning rates in manufacturing. They find a median learning rate of 19-20%. For a smaller sample of energy technologies, McDonald and Schrattenholzer (2001) report a median of 16-17%, so that 15% is chosen as a rather conservative assumption.

impact is negligible. The reason is that for a certain share of consumers, refuelling remains a critical issue and they are served by a few producers that establish a successful temporary niche.[32]

In the case of the gradual tax, the tax level does not have to reach the full level of 40% before (on average) some producers start producing FCVs. Public infrastructure build-up seems to have a much more important influence on the diffusion, because, as Figure 6.3 shows, already about 10% of the filling stations are equipped with hydrogen until the tax reaches a level that forces producers to switch. However, even with a major public infrastructure program, the gradual tax scenario leads to much slower diffusion compared to the shock tax. In the year 2030, 10 years after the gradual tax reached its maximum, the share of FCVs within newly registered cars is 40%, a level reached with the shock tax (without public infrastructure) in less than five years.

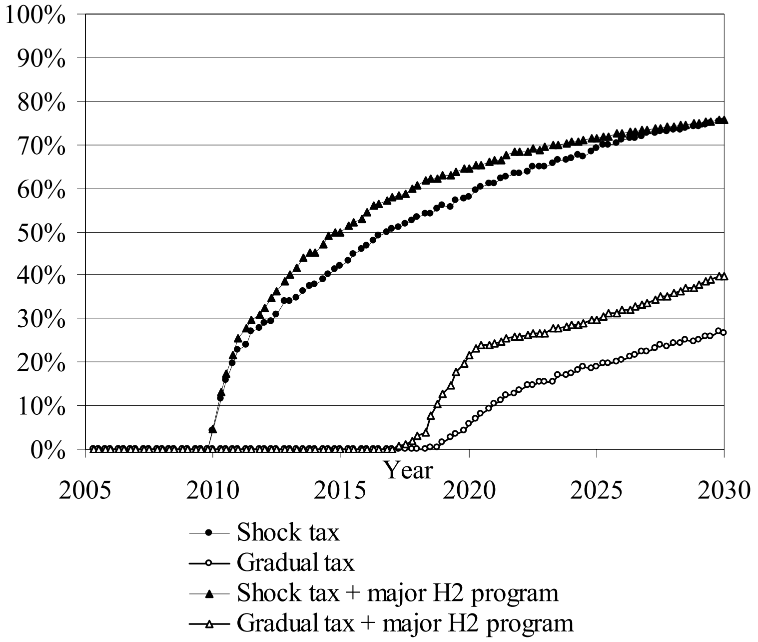

Figure 6.2 Share of FCVs within the newly registered vehicles for different policy scenarios

---

[32] The niche is only temporary, because as soon as there is full infrastructure coverage, all producers will switch to FCVs.

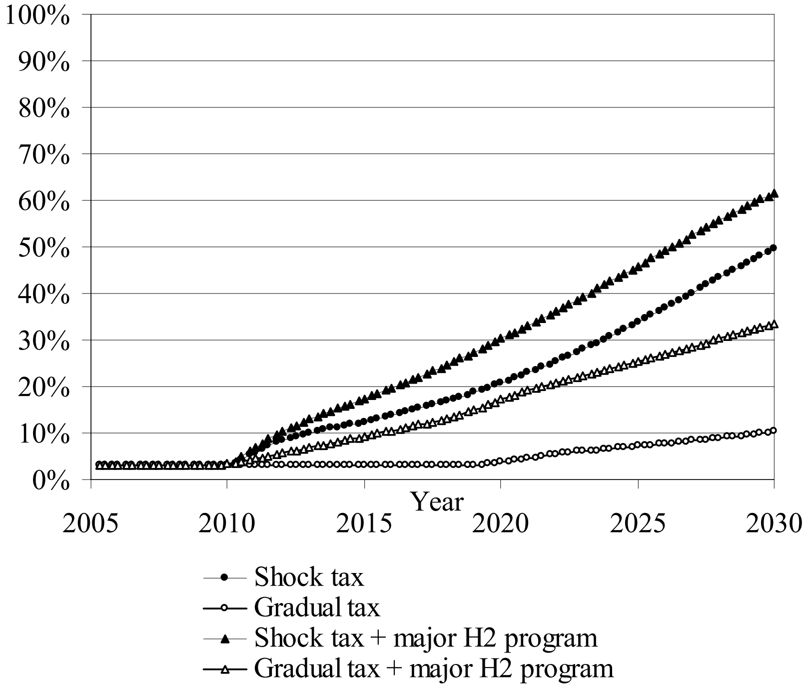

Figure 6.3 *Development of the share of filling stations with hydrogen outlet*

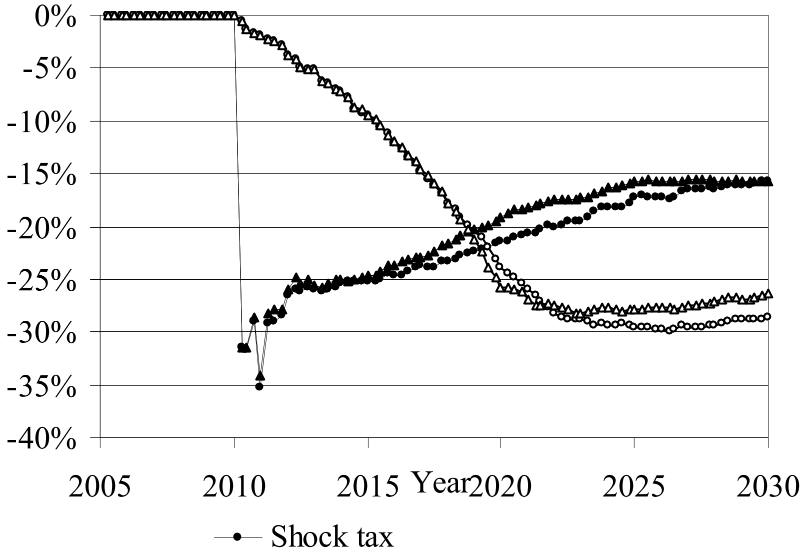

Figure 6.4 *Change in the sum of all cars sold relative to the development without a tax*

