Essays on transition challenges for alternative propulsion vehicles and transportation systems

by
Jeroen J.R. Struben

Master of Science in Physics,
Delft University of Technology (1996)

Submitted to the Alfred P. Sloan School of Management
in partial fulfillment of the requirements for the Degree of

Doctor of Philosophy in Management

at the

Massachusetts Institute of Technology
September 2006

Signature of Author………………………………………………………………
Department of Operations Management and System Dynamics, September 2006

Certified by ……………………………………………………………………
John D. Sterman
Professor of Management
Thesis Supervisor

Accepted by ……………………………………………………………………
Birger Wernerfelt, Chair, Ph.D. Committee
Sloan School of Management

Essays on transition challenges for alternative propulsion vehicles and transportation systems

by
Jeroen J.R. Struben

Submitted to the Alfred P. Sloan School of Management
On August 22, 2006 in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Management

Abstract

Technology transitions require the formation of a self-sustaining market through alignment of consumers' interests, producers' capabilities, infrastructure development, and regulations. In this research I develop a broad behavioral dynamic model of the prospective transition to alternative fuel vehicles.

In Essay one I focus on the premise that automobile purchase decisions are strongly shaped by cultural norms, personal experience, and social interactions. To capture these factors, I examine important social processes conditioning alternative vehicle diffusion, including the generation of consumer awareness through feedback from driving experience, word of mouth and marketing. Through analysis of a simulation model I demonstrate the existence of a critical threshold for the sustained adoption of alternative technologies, and show how the threshold depends on behavioral, economic and physical system parameters. Word-of-mouth from those not driving an alternative vehicle is important in stimulating diffusion. Further, I show that marketing and subsidies for alternatives must remain in place for long periods for diffusion to become self-sustaining. Results are supported with an analysis of the transition to the horseless carriage at the turn of the 19th century.

In the second Essay I explore the co-evolutionary interdependence between alternative fuel vehicle demand and the requisite refueling infrastructure. The analysis is based on a dynamic behavioral model with an explicit spatial structure. I find, first, a bi-stable, low demand equilibrium with urban adoption clusters. Further, the diffusion of more fuel efficient vehicles, optimal for the long run, is less likely to succeed, illustrating the existence of trade-offs between the goals of the early stage transition, and those of the long-run equilibrium. Several other feedbacks that significantly influence dynamics including, supply and demand, and supply-coordination behaviors, are discussed.

In Essay three I examine how technology learning and spillovers impact technology trajectories of competing incumbents - hybrid and radical entrants. I develop a technology lifecycle model, with an emphasis on technology heterogeneity. In the model,

spillovers can flow to the market leader and can be asymmetric across technologies. I find that the existence of learning and spillover dynamics greatly increases path dependence. Interaction effects with other feedbacks, such as scale economies, are very strong. Further, superior radical technologies may fail, even when introduced simultaneously with inferior hybrid technologies.

Thesis Supervisors:

Charles H. Fine
Professor of Management

Ernest J. Moniz
Professor of Physics and Engineering Systems

John D. Sterman (Chair)
Professor of Management

Acknowledgments

Where to begin? While the PhD thesis might appear to be the result of one individual, it is not. Many people were integral to this process as coach, friend, family, teacher, colleague, or fellow student. Some fulfilled several roles. Many were behind the scene during the years leading up to this. It is therefore impossible to render a sufficiently accurate reconstruction of all the support I received, in order to be able to tell who to blame for what.

Uncontested is the enormous inspiration, coaching, humor, critical thinking, and perfectionism that John Sterman unloaded upon me during the past years. The combination of space for exploration and direct critiquing served me well. My other two committee members, Charlie Fine and Ernie Moniz, helped me approach this research from different angles. Putting these three grey-haired men together at one table resulted in a sparkling discussion at the defense. I learned a lot, and look forward to continuing to work with each.

The local System Dynamics and Sloan community provide a good example of the value of intertwining academic and personal relationships. I think back with particular pleasure to the many thought experiments Hazhir and I undertook and analyzed together, and to our summer seminar in 2003. I enjoyed my dinner discussions with Gokhan and his straight and clear thinking. I also have excellent memories of working with Anjali, and although less frequently, with Nelson. I shared many good experiences with Jim Hines (yes, Jim I am also cool now), Kim, Rad, Joe, John, Tim, Rafel, the Albany folks (in particular George and David), Brad, Sally, Paolo, and Etienne.

These years, pap en mam have always been much closer to me than the geographical distance suggests. Vincka, Olivier, Camilla, and the wider family too! One of the highlights of these years were Opa Api's visits. Api also reminds me of the progress of time, and technology. In the 1940s my Oma typed Api's PhD thesis. In the 1980s Peter Senge's wife, Diane, drew the graphics for many SD theses. In the 21st century my partner edited my thesis.

This brings me to Ruthie, who filled all the roles mentioned, and many more. Now that I am out of the zone, we can continue tasting our small, wonderful chunks of life, with a special one underway.

And then there are all these special Giants on whose shoulders I stand:

Christine Liberty; Daniel (for all the help in the library, too bad the system caught up); Joop Kuyper for his valuable and inspirational support; MIT; Yuli Nazarov, Kung-Ann Chao and Gerrit for generating early academic interest; a fantastic community of friends from Delft (their loyalty represented at the defense by Eric-Jaap); the unfortunate Robert; Eric Dolphy; the infinite corridor; drivers in Boston and Cambridge who mostly did not

hit me while I cycled through these mean streets (some did); Jan and Annetine; Florian for relaxation; Sol LeWitt (especially his #869C); Ivan, Yuri, and Michael (my Russian chess, boxing, and brandy partners); Andrea and Arretti; Cory Welch; Derek Supple; Peet's, Wally's, Emma's,…

This was, and is, most fulfilling. Thank you!

Table of Contents

Essay 1 - Transition challenges for alternative fuel vehicles: Consumer acceptance and sustained adoption

Essay 2 - Identifying challenges for sustained adoption of alternative fuel vehicles and infrastructure

Essay 3 - *Alternative fuel vehicles turning the corner?*: A product lifecycle model with heterogeneous technologies

Essay 1

Transition challenges for alternative fuel vehicles: Consumer acceptance and sustained adoption

Abstract

Automobile firms are now developing alternatives to internal combustion engines (ICE), including hydrogen fuel cells and ICE-electric hybrids. Adoption and diffusion dynamics for alternative fuel vehicles (AFVs) are more complex than those typical of most new products due to the size and importance of the automobile industry, the size and impact of the vehicle fleet, and the presence of various forms of increasing returns to scale. This paper describes a model examining the diffusion dynamics for and competition among AFVs, focusing on the generation of consumer awareness of alternative propulsion technologies through feedback from driving experience, word-of-mouth, and marketing. Through detailed model analysis the existence of a critical threshold for sustained adoption of new alternative technologies is shown. Word-of-mouth from those not driving an alternative vehicle is identified as important in stimulating adoption. The reduced form treatment of network effects and other positive feedbacks are analyzed. The model is discussed in light of the transition to the horseless carriage at the turn of the 19th century. As with 19th century vehicles, the combination of scale effects and familiarity are the key mechanisms for adoption and stagnation and they pose serious challenges for the diffusion of AFVs.

Introduction

In the 1860s the first self-propelling steam vehicles in the United States were banned

from the turnpikes, because of their reckless speed, noise, and explosions. Twenty-five

years later, New York, Boston, and Philadelphia were among cities that provided a warm

welcome to the clean, silent electric "horseless carriages" as alternatives to the polluting

horse-drawn carriage (Kirsch 1996). There was great enthusiasm among inventors,

including Thomas Edison, for the potential of electric vehicles (EVs). In 1899, an EV set

the world speed record of 61 mph (Flink 1988) and during that time the professional elite debated the relative efficacy of the various platforms, including steamers, EVs and internal combustion engines (ICE) (Schiffer et al. 1994). However, eventually the installed base of ICE developed most rapidly and shaped the standard of driving, becoming the dominant design.

Today, motivated by environmental pressures and increasing constraints on energy resources, we face another potential transition - away from fossil-powered ICE vehicles. Various alternatives, compressed natural gas vehicles (CNG), hydrogen fuel cell vehicles (HFCVs), or hybrids, are expected to compete with each other and with ICE. Market formation for AFVs harbors many uncertainties and the successes of past introductions are limited. In the United States and elsewhere, diesel and CNG vehicles have failed to create self-sustaining markets despite initial subsidies. Many other AFVs, such as EVs, have failed to take of at all despite repeated attempts in many countries (Callon 1986; Schiffer et al. 1994; Cowan and Hulten 1996; Kirsch 2000; Mom and Kirsch 2001).

The failure of new technologies to take off, despite an anticipated potential, is often attributed to the existence of increasing returns to scale. Arthur (1989), David (1985), and Katz and Shapiro (1985, 1986) developed and analyzed arguments about lock-in through increasing returns. Whether the take-off mechanism involves economies of scale, scope, or R&D, complementarities, or network externalities, an increase in adoption raises the installed base, and subsequently improves the attractiveness of the technology. As this technology gains an even larger market share, it wins further opportunities to improve its

performance, further increasing its market share. Such mechanisms are indeed central to industry dynamics generally and the automobile industry specifically as it is subject to many such positive feedbacks such as learning-by-doing in production and maintenance, scale- and scope economies, and complementarities, especially fueling infrastructure.

There are competing ideas about market share capture. Many new technologies do break through despite such entrance barriers. Scale effects eventually exhibit diminishing returns, and once an entrant technology does get traction, it can be expected to catch up with the established technology. Christensen (1997) describes this mechanism: disruptive technologies often emerge in a neighboring market and compete on dimensions of merit that were previously unavailable. As the experience of the entrant grows, its' superior performance on the new attributes allows the entrant to outplay the incumbent.

While important, explanations that focus primarily on objective technology efficacy do not provide full explanations of the patterns. Further, there is much variation in take off of identical technologies in different contexts: diesel vehicles have taken off since the 1980s in several European countries, but failed to do so in the United States. Similarly, CNG vehicles gained traction in Argentina, but sizzled and then fizzled in New Zealand and in Canada, and stagnated at low levels in Europe and the United States. These examples demonstrate that otherwise technologically promising and economically viable products face strong resistance and sometimes fail to take-off at all.

A new technology's efficacy is often ambiguous and debates about them are often incommensurable with each other and shaped by series of social events (Bijker et al. 1987). Part of a wave of non-traditional car designs in the 1930s, the super efficient Dymaxion car, oddly shaped into a tear-drop, received large attention and stirred public debates. However, soon after its concept car was involved in fatal incident during a test-drive at the 1933 Chicago World fair the public opinion tipped against it, despite that the accident was unrelated to the car's design (Kimes and Clark 1996). The role of public understanding in the failed or delayed diffusion is acknowledged to be important for non-automotive examples such as rigid airships (Botting 2002), nuclear energy (Gamson 2001), and renewable energy (Krohn 1999).

Consumers need to learn about the existence, availability, and relevance of a new technology. Automobile purchase decisions are not the result of "cold" economic calculation. Cars are an important symbol in society and a source of personal identity, status, and emotional resonance (Urry 2004). Efficacy and safety of designs and their features are shaped by historic events, experience, and social interactions (Miller 2001). New technologies need to become accepted as a viable alternative, yet in the early stages it is unclear what a new technology may bring. Many technologies, in particular automobiles, are complex and are evaluated along many dimensions of merit. Besides price, new platforms need to establish themselves on attributes such as safety, performance, reliability, and comfort. Thus, awareness of a new technology is not sufficient. Considering a new technology to be a viable alternative, in comparison with the more familiar and trusted alternatives, will require more knowledge and exposure. In

the case of automobile platforms, such considerations will have to build up under stringent competition for attention from various alternatives.

The existence of consumer uncertainty in the early developmental stages of a technology is not new. Abernathy and Utterback (1978) identified the role of uncertainty in the early stages of a product life-cycle. Word-of-mouth is the basic mechanism of spread of information about a product in the diffusion literature (Rogers 1962; Bass 1969). However, while suggestive (Gladwell 2000), the importance of the process of consumer acceptance and learning in the success or failure of a new technology is suggestive, it has received only limited attention.

This paper focuses on adoption generated by consumer awareness and learning through feedback from driving experience, word-of-mouth, and marketing. I model the process of consumer acceptance of a new technology and its role as a critical factor for the successful diffusion and sustained adoption of AFVs. In the prospective transition to AFVs, the social diffusion interacts strongly with other scale effects, creating further barriers to entry. I examine the various adoption patterns that can emerge in the context of competing vehicle platforms. Further, I argue, supported by analysis of the formal model, that the effects that are considered important for transition dynamics, such as learning-by-doing and complementarities, should not be studied in isolation but in interaction with other take-off mechanisms.

Understanding these mechanisms is critical for successful transitions towards an AFV-based transportation system. In other the other two essays I treat the role of learning, research, spillovers, infrastructure, and supply-demand dynamics more explicitly and in detail. Here I will illustrate the importance of interaction effects of the various mechanisms in a more aggregate way, which provides focus on the general patterns.

In what follows I first discuss the model's scope in relation to the existing literature on consumer choice and social exposure. Next, I discuss the core model, followed by detailed analysis of its dynamics. The subsequent section expands the model to include richer dynamics that result from the interaction with scale effects and consumer learning. This is followed by a discussion of this larger model in light of the 19th century transition to ICE. Based on the understanding that is achieved here, I discuss the efficacy of potential policies to stimulate successful diffusion. I end with a conclusion and discussion of the implications.

Modeling consumer choice

Conceptual and formal models of the product life cycle are useful starting points for considering the possible transition to alternative vehicles. Abernathy and Utterback (1978) emphasized the role of uncertainty in consumer choice and Klepper (1996) introduced a formal model that incorporates learning and scale economies. Arthur (1989)[1] examined factors such as learning and externalities that drive self-reinforcing

---

[1] Arthur's original manuscript was in circulation in 1983 but was not published until 1989.

mechanisms that have become part of mainstream organizational and industrial literature. For example, Katz & Shapiro (1985) examine the formation of standards and the role of expectations and Loch & Huberman (1999) discuss technological diffusion in the context of network externalities.

I draw on innovation diffusion models (Bass 1969; Norton and Bass 1987; Mahajan et al. 1990; Mahajan et al. 2000), and their applications in the auto industry (Urban et al. 1990; Urban et al. 1996). Models of innovation diffusion date from the 1950's (Griliches 1957; Rogers 1962; Bass 1969). The basic Bass (1969) model, with endogenous word-of-mouth from adopters, has been extended by to include other effects: marketing and media attention (Mahajan et al. 2000); uncertainty about the value of the innovation (Kalish 1985), substitution among successive technology generations (Norton and Bass 1987); repurchases (Sterman 2000). All these models yield an S-shaped growth curve for the introduced product and are widely used, as many new products follow that basic pattern. However, the models do not account for other patterns of diffusion, including "rise and demise," stagnation at low penetration levels, or fluctuations. One exception is Homer (1987), who develops a diffusion model with endogenous technology, learning-by-doing, and adoption applied to medical innovations. Homer shows that the model can explain a wide range of diffusion patterns: the classic success (S-shaped); boom and bust; boom, bust, and recovery.

This paper builds on the family of Bass-diffusion models with significant modifications. These traditional diffusion models confound exposure, familiarity, and the purchase

decision. However, vehicles are complex and involve many experience attributes (Nelson 1974), such as vehicle power, reliability, operating cost, and fuel economy, that can only be determined after purchase or usage, or extensive exposure. Thus familiarization requires sufficient exposure. This exposure must continue over long periods of time as vehicles are semi-durable goods and individuals take on average about a decade between two purchase decisions. As a result, a much more careful delineation of the social exposure mechanisms may be needed to capture critical dynamics. This supposition leads me to include six concepts in the model, most of which have been discussed in various separate bodies of literature, but not brought together, which is critical for exploration of transition dynamics.

The first concept involves decoupling of the process of familiarization from adoption and replacement decisions. This is important, especially for novel semi-durable goods but rarely done (for a notable exception, though for different purpose, see Kalish 1985). Second, I explicitly capture the different channels through which consumers bring an alternative into their consideration set, a concept discussed by Hauser et al. (1993). Third, because of the competition for attention, in the absence of any subsequent purchase, consideration of a new alternative is gained only slowly, and can degrade or be forgotten (Dodson and Muller 1978). This is especially important because of complexity in weighing the different attributes, their ambiguous role in the functioning of the car, and the emotion and social pressure involved in purchasing a vehicle. Fourth, regarding the individual attributes, learning about their relevance and performance is a lengthy process that requires confirmation from various sources (Gladwell 2000). Fifth, I explore multi-technology competition. To do so, I integrate the traditional diffusion concept with

discrete consumer choice models (McFadden 1978; Ben-Akiva and Lerman 1985) that are often applied to transport mode choice (Domencich et al. 1975; Small et al. 2005), and automobile purchases (Berry et al. 2004; Train and Winston 2005), including alternative vehicle choice (Brownstone et al. 2000; Greene 2001).

In the analysis, I characterize the global dynamics and parameter space of the model rather than estimation of parameters for particular AFVs, since i) these are highly uncertain, and ii) identifying which parameters are sensitive guides subsequent efforts to elaborate the model and gather needed data.

The model

We begin with the fleet and consumers' choice among vehicle platforms. The total number of vehicles for each platform $j = \{1,...,n\}$, V_j, accumulates new vehicle sales, s_j, less discards, d_j:

$$\frac{dV_j}{dt} = s_j - d_j \tag{1}$$

Sales consist of initial and replacement purchases. Discards are age-dependent. Initial purchases dominated sales near the beginning of the auto industry, and do so today in China, but in developed economies replacements dominate. Appendix 2a treats age-dependent discards and appendix 2b treats initial purchases; for simplicity, I assume the fleet is in equilibrium and focus here on replacement purchases:

$$s_j = \sum_i \sigma_{ij} d_i \tag{2}$$

17

where σ_{ij} is the share of drivers of platform i replacing their vehicle with platform j. The share switching from i to j depends on the expected utility of platform j as judged by the driver of vehicle i, u^e_{ij}. Because driver experience with and perceptions about the characteristics of each platform may differ, the expected utility of, for example, the same fuel cell vehicle may differ among those currently driving an ICE, hybrid, or fuel cell vehicle, even if these individuals have identical preferences. Hence,

$$\sigma_{ij} = \frac{u^e_{ij}}{\sum_j u^e_{ij}} \tag{3}$$

Expected utility depends on two factors. First, while drivers may be generally aware that a platform, such as hybrids, exists, they must be sufficiently familiar with that platform for it to enter their consideration set. Second, for those platforms considered, expected utility depends on perceptions of various vehicle attributes. To capture the formation of a driver's consideration set we introduce the concept of familiarity among drivers of vehicle i with platform j, F_{ij}. "Familiarity" captures the cognitive and emotional processes through which drivers gain enough information about, understanding of, and emotional attachment to a platform for it to enter their consideration set. Everyone is familiar with ICE, so $F_{i,ICE} = 1$, while $F_{ij} = 0$ for those completely unfamiliar with platform j; such individuals do not even consider such a vehicle: $F_{ij} = 0$ implies $\sigma_{ij} = 0$. Hence

$$u^e_{ij} = F_{ij} * u_{ij} \tag{4}$$

where utility, u_{ij}, depends on vehicle attributes for platform j, as perceived by driver i.

For an aggregate population average familiarity varies over the interval [0, 1]. Familiarity increases in response to social exposure, and also decays over time:

$$\frac{dF_{ij}}{dt} = \eta_{ij}\left(1 - F_{ij}\right) - \phi_{ij}F_{ij}$$ (5)

where η_{ij} is the impact of total social exposure on the increase in familiarity, and ϕ_{ij} is the fractional loss of familiarity about platform j among drivers of platform i. The full formulation accounts for the transfer of familiarity associated with those drivers who switch platforms (see appendix 2c).

Total exposure to a platform arises from three components: (i) marketing, (ii) word-of-mouth contacts with drivers of that platform, and (iii) word of mouth about the platform among those not driving it, yielding:

$$\eta_{ij} = \alpha_j + c_{ijj}F_{jj}\left(V_j/N\right) + \sum_{k \neq j}c_{ijk}F_{kj}\left(V_k/N\right)$$ (6)

Here α_j is the effectiveness of marketing and promotion for platform j. The second term captures word of mouth about platform j - social exposure acquired by seeing them on the road, riding in them, talking to their owners. Such direct exposure depends on the fraction of the fleet consisting of platform j, V_j/N, and the frequency and effectiveness of contacts between drivers of platforms i and j, c_{ijj}. The third term captures word of mouth about platform j arising from those driving a different platform, $k \neq j$ – for example, an ICE driver learning about hydrogen vehicles from the driver of a hybrid.[2]

---

[2] Eq. 6 can be written more compactly as $\eta_{ij} = \alpha_j + \sum_k c_{ijk}F_{kj}\left(V_k/N\right)$; we use the form above to emphasize the two types of word of mouth (direct and indirect).

It takes effort and attention to remain up to date with new vehicle models and features. Hence familiarity erodes unless refreshed through social exposure. The loss of familiarity is highly nonlinear. When exposure is infrequent, familiarity decays rapidly: without marketing or an installed base, the electric vehicle, much discussed in the 1990s, has virtually disappeared from consideration. But once exposure is sufficiently intense, a technology is woven into the fabric of our lives and "automobile" implicitly connotes "internal combustion". Familiarity with ICE =1 and there is no decay of familiarity. Thus the fractional decay of familiarity is:

$$\phi_{ij} = \phi_0 f\left(\eta_{ij}\right); \qquad f(0) = 1, f(\infty) = 0, f'(\cdot) \le 0. \tag{7}$$

Familiarity decays fastest (up to the maximum rate ϕ_0) when total exposure to a platform, η_{ij}, is small. Greater exposure reduces the decay rate, until exposure is so frequent that decay ceases. I capture these characteristics with the logistic function

$$f\left(\eta_{ij}\right) = \frac{\exp\left(-4\varepsilon\left(\eta_{ij} - \eta^*\right)\right)}{1 + \exp\left(-4\varepsilon\left(\eta_{ij} - \eta^*\right)\right)} \tag{8}$$

where η^* is the reference rate of social exposure at which familiarity decays at half the normal rate, and ε is the slope of the decay rate at that point. Varying η^* and ε enables sensitivity testing over a wide range of assumptions about familiarity decay.