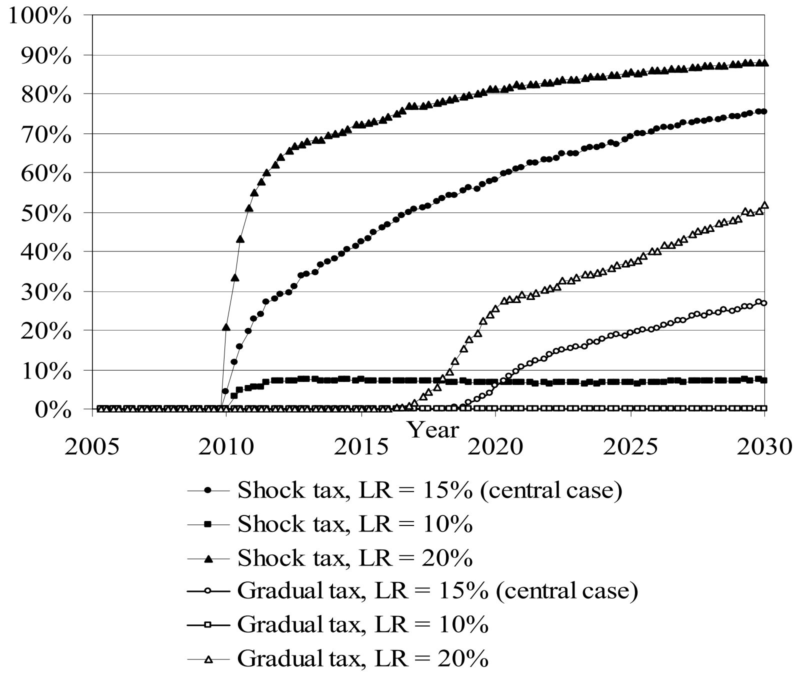

Figure 6.5 Share of FCVs within the newly registered vehicles: Sensitivity to the learning rate

The reason why the shock tax leads to faster diffusion can be seen in Figure 6.4, which shows the number of cars produced relative to the development without a tax. There is a sharp drop in sales due to an increase in after-tax prices in the shock tax case. However, producers, who change production directly due to the shock tax, still include, to some degree, the relatively high pre-tax production levels into their decision. Thus, they expect comparatively high LBD cost reductions and are, therefore, more likely to switch. In the gradual tax case, demand goes down steadily, so that if the tax rate reaches levels that make switching considerable, demand is very low. Producers' expectations about LBD are therefore also low and production of FCVs is postponed.

In Figure 6.5 the relationship between tax scenario and LBD expectations is further illustrated. Sensitivity results are shown for different learning rates (LR). With a rather low learning rate (LR = 10%) only the shock tax leads to diffusion.[33] The gradual tax signal is not sufficient to stimulate diffusion. With higher learning (LR = 20%) the shock tax generates an instantaneous introduction of the new technology, with more than 50% market share reached within the first year. Even in that case, the gradual tax only leads to a smooth introduction and it takes until 2030 (ten years after the 40% tax level is reached) until every second car sold is an FCV. The strong impact of the learning rate on the speed of diffusion also demonstrates how important knowledge of potential learning processes is. The changes in production (respectively sales) in Figure 6.4 illustrate that production only recovers after diffusion begins. Thus, a substantial tax that is not sufficient to actually promote switching, because learning effects have been overstated, would be extremely destructive.

---

[33] 65% of the observed learning rates presented in McDonald and Schrattenholzer (2001) and 82% of those presented in Dutton and Thomas (1984) exceed 10%.

6.4.2 Asymmetric impacts on agents

The production figures are basically a mirror image of the after-tax price development shown in Figure 6.6. Thus, Figure 6.4 and Figure 6.6 together give an impression of how consumers are affected by the tax. Consumers face considerable price increases in every policy scenario. With the shock tax, these increases happen directly after the introduction, but later on price levels are actually lower than in the gradual tax cases. Thus, consumers would be (relatively) better off in the long run. However, consumers would suffer from the drastic price increase right at the beginning, and such a policy would, therefore, be rather difficult to implement. In any case, consumers would benefit from a major infrastructure program via lower car prices.[34] The price effect is due to the generally faster diffusion that implies LBD cost reductions, and these cost reductions are at least partly passed on to the consumers.