These channels of awareness generation create positive feedbacks that can boost familiarity and adoption of AFVs (**Figure 1**). First, a larger alternative fleet enhances familiarity as people see the vehicles on the roads and learn about them from their drivers. Greater familiarity, in turn, increases the fraction of people including AFVs in their consideration set and, if their utility is high enough, the share of purchases going to AFVs (the reinforcing Social Exposure loop R1a). Further, as the AFV fleet grows, people driving other platforms increasingly see and hear about them, and the more socially acceptable they become, suppressing familiarity decay (reinforcing loop R1b).

Second, familiarity with AFVs among those driving ICE vehicles increases through word of mouth contacts with other ICE drivers who have seen or heard about them, leading to still more word of mouth (reinforcing loops R2a and R2b). The impact of encounters among non-drivers is likely to be weaker than that of direct exposure to an AFV, so $c_{ijj} > c_{ijk}$, for $k \neq j$. However, the long life of vehicles means AFVs will constitute a small fraction of the fleet for years after their introduction. The majority of information conditioning familiarity with alternatives among potential adopters will arise from marketing, media reports, and word of mouth from those not driving AFVs. Word of mouth arising from interactions between adopters and potential adopters will become significant only after large numbers have already switched from ICE to alternatives.

This concludes the exposition of the core model. The formulation differs from those of the standard Bass models through the decoupling of exposure, familiarity and the adoption decision, the word of mouth through non-users and the discrete choice replacement, for durable goods. Appendix 3a describes how we can recover the Bass model, under special conditions, and interpret its parameters in terms of those used in familiarity model. However, the dynamics are expected to differ considerably from the deterministic S-shape. We will now analyze its fundamental dynamics.

Analysis of the principal dynamics

For analytic purposes I will hold driver population and vehicles per household, and thus total installed base, constant. This assumption simplifies the potential dynamics and

analysis and draws attention away from the specific, uncertain parameters and numerical outcomes, towards the different *patterns* of behavior that are generated and their causes. Further, much of the critical dynamics will occur early on, say, in the first two decades after introduction, during which the growth pattern will not have a significant impact on the dynamics.

A first-order model: familiarity

The model generalizes to any number of vehicle platforms and constitutes a large system of coupled differential equations. To gain insight into the diffusion of alternative vehicles, we analyze a simplified version with only two platforms, ICE (j=1) and an AFV (j=2). We assume constant driver population and vehicles per driver, so the total fleet, $N = \sum_j V_j$, is constant. Familiarity with ICE can reasonably be assumed to remain constant at 1 throughout the time horizon. Further, AFV drivers are assumed to be fully familiar with their own AFVs. Thus

$$\mathbf{F} = \begin{vmatrix} 1 & F_{12} \\ 1 & 1 \end{vmatrix},$$

(9)

significantly reducing the dimensionality of the model.

Long vehicle life means the composition of the fleet will remain roughly fixed in the first years after alternatives are introduced. Assuming the fleet of each platform is fixed reduces the model to a first-order system where the change in familiarity with AFVs among ICE drivers, dF_{12}/dt, is determined only by the level of familiarity itself and constant effects of marketing and social exposure to the small alternative fleet.

Figure 2 shows the phase plot governing familiarity for a situation with a strong marketing program for AFVs and a modest initial fleet (Table 1 lists model parameters). When familiarity with the alternative is low, word of mouth from non-drivers is negligible, and the gain in familiarity comes only from marketing and exposure to the few AFVs on the road. Since the total volume of exposure is small, the decay time constant for familiarity is near its maximum. As familiarity increases, word of mouth about AFVs among ICE drivers becomes more important, and increasing total exposure reduces familiarity loss.

The system has three fixed points. There are stable equilibria near $F=1$, where familiarity decay is small, and near $F=0$, where word of mouth from non-drivers is small and familiarity decay counters the impact of marketing and exposure to the small alternative fleet. In between lies an unstable fixed point where the system dynamics are dominated by the positive feedbacks R2a and R2b. The system is characterized by a threshold, or tipping point. For adoption to become self-sustaining, familiarity must rise above the threshold; otherwise, it (and thus consumer choice) will tend toward the low-consideration equilibrium. The existence and location of the tipping point depends on parameters. Sensitivity analysis shows the low-familiarity equilibrium increases, and the tipping point falls, as i) the magnitude of marketing programs for AFVs, α_2, rises; (ii) the impact of word of mouth about AFVs between AFV and ICE drivers, c_{122}, increases; iii) the size of the initial alternative fleet grows; iv) the impact of word of mouth about AFVs within the population of ICE drivers, c_{121}, increases; and v) as familiarity is more durable (smaller ϕ_0 and η^* and larger ε). Continuing these parameter changes causes the unstable

fixed point to merge with the lower stable equilibrium, and then to a system with a single stable equilibrium at high familiarity.

A second-order model: familiarity versus adoption

We now relax the assumption that the share of alternative vehicles is fixed, adding the social exposure loops R1a and R1b. We simplify the dynamics of fleet turnover by aggregating each fleet into a single cohort with constant average vehicle life $\lambda_j = \lambda$, yielding

$$d_j = V_j / \lambda \tag{10}$$

Since $V_2 = N - V_1$, fleet dynamics are completely characterized by the evolution of the alternative, which, from eq. 1 and 2, is

$$\frac{dV_2}{dt} = \left(\sigma_{22} V_2 + \sigma_{12} (N - V_2) \right) / \lambda - V_2 / \lambda. \tag{11}$$

By eq. 3 and 4, the fraction of drivers purchasing an AFV is

$$\sigma_{i2} = F_{i2} u_{i2} / \left(F_{i1} u_{i1} + F_{i2} u_{i2} \right) \tag{12}$$

As before, we assume AFV drivers are fully familiar with their AFVs, and that everyone is familiar with ICE. Assuming for now that the perceived utilities u_{ij} are also constant, σ_{22} is constant at $u_{22}/(u_{22} + u_{21})$ and

$$\sigma_{12} = F_{i2} u_{i2} / \left(1 \cdot u_{11} + F_{12} u_{12} \right). \tag{13}$$

With the equation governing familiarity, the system reduces to a pair of coupled differential equations with state variables V_2 (the AFV fleet) and F_{12} (the familiarity of ICE drivers with AFVs).

Figure 3 shows the phase space of the system for several parameter sets. In all cases, the utilities of the two platforms are assumed to be equal so that AFV purchase share is 0.5 at full familiarity. Table 1 shows other parameters. The nullclines (dashed lines) are the locus of points for which the rate of change in each state variable is zero. Fixed points exist where nullclines intersect. With moderate marketing and no non-driver word of mouth (**Figure 3**a) there are three fixed points, as in the one-dimensional case, and the state space is divided into two basins of attraction. For small initial alternative fleets, familiarity and the fleet decay to low levels, even if initial familiarity is high. On the other side of the separatrix dividing the basins, familiarity rises and more ICE drivers switch to AFVs, further increasing familiarity and triggering still more switching. **Figure 3**b shows a case with no marketing but moderate non-driver word of mouth. As in the one-dimensional case, indirect word of mouth among ICE drivers shrinks the basin of attraction for the low adoption equilibrium. In **Figure 3**c marketing and non-driver word of mouth are large enough that there is only one fixed point, with high familiarity and diffusion.

In **Figure 3** marketing impact is constant. In reality, marketing is endogenous. Successful diffusion boosts revenues, enabling marketing to expand, while low sales limit resources for promotion. Declining marketing effort lowers α_2, moving the low-diffusion equilibrium toward the origin and enlarging its basin of attraction. Figure 4 illustrates a set of simulations beginning with no familiarity or installed base for the alternative. An

aggressive marketing campaign (α_2=0.02) begins at t=0. The campaign ends after T years, $10 \leq T \leq 50$ years. When the campaign is "short" -- only about a decade -- familiarity and market share drop back despite initial success: the campaign does not move the system across the basin boundary. With the assumed parameters, aggressive promotion must be maintained for roughly 30 years before diffusion becomes self-sustaining.

The long time required to move the system across the basin boundary to where adoption is self-sustaining depends, of course, on parameters. However, the key time delays, particularly the long lives of vehicles, are not in doubt. Vehicle lifetime affects the slow transition dynamics in two ways: first, the long lifetime constrains the physical diffusion speed. Second, the long replacement times require a much larger utility or total market volume to overcome the forgetting dynamics at low exposure (that is to overcome dominance of loop B1 in Figure 1). This would suggest that more durable products and systems (cars, energy deployment) are especially affected by such dynamics. This distinction does not come out in the classic diffusion models, or the standard demand models. Collapses after initial take-offs have been observed. For example, attempts to introduce CNG vehicles faltered in Canada, and in New Zealand after initial subsidies expired, despite some initial diffusion; and stagnated at low penetration in Italy, even with continued subsidies (Cowan and Hulten 1996; Di Pascoli et al. 2001; Sperling and Cannon 2004; Energy Information Administration 2005). This concludes the discussion of the fundamental social exposure dynamics. We will now analyze the implications of interactions with other increasing returns to scale mechanisms and the role of competition.

Social exposure, endogenous performance and competition

The analysis above illustrates the fundamental dynamics generated by this structure. In real life, however, technologies are engaged in a dynamic competition with each other, with the performance of each technology improving over time. In addition, consumers have to learn about each technology's efficacy over time. Here we want to illustrate the relevance of consumer learning in such settings. To do this, we expand the model to include endogenous technology improvement through learning-by-doing and consumer learning about performance and focus our analysis on the competitive interaction between alternatives.

Figure 5 illustrates how the dynamics are now shaped by additional feedback loops that have highly nonlinear characteristics: increased sales allow improving the technology, which increases attractiveness and market share. New entrants can close the gap, especially because such feedbacks, which exhibit diminishing returns, are very strong for those with little prior experience (R3). However, early during the diffusion, new entrants receive limited effective exposure, suppressing market shares, as discussed above, which reduces learning-by-doing (R1 and R2 interacting with R3). Consumers will learn fast about the actual performance of established technologies (B2), but for novel technologies consumers get much less opportunities to learn about the state of the art or its potential (R4), which suppresses learning and perceived utility, further limiting sales and exposure. Such interactions between actual improvement, the consumers' perception of the performance, and their consideration for purchase of a technology were critical during the

failed introduction of diesels in the US during the late 70s. The early models had weak performance, which resulted in large-scale rejection, making them unable to occupy a critical niche that would allow them to improve. Further, despite substantial improvements of diesel technology over the years, and its large-scale acceptance in Europe, in the US sales and perception remain poor (Moore et al. 1998). Finally, the competitive dynamics between various AFVs (indicated in **Figure 5** with the layered stocks) limit their market share, and thus production volumes and hence exposure and learning as well.

Model expansion: endogenous performance and consumer perception

Utility takes the reference value u^* when expected performance P_{ijl}^e equals a reference value P_l^*:

$$u_{ij} = u^* \exp\left(\beta_l P_{ijl}^e / P_l^*\right) \tag{14}$$

where β is the sensitivity of utility to performance. The exponential utility function means the share of purchases going to each platform follows the standard logit choice model. A person's assessment of a platform's attractiveness depends on her perceptions of the vehicle characteristics, including purchase cost, fuel efficiency, power, features, and range, here aggregated into a single attribute denoted "vehicle performance."

The actual states of the attributes are not observed directly, but learned over time. For instance, Urban et al. (1990) find that users gradually update their assessment of the attractiveness of the latest model of a platform through social interactions with other people (drivers and non-drivers) and exposure to marketing and media. It took years

28

before the consumer perception of diesels caught up with the reality of improved

technology. We model this by allowing perceived performance of attribute l for platform

j by a driver of platform i, P_{ijl}^e, to be updated through channels similar to those affecting

familiarity: i) marketing, through which drivers learn about the performance of the new

technology; ii) P_{jl}^n; drivers of platform j, through which they learn about their experience

with the platform P_{jl}^e; and, iii) non-drivers of platform j, through which they learn about

their perceived performance P_{il}^e. On top of that, drivers of j experience the current

performance P_{jl}^f themselves, at an experience adjustment rate ε_{jl}. Hence,

$$\frac{dP_{ijl}^e}{dt} = \alpha_{jl}\left(P_{jl}^n - P_{ijl}^e\right) + c_{ijjl}\left(P_{jl}^e - P_{il}^e\right)v_j + \sum_{k \neq j} c_{ijkl}\left(P_{kl}^e - P_{il}^e\right)F_{jk}v_k + \varepsilon_{jl}\left(P_{jl}^f - P_{jl}^e\right)\delta_{ij} \quad (15)$$

Effectiveness of contacts c_{ijkl} and marketing depend on the attributes. For instance, more

complex products will require more time to be fully comprehended. Such parameters

cannot be observed but can be approximated through calibration. When a person switches

from one platform to another, she will take her perception of the state of attractiveness

with her (see appendix 1c for the treatment of co-flows).

I capture the technological improvement of platforms with standard learning curves. In

reality this happens differently for different attributes P_{jl}^n, but here we aggregate all into

one (see Appendix 3a of Essay 3 for a more detailed discussion of the individual

attributes). Performance of the new vehicles follows a standard learning curve, rising as

relevant experience with the platform, E, improves,

$$P_j^n = P_j^0 \left(E_j / E^0 \right)^\gamma \tag{16}$$

where performance equals an initial value P^0 at the reference experience level E^0.

We proxy a platform's experience and learning from all sources with cumulative sales:

$$\frac{dE_j}{dt} = s_j \tag{17}$$

Finally, the performance of the installed base, P_j^f, improves gradually upon selection by consumers. This is modeled through a co-flow structure, with sales and discards of vehicles (see appendix 1c for the treatment of co-flows).

This concludes the exposition of the learning-by-doing and learning of performance. We now examine how the social diffusion processes interact dynamically with learning-by-doing, consumer learning, and platform competition.

Analysis of performance and competition dynamics

In this section we will examine the dynamics of the expanded model through simulation of the introduction of entrants in an environment with a mature technology that has full market share. To simplify, we can assume that consumers have a constant perceived utility of not adopting an entrant, unless otherwise stated, $u^o = 0.5$. Up to two entrants are introduced, with equivalent technology potential $\left(P_1^0 = P_2^0 = 1 \right)$. This means that, in the absence of learning, there exists a theoretical equilibrium in which both entrants reach

40% of the market. [3] Other parameters are specified in or discussed for each graph

separately. The simulation over time, shown in Figure 6, provides an illustrative example

of the typical dynamics involved. In this simulation, the learning curve strength has a

value that is typical for production technologies $(\gamma = 0.3)$ (Argote and Epple 1990;

Zangwill and Kantor 1998); performance of the technology is directly observed by all

consumers and valued equally $\left(P_{ij}^e = P_j \; \forall i, j \right)$. Time zero is defined as the introduction of

entrant 2, ten years after the first entrant. After introduction, marketing effectiveness for

each is held constant at a base level $\left(\alpha_j^0 = 0.01 \right)$, except for the first 10 years after launch,

during which they receive an additional exposure boost from an introductory marketing

program α_j^p. Such a marketing boost can be seen as an initial period of "free" media

attention and public interest, because of its novelty. Its effectiveness will depend on many

social factors, among others, such as how usage of the technology differs from the

existing habits, its complexity, and how it fits with the existing norms. In this simulation

the effectiveness of the introductory campaign for entrant 2 is slightly higher than for

entrant 1 $\left(\alpha_1^p = 0.04; \alpha_2^p = 0.035 \right)$, yielding the dotted lines for the total market

effectiveness for each. We see that consumers' familiarity with platform 1 grows upon

introduction, together with the installed base. However, familiarity drops after its

intensive marketing program ends, resulting in stagnation of the installed base. At the

same time, the share of the second entrant starts to grow in combination with potential

---

[3] The entrants' attractiveness, used for the MNL equals in this case $e^{\beta(Po-1)}=1$, thus its share will be
1/(1+1+0.5)=50%. Note further that the attractiveness of the mature technology equal to 0.5 corresponds
with a performance of the mature technology equal to $P^o=1-ln(u^o)=0.37$.

consumers' familiarity. However, when the marketing program of entrant 2 ends, its installed base is still below that of entrant 1. Despite this lag and the fact that the 2nd entrant certainly has less experience and its technology is less attractive than entrant 1's, entrant 2 can gain a larger market share. Even though its program is only marginally more effective than that of the first entrant, it allows building critical exposure from early adopters of the platform and those who do not drive it, which further increases the consumers' consideration of the platform. Once it expands, the attractiveness of the platform grows, even compared to the first entrant.

As is shown by the dashed line, in the absence of a 2nd entrant, entrant 1 would have taken off, despite early stagnation. While familiarity will also decline in this case after the program, absent any serious competition, exposure from the existing installed base and familiarity are sufficient to overcome this. However, under more sever and increasing competitive pressure, the installed base does not grow sufficiently to overcome the decay in familiarity. The attractiveness of a second growing alternative allows less and less opportunity for entrant 1 to build its own products. These mechanisms allow a lagging entrant to overtake an earlier, equivalent technology. In the next simulations we explore in more detail under what conditions the late entrant will take over, and what other equilibria can result.

 I now explore the role of positive attention, learning, and technological performance on the competitive diffusion dynamics. I start by examining how the simulation result of **Figure 6** changes, when the strength of the initial marketing program and the learning

curve parameter are varied for each. As before, we first set the technology potential of the

two entrants identical to each other. We then examine the impact of the learning-curve

strength in isolation. Figure 7a shows, under constant and full familiarity, the effect of

learning-curve strength, $\gamma \in [0.0.5]$, and the effect of a head start by entrant 1,

represented by varying non-zero installed base at $t=0$ for the first entrant 1, $\upsilon_1^0 = [0,0.2]$.

We see that the equilibrium share for the lagging entrant 2 is barely affected by a change

in the learning-curve strength, or the lead of entrant 1 is increased. Generally, both

entrants 1 and 2 capture a little over 40% of the market. Only the situation of an

extremely large learning-curve strength and introduction lag for entrant 2 will have long-

term impact on entrant 2's equilibrium market share.

Thus, under these conditions that we will maintain for the rest of this analysis, the

learning-curve dynamics and introduction lags by themselves exhibit only very weak

selection effects between technologies. Further, a variation of these parameters triggers

only a gradual response. Of course the response depends also on the sensitivity of

consumers to a change in performance. This value is represented in the model by β and

set to equal to the neutral value of 1 throughout (Table 1).[4]

We now continue with the set up of Figure 6. Figure 7b-d show the equilibrium installed

base share for entrant 2 when, as before, entrant 2 is introduced with a 10-year lag after

entrant 1. We analyze the impact on this equilibrium share of effectiveness of the

---

[4] The sensitivity being equal to 1 yields a demand elasticity of 0.5 and can be interpreted that at normal
performance and 50% market share in equilibrium, a 1% increase in performance will yield a 0.5% increase
in its equilibrium market share. The factor of 0.5 is the result of the market saturation effect.

marketing programs for both entrants during the first 10 years after introduction,

$\alpha_i^p \in [0, 0.06]$. Figure 7b shows the results in the absence of any learning-by-

doing $(\gamma = 0)$ with perceived and actual performance of both platforms still being equal.

Even in this highly technologically stable environment, three equilibria can be identified:

i) The two attain equal market shares. This occurs when both programs are very effective.

In this case, the late entrant can overcome the burden of limited exposure early on. ii) The

late entrant does not achieve a sustainable market share. This occurs typically when the

marketing effectiveness of the second entrant is constrained and cannot be considerably

larger than that of the first. In the case of very effective programs, the threshold is

independent of the program of the first entrant. iii) full tipping towards the late entrant.

This occurs when the program of entrant 2 is significantly stronger than a weak program

of a first entrant. Thus, dynamic paths depend critically on the early stages of the

platform competition. For instance, there are conditions where a strong program fails,

while a weaker one leads to a sustainable market share.

Figure 7c shows the results in the presence of learning-by-doing $(\gamma = 0.3)$. While the

tipping regions are maintained, it is now much harder for the late entrant to catch up

under moderate marketing of the second entrant. On the other hand, under these

conditions, the opportunity to out compete the first entrant is also larger, as in this case

earlier entrants also struggle to attain a reasonable market share (moderate market

effectiveness for entrant 1, large for entrant 2). Underlying each time are the same

mechanisms as discussed under Figure 6. Note further that the maximum market share is

now not constrained to be 65% of the market, as learning-by-doing, at a high production

rate allows for improvement well above the reference performance. In Figure 7d we relax the assumption that consumers observe the actual performance. Using parameters for consumer learning (equation 15), with a strength of 0.5 times of those used for social exposure (see equation 15) and an experience adjustment rate for actual drivers of 0.5, we see that the conditions under which the late entrant can catch up are greatly reduced.

In Appendix 3b I illustrate that the rich tipping dynamics observed here depart radically from the traditional lower order diffusion models.

The analysis thus far was done with a fixed attractiveness of an alternative, with $u^o = 0.5$ in each case. However, in some markets no viable alternative exists, which would correspond with a very small u^o; in others, the alternatives are introduced at par with the incumbent technologies. Diffusion dynamics will depend on such differences, but how is not clear. Figure 8 shows how the equilibrium entrant share of entrant 2 varies as a function of its marketing effectiveness during introduction (as before) and a function of the mature technology's attractiveness (u^o). When rescaled to $\sigma^{o*} = u^o / (u^o + 1)$, this last axis can be interpreted as the non-linear scaled market share that the mature technology would attain, when one entrant would enter the market. The introductory marketing effectiveness of the first entrant is held constant at moderate levels $\alpha_1^P = 0.03$. The thick line indicates correspondence with the output of **Figure 7**c, for a marketing effectiveness at 0.03. We see that, even given a very low attractiveness of the incumbent, an aggressive marketing program is needed to reach a sustainable market share since, in this case, the first entrant has also been able to establish itself. Because of that, the second

entrant can at best reach an equivalent share of the market to the first entrant. When attractiveness of the incumbent is high, a similar aggressive marketing campaign is needed, but now only one entrant will survive with its equilibrium market share gradually decreasing as the mature technology becomes more superior. In between, with moderate attractiveness of the incumbent, a sweet spot exists at which the market can be penetrated with more modest marketing efforts.

Early on the diffusion dynamics are critical, and highly path-dependent. Of course the dynamics are parameter-dependent and detailed calibration can provide insights concerning under what conditions the different dynamics are more likely. However, the type of the dynamics seem to correspond with observations, for instance about earlier AFV introductions. Detailed analysis of historic cases will provide insights to what extend the insights hold. To illustrate this, we now discuss the 19th century transition to the horseless carriage.

The transition to the horseless carriage

The early transition to the current ICE-dominated system in the late 19th century serves as a good illustration of the thesis of this essay. In 1900 there were about 18 million horses in the United States and 8000 registered vehicles for a population of 76 million. Twenty-five years later, 125 million Americans drove 26 million ICE vehicles and held just 11 million horses (US Census 1976). While such numbers correspond to an S-shaped diffusion pattern typical when a new, superior technology replaces an inferior one, a closer look reveals a dramatically different story. I discuss the role of consumer learning

and socialization regarding technological alternatives in the context of this transition towards ICE in two phases: the slow emergence of the automobile before 1895 with take-off around 1900; the failed revival of EVs in the 1910s.

The slow emergence of the automobile

While ICE vehicles did not take off before the end of the 19th century, early automobiles had already been introduced in the United States since the 1850s. Figure 9 shows the distribution of auto types between 1876 and 1942, measured by the share of firms producing each type of vehicle. In these early days, engineers, carpenters, and hobbyists devoted their time mainly to developing self-propelling steam and electric machines. ICE did not seriously enter the market before Carl Benz demonstrated the first operating ICE vehicle in 1885 (Westbrook 2001). Even around 1900, of the 4200 vehicles sold, 3200 were equally shared among EVs and steamers, with only about 1000 ICE vehicles (Geels 2005).

The pre-1890 steamers, developed by individual entrepreneurs, were tested and could be seen in small villages in various regions of the United States. The steamers delivered greater speed, lower operating costs, and less pollution than horses. As Robert Thurston, president of the American Society of Mechanical Engineers, suggested in his inaugural address in 1881, the triumph of the steam vehicles was imminent (McShane 1994). Further, many technological improvements, such as more efficient boilers, were available, but not widely implemented (Kimes and Johnson 1971). However, at this stage, the public that came to hear about these powered road-running vehicles feared that they

37

would destroy the traditional non-travel functions of urban streets: social activities (such as meeting places and provision of safety); walking; economic activities (through open air markets, bartering, and vendors with push carts or horse-drawn carriages) (Jacobs 1961, McShane 1994). Reflecting this, local regulators even went so far as to ban steamers because of "their speed, smoke, steam exhaust, and potential for explosion" (Schiffer et al. 1994). Its application as a device of travel, for which the railroad was perceived as sufficient, was not really considered. Hiram Percy Maxim, an early and respected experimenter with ICE vehicles, states that "*It has been the habit to give the gasoline engine all the credit for bringing the automobile, as we term the mechanical road vehicle today. In my opinion this is a wrong explanation. We have had the steam engine for over a century. We could have built steam vehicles in 1880, or indeed in 1870. But we did not. We waited until 1895*" and providing argument other transportation developments such as the bicycle "*had not directed men's mind to the possibilities of independent, long-distance travel over the ordinary highway*" (Jamison 1970).

Thus, while inherent performance was not an issue, in the early days there existed no organized groups of stakeholders that could mount campaigns to promote the vehicle as a viable alternative. On the contrary, led by threatened groups, much of the media attention and word-of-mouth stressed the negative side-effects of this new technology, and the resistance to it suppressed any diffusion and limited exposure and growth of familiarity with the automobile. There was no opportunity to gain the experience necessary for testing other applications, to improve the vehicle, or to increase the talk-of-the-town that could ignite a serious competition with the horse.

While in the early days the horseless carriage was often considered nothing more than a fad or toy, and horse traffic was initially protected by regulation against the "race-devils" (Beasly 1988), this negative attitude towards the vehicle changed considerably during the roaring 1890s (Westbrook 2001, Beasley 1988). Early experiments with electrics and steams had spilled over to new forms of public transportation. For example, between 1880 and 1895 a battery-trolley boom resulted in 850 new systems throughout the US, transporting over one hundred million passengers annually (Geels 2005). Electric trams carried passengers much faster and farther than horse-drawn trolleys (Schiffer et al. 1994). Gradually, even the middle class could afford living farther away from the workplace, setting off a trend of suburban life. Concomitantly, bicycles emerged as a form of personal, speedy transportation (Geels 2005). During this period the population gradually got accustomed to the idea of mechanized personal transportation.

A truly big breakthrough for the automobile was the 1895 Chicago *Times-Herald* contest. Much more than a contest, it was a show of comparison of the car with the horse: performance indicators that were considered critical in the contests included responsiveness, tractability, economy of maintenance, power, and docility, with speed being much less emphasized (Kimes and Johnson 1971). Besides its most practical and public scientific testing of the automobile, it brought the auto pioneers and enthusiasts together for the first time. While the general public had never learned the potential value of early steamers, the *Times-Herald* contest was considered a great success and proved an enormous stimulus for other events. Automobile periodicals started to make their debut,

as with the appearance of *The Motorcycle* and *Horseless Age* in 1895 (Flink 1998). This was also the year of the first US automobile advertising, placed in *The Motorcycle* by Carl Benz (McShane 1997). In 1896 William Jennings Bryan conducted his presidential campaign throughout Illinois in an automobile (Kimes and Johnson 1971). The automobile industry as a whole benefited enormously from the positive attention, and from the learning by consumers about applications and preferences. Around 1895 public interest was great and the number of entrepreneurs developing EVs, steamers, and ICE vehicles exploded, as illustrated by Figure 9. From the engineering side the main focus was on steamers and electrics, with their known technologies finding it easier to attract capital, while ICE developers were mainly individual entrepreneurs with more limited capital (McShane 1994).

Thus, this period is associated with gradual learning about the potential function of self-propelled vehicles by both potential adopters and developers. Although initially only the wealthy could afford a vehicle (Epstein 1928), publicity was enormous, building a potential for a much broader consumer base.

Besides the aggregate shaping of the idea of the automobile, this was also a period when platforms became identified with particular attributes. ICE vehicles were complex to operate and noisy, but good for long trips. Steamers were much faster and held the speed record until 1906, taken away from EVs in 1899. On the other hand, they required a longer start-up time, and more fuel (as well as water), and were likewise noisy. EVs, on the other hand, were simple and quiet but had heavy battery packs and could not bring the

40

tourist to remote areas. Their most dominant early applications were as taxicabs in the bigger cities. Debates about vehicles' current and potential advantages were often incommensurable because of the many considerations. Camps of partisans emerged and respected journals backed different technologies. For instance, *Scientific American*, which had first supported steamers, favored EVs, while the more mundane but influential journal *Horseless Age* was more favorable to ICE vehicles (Schiffer et a. 1994; Horseless Age 1895-1899). In those days, for the public, selecting one kind of car with confidence was far from easy.

While decisions based on actual performance were difficult, the media and regulators were becoming amenable to the automobile as a form of personal transportation. However, because steamers were still feared by the general public, haunted by their bad reputation for danger (drivers of steamers were still required to obtain a boiler license), interest in ICE vehicles and EVs grew. Even though steamer performance was superior, many entrepreneurs shifted away from steam to ICE and EVs (McShane 1994, Kimes and Clark 1996) and cities started to allow ICE vehicles.

These dynamics were reinforced by the technological improvements that started to take shape. With the take-off in demand, each platform began to introduce many new concepts. For example, steamers could now be seen with flash-boilers and improved hulls for boilers, making them more efficient. ICE, with its more complex technology benefited particularly from other industrialization developments in the US, such as the experience gain in exchangeable parts production (Flink 1970). ICE was unhindered by

constraints of public acceptance or complementarities, making it a favorite for investors. Growing production and driving experience and R&D led to spectacular advances in ICE vehicles and they soon outgrew the others. In addition, ICE received much more exposure, further strengthening its position as the standard choice. EVs and steam were unable to keep up. By 1905, 85% of automobile sales consisted of ICE vehicles.

The dynamics of the early transition to the horseless carriage illustrate the concepts analyzed in this essay including: consumer learning about existence and performance of attributes of multiple platforms; the process of familiarization through social exposure and experience; the complication of inter-platform competition and learning-by-doing. To illustrate the importance of all these factors, Figure 10 shows the dynamics resulting from different hypothesis about key drivers of the dynamics for the competition. For analytical clarity we focus on ICE and steam alone. Parameters are set equal to the base parameters (Table 1), unless otherwise stated. Attractiveness of the mature mode of transportation (horses) is considerably lower than steam and ICE: $u^o = 0.2$.[5] Further, I include a typical learning-curve-strength $(\gamma = 0.3)$, unless otherwise stated. Steam is introduced 10 years prior to ICE. Increased marketing effectiveness is active for both between year 0 and 10, representing the increased interest in the automobile. Further, to represent more favorable exposure to ICE, I set the marketing effectiveness during that time higher for ICE $\left(\alpha_{ICE}^{p} = 0.04; \alpha_{Steam}^{p} = 0.03\right)$.

---

[5] Which would lead to 80% adoption of Steam or ICE, if only one of them would break through, and in absence of learning.

I examine, a), the dynamics resulting from the full model structure introduced in this Essay, and compare this with: b) a reduced model that emphasizes the superiority of ICE, setting familiarity equal to 1 for both; and c) basic word-of-mouth diffusion. In this last scenario I confound exposure, familiarity and adoption into one instantaneous adoption variable, ignoring the accumulation of familiarity, as well as the role of non-drivers in social exposure. I focus on the qualitative patterns of behavior of each scenario. Figure 10a1 shows the results.

The qualitative reference mode is easily represented by the full model structure, and, explanations for the behavior match, as for instance discussed with Figure 6, with the observations around the early stages of the automobile transitions. Clearly, performance is not a necessary explanation: even with steam and ICE equivalent, ICE can take over. Further, we can hypothesize that steam was a viable candidate, in absence of ICE (Figure 10a2). Figure 10b, ICE superior, shows dynamics under full familiarity, and with ICE 4 times higher performance than steam. We see direct and fast adoption for steam, even when learning is included and for low steam performance. The first order adjustments result in very inert dynamics and do not offer room for a strong narrative about the transition. Finally we illustrate the dynamics for "instantaneous word-of-mouth", we set familiarity equal to its equilibrium value, ignoring the role of non-drivers $\left(c_{ijk} = 0 \ \forall j \neq k \right), F_{ij}^{B} \equiv \kappa^{B} \left[\alpha_{j} + c \left(V_{j} / N \right) \right] \left(N - V_{j} \right) \big/ \phi_{0}$, with κ^{B} a free parameter with which we can generate a best case for the word-of-mouth scenario (Appendix 3a recovers the Bass expression from this familiarity model). For the equivalent technology scenario (Figure 10c1), $\kappa^{B} = 0.15$, and the marketing effectiveness for entrant 2 is to be

increased to 0.2 to generate this best result. The key observation is that the tipping point are much less pronounced (but still there due to the competitive dynamics). Without strong feedbacks active around the tipping point, Steam does not die out.

The failed revival of EVs

The role of consumer learning in a dynamic environment in which platforms enter and technologies change over time is well illustrated by the short-lived revival of EVs during the 1910s. Before that, while expectations and potential were high, their performance improvements lagged ICE and driving range was short. These factors can easily be assumed to be the main reason for their demise. However, I argue that such learning and network externality dynamics are strongly conditioned by social processes around the adoption decision. I discuss how the process of acceptance or rejection plays an important role in opening, and closing, a window of opportunity for technology introduction and its diffusion.

Around 1900, along with the contests and touring enthusiasm, a few attributes emerged as dominant for the social group of young affluent men, who had the required purchasing power at that time: vehicle speed and capacity for long-range touring. Those attributes were especially well provided by ICE and steam. On the other hand, EVs became increasingly criticized for their lack of "active radius." More subtly, EVs' use for short trips in urban areas offered little incentive to develop recharging stations in remoter areas. The lack of recharging stations or standardization fed back to limit the appeal of the electric cars in those areas, slowing diffusion further. Finally, the heavy EVs could easily

get stuck on unpaved roads (at this time still comprising 99% of all roads) and thus provided little opportunity to gain experience in using batteries over longer ranges.

Steamers and ICE encountered fewer such problems, as their driving range was typically larger. More importantly, their fueling infrastructure was much less constrained, due to wide availability of fuel at retail outlets. This allowed for gradual growth in the sophistication of its fuel-distribution network, supported by a growing petroleum lobby (Kirsch 2000). Moreover, repairs for ICE and steamers in rural areas were much easier, as experience with engines (especially gas engines) was growing. Including such feedbacks further strengthens the tipping dynamics. This difference in application also provided those vehicles visibility almost everywhere, with media attention on spectacular and heroic long-distance tours, while EVs were mainly used for unexciting taxi rides in the city. After 1900 EVs were generally perceived as losing ground.

Nevertheless, EVs kept developing, and a race for battery improvement started between Thomas Edison and the Electric Vehicle Company, leading to a series of significant improvements by 1910 (Geels 2005, Kirsch 1996). Other advances around that time included infrastructure improvements, such as reliable boost charging, developed by Edison, and curbside recharging technologies (Schiffer et al. 1994). Finally, with the battery improvements, central stations realized that revenues could be made by providing off-peak charging for a larger installed base of EVs. They started to provide that and other services and built their own EV fleets, with leasing and rental services. The sudden spurt in advertisements for EVs in some magazines was no accident. Car makers,

managers of the larger central stations, and battery companies, perhaps convinced that the missing link was solved, agreed that EVs had a bright future if the public were educated, through advertising, about their many advantages. In those years much investment, research, and collaborative efforts went into the EV, including collaborations between Ford and Edison, and, as this discussion illustrates, resulted in organized efforts to revive the EV (Schiffer et al. 1994; Kirsch 2000; Westbrook 2001). While sales did indeed increase during the "golden age" of the EV, and picked up for professional fleets, public interest remained moderate, suggesting that while performance did indeed improve Americans still harbored prejudices against EV performance (Schiffer et al. 1994).

Throughout the 20th century, and even now, the EV has been repeatedly introduced into the market by confident entrepreneurs, OEMS, and engineers, (Callon (1986);Schiffer et al. (1994);Westbrook (2001);MacLean and Lave (2003); MacCready (2004)), but so far has never found a way to overcome the burden of history (Kirsch 2000). We saw that early in the transition ICE was able to surpass steamers, because the automobile was still novel and steamers faced a burden of negative association. However, once ICE took off, there remained only limited attention for organized efforts to reintroduce the EVs. The improved EVs and EV infrastructure arrived too late, as consumers and investors were already comfortable with ICE. AFVs today face an even more established system of users, producers, and suppliers, illustrating the importance of explicit consideration of social exposure affecting consumer perception and choice.

Exploring policy levers

To illustrate that capturing the richness of the various mechanisms allows for exploration of high-leverage policies, I discuss a series of proposed policies to stimulate sustained adoption of an entrant. I do this with the help of Figure 11, which shows a base run with a failed introduction of an entrant and various policies. In the base run one entrant is introduced into the market with an established technology that is equivalent in potential performance $\left(P_1^0 = P_2^0 = 1 \right)$. The learning-curve strength is set equal to a typical value $\left(\gamma = 0.3 \right)$ and, to simplify dynamics, we assume that performance of the technology is directly observed by the consumers $\left(P_{ij}^e = P_i \right)$. As before, marketing effectiveness is held constant at a base level $\left(\alpha_i^0 = 0.01 \right)$, except for the first 10 years after launch of the entrant, during which the additional market effectiveness is moderate $\left(\alpha_1^p = 0.02 \right)$.

A first policy to consider is extending the marketing program in duration (run 1). However, even an extension to a 30-year program does not lead to adoption. The feasibility of such an effort is highly questionable, even if 30 years would lead to take-off, given far shorter political and industrial decision cycles. A slightly different approach involves increasing the marketing effort at introduction. Simulation 2 shows the doubling of the effectiveness $\left(\alpha_1^p = 0.04 \right)$, the lowest value that leads to sustained adoption. While this can be successful, increasing effective exposure for a considerable time is very difficult and costly, probably more than linearly. A third consideration is the value proposition for consumers. The simulation succeeds here for $P_2^0 / P_1^0 = 3$, a significant

difference. This could be done, for instance by developing complementary applications

for, say, HFCVs. While eventually effective, it takes a long period before the higher

market share is realized. This is because little is done to improve familiarity, while the

established technology has comparable performance for the first 20 years. Including the

perception lags of consumers would pose even more of a challenge for this approach.

The fourth strategy is an increase of exposure effectiveness, which leads to a successful

diffusion for $c_{ij} = 0.6$, double the original value. However, the effect takes a while,

because of the fundamentally slow replacement dynamics. Increasing the influence of

non-drivers (simulation 5) has a much faster effect. Finally, reducing the replacement

time is explored. In the simulation, this is done by reducing τ^d for the incumbent by a

factor 3, for the first 15 years only. This could be realized through buy-back programs of

old vehicles and fleets, leasing, and targeted early adopters. As shown, this offers

enormous leverage, especially in the early years.

The policy analysis served to show that including the richer model, allows exploring and

testing policies and strategies that are out of scope otherwise. In particular, the last two

policy scenario's make use of the fundamental dynamics explored in this essay.

Discussion and conclusion

Understanding the dynamic challenges of a competition with an existing system is critical

for achieving self-sustaining alternative-fuel-based markets. Recent attempts to introduce

alternative fuel vehicles, such as CNG and diesel vehicles, have yielded mediocre results, and their general slow diffusion or complete failures illustrate the complexity of any such transition in modern transportation. This essay focused on one of the critical aspects of such a transition: the social processes that shape acceptance of and learning about the efficacy of new technologies. The importance of these processes is illustrated by the fifty years that stood between the introduction of the first steam automobiles in the US, and the actual take off of the automobile industry.

Especially in the early stages of a prospective transition, social exposure plays a dominant role in the success and failure of technologies. First, consumers -- and producers and other stakeholders -- need to become familiar with the alternatives, whose acceptance are constrained by established habits and socialized preferences. In addition, learning from experience a technology's attributes and performance requires significant exposure, which takes considerable time. Such mechanisms do not form a barrier when a new technology can be swiftly introduced in a new market, such as consumer electronics, or the movie industry. In fact, in such markets exposure mechanisms are often utilized to set high expectations. Setting high expectations can be used to drive up sales that help in achieving a critical mass for network externalities, as with i-Pods, or before potential feedback from consumer experience corrects the expectations, as with blockbuster movies. However, when a new technology is introduced into a market of semi-durable goods, as with automobiles -- where there are established consumer preferences, experiences, networks, and complementarities -- the challenges are enormous. Further, more complex products are particularly subject to these dynamics. A larger set of

attributes implies that multiple exposures are required until perception has caught up with an alternative's efficacy (Centola and Macy 2005).

Here I developed and analyzed a formal framework showing how the process of acceptance plays an important role in technology introduction, in particular in combination with scale effects. The model introduced here extends Bass diffusion models by incorporating two important characteristics of automobile purchase decisions. First, the process of consumer awareness is decoupled from the sales process. The maximum diffusion rate is slow due to the physical characteristics, thus making competition for attention more complicated than usual. Second, adopters face a choice among a variety of technologies, of which performance is endogenously affected by adoption. The fundamental tipping point dynamics could be captured in a 1st order analytically tractable model. Word-of-mouth through non-users and endogenous media attention is represented explicitly and is important for take-off in the early stages. However, to capture the essence of the structural characteristics behind our hypothesis a broader model is needed, and this was explored in depth.

The analysis offers significant insights into the typical challenges that lie ahead for contemporary vehicle propulsion system transitions. First, this analysis revealed a tipping point in consumer acceptance and the adoption of novel technologies, purely determined by the social exposure dynamics. Second, these mechanisms suggest slower take-off, beyond what is expected from learning or replacement dynamics which are especially

slow for more durable goods. Third, competition between entrant technologies strongly increases the path-dependence of the system.

Obviously the scope of the current model is limited and further work is needed. One challenge is to further understand the interplay between exposure, socialization, and formation of choice sets. These socio-cognitive interrelations that are involved in the formation of new markets or product portfolios (Garud and Rappa 1994) are not directly observable and are difficult to estimate. Parameter sensitivity tests based on various real-life cases of adoption will help improve this basic structure. The model has been formulated so that this is feasible.

The performance of a technology is not objective but a complex process of socialization that involves both learning about its qualities and the evolution of consumer preferences, while actual adoption drives the development of complementary assets that in turn feed back to attractiveness. Current exposure to the HFCVs and, in a broader context, to the hydrogen economy is growing. But even when an AFV's expected future performance is superior to alternatives there are several challenges to be met before it can take off. First, as this research suggests, there seem to be few benefits to generating costly early awareness, if subsequent potential adoption rates are low. Acceptance will demand a process that allows building "trust" in the new technology and "confidence" through actual experience and intimate exposure. The debates about, and camps formed behind the different platforms during the early transition towards the horseless carriage is

illustrative here. There are strong feedbacks that operate around perceived performance: when expectations are high, the platform experiences a lot of "free" media attention. This effect will gradually disappear as performance improvement slows. Eventually, the focus could even be directed to incidents of failure, thus turning a virtuous cycle in a vicious one.

Second, limited performance improvement can constrain adoption, which in turn further reduces the opportunity for learning-by-doing. A 2003 MIT report by Heywood et al. estimates that performance of HFCVs will not equal that of ICE, gas-electric hybrids, or diesel engines for 20 years. In the meantime, the dominant internal combustion engine has the opportunity to "free ride" on innovative ideas that emerge out of research on the hydrogen platform. A combination of consumer acceptance and scale effects such as learning-by-doing and spillovers can lead to perverse dynamics, even under relatively benign initial conditions, such as the current high gasoline prices, security concerns about the current energy systems, environmental pressures, and the existence of potentially efficacious technologies.