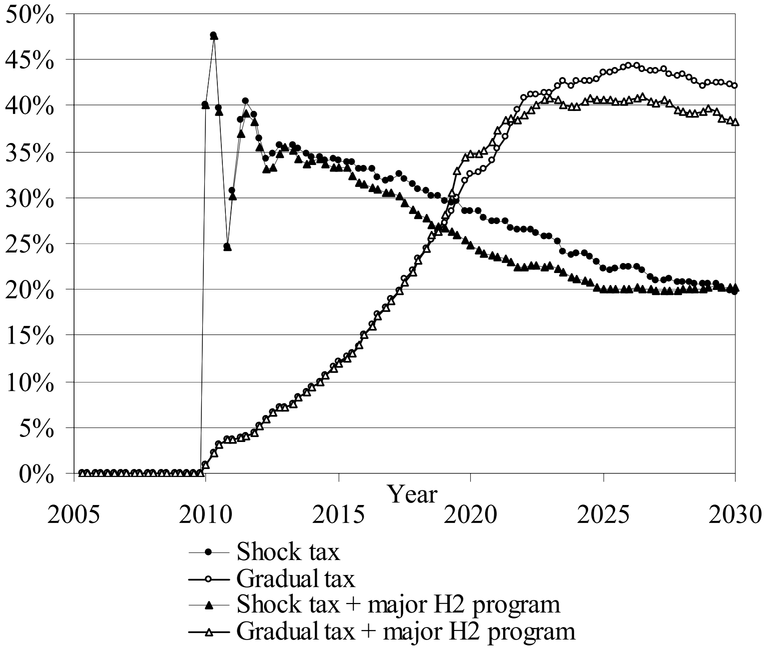

Figure 6.6 Change in average after-tax price relative to the development without a tax

It might be surprising to see that prices increase, at least temporarily, by a higher percentage than the tax. This is partly due to the optimization routine of the producers, which focuses on relative income and market share and not absolute, so that they might fully shift the tax burden to the consumers. But another reason is that the large producers have the advantage of predicting demand changes better during changes in the tax rate, as they have a higher influence on average prices. This leads to a noticeable increase in concentration, as one can see from the Herfindahl-index displayed in Figure 6.7. This index is constructed by summing the square of market shares for all firms and lies between 0 (perfect competition) to 1 (monopoly). The index jumps up from 0.13 to 0.17 for the shock

---

[34] Note that the model neglects the costs of the infrastructure program and a potential use of the tax revenues for infrastructure investments. Thus, the model does not allow a full cost-benefit analysis. However, the development of the after-tax car price is considered to be a good proxy for potential resistance to certain policies.

tax.[35] There is a temporary backlash (due to price cuts from small producers in order to survive), but the concentration index remains high afterwards. In the gradual tax cases, there is a steady increase in concentration. In both tax scenarios, the higher concentration, which implies greater market power, leads to higher prices.[36]

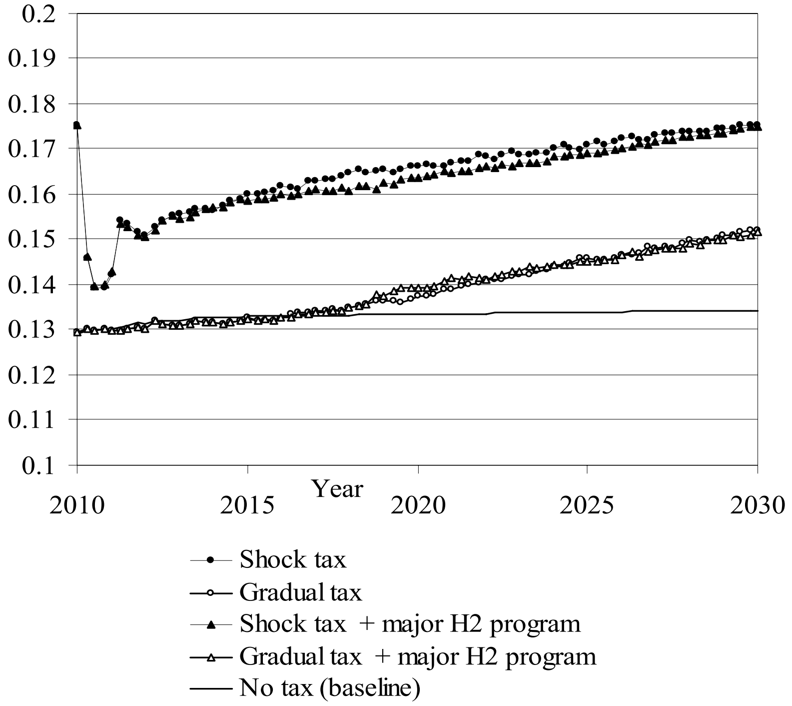

Figure 6.7 Development of the Herfindahl-index after introduction of the tax

This result already suggests that producers are differently affected by the policies, depending on their size. Figure 6.8 and Figure 6.9 show the change in the sum of profits of the three largest and three smallest producers relative to the no tax baseline. The profits of the large producers are substantially hit by the shock tax, but then recover very quickly (within about a year) before they actually exceed profits without the tax. The increase in profits then continues, following the development of the concentration discussed above. In the gradual tax cases, profits do not decrease that much, but stay below the level without a tax for almost a decade until the first producers switch to the production of FCVs. Thus, large producers would actually be better off with the shock tax that promotes diffusion immediately and quickly raises their profits above the level without the tax.[37]

---

[35] A Herfindahl-index of 0.13 applies to a market in which 7 to 8 firms compete with an equal market share. With 0.17 this number drops to 5-6, thus, the concentration increase is substantial.

[36] Note that the possibility of (foreign) entry is ignored.

[37] Only for an unrealistically high discount rate large producers would be better off with the gradual tax, as it does not imply the drastic profit reduction right after implementation.

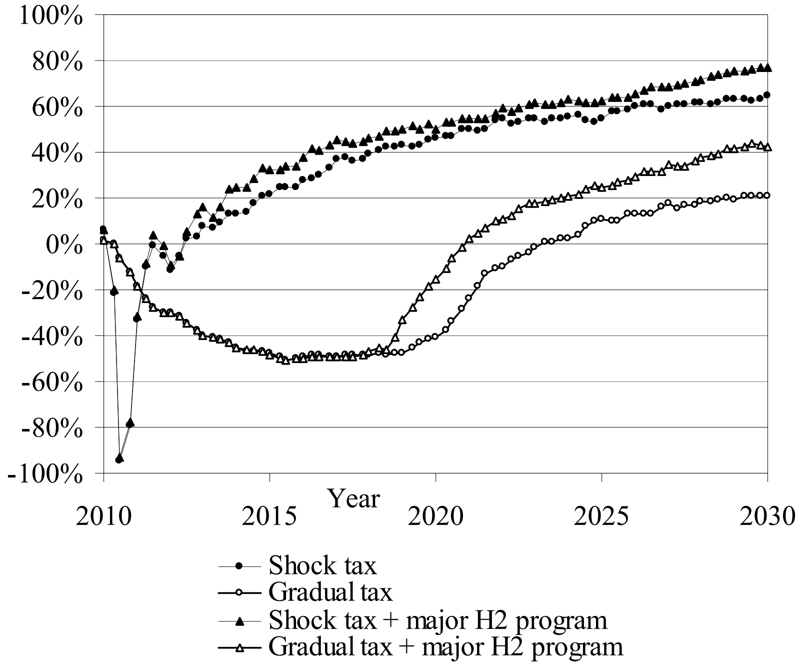

Figure 6.8 Change in the sum of the profits of the three largest producers relative to the development without a tax

Small producers suffer from any of the policies, but particularly from the shock tax. In that case, the infrastructure program makes them even worse off, which suggests that small producers are the losers especially of those policies that generate fast diffusion of the new technologies. Thus, strong resistance to FCV supporting policies can be expected. On the other hand, large producers additionally win from infrastructure investments. This might actually let them consider side payments to fuel suppliers to support fast infrastructure build-up.

6.4.3 Sensitivity with respect to different buying decisions

Schwoon (2006a) presents a wide range of results from sensitivity analyses, identifying parameters that crucially determine the speed of diffusion. The most important parameters are the price elasticity of consumers, the distribution of weights between individual preferences and social needs in the consumer population, and also the weight between expected income share and market share in the objective of the producers. Here, I only show how the speed of diffusion is affected if some of the consumers do ‘social comparison’ as defined in Janssen and Jager (2002). Applying this different consumer behaviour is an example for a change in the consumer module that can be implemented independently from the rest of the model, reflecting the advantage of the modular set-up of the simulation model.

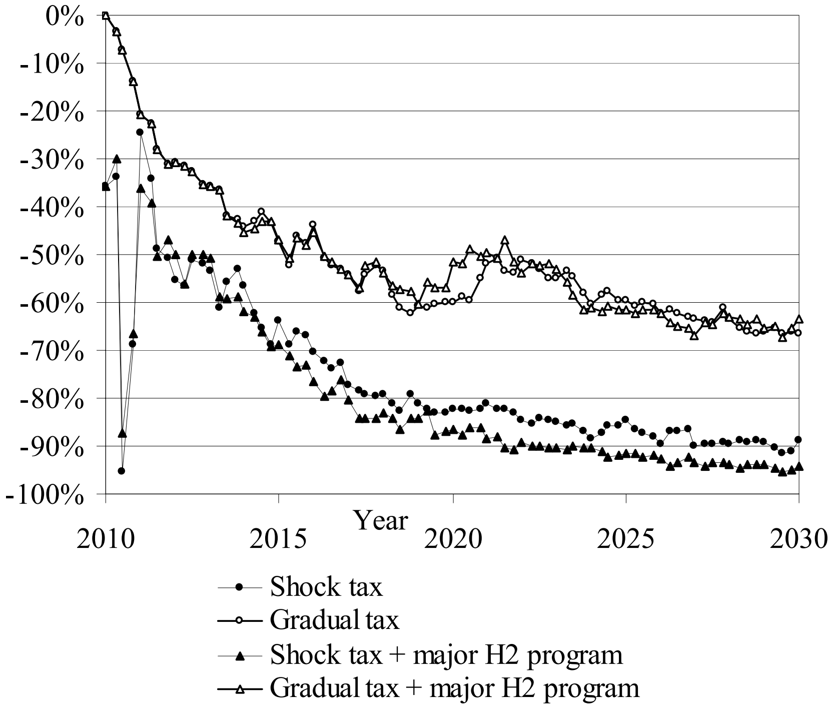

Figure 6.9 Change in the sum of the profits of the three smallest producers relative to the development without a tax

Consumers might be uncertain in judging car characteristics or, e.g., operating costs and so on, so they consider themselves unable (or not willing due to information costs) to evaluate all the cars available on the market. In that case, they are assumed to compare only the utility associated with the car that is driven by the majority of their neighbours with the utility from buying the latest version of their old car again. This means that they reduce their decision space to two directly perceivable products.