Managing the transition trajectories of these socio-technical systems is difficult. Without a "fertile supportive environment," early marketing and media attention will not be a leverage point for replacement. High prospective performance is not a guarantee for success. The durable enthusiasm of engineers, suppliers, and producers for EVs in the early 20th century and the limited success of AFV introductions illustrate the enormous

misconceptions in understanding the power of social factors and their co-evolutionary

dynamics with other scale effects in determining the success or failure of innovations.

References

Abernathy, W. J., & Utterback, J. M. 1978. Patterns of Industrial Innovation. *Technology Review*, **80**(7), 40–47.

Arthur, W. Brian. 1989. Competing Technologies, Increasing Returns, and Lock In by Historical Events. *Economic Journal*, **99**, 116–131.

Bass, F. M. 1969. A New Product Growth Model for Consumer Durables. *Marketing Science*, **15**(5), 215–227.

Beasley, David R. 1988. *The suppression of the automobile : skulduggery at the crossroads*. Contributions in economics and economic history, no. 81. New York: Greenwood Press.

Ben-Akiva, M. E. and S. R. Lerman (1985). *Discrete choice analysis : theory and application to travel demand.* Cambridge, Mass., MIT Press.

Berry, S., J. Levinsohn, et al. (2004). "Differentiated products demand systems from a combination of micro and macro data: The new car market." *Journal of Political Economy* **112**(1): 68-105.

Bijker, W. E., T. P. Hughes, et al. (1987). *The Social construction of technological systems : new directions in the sociology and history of technology.* Cambridge, Mass., MIT Press.

Botting, D. (2001). *Dr. Eckener's dream machine : the great zeppelin and the dawn of air travel.* New York, H. Holt.

Brownstone, D., D. S. Bunch, et al. (2000). "Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicles." *Transportation Research Part B-Methodological* **34**(5): 315-338.

Callon, M. (1986). The Sociology of an Actor-Network: The Case of the Electric Vehicle. *Mapping the dynamics of science and technology : sociology of science*

in the real world. M. Callon, J. Law and A. Rip. Basingstoke, Macmillan: xvii, 242 p.

Census, U.S. Bureau of the. (1975). Historical Statistics of the United States-bicentennial edition. Tech. rept.

Cowan, R. and S. Hulten (1996). "Escaping lock-in: The case of the electric vehicle." *Technological Forecasting and Social Change* **53**(1): 61-79.

Christensen, C. M. (1997). *The innovator's dilemma : when new technologies cause great firms to fai*l. Boston, Mass., Harvard Business School Press.

Cohen, W. M., & Levinthal, D. A. 1989. Innovation and Learning -the 2 Faces of R-and-D. *Economic Journal*, 99(397), 569–596.

David, P. A. (1985). "Clio and the Economics of Qwerty." *American Economic Review* **75**(2): 332-337.

Di Pascoli, S., A. Femia, et al. (2001). "Natural gas, cars and the environment. A (relatively) 'clean' and cheap fuel looking for users." *Ecological Economics* **38**(2): 179-189.

Dodson, J. A. and E. Muller (1978). "Models of New Product Diffusion through Advertising and Word-of-Mouth." *Management Science* **24**(15): 1568-1578.

Domencich, T. A., D. McFadden, et al. (1975). *Urban travel demand : a behavioral analysis : a Charles River Associates research study.* AmsterdamNew York, North-Holland Pub. Co. ;

Farrell, A. E., Keith, D. W., & Corbett, J. J. 2003. A strategy for introducing hydrogen into transportation. *Energy Policy*, 31(13), 1357–1367.

Flink, J. (1970). *America adopts the automobile, 1895-1910*. Cambridge, Mass.,, MIT Press.

Gamson, W. A. and A. Modigliani (1989). "Media Discourse and Public-Opinion on Nuclear-Power - a Constructionist Approach." *American Journal of Sociology* **95**(1): 1-37.

Garud, R. and Rappa, M.A. (1994) 'A socio-cognitive model of technology evolution: the case of cellular implants', *Organization Science*, **5** (3) 344-62

Geels, Frank W (2005) *Technological Transitions and Systems Innovations – A Co-evolutionary and Socio-Technical Analysi*s. Northampton MA Edward Elgar

Griliches, Z. 1957. Hybrid Corn: an exploration in the economics of technological
change. *Econometrica*, 25(4), 501.

Gladwell, M. (2000). *The tipping point : how little things can make a big difference*.
Boston, Little, Brown.

Hauser, J. R., G. L. Urban, et al. (1993). "How Consumers Allocate Their Time When
Searching for Information." Journal of Marketing Research 30(4): 452-466.

Heywood, John B., Weiss, Malcom A., Schafer, Andreas, Bassene, Stephane A., &
Natarajan, Vinod K. 2003. *The Performance of Future ICE and Fuel Cell
Powered Vehicles and Their Potential Fleet Impact*. Tech. rept. LFEE 2003-004
RP. laboratory for energy and the environment -Massahusetts Institute of
Technology.

Homer, J. B. 1987. A Diffusion-Model with Application to Evolving Medical
Technologies. *Technological Forecasting and Social Change*, 31(3), 197–218.

The Horseless Age Journal. Editions November 1895-October 1900.

Jacobs, J. (1961). The death and life of great American cities. New York, Random House.

Jamison, A. (1970). The steam-powered automobile; an answer to air pollution.
Bloomington,, Indiana University Press.

Jovanovic, B., & Macdonald, G. M. 1994. Competitive Diffusion. *Journal of Political
Economy*, 102(1), 24–52.

Kalish, S. (1985). "A New Product Adoption Model with Price, Advertising, and
Uncertainty." Management Science 31(12): 1569-1585.

Katz, M. L. and C. Shapiro (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.

Kimes, Beverly R., & Jr., Clark Henry A. 1996. *Standard Catalog of American Cars
1805-1942*. 3rd edition edn. Iola, WI USA: Krause Publications.

Kimes, Beverly R. and S. E. Johnson, Eds. (1971). The American Car since 1775. New
York, NY, L. Scott Bailey.

Kirsch, David A. 2000. *The electric vehicle and the burden of history*. New Brunswick,
N.J.: Rutgers University Press.

Klepper, S. 1996. Entry, exit, growth, and innovation over the product life cycle. *American Economic Review*, 86(3), 562–583.

Loch, C. H., & Huberman, B. A. 1999. A punctuated-equilibrium model of technology diffusion. *Management Science*, 45(2), 160–177.

MacCready, P. B. (2004). The Case for the Battery Electric Vehicles. *The Hydrogen Economy Transition: Moving Toward the Post Petroleum Age in Transportation.* D. Sperling and J. S. Cannon.

Mahajan, Vijay, Muller, Eitan, & Wind, Yoram. (2000). *New-product diffusion models.* International series in quantitative marketing ; 11. Boston: Kluwer Academic.

Mahajan, V., E. Muller, et al. (1990). "New Product Diffusion-Models in Marketing - a Review and Directions for Research." *Journal of Marketing* **54**(1): 1-26.

MacLean, H. L. and L. B. Lave (2003). "Evaluating automobile fuel/propulsion system technologies." *Progress in Energy and Combustion Science* **29**(1): 1-69.

McFadden, D. (1978). Econometric Models of Probabilistic Choice. *Structural Analysis of Discrete Data with Econometric Applications.* C. Manski and D. McFadden. Cambridge, MA, MIT Press.

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

Miller, D. (2001). *Car cultures. Oxford* ; New York, Berg.

Mintz, Marianne, Molburg, John, Folga, Stephen, & Gillete, Jerry. 2003. *Hydrogen Distribution Infrastructure*. Tech. rept. Center for Transportation Research.

Mom, G. P. A. and D. A. Kirsch (2001). "Technologies in tension - Horses, electric trucks, and the motorization of American cities, 1900-1925." *Technology and Culture* **42**(3): 489-518.

Moore, J. S., J. D. Maples, et al. (1998). "Light-duty diesels - Consumer perspectives and US energy supply issues." *Energy, Air Quality and Fuels 1998*(1641): 19-26.

Nelson, P. (1974). "Advertising as Information." *Journal of Political Economy* **82**(4): 729-754.

Norton, J. A., & Bass, F. M. 1987. A Diffusion-Theory Model of Adoption and Substitution for Successive Generations of High-Technology Products. *Management Science*, **33**(9), 1069–1086.

Rogers, Everett M. 1962. *Diffusion of innovations*. New York: Free Press.

Schiffer, M. B., T. C. Butts, et al. (1994). *Taking charge : the electric automobile in America.* Washington, Smithsonian Institution Press.

Sterman, John. (2000). *Business dynamics : systems thinking and modeling for a complex world.* Boston, Irwin/McGraw-Hill.

Urban, G. L., Hauser, J. R., & Roberts, J. H. (1990). Prelaunch Forecasting of New Automobiles. *Management Science*, 36(4), 401–421.

Urban, G. L., B. D. Weinberg, et al. (1996). "Premarket forecasting really new products." *Journal of Marketing* **60**(1): 47-60.

Urry, J. (2004). "The 'system' of automobility." *Theory Culture & Society* **21**(4-5): 25-+.

Westbrook, Michael H. 2001. *The Electic Car -Development and future of battery hybrid and fuel-cell cars*. The Power and Energy Series 38. Institution of Electrical Engineers.

Zangwill, W. I., & Kantor, P. B. (1998). Toward a theory of continuous improvement and the learning curve. *Management Science*, **44**(7), 910–920.
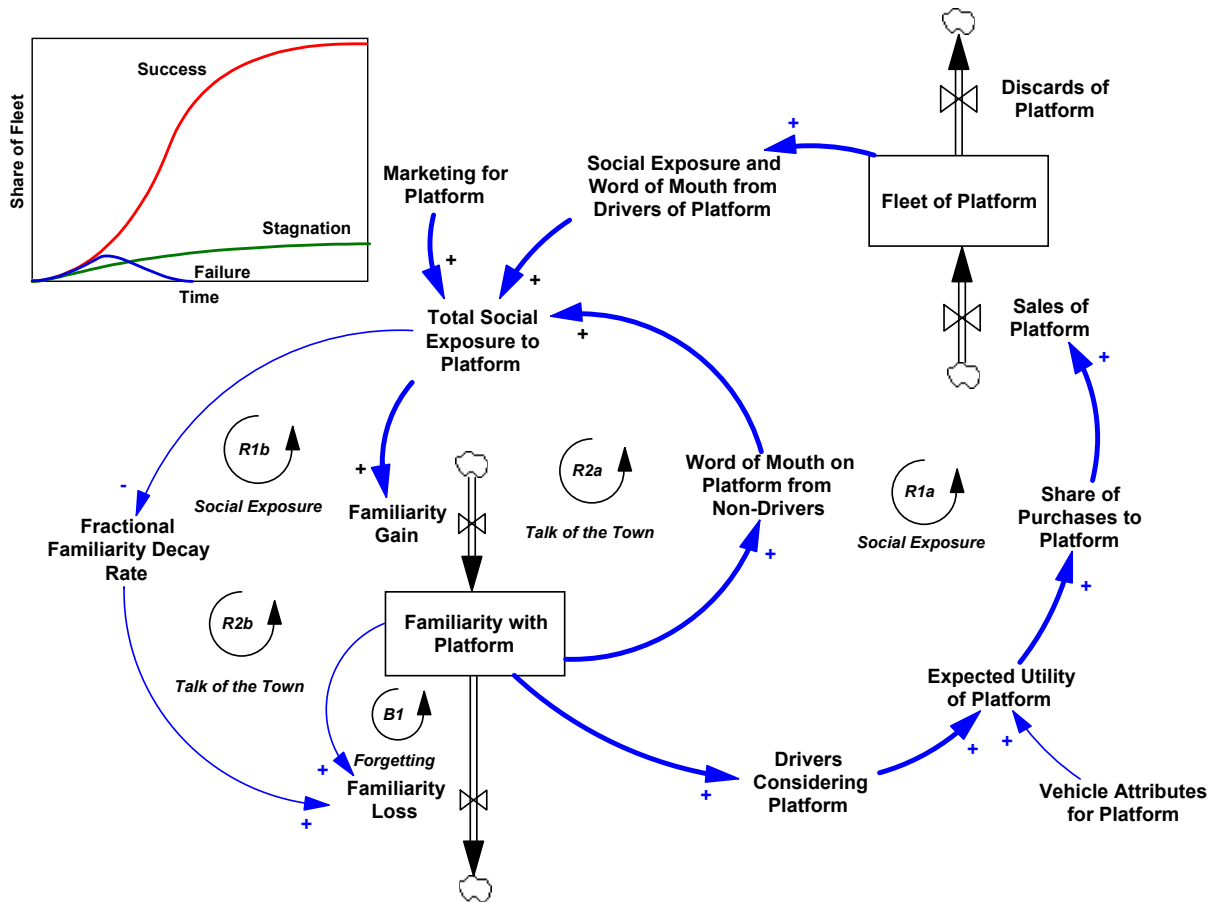
Figures

Figure 1 Principal positive feedbacks conditioning familiarity and consumer choice, with expected modes of behavior.
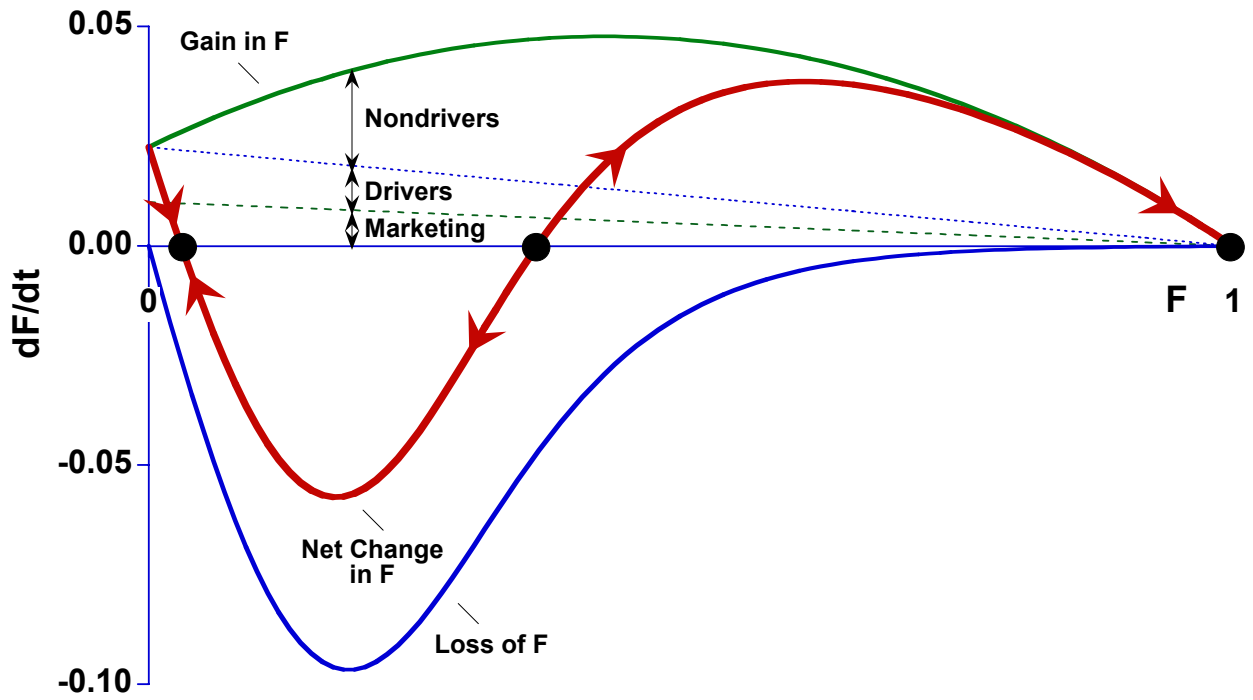
Figure 2 Phase plot for one-dimensional system showing two stable and one unstable fixed points for familiarity of ICE drivers with AFVs (parameters in Table 1).
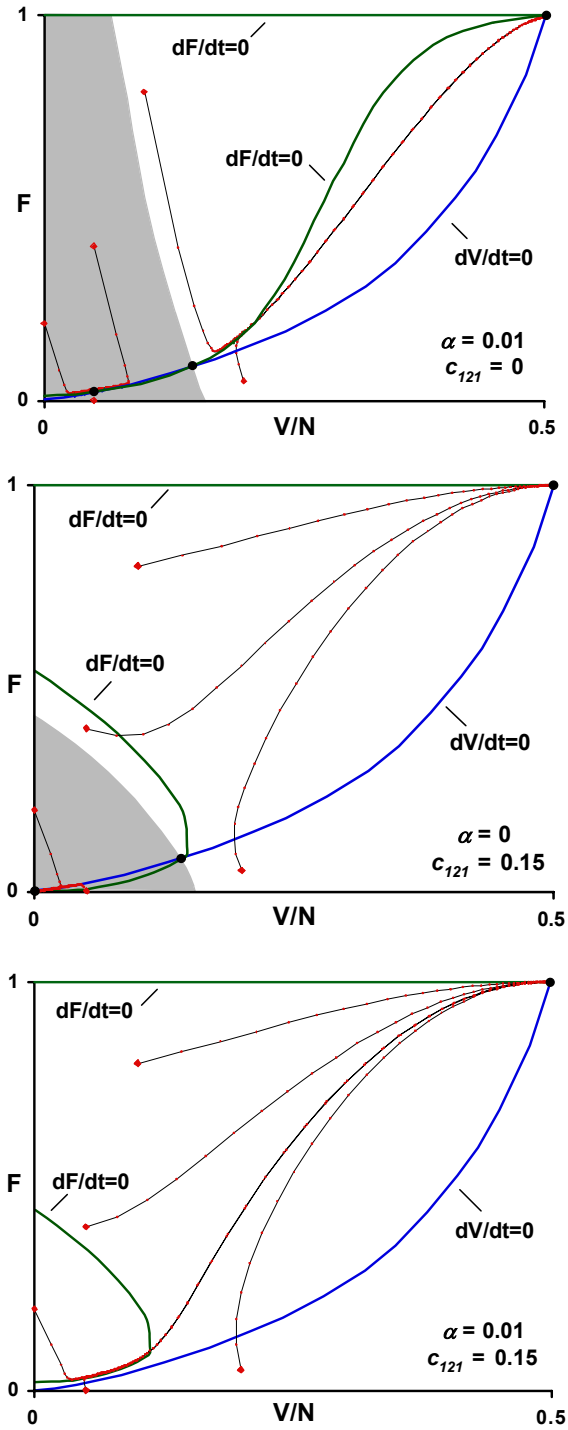
Figure 3 Phase space for two-dimensional system with endogenous familiarity and fleet. Fixed points exist at intersections of nullclines; sample trajectories shown as dots. Grey area shows basin of attraction for the low-diffusion equilibrium. Strength of marketing and non-driver word of mouth as shown. Other parameters as in Table 1.

Figure 4 Alternative vehicle familiarity and fleet with an aggressive marketing and promotion program. Duration of high marketing impact ($\alpha_2 = 0.02$) varies between 10 and 50 years.