In the social comparison cases in Figure 6.10, on average some 50% of the consumers actually do social comparison. The small decision space increases the speed of diffusion at the beginning. Consumers stick to their brand or choose that of their neighbours even if it is now only available as a FCV. But later on this effect of a continuation of previous behaviour leads to resistance to full diffusion, so that by the year 2030 the share of newly registered FCVs is lower than without social comparison. Note that these results are driven by the fact that producers radically switch to producing the new technology. Thus, consumers sticking to their ‘old product’ might actually be forced to buy a FCV. In a more realistic model that allows producers to offer the same car with different drive trains, social comparison is likely to lead to much slower diffusion in the beginning, because consumers doing social comparison would hardly be exposed to the new technology and, therefore, not consider them at all. They would generally drop out as potential initial adopters, even if, e.g., their individual driving behaviour militates in favour of adoption. The number of initial adopters, though, is critical for producers to introduce FCVs. Thus, a large share of consumers doing social comparison might actually prevent a successful introduction of FCVs.

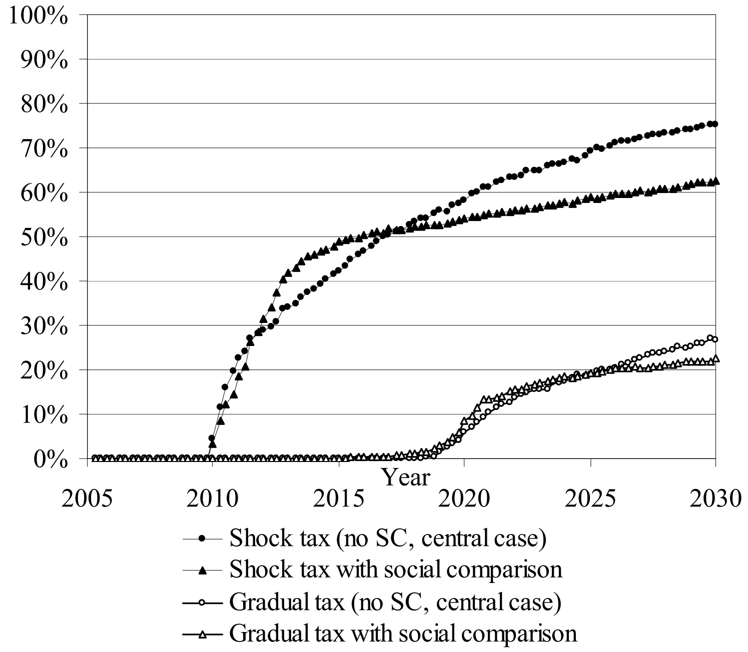

Figure 6.10 *Impact of the share of FCVs within newly registered vehicles if 50% of the consumers do social comparison*

6.5 Conclusion

In this paper, I described the core features of an agent-based computational model that has been developed to understand the dynamics of a policy driven transition to FCVs fuelled with hydrogen. The model combines, in a modular way, an existing producer and consumer model and adds a fuel supplier component. The modules operate relatively independently as they have only a few (but dynamically important) connections. Future improvements or experiments within one of the modules are, therefore, easily implemented. The model is used to evaluate certain tax and infrastructure-policy scenarios that enter the model as exogenous drivers. It incorporates several features separating it from neoclassic approaches towards technology adoption in the presence of adoption externalities. Agents are heterogeneous in several characteristics and in behaviour. They are myopic, and their decisions (as reactions to the policies) are driven by individual interactions with other agents.

The model is specified and calibrated to represent the dynamics of a policy driven introduction of FCVs in the German compact car market. However, the structure can generally be applied to new car technologies that require a specific fuel that is rarely available. Results are shown for a shock tax and a gradual tax scenario, so as to represent extreme cases. The taxes may or may not be combined with additional public infrastructure investments that increase the share of filling stations that offer the new fuel. The shock tax initiates a diffusion of FCVs in terms of the share of newly registered cars right after the introduction of the tax. With a gradual tax, it takes several years until a tax level is reached that forces producers to switch to the production of FCVs. But even from that later point in time, diffusion is much slower compared to the shock tax. If the learning rate of fuel cell technologies is rather low, the gradual tax might even be insufficient to stimulate a single producer to switch.

However, in the central case parameterization with a learning rate for fuel cell technology of 15% the taxes are able to overcome the chicken-and-egg-problem usually associated with the introduction of FCVs and hydrogen infrastructure.

The different policy scenarios have substantially different impacts on the agents. Consumers are likely to prefer a gradual tax that leads to slowly increasing prices. In any case, they would be in favour of a major infrastructure program, because it promotes faster diffusion and higher learning cost reduction that keep average car prices comparatively low. Both tax scenarios increase concentration in the market. Large producers benefit from higher profits in the medium to long term, particularly in the shock tax case. The benefits are at the expense of small producers, who are likely to oppose any diffusion policy, as they would suffer substantial losses. The faster the diffusion, the more profitable are large producers. Therefore, they would also be the winners of a major infrastructure program.

There are two types of modelling issues that limit the validity of the results: model-specific simplifications and problems of simulation models in general. Simplifications are necessary to keep the already rather complex model manageable, so that it does not become a ‘black box’, in which too many parameters and behavioural equations tend to obscure results. A major simplification, however, is that producers only have the option to fully switch to the new technology. In reality, producers are more likely to introduce the new technology in certain product lines. Moreover, the model is restricted to a single market segment. But the tax might force consumers, e.g., to switch to cars in a cheaper segment rather than to adopt the new technology. A more realistic model would also call for a more detailed representation of fuel suppliers, including investment decisions with relatively long payback periods. Another drawback is that the consumer model does not allow for a computation of consumer rents[38], so that efficiency costs of the tax cannot be investigated. This would be necessary to derive the environmental performance relative to the tax burden (or relative to infrastructure expenditures). The results already indicate that environmental performance of the policies, at least over the simulated time period, is not straightforwardly computed. The taxes lead to declines in sales of newly registered cars, suggesting that old cars tend to be driven longer. This might imply adverse environmental effects under the assumption that environmental performance of new cars (FCVs and Hy-ICEVs) is generally higher than that of the average car in the car population.

Apart from the explicit limitations of the model, there are problems related to the methodology of simulations as such. The simulations underlie parameter uncertainty together with uncertainty of behavioural assumptions. Uncertainty is particularly large, because the model addresses a very specific technological transition that is unprecedented in history, so that standard calibration/validation cannot be applied. These issues can be summarized as model uncertainties. The only way to deal with it is sensitivity analysis in order to identify those parameters (or behavioural equations) that have the most severe impact on results. In addition to model uncertainty, the model itself generates uncertainty as a simulation of reality, in which decisions are at least partly driven by random events. Random events that drive the results are controlled for by comparing averages over hundreds of simulations. However, the future will not follow an ‘average path’ but will be, so to say, a singular chain of events.

---

[38] The reason is that consumers compare utilities from heterogeneous products and, therefore, do not have a specific willingness to pay that could be used to compute an aggregate demand function. Note that aggregate producer rents are straightforwardly derived.

Model uncertainties together with model-inherent stochastic developments rule out that simulation results can be interpreted as forecasts. But the model results are the key to understanding the main dynamics of a complex technological system. For the introduction of FCVs, lessons learned independently of actual magnitudes are that a high immediate taxation of conventional cars promotes fast diffusion, but at the price of not only declines in sales but also increasing market power of already large producers. In addition, large producers would be the beneficiaries of a major public infrastructure program, whereas small producers would actually suffer. Impacts on consumers and industry performance have been so far ignored by studies that address the costs of switching to hydrogen based individual transport. The introduction of a major tax and/or infrastructure program is likely to face resistance of certain consumer and industry councils. Therefore, a better understanding of transition dynamics helps developing strategies that keep disruptive impacts as small as possible. Identifying the resulting winners and losers of the policies in advance, would also allow for compensation policies that might reduce resistances.

References

Bentley, R.W. (2002). Global oil & gas depletion: an overview. *Energy Policy* 30: 189-205

Brynjolfsson, E. and C.F. Kemerer (1996). Network Externalities in Microcomputer Software: An Econometric Analysis of the Spreadsheet Market. *Management Science* 42 (12): 1627-1647

Bunch, D.S., M. Bradley, T.F. Golob, R. Kitamura, and G. Occhiuzzo (1993). Demand for Clean-Fuel Vehicles in California: A Discrete-Choice stated preference pilot project. *Transportation Research Part A* 27 (3): 237-253

Burnham, J. (2001). *Review for the Commission for Integrated Transport of the report for the Scotland Office* by Colin Buchanan and Partners 'Comparison of motoring taxation costs across Europe'. Middlesex University
Avalaible at: http://www.cfit.gov.uk/docs/2001/scot0122/scot0122/index.htm.

Dutton, J.M. and A. Thomas (1984). Treating progress functions as a managerial opportunity. Academy of *Management Review* 9: 235-247

Economides, N. and C. Himmelberg (1995). Critical Mass and Network Evolution in Telecommunications. In: Brock, G. (ed.). *Toward a Competitive Telecommunications Industry: Selected Papers from the 1994 Telecommunications Policy Research Conference*
Available at: http://pages.stern.nyu.edu/networks/tprc.pdf

EEA (2006). *Transport and environment: facing a dilemma*. EEA Report 3/2006.
Available at: http://www.eea.eu.int

EC-JRC (2006). *Well-to-Wheels Analysis of Future Automotive Fuels and Powertrains in the European Context.* Well-to-Wheels Report, Version 2b, May 2006
Available at: http://ies.jrc.ec.europa.eu/wtw

EPA (1996). *Greenhouse Gas Emissions form the U.S. Transportation Sector 1990-2003.*
Available at: http://www.epa.gov/otaq/climate.htm

Farrell, A.E., D.W. Keith and J.J. Corbett (2003). A strategy for introducing hydrogen into transportation. *Energy Policy* 31: 1357-1367

Farrell, J., and G. Saloner (1986). Installed Base and Compatibility: Innovation, Product Preannouncements, and Predation. *American Economic Review* 76 (5): 940-955

Farrell, J., and G. Saloner (1985). Standardization, compatibility, and innovation. *The RAND Journal of Economics* 16 (1): 70-83

Friedrich, R., and P. Bickel (eds) (2001). *Environmental External Costs of Transport*. Springer-Verlag, Berlin, Heidelberg, New York

Gandal, N. (1994). Hedonic Price Indexes for Spreadsheets and an Empirical Test for Network Externalities. *The RAND Journal of Economics* 25: 160-170

Greene, D.L. (1998). Survey Evidence on the Importance of Fuel Availability to the Choice of Alternative Fuels and Vehicles. *Energy Studies Review* 8 (3): 215-231

Hart, D. (2005). *The Economics of a European Hydrogen Automotive Infrastructure - A study for Linde AG. E4tech study.* Available at: http://www.braunschweig.ihk.de/innovation_umwelt/nachrichten/april_05/The%20Economics%20of%20a%20European%20Hydrogen%20Automotive%20Infrastructure.pdf

Hoed, R. van den (2005). Commitment to fuel cell technology? How to interpret carmakers' efforts in this radical technology. *Journal of power sources* 141 (2): 265-271.