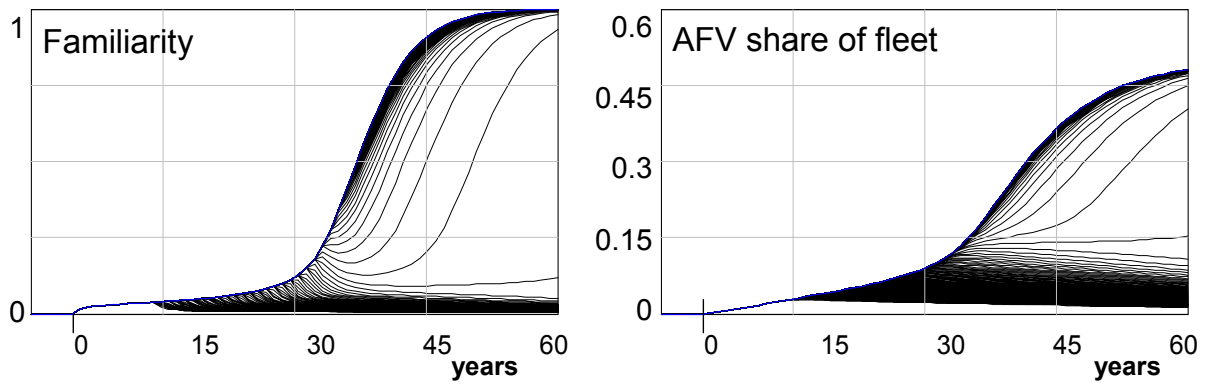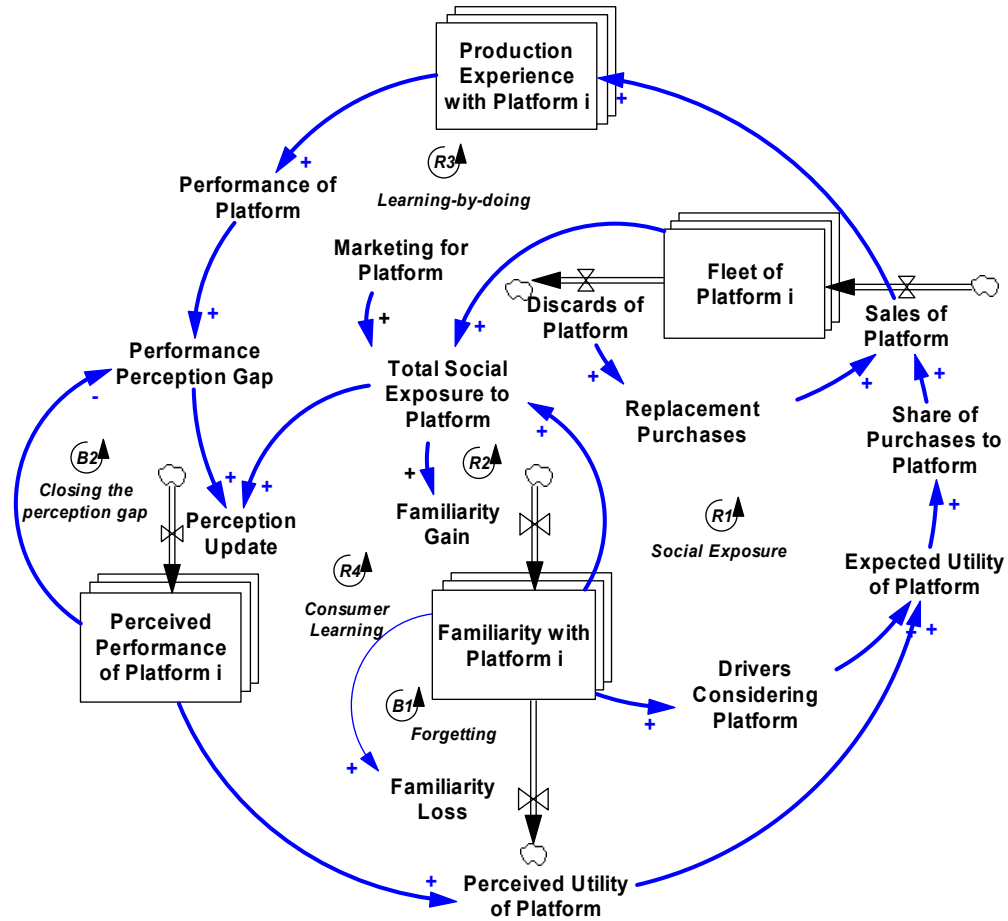**Figure 5** Exploring other feedbacks: social exposure interacts tightly with learning-by-doing, consumer learning about the vehicle efficacy, and with the competitive dynamics (indicated by the layering of the stocks).
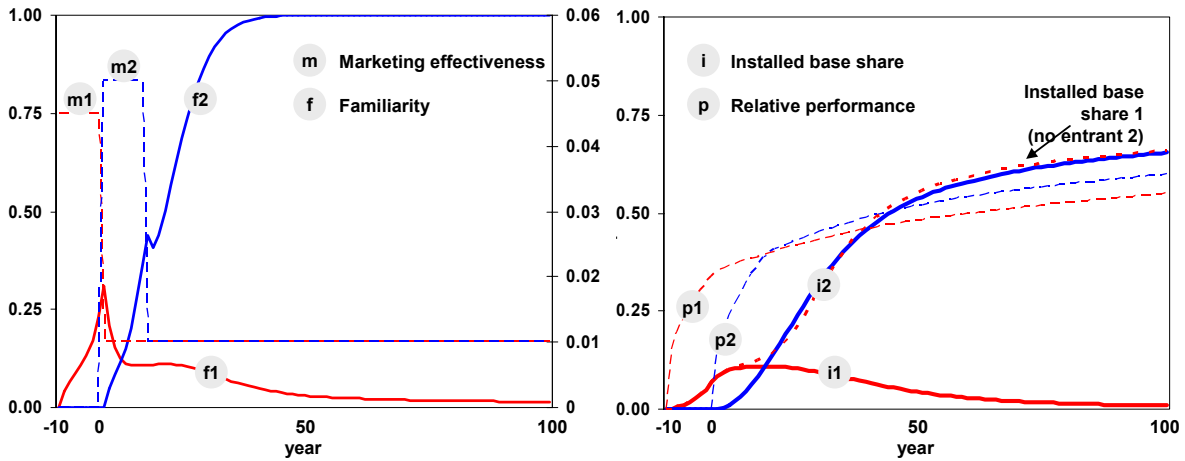
Figure 6 Competitive diffusion dynamics for two entrants. Entrant 2 is introduced at time 0 with a 10-year lag behind entrant 1. Marketing programs at introduction are different for first 10 years: a) exogenous marketing effectiveness (dotted, labeled as m1, m2 for the corresponding entrants) and familiarity of both entrants (thick); b) installed base share (thick) and relative performance of both entrants (dashed). Also shown is the hypothetical installed base share for entrant 1 in the absence of a second entrant (dotted).
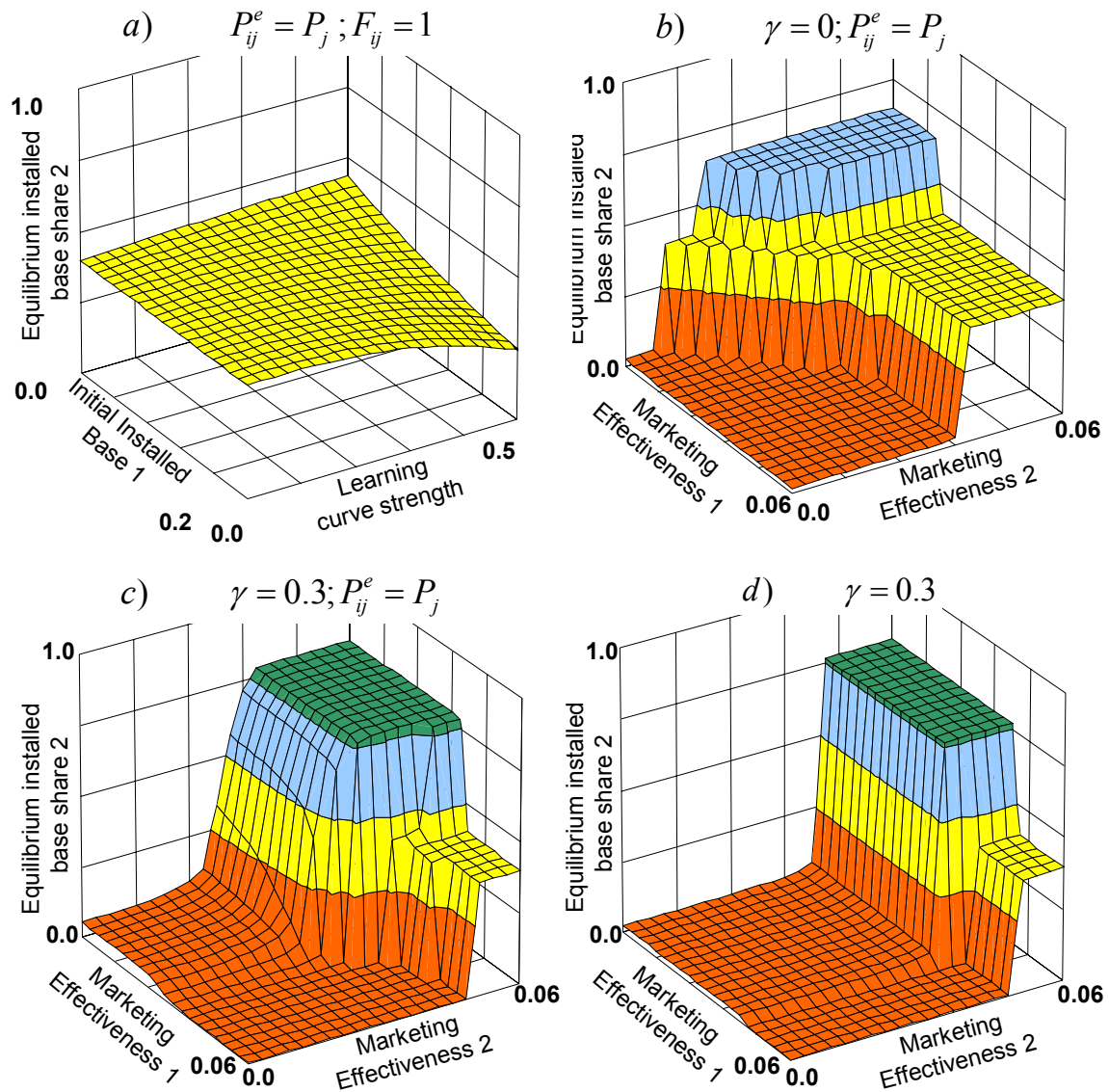
Figure 7 Competitive diffusion dynamics for two entrants. Each graph shows the equilibrium installed base share of entrant 2. a) shows, under constant and full familiarity, the effect of learning-curve strength and head-start installed base of the first entrant. b)-d) show the results including the endogenous familiarity, with entrant 2 introduced with a 10 year lag. Shown are the results of marketing effectiveness in the first 10 years after introduction for both entrants. Conditions in the different graphs are: b) absence of learning-by-doing, c) normal learning-by-doing, d) normal learning-by-doing and consumer learning about performance.
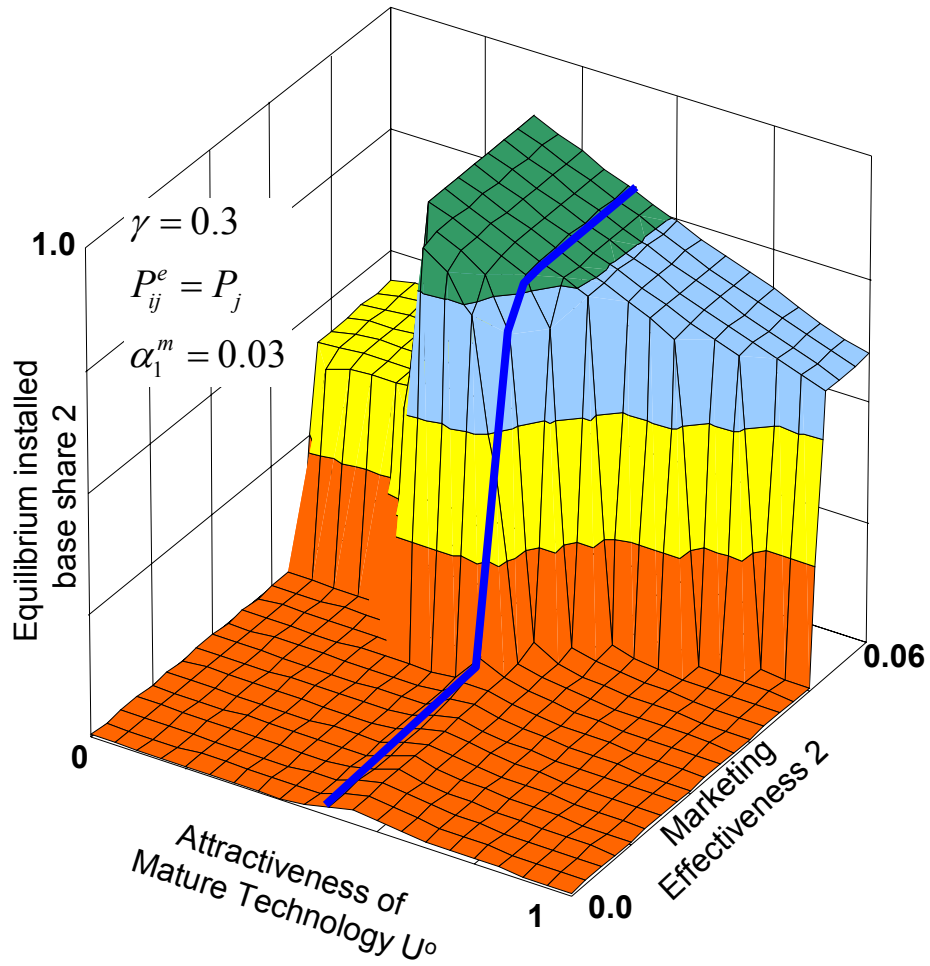
Figure 8 Competitive diffusion dynamics for two entrants. Equilibrium installed base share of entrant 2 is graphed as a function of marketing effectiveness of entrant 2 during the 10-year introduction period and as a function of attractiveness of not adopting. The thick line indicates correspondence with the output of graph 6c for marketing effectiveness at 0.03.
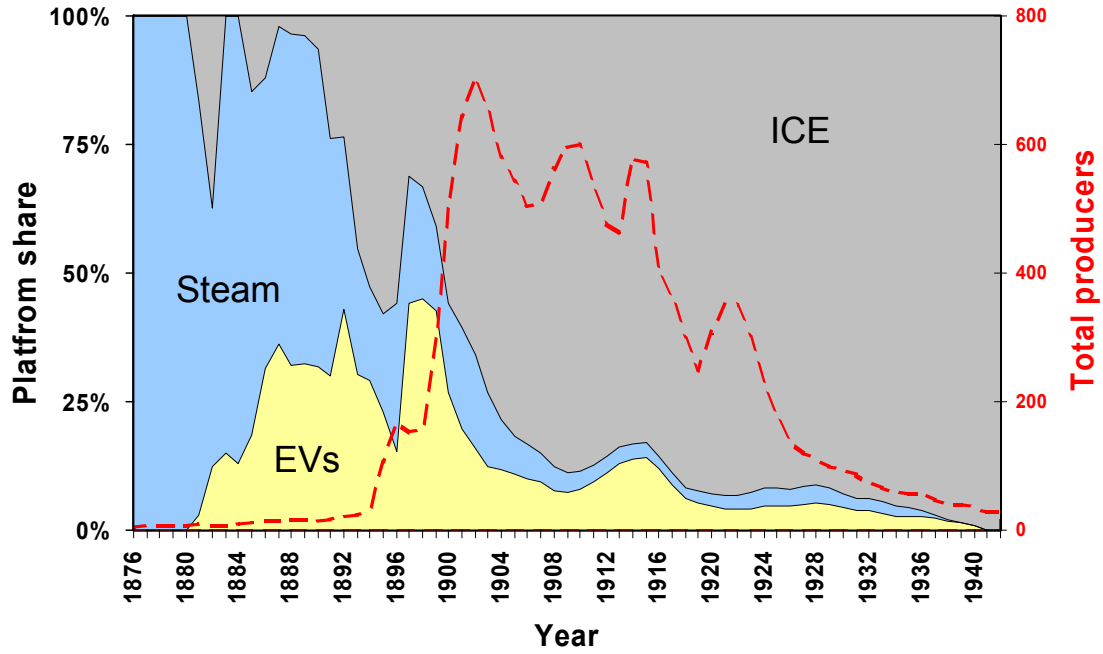
Figure 9 Producers by platform. Source: Compiled by author from Kimes and Clark (1996), crosschecking with Epstein (1928), Geels (2005), Kirch (2000) and others. Excluded are other platforms, such as spring-powered, compressed air, or hybrids that constitute small numbers.
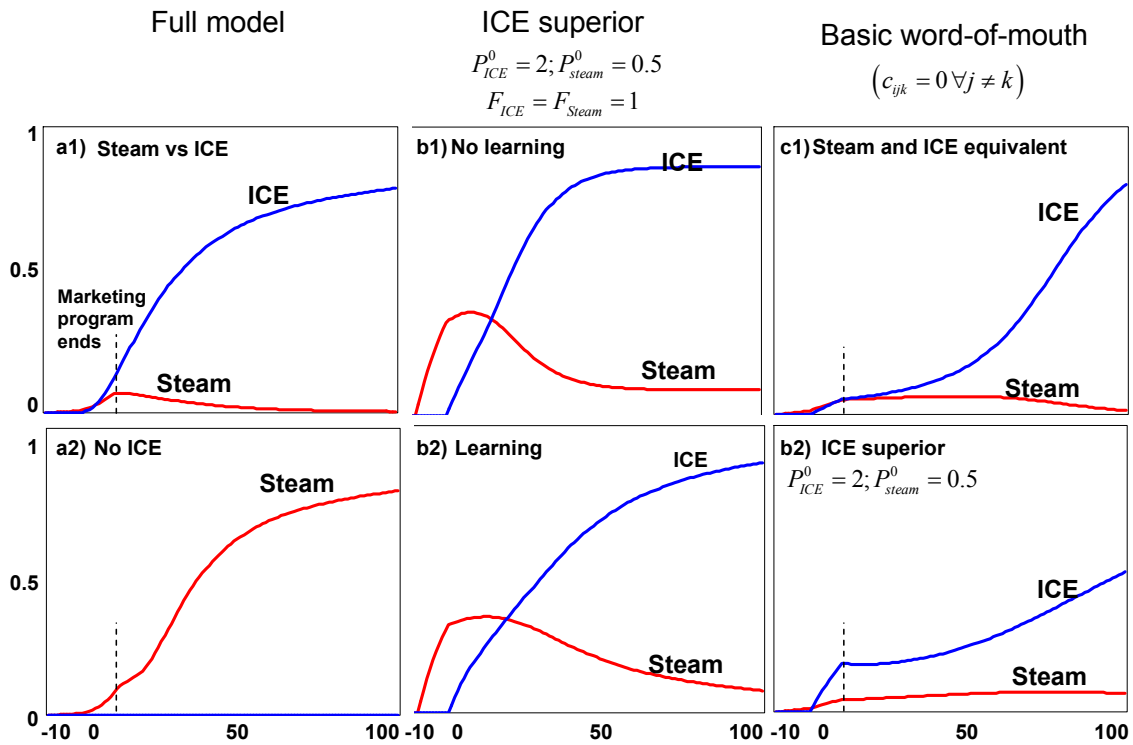
Figure 10 The early transition towards the automobile using competing models; a) dynamics resulting from inclusion of full structure; b) focusing only on technology, using superiority of ICE as an hypothesis; c) using traditional word-of-mouth dynamics.

AFV installed base response to policies
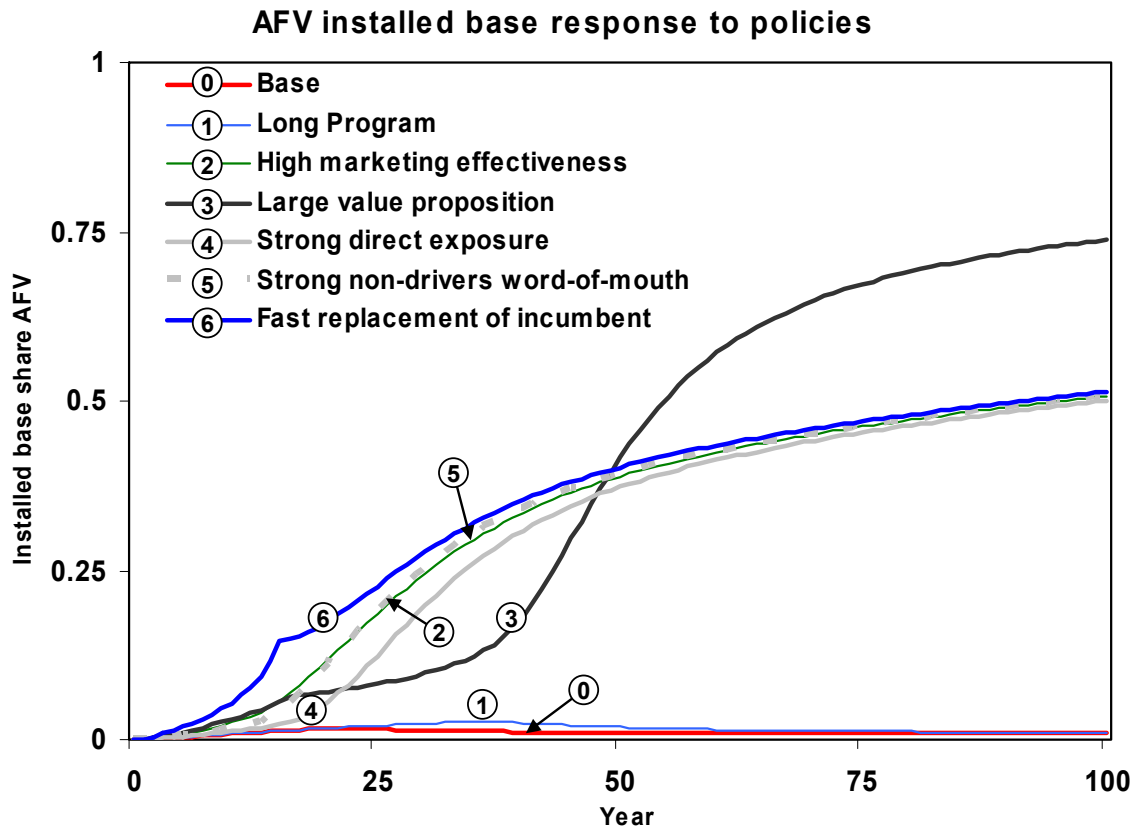
Figure 11 Exploration of policies.

Tables

Table 1 Parameters used in simulations.

	Definition	unit	value
α_2	AFV marketing effectiveness	1/year	0.01
c_{122}	Strength of word of mouth about AFVs for contacts between ICE and AFV drivers	1/year	0.25
c_{121}	Strength of word of mouth about AFVs for contacts between ICE and other ICE drivers	1/year	0.15
ϕ_0	Maximum familiarity loss rate	1/year	1
η^*	Reference rate of social exposure	1/year	0.05
ε	Slope of decay rate at reference rate	year	20
λ	Average vehicle life	year	8
u^*	Reference utility	dmnl	1
β	Sensitivity of utility to performance	dmnl	1
γ	Learning curve strength	dmnl	0.379[*]
E_0	Reference years of effective experience	years	20

---

[*] The learning curve exponent γ is calculated from the assumed fractional performance improvement per doubling of knowledge, $(1 + \Delta)P_0 = P_0(2K_0/K_0)^\gamma$, or $\gamma = \ln(1+ \Delta)/\ln(2)$. We assume a 30% learning curve, $\Delta = 0.3$, so $\gamma=0.379$.

Technical appendix accompanying Essay 1

1 Introduction

The model described in this Essay is designed to capture the diffusion of and competition among multiple types of alternative vehicles, along with the evolution of the ICE fleet. For example, the model can be configured to represent ICE and alternatives such as ICE-electric hybrid, CNG, HFCV, biodiesel, E85 flexfuel, and electric vehicles. However, the Essay focuses on intuition about the basic dynamics around the diffusion of alternatives to ICE by considering two platforms, ICE and an alternative vehicle, and makes a number of other simplifying assumptions that allow us to explore the global dynamics of the system. In this appendix I discuss additional components of the full model, highlighting those structures required to capture the competition among multiple alternative platforms.

The appendix is divided into three sections. The first section provides elaborations on the model. The second section provides connections to, and differences from the original Bass structure as discussed in the Essay. The model and analyses can be replicated from the information provided in the Essay. The last section provides a link to the full model and analysis documentation.

2 Elaborations on the model

This section elaborates segments of the model that were highlighted in the paper but not fully expanded due to space limitations.

a) Vehicle fleet aging chain

For simplicity, the age structure of the fleet is not treated in the paper. Below we lay out how this is incorporated in the full model.

The total number of vehicles for each platform j, $j=\{1,…, J\}$, of each age cohort m, V_{j_m}, accumulates net vehicle replacements and aging (see Figure 1):

$$\frac{dV_{j_m}}{dt} = v^r_{j_m} + v^a_{j_m} \tag{A1}$$
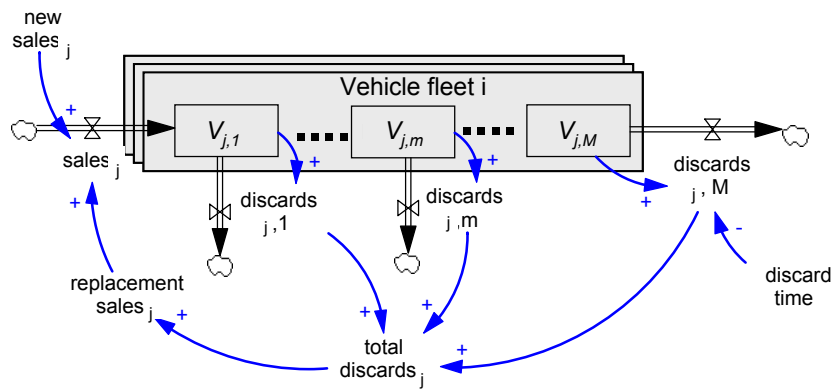
Figure 1 Vehicle replacement with aging chain.

Aging captures vehicles coming from a younger cohort less those aging into the next cohort:

$$v_{j_m}^{a} = v_{j_m}^{a+} - v_{j_m}^{a-} \tag{A2}$$

with

$$v_{j_m}^{a+} = \begin{cases} 0 & m=1 \\ v_{i_{m-1,m}} & m>1 \end{cases} \qquad v_{j_m}^{a-} = \begin{cases} v_{i_{m,m+1}} & m \le M \\ 0 & m=M \end{cases} \tag{A3}$$

while

$$v_{i_{m,m+1}} = f_{i_m}^{s} V_{i_m} / \tau^{c} \tag{A4}$$

Where $f_{i_m}^{s}$ is the survival fraction for each cohort.[6,7]

Net vehicle replacements are new vehicle sales, s_{j_m}, less age dependent discards, d_{j_m}:

$$v_{j_m}^{r} = s_{j_m} - d_{j_m} \tag{A5}$$

We do not consider the used car market here. New vehicle sales enter the first age cohort, thus:

$$s_{j_m} = \begin{cases} s_j & m=1 \\ 0 & m>1 \end{cases} \tag{A6}$$

Total sales for platform j, s_j, consist of initial and replacement purchases:

$$s_j = s_j^{n} + s_j^{r} \tag{A7}$$

The full model allows for growth in the fleet as population and the number of vehicles per person grow. In the paper population and the number of vehicles per person are

---

[6] Annual survival (and/or scrappage) rates by model year can be derived from registration data (e.g. by L. Polk &Co, AAMA).

[7] In equilibrium average vehicle life λ^{v} is found by: $\lambda_j = \sum_{m=1}^{M-1}\left(\prod_{m'=1}^{m-1} f_{j_{m'}}^{r}\right)\lambda^{c} + \prod_{m'=1}^{M-1} f_{j_{m'}}^{r} \lambda^{c_M}$

assumed constant, implying the total fleet is in equilibrium and initial purchases are zero.

Vehicles sales for platform j arise from the replacement of discards from any platform i

and cohort m, $d^r_{i_m}$:

$$\sum s^r_j = \sum_{i,m'} \sigma_{i_{m'}j} d^r_{i_{m'}} \qquad (A8)$$

where σ_{i_nj} is the share of drivers of platform i cohort n replacing their vehicle with a new

vehicle of platform j. The share switching from i to j depends on the expected utility of

platform j as judged by the driver of vehicle i, cohort n, $u^e_{i_nj}$, relative to that of all options

$u^e_{i_nj'}$.

Thus:

$$\sigma_{i_mj} = \frac{u^e_{i_mj}}{\sum_{j'} u^e_{i_mj'}}. \qquad (A9)$$

To capture a driver's consideration set we introduce the concept of familiarity among

drivers of vehicle i with platform j. The model can be elaborated to include cohort-

specific levels of familiarity, recognizing that drivers of, say, a 10 year old ICE vehicle

have a different (presumably lower) familiarity with new ICE vehicles than the driver of

a 1 year old vehicle. Such distinctions may matter when vehicle attributes change

rapidly, as is likely for early AFVs as experience and technology rapidly improve.

(Further disaggregation would eventually lead to an agent-based representation where

each driver has an individual-specific level of familiarity with different platforms).

These issues will be treated in future work. For simplicity I assume here that familiarity

is equal across all cohorts of a given platform and remains F_{ij}, thus expected utility is:

$$u^e_{i_m j} = F_{ij} * u_{i_m j}.$$
(A10)

b) Initial purchases and fleet growth

New car sales for fleet j are:

$$s^n_j = \sigma^n_j s^n$$
(A11)

where the share σ_j is equal to the share of replacement sales: $\sigma^n_j = s^r_j / \sum_i s^r_i$.

Total new car sales allow the total fleet $V = \sum_{j,m} V_{j_m}$ to adjust to its indicated level V^*:

$$s^n = \frac{\max\left[0, \left(V^* - V\right)\right]}{\tau^v}$$
(A12)

where total desired vehicles $V^* = \rho^v * H$ is product of the target or desired number of

vehicles per household ρ and total households H, and τ^v is the fleet adjustment time. The

max function ensures sales remain nonnegative in the case where V^* falls below V (a

possibility if there is a large unfavorable shift in the utility of AFVs when the installed

base is small).

Discards, d_{j_m} are found by:

$$d_{j_m} = \begin{cases} (1 - f^s_{j_m}) V_{j_m} / \lambda^c & m < M \\ V_{j_m} / \lambda^{cM} & m = M \end{cases}$$
(A13)

where λ^c is the cohort residence time; λ^{cM} is the residence time of the last cohort.

The number of discards people choose to replace is give by:

$$d^r_{j_m} = f^r d_{j_m}$$
(A14)

74

where f^r is the nonnegative part of the difference between total discards and the

indicated contraction rate as a fraction of the total discard rate:

$$f^r = \frac{\max\left[0, d - v^{c*}\right]}{d} \tag{A15}$$

Here $d = \sum_{i,m} d_{i_m}$ is total discards, and $v^{c*} = \frac{\max\left[0, V - V^*\right]}{\tau^v}$ is the indicated fleet

contraction rate. The fleet of a particular platform can contract when, for example, the

perceived utility of that platform suddenly falls (say, due to unfavorable shifts in fuel

costs or perceived safety, reliability, or costs) and if the existing installed base is small

enough and young enough so that discards from normal aging are small.

c) Co-flows

The model accounts for transfer of familiarity and perceived performance associated with

those drivers who switch platforms. I will capture this through the co-flow structure

(Sterman 2000). The formal structure is identical for both and I will discuss the

familiarity co-flow as an example. The familiarity of drivers of platform i with platform j

is updated through social exposure, as discussed in the paper. When a driver switches

from platform i to k, their familiarity with platform j is transferred from F_{ij} to F_{kj}. For

example, consider a model in which three platforms are portrayed, say, ICE, hybrids, and

HFCVs (denoted platforms 1, 2, and 3, respectively). When an ICE driver switches to a

hybrid, the familiarity of that driver with HFCVs, previously denoted F_{13}, now becomes

F_{23}. In the two platform simulations considered in the paper these dynamics do not matter

since all drivers are assumed to be fully familiar with ICE, and AFV drivers are assumed

fully familiar with AFVs, so the only dynamic relates to the growth of familiarity of ICE

drivers with AFVs (F_{12}).

To model the transfer of familiarity as drivers switch platforms, it is convenient to

consider the evolution of familiarity at the population level:

$$\frac{d(F_{ij}V_j)}{dt} = V_i \frac{dF_{ij}}{dt} + F_{ij} \frac{dV_i}{dt} = f_{ij}^u + f_{ij}^t \tag{A16}$$

where the first term, which we call f_{ij}^u, captures updating of familiarity with platform j

by drivers of platform i, as discussed in the paper. The second term, denoted f_{ij}^t, captures

the transfer of familiarity arising from drivers who switch platforms. When familiarity is

updated much faster than fleet turnover (and therefore switching), the second term has

limited impact on the dynamics of familiarity. On the other hand, when fleet turnover is

very fast, the transfer of familiarity as drivers switch platforms can be important.

Familiarity updating is formulated as described in the paper: updating of total familiarity

is the average update from social exposure, including familiarity decay (equation 5 of the

paper), over the total number of drivers V_i:

$$f_{ij}^u = \left[\eta_{ij} \left(1 - F_{ij} \right) - \phi_{ij} F_{ij} \right] V_i \tag{A17}$$

where η_{ij} is the total impact of total social exposure to platform j on the increase in

familiarity for drivers of platform i, and ϕ_{ij} is the fractional loss of familiarity about
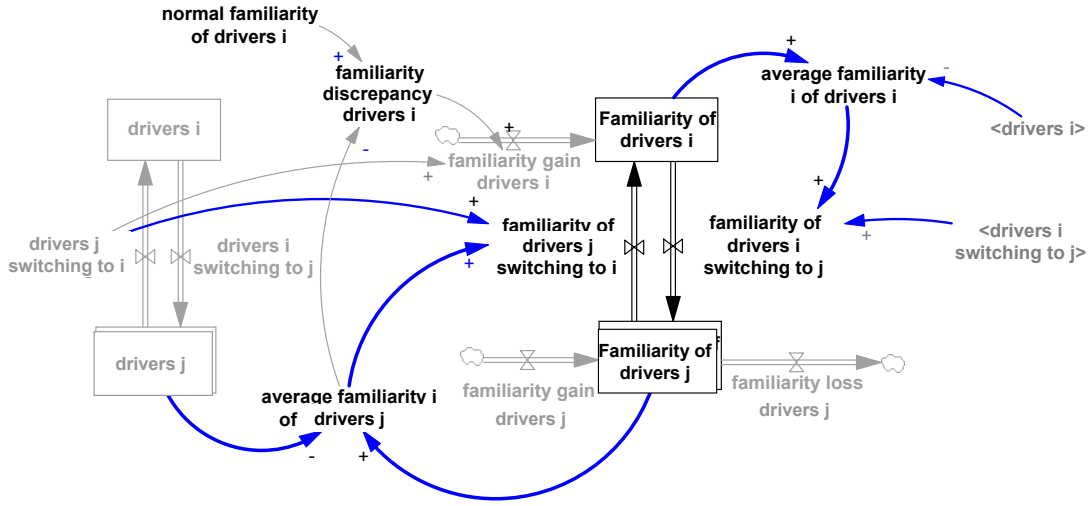
platform j.

Figure A2 Familiarity change for drivers that switch between platforms

The transfer term captures two that track the movement of the familiarity of a driver of platform *i* with platform *j*, one arising from vehicle sales and one arising from discards:

$$f_{ij}^t = f_{ij}^s - f_{ij}^d . \tag{A18}$$

The first term, f_{ij}^s, captures the transfer of familiarity through sales:

$$f_{ij}^s = s_i^n F_{ij} + \begin{cases} \sum_k s_{ki}^r F_{kj} & i \neq j \\ \sum_{k \neq j} s_{ki}^r & i = j \end{cases} \tag{A19}$$

This term contains the flow of new drivers purchasing platform *i*, and their average familiarity with platform *j*, assumed to equal the familiarity of current drivers of *i* with platform *j*. The second term is the transfer of familiarity associated with the flow of drivers of platform *k* replacing their vehicles with one of platform *i*. The average familiarity of these drivers with platform *j* is transferred as they switch. We assume drivers become fully familiar with the platform they are driving, so those who purchase a

vehicle of platform j (the case $i=j$) achieve full familiarity with platform j (in a time much shorter than the other time constants).

The second term in equation (A18) captures the transfer of familiarity with platform j associated with drivers of platform i through discards:

$$f_{ij}^d = d_i \, F_{ij} \qquad \text{(A20)}$$

where $d_i \equiv \sum_m d_{i_m}$ is total discards .

The transfer term f_{ij}^t was used in the simulations of the paper, for the relevant cases. The transfer of familiarity as drivers switch platforms has a small but significant contribution to the dynamics: early alternative fuel adopters who switch back from the alternative to ICE have full familiarity with the AFV, and contribute strongly to word of mouth. Technically, a balancing loop is generated, in similar fashion as marketing effectiveness, with strength $\left[1 - u_{jj} / \sum_i u_{ji} \right] / \lambda^v$. However, a more complicated result emerges when learning about performance through social exposure is involved, as early adopters might learn about mediocre performance. Hence, their word of mouth results in lower perceived attractiveness of alternatives among others.

Other co-flows that follow the same logic are those guide an adjustment of installed base performance, P_j^f to the new vehicle performance P_j^n, and those that allow the perceived performance P_{ij}^e to be updated when drivers switch platforms. The first one is a simple co-flow that only changes with sales and discards. The perceived performance updated

implies taking all influences of Equation 17 into account. This is done by separately capturing drivers of a platform j P^e_{jj} , and non-drivers of platform j, $P^e_{ij}, i \neq j$.

3 Stipulations

a) Equivalence to Bass

Here we recover the Bass equation from the familiarity model in this paper, for durable goods, with validity for low familiarity. The formulation differs from those of the standard Bass models through the decoupling of exposure, familiarity and the adoption decision, the word of mouth through non-users and the discrete choice replacement, for durable goods.

The original Bass model is describes diffusion of isolated technologies and is specified as follows:

$$\frac{dV^B}{dt} = \left(\alpha^B + c^B \left(V^B / N \right) \right) \left(N - V^B \right) \tag{A21}$$

Where the marketing effectiveness α^B and contact rate c^B have the same interpretation as in the familiarity model. The functional form is a the logistic growth and the associated dynamics yield an S-shape curve.

To recover the Bass model, we ignore population change, which is sensible with the shorter time horizons of product replacements in usual Bass settings), aging chains (the

arguments below can be easily expanded), heterogeneity in contact effectiveness, and word of mouth through non users, and denote this simplified version '1':

$$\frac{dF_{ij}^1}{dt} = \eta_{ij}^1 \left(1 - F_{ij}^1\right) - \phi\left(\eta_{ij}^1\right) F_{ij}^1 \qquad (A22)$$

We ignored here the higher order terms that involve transfer of familiarity through sales and discards. Simplifying further, setting contact effectiveness between drivers of platforms j with non-drivers equal for all j and $i \neq j$, with $c^d = c_{ijj}$, for all $i \neq j$:

$$\eta_{ij}^1 = \alpha_j + c^1\left(V_j^1 / N\right) \qquad (A23)$$

Now we specify the sales rate for a platform j, which is identical to the actual familiarity model $s_j = \sum_{i \neq j} \sigma_{ij} V_i / \lambda$, with λ being the vehicle life.

We further assume perceived utility to equal actual utility and derive the Bass equation for durable goods, with validity for low familiarity. When the product of familiarity and relative attractiveness is low, we can make a first order approximation for the share going to i from j:

$$\sigma_{ij} = \begin{cases} \dfrac{F_{ij}u_j}{u^o + u_i + \sum\limits_{j' \neq i} F_{ij'}u_{j'}} \approx \tilde{u}_{-j}F_{ij} & i \neq j; \tilde{u}_{-j} \equiv \dfrac{u_j}{u^o + u_i} \\[4ex] \dfrac{u_j}{u^o + u_j + \sum\limits_{j'} F_{jj'}u_{j'}} \approx \tilde{u}_j & i = j; \tilde{u}_j \equiv \dfrac{u_j}{u^o + u_j} \end{cases} \qquad (A24)$$

Then, letting all alternatives to j yield the same utility, the net sales rate equals the new vehicle sales minus the discards that are not replaced:

$$\frac{dV_j}{dt} = s_j - d_j \approx \left[\sum_{k \neq j} \tilde{u}_{-j} F_{kj}\right]\frac{1}{\lambda} - \left(1 - \tilde{u}_j\right)\frac{V_j}{\lambda}$$

$$= \frac{1}{\lambda}\left(\tilde{u}_{-j} F_j \left(N - V_j\right) + \left(1 - \tilde{u}_j\right)V_j\right) \tag{A25}$$

Further, with adoption dynamics slow relative to the familiarity dynamics, which is justified for durable goods, we use the steady state familiarity as a function of the number of adopters. Ignoring the word of mouth through non-drivers, and the higher order terms that include that include transfer of familiarity through discards:

$$\frac{dF_{ij}^1}{dt} = \eta_j^1\left(1 - F_{ij}^1\right) - \phi\left(\eta_j^1\right)F_{ij}^1 = 0 \tag{A26}$$

Using a piecewise linear expression for equation (7):[8]

$$\phi\left(\eta_j\right) = \begin{cases} 0 & \eta_j < \eta_0 - \dfrac{0.5}{\varepsilon} \\ \phi_0\left[0.5 + \varepsilon\left(\eta_0 - \eta_j\right)\right]F_{ij} & otherwise \\ \phi_0 & \eta_j > \eta_0 + \dfrac{0.5}{\varepsilon} \end{cases}$$

We get for the equilibrium familiarity:

$$F_{ij}^{1*} = \begin{cases} \dfrac{\eta_j^1}{\eta_j^1 + \phi_0\left(0.5 + \varepsilon\left(\eta_0 - \eta_j^1\right)\right)} & \eta_j^1 \leq \eta_0 + \dfrac{0.5}{\varepsilon} \\ 1 & \eta_j^1 > \eta_0 + \dfrac{0.5}{\varepsilon} \end{cases}$$

Then, with η_j^1 small compared to ϕ_0, and, and by definition of the interesting case where familiarity is not saturated yet, $\eta_j^1 \ll \eta_0 + 0.5/\varepsilon$, and thus:

---

[8] This functional form for forgetting leads to results that are indistinguishable from the non-linear form used in the paper.

$$F_{ij}^{1*} \approx \eta_j^1 \big/ \phi = \left(\alpha_j + c^1 \left(V_j^1 / N \right) \right) \Big/ \phi_0 \qquad \text{(A27)}$$

and thus, combining with equation (A25):

$$\frac{dV_j^1}{dt} \approx \left[\frac{\tilde{u}_{-j}}{\phi_0 \lambda} \left(\alpha_j + c^1 \left(V_j^1 / N \right) \right) \left(N - V_j^1 \right) - \frac{\left(1 - \tilde{u}_j \right)}{\lambda} V_j^1 \right]$$

Which can be rewritten as:

$$\frac{dV_j^{1*}}{dt} = \kappa_j \left[\alpha_j' + c \left(V_j / N \right) \right] \left(N - V_j \right) - y_j \qquad \text{(A28)}$$

Where, $\kappa_j \equiv \tilde{u}_{-j} \big/ \phi_0 \lambda$ is the conversion parameter between Bass and the familiarity model

that captures the relative attractiveness, replacement rate, and forgetting rate are

convoluted in the Bass model, but explicit in the familiarity model. Further,

$\alpha_j' \equiv \alpha_j + \left(1 - \tilde{u}_j \right) \big/ \tilde{u}_{-j} \phi_0$ is the adjusted marketing effect, of which the second term, in

multiplication with the conversion parameter captures the "free marketing" exposure that

derives from drivers who discard their vehicles and become non-drivers (which are not

included in the original Bass model). Finally, $y_j = \left(1 - \tilde{u}_j \right) N / \lambda$ is a constant adjustment

that accounts for discards, offsetting any adoption. Note that when drivers are zero, the

last two effects cancel out, naturally preserving non-negativity.

With equation (A28) we have derived at the original Bass model, except for a correction

term. Note that the connection implies structural equivalence, this only held under

specified conditions, for instance assuming equilibrium familiarity and ignoring the role

of non-drivers. Because of the equilibrium assumption, the complex dynamics have been

filtered out. This derivation illustrates the connection of the parameters of the two

models, as well as an interpretation of the Bass parameters in the context of competitive platforms. This interpretation will also be used in the analysis section of the Essay.

b) Platform competition: The familiarity model compared to Bass

The Essay illustrates the strong tipping dynamics that the familiarity model reveals for competing entrants, as a function of their respective marketing programs (Figure 7b and 7c).Here we compare the dynamics of platform competition to that what can be generated by Bass models. We proxy the Bass model with platform competition, by deriving the equilibrium familiarity and absence of word-of-mouth from non-drivers (see Appendix 2a), and combine this with the multiplatform logit decision structure. Figure A3 shows the results, using exactly the same scenario as in Figure 7c and 7d of the Essay.
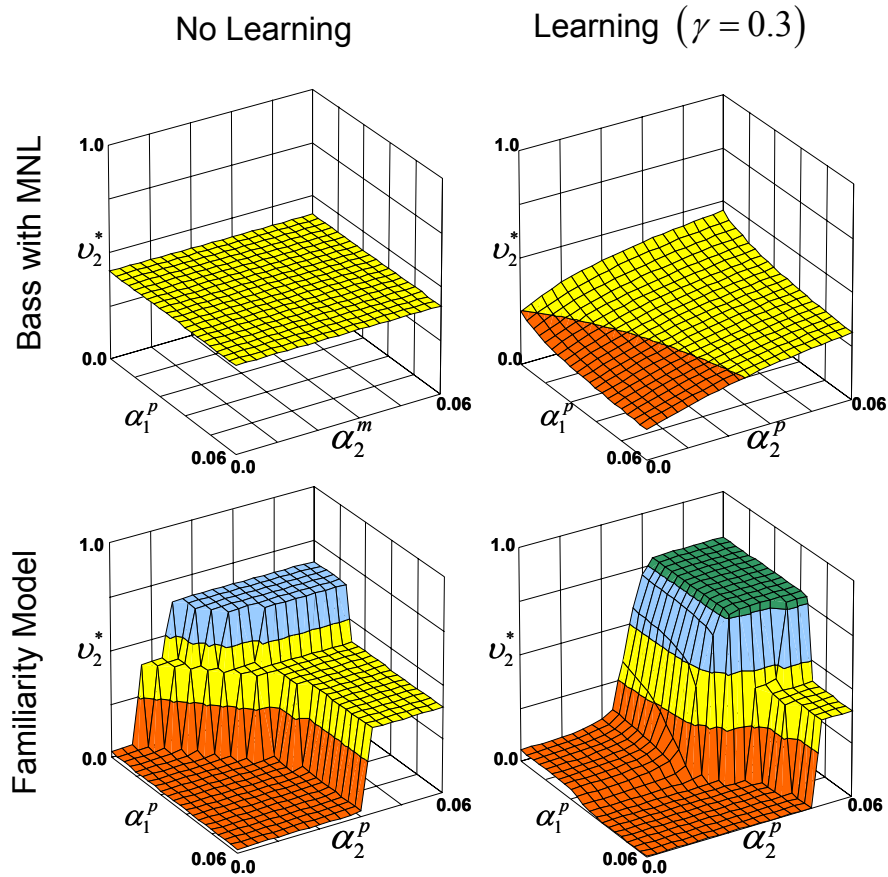
Figure A3 Multiplatform competition: comparing the Familiarity model (Bottom – identical to Figure 7b and c of the paper) with the Bass representation the Bassrepresentation, combined with the MNL decision structure (Top).