Jager, W. (2000). *Modelling consumer behaviour.* PhD thesis, University of Groningen, NL
Available at: http://www.bdk.rug.nl/medewerkers/w.jager/

Janssen, M.A. and W. Jager (2002). Stimulating diffusion of green products - Co-evolution between firms and consumers. *Journal of Evolutionary Economics* 12 (3): 283-306

Junginger, M., Faaij, A., and Turkenburg, W.C. (2005). Global experience curves for wind farmes. *Energy Policy* 33: 133-150

Katz, M. and Shapiro, C. (1985). Network Externalities, Competition, and Compatibility. *American Economic Review* 75 (3): 424-440

Katz, M.L. and C. Shapiro (1986). Technology Adoption in the Presence of Network Externalities. *Journal of Political Economy* 94 (4): 822-841

Kwasnicki, W. (1996). *Knowledge, innovation and economy: an evolutionary exploration*. Edward Elgar Publishing Limited; Cheltenham (UK)

Kwasnicki, W. and H. Kwasnicka (1992). Market, innovation, competition - An evolutionary model of industrial dynamics. *Journal of Economic Behaviour and Organization* 19: 343-368

Lipman, T. and D. Sperling (1999). *Forecasting the Costs of Automotive PEM Fuel Cell Systems - Using Bounded Manufacturing Progress Functions*. Proceedings of the IEA International Workshop at Stuttgart. Publication No. UCD-ITS-RP-00-6.

Mackay, R.M. and S.D. Probert (1998). Likely market-penetrations of renewable-energy technologies. *Applied Energy* 59 (1): 1-38

McDonald, A. and L. Schrattenholzer (2001). Learning rates for energy technologies. *Energy Policy* 29: 255-261

Mercuri, R., A. Bauen and D. Hart (2002). Options for refuelling hydrogen fuel cell vehicles in Italy. *Journal of Power Sources* 106: 353-363

Moore, R.B. and V. Raman (1998). Hydrogen Infrastructure for Fuel Cell Transportation. *International Journal of Hydrogen Energy* 23 (7): 617-620

Neij, L., (1997). Use of experience curves to analyse the prospects for diffusion and adoption of renewable energy technology. *Energy Policy* 23 (13), 1099-1107

Neij, L., P.D. Andersen, M. Durstewitz, P. Helby, Hoppe-Kilpper and M., Morthorst, P.E. (2003). *Experience Curves: A Tool for Energy Policy Assessment. Environmental and Energy Systems Studies* Sweden. Available at: http://www.iset.uni-kassel.de/extool/.

Ogden, J.M. (1999). Developing an infrastructure for hydrogen vehicles: Southern California case study. *International Journal of Hydrogen Energy* 24 (8): 709-730

Ogden, J.M. (2002). *Modelling Infrastructure for a Fossil Hydrogen Energy System with CO_2 Sequestration*. Paper prepared for the Sixth Greenhouse Gas Control Technologies Conference, Fall 2002, Kyoto, Japan

Oi, T. and K. Wada (2004). Feasibility study on hydrogen refuelling infrastructure for fuel cell vehicles using the off-peak power in Japan. *International Journal of Hydrogen Energy* 29: 347-354

Rogner, H.H. (1998). Hydrogen Technologies and the Technology Learning Curve. *International Journal of Hydrogen Energy* 23 (9): 833-840

Schwoon, M. (2006a). Simulating the Adoption of Fuel Cell Vehicles. *Journal of Evolutionary Economics* 16 (4): 435-472

Schwoon, M. (2006b). *Learning-by-doing, Learning Spillovers and the Diffusion of Fuel Cell Vehicles*. FNU Working Paper #112. Available at: http://www.uni-hamburg.de/Wiss/FB/15/Sustainability/WP-FNU-112FCV_learning_by_doing.pdf

Sørensen, B., A.H. Petersen, C. Juhl, H. Ravn, C. Sondergren, P. Simonsen, K. Jorgensen, L.H. Nielsen, H.V. Larsen, P.E. Morthorst, L. Schleisner, F. Sørensen and T.E. Pedersen (2004). Hydrogen as an energy carrier: scenarios for future use of hydrogen in the Danish energy system. *International Journal of Hydrogen Energy* 29: 23-32.

Steinberger-Wilckens, R. (2003). Not cost minimisation but added value maximization. *International Journal of Hydrogen Energy* 28: 763-770

Stromberger, R. (2003). *Building up a gas station infrastructure to supply vehicles with hydrogen as an alternative fuel, taking technical, economic and ecological aspects into consideration*. Abridged version of dissertation, Vienna University of Technology.
Available at: http://www.hydrogenhighway.ca.gov/tt/economy/stromberger.pdf

Thomas, C.E., Kuhn, I.F., James, B.D., Lomax, F.D., and Baum, G.N. (1998). Affordable Hydrogen Supply Pathways for Fuel Cell Vehicles. *International Journal of Hydrogen Energy* 23 (6): 507-516.

Tsuchiya, H. and Kobayashi, O. (2004). Mass production cost of PEM fuel cell by learning curve. *International Journal of Hydrogen Energy* 29: 985-990.

Valente, M. (1999). *Evolutionary Economics and Computer Simulations - A Model for the Evolution of Markets*. PhD Thesis, University of Aalborg, Denmark.

Valente, M. and Andersen, E.S. (2002). A hands-on approach to evolutionary simulation: Nelson-Winter models in the Laboratory for Simulation Development. *The Electronic Journal of Evolutionary Modelling and Economic Dynamics* 1, 1003
Available at: http://beagle.u-bordeaux4.fr/jemed/1003/index.php.

Verspagen, B. (2006). *The use of modelling tools for policy in evolutionary environments*. Chapter 2 in this report.

Wene, C.O. (2000). *Experience Curves for Energy Technology Policy*. IEA/OECD, Paris.
Available at: http://www.iea.org/books/studies/2000/curve2000.pdf.

Annex 1: Programme workshop ‘Environmental policy and modelling in evolutionary economics’

Thursday 18 May 2006

11:45-17:30

	Special session 11h45-13h00: open to all participants of the conference (Turing room)
11:45 – 12:05	Paul Windrum (Manchester Metropolitan University) *Technology successions and policies for promoting more environmentally friendly technologies*
12:05 – 12:25	Maïder Saint-Jean (University of Bordeaux) *Environmental innovation and policy: Lessons from an evolutionary model of industrial dynamics*
12:25 – 12:45	Wander Jager (University of Groningen) *Simulating consumer behaviour: a perspective*
12:45 – 13:00	Jeroen van den Bergh (Free University, Amsterdam) *Discussant*
13:00 – 14.15	**Lunch break**

	Workshop 14h15-17h30: upon invitation only (Room M279)
14.15 – 14.25	Fred Langeweg (MNP - Netherlands Environmental Assessment Agency) *Mission statement* (http://www.mnp.nl/en/aboutmnp/index.html) *Questions from the Netherlands Environmental Assessment Agency to workshop participants*
14.25 – 14.45	Bart Verspagen (Eindhoven Technical University) *The use of modelling tools for policy in evolutionary environments*
14.45 – 15.15	*Working session 1 (three parallel groups): If the Netherlands Environmental Assessment Agency will work on an evolutionary model of transitions towards sustainable technologies:* • Which building blocks are already available? Where to start? • What should be the objective(s) of this model? • What should be included and what should be excluded? • What type of lessons could be learned from the model? • How to validate the model?
15.15 – 15.45	Presentation of the results of the three groups: 10 minutes per group
15.45 – 16.00	**Coffee/tea break**
16.00 – 16.30	*Working session 2 (three parallel groups): Three policy dilemma’s:* • Should a government pre-select sustainable technologies or rely on market selection, and how can the government influence the market without pre-selecting technologies? • Should a government stimulate diversity of technologies, firms, products and strategies and can a diversity-oriented policy be compatible with the now fashionable policies of stimulating collaboration and the creation of a ‘critical mass’? • Should a government apply short-term standards triggering firm entry and radical innovation or long-term standards hereby creating a level-playing field? • Should a government try to influence consumer behaviour in an age of consumer sovereignty ideology, and, if so, how? Per dilemma: • Which policy advice can be given at this moment on the dilemma? • What is the foundation from evolutionary economics for your advice? • Which further research questions should be addressed in relation to each dilemma?
16.30 – 17.00	Presentation of the results of the three groups: 10 minutes per group
17.00 – 17.25	Open discussion
17.25 – 17.30	Koen Frenken (Utrecht University) and Annemarth Idenburg (DHV) *Conclusions*
17.30 – 18.00	**Drinks**

Annex 2: List of participants

Name	Affiliation	Email
Floortje Alkemade	Utrecht University	f.alkemade@geo.uu.nl
Jeroen van den Bergh	Free University Amsterdam	jbergh@feweb.vu.nl
Corjan Brink	MNP - Netherlands Environmental Assessment Agency	corjan.brink@mnp.nl
Alessio Delre	University of Groningen	sadelre@bdk.rug.nl
Albert Faber	MNP - Netherlands Environmental Assessment Agency	albert.faber@mnp.nl
Koen Frenken	Utrecht University	k.frenken@geo.uu.nl
Julian Garcia	Free University Amsterdam	j.garcia@feweb.vu.nl
Annemarth Idenburg	DHV Consultancy and Engineering	annemarth.idenburg@dhv.nl
Wander Jager	University of Groningen	w.jager@rug.nl
Peter Janssen	MNP - Netherlands Environmental Assessment Agency	peter.janssen@mnp.nl
Fred Langeweg	MNP - Netherlands Environmental Assessment Agency	fred.langeweg@mnp.nl
Frans Oosterhuis	Free University Amsterdam	frans.oosterhuis@ivm.falw.vu.nl
Jan Ros	MNP - Netherlands Environmental Assessment Agency	jan.ros@mnp.nl
Karolina Safarzynska	Free University Amsterdam	karolina.safarzynska@ivm.vu.nl
Maïder Saint-Jean	University of Bordeaux	saintjea@u-bordeaux4.fr
Malte Schwoon	Hamburg University	m.schwoon@gmx.de
Bart Verspagen	Eindhoven Technical University	b.verspagen@tm.tue.nl
Bert de Vries	MNP - Netherlands Environmental Assessment Agency	bert.de.vries@mnp.nl
Paul Windrum	Manchester Metropolitan University	p.windrum@mmu.ac.uk