We see that the dynamics depart considerably. In absence of learning, in the Bass model there is always convergence to the equal equilibrium share (as the background marketing is nonzero, equaling 0.01). When we include learning, we see that some path-dependency is created in the Bass representation, albeit very smoothly. The results from the Familiarity model contrast greatly to this (Bottom).

4 Model and analysis documentation

The model and analyses can be replicated from the information provided in the Essay.

In addition model and analysis documentation can be downloaded from

http://web.mit.edu/jjrs/www/Thesis Documentation.htm

5 References

Sterman, J. (2000). *Business dynamics : systems thinking and modeling for a complex world.* Boston, Irwin/McGraw-Hill.

Essay *2*

Identifying challenges for sustained adoption of alternative fuel vehicles and infrastructure

Abstract

This paper develops a dynamic, behavioral model with an explicit spatial structure to explore the co-evolutionary dynamics between infrastructure supply and vehicle demand. Vehicles and fueling infrastructure are complementarities and their "chicken-egg" dynamics are fundamental to the emergence of a self-sustaining alternative fuel vehicle market, but they are not well understood. The paper explores in-depth the dynamics resulting from local demand-supply interactions with strategically locating fuel-station entrants. The dynamics of vehicle and fuel infrastructure are examined under heterogeneous socio-economic/ demographic conditions. The research reveals the formation of urban adoption clusters as an important mechanism for early market formation. However, while locally speeding diffusion, these same micro-mechanisms can obstruct the emergence of a large, self-sustaining market. Other feedbacks that significantly influence dynamics, such as endogenous topping-off behavior, are discussed. This model can be applied to develop targeted entrance strategies for alternative fuels in transportation. The roles of other powerful positive feedbacks arising from scale and scope economies, R&D, learning by doing, driver experience, and word of mouth are discussed.

Introduction

In response to environmental, economic, and security related pressures on our current

energy system, automakers are now developing alternatives to internal combustion

engines (ICE). A diverse set of alternatives are considered ranging from promoting

existing possibilities that run on alternative fuels, such as compressed natural gas (CNG),

bio-fuels (such as E85), and diesel, to radically different hydrogen fuel cell vehicles

(HFCVs), and to hybrid forms, such as hybrid electric-ICE vehicles (HEV-ICE). Current

87

perspectives on the possibility of a successful transition to various alternative fuel

vehicles (AFVs) are diverse. For example, concerning HFCVs, Lovins and Williams

(1999) emphasize their long-term socio-economic advantages, while Romm (2004)

stresses the current costs and performance factors that disadvantage hydrogen. Central to

these debates are the various so-called chicken-egg dynamics "that need to be overcome"

(National Academy of Engineering 2004) For example, drivers will not find HFCVs

attractive without ready access to fuel, parts, and repair services, but energy producers,

automakers and governments will not invest in HFC technology and infrastructure

without the prospect of a large market (e.g. Farrell et al. 2003, National Academy of

Engineering 2004). The non-compatibility of an infrastructure with that of the existing

gasoline network is a major issue for most alternatives and past introductions of AFVs

have yielded mediocre results, despite subsidies and promotions. Ethanol in Brazil, CNG

in Argentina, and diesel in Europe are examples of large scale penetration and potentially

self-sustaining markets. In contrast diesel in the United States and CNG in Canada and in

New Zealand have fizzled after an initial period of sizzle. Most commonly however,

whether they are gaseous-, liquid-, or flex-fuel vehicles or electrics (EVs), alternatives

fail to exceed penetration levels of a few percent (Cowan and Hulten 1996; Di Pascoli et

al. 2001; Sperling and Cannon 2004; Energy Information Administration 2005).

The underlying dynamics are much more complex than simple chicken-egg analogies

suggest. **Table 1** lists various sources for dynamic complexity for AFVs. First,

competitive dynamics are determined by the interplay of several feedbacks: a transition

towards any AFV, but especially towards HFCVs, involves building of consumer

acceptance, automotive learning-by-doing that improves with production experience, co-development of complementarities, especially maintenance and fueling infrastructure, and investment synergies with non-automotive applications. Further, these interactions play out under a system of government incentives, but also in concert with public interest and media attention. Second, the system is distributed in various ways: a multiplicity of stakeholders has varying perceptions and conflicting goals (Bentham 2005); the adoption population is heterogeneous in physical and socio-economic space; and the alternative options for technology deployment are many and diverse. Third, elements in the system change with large time delays. Some of those elements are tangible, such as consumers' vehicle replacement times, while others are more difficult to observe, such as adjustment of consumers' perceptions of value, or of their familiarity with the technologies. Finally, many of these relationships are highly non-linear. For example, in the very early stages when there are few fueling stations, the marginal benefit of one or two additional fueling locations is very low for consumers but increases dramatically as the number of stations increases and returns to zero when stations are found on every corner.

The existence of such dynamic complexity in the early stage of a market formation process suggests that the evolution of new technologies such as these is likely to be strongly path dependent (David 1985; Arthur 1989; Sterman 2000). In such environments policymakers' and strategists' efforts to stimulate adoption can contribute to its failures. Consequently, in order to understand how policy can effectively stimulate adoption on a large scale, it is essential to have a quantitative, integrative, dynamic model with a broad boundary, long time horizon, and realistic representation of decision making

by individuals and other key actors. Such a model should take economic, social and cultural, but also technical and physical parts of the system into account. This thesis lays the groundwork for a behavioral, dynamic model to explore the possible transition from ICE to AFVs such as hybrids, CNG, and HFCVs. **Figure 1** shows a conceptual overview of the main feedbacks in the model. The approach emphasizes a broad boundary, endogenously integrating consumer choice, as conditioned by product attributes, driver experience, word of mouth, marketing, and other channels, with scale economies, learning through R&D and experience, innovation spillovers, and infrastructure. The full scope for such a model is discussed in more detail in Struben and Sterman (2006).

In this essay I analyze one of the mechanisms in depth: the dynamics resulting from interactions between AFVs' adoption and the necessary fueling infrastructures. To support my analysis of the critical mechanisms, I develop a dynamic behavioral spatial simulation model. A full policy analysis requires a model that integrates infrastructure dynamics with the other feedbacks. However, such an integrated model will be complex and its behavior difficult to understand. This essay builds an understanding of the complex dynamics surrounding the infrastructure question as a foundation for an integrated analysis. Similarly, Essays 1 and 3 analyze other key feedbacks: Essay 1 focuses on key interactions between consumer familiarity and adoption; Essay 3 focuses on the dynamics of performance improvement of alternative fuel vehicles through learning-by-doing and R&D, and spillovers between them. The analysis in this essay as well as in the others provides an understanding of the dynamics that are associated with the integrated framework.

Understanding the dynamics that result from the interdependency of vehicle adoption and development of fueling infrastructure is critical for achieving the successful introduction of various AFVs. Infrastructure development is considered to be one of the biggest challenges for HFCVs (Farrell et al. 2003, National Academy of Engineering 2004, Ogden 2004), but is also central to diffusion of other AFVs, whether CNG (Flynn 2002), prospective bio-fuel vehicles, or even plug-in hybrids. While the dynamics result from demand externalities that lay behind the complementary character of vehicles and their fueling infrastructure, the actual underlying mechanisms are more subtle. Ascertaining when the market can be self-sustaining, or when incentives or coordination are critical - and if so, to what extent, and how - requires knowing how demand for fuel, vehicle adoption multiplied with desired travel behavior, grows with infrastructure as well as the economics of infrastructure supply in the early transition.

An earlier transition, from the horse-driven to the horseless carriage at the turn of the 19th century, with ICE as the eventual winner, can serve as a useful starting point for building an understanding of the co-evolutionary dynamics between vehicle demand and fueling infrastructure. In those days ICE vehicles and the fueling infrastructure co-evolved gradually over time. Slow evolution was possible because the need for long-distance automotive travel had not developed. First, long-distance travel services were provided by the rail network, while proper roads, especially between settlements were virtually nonexistent. Second, there existed limited experience and familiarity with the idea of driving for pleasure. Third, cars frequently broke down. Together these conditions hardly

provided incentives to extend the road network. Further, as touring by individual

transport was a novelty in the early days of the automobile, the initial adopters were

adventurous and willing to put up with inconveniences, such as the problem of finding

fuel. Thus, early on proper refueling facilities were only required in urban settlements.

Later, around 1900, gasoline also became available at local retail shops all over the

country, allowing, in a period where touring became ever more popular and road

construction grew, for a gradual diffusion of demand to more remote areas (Geels 2005).

Thus, the emergence of a gasoline fueling network through local pockets that gradually

connected to each other was a viable, though slow path for ICE in the early 1900s.

In contrast to this, contemporary consumers are accustomed to a dense, high-performing

network of fueling infrastructure. Consumers demand high levels of service along the

dimensions of availability, speed and convenience for all their trips. Such demands

greatly constrain the viability of an alternative transportation fuel when the infrastructure

is developing. Figure 2 (on left) illustrates the feedback that lays behind this, and what

policymakers term the "chicken-egg" problem (Farrell et al. 2003). To increase the

attractiveness to drive, the availability of fuel needs to be sufficient, and likewise, without

considerable expectations about demand, investors have no confidence to invest in and

commit to building and expanding a significant fueling infrastructure. Figure 2 (on right)

illustrates the conditions for such a tipping point graphically. It depicts vehicle demand as

a function of the number of stations. Starting with only one fueling station, no one is

willing to adopt, or drive. When the fueling infrastructure grows, demand grows at

increasing rate as more factors favor adoption: initially only short trips for a few are

covered, subsequently some people can also make longer trips and trips fpr those already covered are more convenient. This encourages more adoption and more consumption per vehicle. Demand growth flattens when the average station distance becomes small enough, not bringing significant additional benefits to drivers, and eventually demand becomes irresponsive to an increase in the number of fueling stations. Assuming cumulative industry costs of fuel supply grow linear with infrastructure, the S-shaped demand curve intersects the cost curve at a critical point, above which the industry is profitable and the market is self-sustaining.

In order to test this hypothesis, I analyze the detailed mechanisms underlying the, thus far high-level, concept of chicken-egg dynamics. Rather than treating fuel station development as independent, various sources of dynamic complexity - feedbacks between demand and supply, distributed decision making, time delays and non-linearities are taken into account. Further, it is critical to appreciate that feedback between fuel supply and demand is mediated through interactions that are non-uniformly distributed in space. For example, households in urban areas will not be satisfied with fuel services limited to their home locations. They also want to make long trips. An urban dweller living in San Francisco, also wants to make an annual trip to the Yosemite national park, or to Las Vegas and they require fueling infrastructure in these distant places. The consumer's utility includes the distribution of stations through space. This interaction in space, across settings with a heterogeneous population distribution, strongly contributes to the non-linear and distributed characteristics of the transition dynamics (Table 1 has the spatial component explicitly listed). In this essay the chicken-and-egg dilemma is

93

explicitly modeled by considering consumers' choice for adoption, driving and refueling, as well as the fuel station entry, exit and capacity adjustments in response to and anticipation of fuel demand developments. These infrastructure developments in turn feedback to change the consumer's trip convenience. [9]

This essay begins with a brief motivation of the modeling approach and an exposition of the conceptual model. Next I present the formulation of the spatial dynamic behavioral model and the analysis that is based on the simulations of the model. While the model is generally applicable, the analysis uses the state of California as a laboratory. I discuss the finding that low adoption levels, with clusters concentrated in urban areas, form a bi-stable equilibrium. I identify and discuss the technical and economic parameters to which the dynamics are particularly sensitive. Finally, the counterintuitive finding that the introduction of more fuel efficient AFVs can yield larger thresholds for a successful transition is discussed. The analysis demonstrates that the behavioral assumptions are critical to understand such phenomena. In the conclusion and discussion section I suggest that relying on the standard assumptions, such as exogenous demand or supply is problematic. To understand how policy can effectively stimulate AFV adoption on a large scale, a quantitative, integrative, dynamic model with a broad boundary, long time horizon, and realistic representation of decision making by individuals and other key

---

[9] An earlier version Struben (2005) generated the insight of clustering through a one dimensional spatial model with a short patch length. The current model develops a much richer structure, provides deeper insight into the dynamics and the role of various other feedbacks, and explores alternative policies and strategies, including supply/demand side subsidies/taxes. Further, it allows for calibration.

actors is essential. The essay ends with a discussion of implications for policy, in particular for the transition challenges for AFVs, and further work.

Modeling spatial behavioral dynamics

The existence of chicken-egg challenges between AFV adoption and fueling infrastructure development are well known (e.g. Farrell 2003; Ogden 2004; National Academy of Engineering 2004). However, a careful analysis of the co-evolutionary dynamics of market formation of AFVs and fueling infrastructure has not been conducted. I conduct such an analysis. Vehicles and their fueling infrastructure are strong complementarities (Katz and Shapiro 1994). However, their short and long-range interactions result in significantly more complex issues than basic hardware-software analogies justify. Before laying out the conceptual model, I discuss briefly existing approaches to problems that have a spatial, behavioral and/or dynamic character.

Transportation and travel research has a long history of modeling demand and supply in space (e.g. Fotheringham 1983). This research has mainly focused on identification of least cost optima (e.g. Collischonn and Pilar 2000), or equilibrium (e.g Lefeber 1958) distributions, and has grown enormously since Dijkstra (1959) published his shortest path algorithm. However, to allow for a detailed computation of trajectories, there is limited room for dynamics. In most of these studies either the demand or the supply side is assumed to be fixed over time. Such approaches are suitable for problems of more static character - explaining the existence of certain equilibria of travel demand -or to study the effect of an optimal solution to marginal changes within an established system - the

impact of a new highway on current traffic flows. This model benefits from particular

concepts developed in this literature, such as shortest path algorithms and gravity demand

models. However, the market formation processes associated with AFV transitions

involve situations of disequilibrium and the potential existence of multiple equilibria

requires focus on the dynamic interrelationship of supply and demand.

Interest in the formation of spatial patterns through reinforcing and balancing feedbacks

took-off after Turing (1952) introduced physio-chemical diffusion reaction structures, or

"Turing Structures". Such patterns are likely to be found where the movement and range

of influence of actors is small compared to the global scale, leading to strong local

correlations. With increases in information processing capabilities, problems throughout

the sciences including problems in statistical physics (e.g. Ising models), material physics

(crystal growth and the process of solidification, or dendrites (Langer 1980)), and organic

surface growth (diffusion limited aggregation, (e.g. Witten and Sander 1981)) were

addressed. Similar trends are found in the social sciences, for example in aggregation and

geographical economics (Krugman 1996).

The field of economic geography has a longstanding history in spatial dynamic problems,

appreciating that the actual location of activity deviates from the optimum location

(Lösch 1940; Christaller 1966). Modern, formal applications focus on the tension

between "centripetal" and "centrifugal" forces regarding geographical concentration

(Krugman 1996). While dynamic, these models seek to filter out core mechanisms from

each of the two competing forces that are perceived to be dominant (Fujita and Krugman

1999, 2004). With little prior understanding of the dynamics of the system, as is the case with technology transitions as the AFVs, a richer set of behavioral feedbacks needs to be included. In this case, it is the combination of spatial heterogeneity with the detailed behavioral feedbacks that gives rise to the dynamic complexity.

The dynamic behavioral spatial model presented in this essay demonstrates that relaxing the assumption that supply and demand directly adjust to clear the market, and including many of the behavioral aspects, leads to transition dynamics that are more diverse than can otherwise be observed. The model captures endogenous driver behavior including decisions regarding the adoption of AFVs, the mode of transportation (AFV, or other) for each trip, where to refuel, and "topping off" behavior. These decisions are influenced by driver concern for the risk of running out of fuel, service times, and how far one has to go out of one's way for refueling. Similarly, on the supply side, decisions about fueling stations, entrance, exit, location and expansion decisions are endogenous. These behaviors mediate interactions that are different over short- and long-distance and could drive dynamics that cannot be observed with mean-field approaches.

Figure 3 provides a spatial representation of the model structure and illustrates at a high level how interactions between supply and demand are captured. For illustration a grid structure is shown overlaying an area representing greater Los Angeles. The area is divided into patches, or zones, the darker ones having a larger household density.

97

Households locations are indicated by index z.[10] Households wish to make trips to various places outside their patch location, for work, leisure, and other purposes. While any set of desired trips can be generated, and thus various types of drivers can be represented, in this essay the distribution of trip destinations z' is assumed to be lognormal in their length l, and randomly distributed in direction θ. I capture boundary constraints properly by disallowing non-feasible trips, such as those that would lead into the ocean. For each zone, the average household's trips are normalized to equal the average vehicle miles of the population. The actual trip choice is endogenous. Drivers will choose whether and how often to travel to a particular location based on their assessment of how difficult the trip will be, including the travel time, the risk of running out of fuel, and the likely extra time and effort involved in finding fuel (the need to go out of their way to find fuel if it is not available on their main route). Similarly, households select between vehicle platforms depending on perceived utility of using it for the trips that they desire to make. The location of fueling infrastructure is also endogenous. Station entry and exit are determined by the expected profitability of each location, for example in zone z', which, in turn, depends on the demand and expected demand for fuel at that location and the density of competition from nearby stations.

For the analyses, I define the patch sizes such that heterogeneity at the scale of typical trip behavior is captured. For more specific analysis, the model can be setup with a finer grid, and with more technical detail, however this will put significant pressure on scarce

---

[10] Throughout this Essay I will use zones and patches interchangeable, the first representing the geographical boundary, and the second being the formal term used in spatial models.

resources, as they dramatically increase computation time, make analysis harder. Finally, lower level provides significant data challenges. Most importantly however, as I will justify in the analysis, a finer level of detail contributes noise, but does not change the fundamental dynamics. For the same reasons the model does not include technical details, such as traffic flows, or highly disaggregated agents, representing large variation of consumer types.

Figure 4 shows a conceptual overview of the main feedback loops in the model that result from behavioral assumptions. Feedback (R1) describes the basic chicken-egg dynamics. An increase in the number of stations of platform i in a zone z, lowers refueling efforts for trips to or through z for households living in a nearby zone z' (depending on their normal trips to/through that area). This increases the attractiveness of driving and raises platform i's market share in that area. A larger number of adopters generates more demand around z, increasing station utilization, sales and finally profitability, contributing to industry-level profits, which increases fuel station entrance for this platform (B1), until fuel station sales and profits are reduced to critical levels. However, those who have already adopted the platform also experience a decrease of trip efforts, induced by a higher number of stations, which leads to an increase of the fraction of trips for which the alternative vehicle is used, rather than a conventional vehicle or other transport modes (R2). High station utilization is good for profitability, but also leads to increased crowding (B2), requiring an increase in the drivers' efforts to refuel, and thus lowering their adoption, and likewise lowering vehicle miles through that region. Finally, within a zone z, higher profitability also leads to a larger share of the

entrants in that particular zone (R3), fewer exits (R4), and capacity expansion (B3) (more pumps), by existing stations. Finally, while not explicitly shown, in response to an inconvenient distribution of fuel along the route, drivers can raise the tank level at which they top off (topping off well before the warning light goes on). For this they trade off an increase in refueling effort for the need to go out of their way to refuel.

These concepts together define the inherent spatial, dynamics between vehicle fleet demand and fueling infrastructure. Combined with other feedbacks, this structure governs the co-evolutionary dynamics among the elements of an alternative-fuel-based transportation system. However, for analytical clarity the model is restricted to the interactions between infrastructure and vehicle demand only.

The Model

In this section I provide an exposition of the model: the demand-side structures for vehicle adoption; the trip, route, and refueling choices. This is followed by a more detailed discussion of the components of trip effort and the supply-side decisions, including entrance, exit, and capacity adjustment.

Adoption

The total number of vehicles for each platform $j = \{1,...,n\}$, in region z, V_{jz}, accumulates new vehicle sales, s_{jz}, less discards, d_{jz}

$$\frac{dV_{jz}}{dt} = s_{jz} - d_{jz} \qquad (1)$$

Ignoring the age-dependent character of discards, and assuming a total fleet in equilibrium, this implies that purchases only involve replacements: [11]

$$s_{jz} = \sum_i \sigma_{ijz} d_{iz} \qquad (2)$$

where σ_{ijz} is the share of drivers of platform i living in location z replacing their vehicle with platform j.

Consumers base their adoption decision on a range of vehicle attributes: vehicle price; power; operation and maintenance; safety; drive range; effort and cost of driving. I capture this by integrating diffusion models with discrete consumer choice theory (McFadden 1978; Ben-Akiva and Lerman 1985). These are often applied to transport mode choice (Domencich et al. 1975; Small et al. 2005), and automobile purchases (Berry et al. 2004; Train and Winston 2005), including alternative vehicles (Brownstone et al. 2000; Greene 2001). Then, the share switching from i to j depends on the expected utility of platform j as judged by the driver of vehicle i, in location z, u_{ijz}^e. Hence,

$$\sigma_{ijz} = u_{ijz}^e \Big/ \sum_j u_{ijz}^e \qquad (3)$$

While drivers may be generally aware that a platform (such as CNGs or HFCVs) exists, they must be sufficiently familiar with that platform for it to enter their consideration set, which I model in Essay 1 by its degree of familiarity F_{ijz}, with $u_{ijz}^e = F_{ijz} u_{ijz}$, where u_{ijz} is the perceived utility of platform j by a driver of platform i in region z. Further, for those platforms considered, expected utility depends on perceptions regarding the set of vehicle attributes a_{ijlz} which represents the performance of platform j with respect to attribute l,

---

[11] See appendix 1a of Essay 1 for the age-dependent structure and appendix 1b of Essay 1 for the initial sales structure.

for a driver of platform i in region z. Driver experience with and perceptions about various characteristics of each platform may differ significantly even if individuals have identical preferences. For example, drivers of HFCVs experience the actual availability of hydrogen fueling stations in their local environment. However, drivers of other platforms who consider buying a HFCV have to learn about these services through various indirect channels, and do not know the exact levels of convenience for their trips. Similar issues relate to attributes associated with vehicle performance. This diffusion process of knowledge about attribute performance is discussed in Essay 1 which shows that it has a significant impact on adoption dynamics. While the socialization dynamics associated with drivers' familiarity and consumers' learning about the performance of the various platforms are important for overall dynamics, here I focus purely on dynamics related to the demand and infrastructure. Therefore I set $F_{ijz} \equiv 1 \forall i, j, z$ and $a_{ijlz} \equiv a_{jlz} \forall i$, where a_{jlz} is the perceived performance of an attribute l to any consumer in z. Consequently, expected utility is identical for all drivers, and equals utility based on the perceived efforts a_{jlz} part of the set L, $u_{ijz}^{e} = u_{jz}$.

Appendix 3a of Essay 3 discusses the general structure capturing the relevant attributes, and their changes, in more depth. Of the many relevant attributes, only the trip convenience is directly affected by the abundance of fuel stations and is thus a central attribute, which yields a utility contribution u_{jz}^{t}. This component will be discussed in the next section. For arguments of consistency, the model must explicitly capture those attributes that are affected by parameters that vary supply and demand elsewhere in the model. For example, the maximum action radius of a vehicle (which correlates with, but

is not identical to, trip convenience), influences not only a consumer's purchase decision, but also influences the number of fuel station visits by drivers, and thus utilization; supply is affected in a non-trivial way. For the same reason, we capture operating cost (which is a function of fuel price that also affects supply) and fuel economy (which affects demand, as well as fuel station visits). We capture these under attributes a_{jlz}. All other attributes by which AFVs may differ, such as vehicle power and footprint, are aggregated under the vehicle-specific term u_{jz}^0. Using the standard multinomial logit formulation we can now state:

$$u_{jz} = u_{jz}^0 u_{jz}^t \exp\left[\sum_l \beta_l \left(a_{jlz}/a_l^*\right)\right] \quad (4)$$

where β_l represents the sensitivity of utility to performance of attribute l.

Trip, route, and refueling choice

Consumers not only decide to purchase vehicles but also how to use them – their driving patterns. Drivers wish to take trips to various places around their home for work, leisure, and other purposes. But trip choice is endogenous. Drivers will choose whether and how often to travel to a particular location based on their assessment of the difficulty of the trip. Drivers select their favorite routes and refueling locations as a function of the availability of fuel.

Determination of refueling effort is explained later. **Figure 5** illustrates how the motivations for consumers' adoption choice, and drivers' trip, route, and refueling choices are captured. The diagram on the left shows the high-level structure: first, as discussed above, consumers in region z, decide on adoption, with share σ_{iz} going to the i-

103

th platform. This share depends, among other factors, on the utility component u_{iz}^{t} to adopt

an AFV. Similarly, the fraction of trips from one's home z to destination z' for which the

AFV is used, $\sigma_{izz'}$, conditional upon prior adoption, depends on their utility to make that

trip $u_{izz'}$. Further, consumers' aggregate utility to drive u_{iz}^{t} depends on the utility derived

from making each trip, weighted by $w_{zz'}$. Going further down the diagram, for each trip,

consumers decide on the route to follow, with share $\sigma_{i\omega_{zz'}}$ depending on the relative

utility for each route one might consider taking, $u_{i\omega_{zz'}}^{t}$. This route utility, also determines

the trip effort of the average consumer in $z\, u_{izz'}^{t}$, weighted by its shares. Finally, in a

similar fashion, drivers decide where to refuel along the route, $\sigma_{i\omega_{zz'}s}$. The refueling effort

is determined by other factors that will be explained later.

The right-hand side of **Figure 5** shows the functional forms that determine the share and

the effort variables. For each choice type the share is determined through a logit-

expression, as listed in column 1. For example, row 1 describes the derivation for the

vehicle adoption share and average efforts to drive that have already been discussed.

Columns 1-2 yield exactly equations (3) and (4) for the vehicle choice decision. A

driver's trip choice involves a driver *i*'s decision on the mode of transportation for a trip

from *z* to *z'*. The fraction of trips that their alternative vehicle is used depends on the

utility for that trip, $u_{izz'}^{t}$ compared to using another mode of transportation that is

available $u_{zz'}^{o}$.

The experienced utility of driving is a non-linear, weighted average of the various trips, as shown by column 3 in **Figure 5**. To represent the effort, several functional forms are possible. The form used and shown in column 3 is the constant elasticity of substitution (CES) function (McFadden 1963; Ben-Akiva and Lerman 1985; see Essay 3 and its Appendix 2e and Appendix 3a of that Essay for expansions on this function). Households' total trips from/to an area (trip generation), combining residential and job locations and trip distribution (location of these trips) are constant. This generates a desired trip frequency distribution per household, $T_{zz'}^{\max}$. The utility to drive is a weighted average over the utility derived from each trip that is part of a driver's desired trip set $T_{zz'}^{\max}$. The weight $w_{zz'}$ can be a function of anything, but I assume it increases with frequency and distance. For example, long-distance trips, while less frequent, could be considered very important (see Appendix 4a):

$$w_{zz'} = g\left(r_{zz'}; T_{zz'}^{\max}\right) \Big/ \sum_{z' \in T_z^{\max}} g\left(r_{zz'}; T_{zz'}^{\max}\right) \tag{5}$$

The parameter μ^t of the CES function can be interpreted as follows: the case where individual consumers make only one unique type of trip corresponds with $\mu^t \to 1$, which means that utility captures the weighted average across all trips, and the expression of vehicle share converges to a standard multinomial logit expression. The case in which individuals make many distinct trips corresponds with $\mu^t < 1$, with the extreme case being $\mu^t \to \infty$, where perceived utility of driving equals that of the individual trip that is perceived to provide the worst utility (in this case trips can be seen as full complements of each other). In the special case $\mu^t = 0$, the aggregate utility equals the utility of the (weighted) average trip.

105

Going further down the hierarchy, in **Figure 5**, the modal choice (Small 1992) of each

trip is endogenous and depends on the fraction of trips between z and z' that are taken

with the alternative fuel platform i, $\sigma_{izz'}$, with the actual frequency of trips for drivers

living in z, owning platform i, with destination in z', with platform i, $T_{izz'} = \sigma_{izz'} T_{zz'}^{\max}$.

Small (1992) offers a long list of factors that influence drive effort, including travel time,

on-time arrival fraction, operating cost, parking. Here we concentrate on the role of fuel

availability. We differentiate i) the normal drive time for a route ω between z and z',

$a_{\omega zz'}^{t0}$ without any refueling; ii) the factors that depend on the availability of fuel, which

include the risk of running out of fuel, and the likely extra time and effort involved in

finding fuel (the need to go out of one's way to find fuel if it is not available on the main

route), which are experienced in the location s where one seeks to refuel, a_{is}^{f} ; iii) all other

factors are aggregated in one effect on trip utility $u_{izz'}^{0}$.

The share of trips between z and z' taken by platform i is derived through a binomial

choice expression, comprising the utility to drive trip $u_{izz'}^{t}$, of driving trip zz' with

platform i, and the combined alternative $u_{zz'}^{o}$ (capturing alternative modes of

transportation and the opportunity cost of not going). A driver's trip utility is the

composite over routes that are part of the route set for trips from z to z', However, in this

case, it is assumed that individual drivers have one favorite route (which can be adjusted),

and $\mu^{\omega} \to 1$. Working our way down **Figure 5**, the perceived effort to drive an individual

trip is experienced on the route. The elasticity parameter β^{ω} represents a driver's

sensitivity to changing routes. If the sensitivity would be large, different drivers would tend to take the same route. The average route effort, $a^t_{i\omega_{zz'}}$, is approximated by the sum of the route effort, in absence of refills $a^{t0}_{\omega_{zz'}}$ and the expected refills per trip, $\phi_{i\omega_{zz'}}$ multiplied by the average effort of refueling (see Appendix 3b), which is the sum over refueling at any location, $a^f_{i\omega_{zz'}} = \sum_{s \in \omega_{zz'}} \sigma_{i\omega_{zz'}s} a^f_{is}$, weighted by the refueling share.

Finally, drivers adjust their refueling behavior and driving, based on variation in perceived effort utility of refueling for trip *zz'*. The refills along the route that locations *s* receive, with share of the total $\sigma_{i\omega_{zz'}s}$, depends on the length of the route that passes through an area $r_{\omega_{zz'}}$, but also on the effort it takes to refuel, within each location.

The ability to select more convenient locations depends critically on refueling behavior. Frequently running the tank down close to empty implies the consumer constrains himself to refueling at locations available when the tank is empty, which would imply refueling shares are constrained to be according to the relative distance that is driven through each location. Such behavior works well when stations are abundant everywhere, as is currently the case with gasoline, and reduces the frequency, and thus total effort, of refueling. At the other extreme, however, when topping-off occurs at extremely higher tank levels (before the warning light goes on), the freedom of choice for refueling becomes limited again. When top-off levels are between these two extremes, the freedom to select those locations that are most attractive for refueling is larger (at the expense of increased refueling frequency). The tank level (converted to miles) available when

consumers refill is referred to as the buffer. The number of miles driven between a full tank and top-off is referred to as the effective range (see Figure 6).

More formally, the effective range between two refills r_{iz}^f equals the maximum range r_i^f minus the average buffer that remains when refueling, r_{iz}^b:

$$r_{iz}^f = r_i^f - r_{iz}^b \tag{6}$$

where $r_i^f = \eta_i^f q_i$, with η_i^f the fuel efficiency and q_i the energy storage capacity of a tank. The refueling sensitivity parameter β^f determines the sensitivity of refueling shares that go to the various locations to a change in the consumers perceived utility to refuel. (see **Figure 5**). Running the tank always empty does not give any freedom of choice to select a more favorable location, thus, $r_{iz}^b \to 0 \Rightarrow \beta^f \to 0$. Reducing the effective tank range too much provide the same constraints, $r_{iz}^b \to r_{iz}^f \Rightarrow \beta^f \to 0$. However, when the buffer and effective range are on the order of the trip length, the freedom of choice is large, or, $r_{iz}^f / r_{i\omega_{zz'}}^t \wedge r_{iz}^b / r_{i\omega_{zz'}}^t \cong 1 \Rightarrow \beta^f \to \beta_{ref}^f$. Then we can state:

$$\beta^f = \beta_{ref}^f g\left(r_{iz}^b / r_{i\omega_{zz'}}^t \right) h\left(\left[r_{iz}^f - r_{iz}^b \right] / r_{i\omega_{zz'}}^t \right); \begin{cases} g(0) = 0; g(1) = 1; g' \geq 0 \\ h(0) = 0; h(1) = 1; h' \geq 0 \end{cases} \tag{7}$$

Where β_{ref}^f is determined by the physical constraints of refueling elsewhere. Typically, the functions h and g can be expected to be concave because of the increasing effect of the physical constraint of refueling. Appendix 2a provides the functional forms used in the model.

Finally, the length of trip equals the sum of the normal route and average distance when refilling, which equals the refills per trip, $\phi_{i\omega_{zz'}}$, multiplied by the average distance one is required to go out of the way for refueling,

$$r_{i\omega_{zz'}}^{t} = r_{\omega_{zz'}}^{t} + \phi_{i\omega_{zz'}} \sum_{s \in \omega_{zz'}} \sigma_{i\omega_{zz'},s}^{f} r_s^{f} \tag{8}$$

See Appendix 3c for the derivation of the refills per trip. This completes the formulation of the consumer decision-making processes regarding adoption, trip choice, route choice, and refueling location. The endogenous component that affects all of these is the trip effort explained below.

Components of trip effort

The normal effort for a route is expressed in time units and is given as

$$a_{\omega_{zz'}}^{t0} = \tau_{\omega_{zz'}}^{d} = \sum_{u} r_{\omega_{zz'},u}^{t} \big/ v_u \tag{9}$$

The speed may depend on the region, for example, the drive time associated with driving an extra mile in a congested urban area is much longer than on a rural highway.

We model the experienced refueling effort in each location as a weighted sum of: (i) the effort to find fuel a_{is}^{d}, which depends on the time spent driving out of one's way to reach a fuel station; (ii) the risk of running out of fuel a_{izs}^{r}, which depends on vehicle range and the location of fuel stations relative to the driver's desired refueling needs; and (iii) servicing time a_{izs}^{x}, which depends on wait times resulting from local demand being higher than the refill capacity at fuel stations. The experienced trip effort in location s is the weighted sum of each of these three components:

$$a_{izs}^{f} = w^{d} a_{is}^{d} + w^{r} a_{izs}^{r} + w^{x} a_{izs}^{x} \qquad (10)$$

The relative value of the weights w^{d}, w^{r} and w^{x} can be interpreted as the relative sensitivity of a driver's utility to a change in these effort components. The out-of-fuel risk involves a cost and time component. The drive and service time both involve time components, but the experience of time is not necessarily the same in each case. A large body of transportation research is devoted to how commuters and other travelers value their time (e.g. Steinmetz and Brownstone 2005); reliability (e.g. Brownstone and Small 2005); and related attributes (e.g. Small 1992; McFadden 1998; Small et al. 2005). The perception of time or cost associated with additional trip efforts may vary considerably by type of trips (recreational, business), individual, and activity (waiting in line to refuel vs. driving to a station). This explicit formulation allows taking the valuing of time into consideration, if it is deemed to be importantly influencing the dynamics. Appendix 5a provides a discussion of the elasticity of utility to a change in the various components.

The effort to find fuel is expressed as the search time, which is the average distance to a station divided by the average driving velocity in region *s:*

$$a_{is}^{d} = \left\langle r_{is}^{d} \right\rangle \Big/ v_{s} \qquad (11)$$

The value of $\left\langle r_{is}^{d} \right\rangle$ depends on fuel station density and can be analytically derived, which is done in appendix 3d.

The second component of driving effort, the perceived risk of running out of fuel within region *s* can be captured by assuming that a combination of experiences and individual

assessments yield results that are qualitatively similar to the expected out-of-fuels per

refill $\langle o \rangle_{izs}$ within a region s:

$$a^r_{izs} = \langle o_{izs} \rangle \tag{12}$$

Expected out-of-fuels is found by integrating over the probability of not reaching a

station within its range, with the refueling buffer r^b_{iz} being the average. The probability

further decreases with station density in region s, and increases with the required distance

driven through that region s. Its full derivation is provided in appendix 3e.

Finally, the service component of the effort attribute is determined by the average

servicing time at the station

$$a^x_{izs} = \tau^x_{izs} \tag{13}$$

Figure 7 shows the main idea of the structure for servicing time. This expression

comprises waiting in line, which depends on station utilization, and the actual refueling

time:

$$\tau^x_{izs} = \tau^w_{izs} + \tau^f_{izs} \tag{14}$$

The refueling time has a variable component of actually operating the pump and a fixed

component (including paying and purchasing ancillary products), $\tau^f_{izs} = \tau^p_{izs} + \tau^0_i$. The

variable component is a function of the quantity demanded and the capacity of the

pumps:

$$\tau^p_{izs} = q_{iz} / k^p_i \tag{15}$$

Average quantity demanded depends on tank capacity, adjusted for the effective top-off

levels (see Equation(6)):

111

$$q_{iz} = q_i \left(r_{iz}^f / r_i^f \right) \tag{16}$$

The wait component in equation (14) depends on the average demand versus capacity.

The expected time that customers must wait depends very non-linearly on the station utilization and the number of pumps, as suggested in **Figure 7**. When the number of pumps is relatively high, say 8, the average wait time will remain low, even for reasonably high utilization. This is because the expected number of empty service points upon arrival remains high. However, when stations have only one or two pumps, for the same station utilization, we are less likely to find an empty pump. Thus, in this case the average wait time for service can be large, even at reasonably low levels of utilization. Representing this relationship is important, especially when we realize that in initial stages, and in particular in those regions where demand is critically low, we might expect stations to be small. This is captured using a simple queuing theory. The wait time depends on the average refill time for that location, τ_{is}^f, given by the mix of demand and equations (15)-(16), the station utilization υ_{is}^f, and the number of pumps per station, y_{is} (discussed below). The resulting mean waiting time is

$$\left\langle \tau_{is}^w \right\rangle = \frac{P_{is}^q}{y_{is} \left(1 - \upsilon_{is}^f \right)} \tau_{is}^f \tag{17}$$

where P_{is}^q is the probability of finding all pumps busy (which is itself a highly non-linear function of average refill time, the utilization, and the arrival rate). Details of how the mean waiting time is derived through application of basic queuing theory and the station utilization are provided in Appendix 3f.

112

It is noteworthy to mention that all expected values and averages expressed in equations

(11)-(17) are derived through probabilistic calculus, as functions of station concentration

or demand in each region and do not involve additional assumptions or parameters

(Appendix 3).

Search time, out-of-fuel risk and service time are based on perceived values of station

density (for search time and out-of-fuel risk), and the wait time at the pump. They adjust

to the actual values with time delay τ^s.

The total vehicle miles driven per year by drivers of platform i equal,

$$m_{iz}^v = \sum_{z's} \phi_{i\omega_{zz'}s} m_{izz'}^v T_{izz'} \left(\upsilon_{iz_u} d_{iz_u} / k_{iz_u} \right) \qquad (18)$$

with utilization υ_{is} and demand d_{is} as derived in Appendix 3f, in the derivation of the

mean waiting time for service. This completes the consumer segment of the model and

the description of how the distribution of fueling stations is influenced by consumers'

decision to adopt a vehicle that is compatible with the fueling infrastructure, as well as

their trip, route, and refueling choices. Supply formation which occurs partly in response

to existing demand is described in the next section.

Fuel Station economics

Before discussing the supply-side decisions, I first set up the basic fuel station

economics. Next, the decisions made by the (potential) fuel station owners, which include

entrance, expansion, and exiting are examined. Stations can serve consumers with various

product mixes. For example, a station with 8 pumps can have 8 gasoline pumps or 6 gasoline and 2 diesel pumps. Throughout this essay, for the purpose of analytical clarity, I ignore explicit modeling of multi-fuel stations and therefore can distinguish stations by the fuel they serve, indexed by v. This is reasonable as a first order approximation, as most of the scale economies do not apply across fuel type. The role of multi-fuel stations will be discussed in later work. Average profits for stations of type v in region s equal revenues r_{vs} minus total cost c_{vs}:

$$\pi_{vs} = r_{vs} - c_{vs} \qquad (19)$$

Revenues equal sales from fuel multiplied by price p_{vs}, and revenues from (net) ancillary sales r_{vs}^a are given by:

$$r_{vs} = p_{vs}s_{vs} + r_{vs}^a \qquad (20)$$

Ancillary sales mainly involve convenience-store items and can account for up to 50% of profits. It might be that ancillary sales opportunities vary by platform. For example, hydrogen fuel stations might be seeking a wider set of services through complementarities with stationary applications, motivated by higher initial capital cost. This is possible for hydrogen because many services, such as maintenance, are not specialized enough, or because of complements with stationary applications. This would, of course, only work in populated areas. In all simulations ancillary sales will be set to a fixed amount per gallon consumed.

Station costs include a fixed, capacity-dependent component, c_{vs}^k, that represents such categories as land rent, equipment, and capital depreciation and a variable component that

114

increases with sales, having unit cost c_{vs}^u. The unit cost comprises feedstock cost c_{vs}^f; and

"other" c_{vs}^o that include electricity, labor, and taxes. Ignoring sunk costs (of starting a

station) and adjustment costs:

$$c_{vs} = s_{vs} c_{vs}^u + c_{vs}^k; \qquad c_{vs}^u = c_{vs}^f + c_{vs}^o \qquad (21)$$

Both fixed cost and unit cost can differ considerably by location, because of the large

contribution of rent, especially in urban areas. Unit cost can be different, because of

gradients in distribution costs. Fixed costs increase with capacity k_{vz} and are equal to

$c_{vs}^{k,ref}$ when the number of pumps y_{vs} are equal to y^{ref}:

$$c_{vs}^k = c_{vs}^{k,ref} f^k \left(y_{vs} / y^{ref} \right); f(0) > 0; f(1) = 1; f' > 0; f'' < 0 \qquad (22)$$

Scale economies are concave in the number of pumps (see Appendix 2b).

Sales are determined by station capacity and utilization υ_{is},

$$s_{vs} = \upsilon_{is} k_{vs} \qquad (23)$$

with station capacity being the product of the number of pumps and pump capacity

$k_{vs} = y_{vs} k_v^p$.

To complete the fuel station economics, price is set at fuel stock cost plus markup:

$$p_{vs} = \left(1 + m_{vs} \right) c_{vs}^f \qquad (24)$$

For simplicity we assume that fuel stock markups are constant.[12]

---

[12] Empirical data between 1960 and 2000 show that the average markups remain virtually constant, they
were reduced only after the first oil shock, when cost of fuel increased dramatically, suggesting a very slow
anchoring and adjustment process. (Sources: U.S. Wholesale Gasoline Price, US Bureau of the Census,
Statistical Abstracts of the United States 1950 &1976 & 1980 & 1994 & 2005; U.S. Retail Gasoline Price,

Supply decisions

Potential station owners also make decisions. **Figure 8** shows the entrance and exit behavior of stations. Potential entrants decide to enter the market based on perceived industry return on investment. Next, entrepreneurs decide where to locate, after which a permitting procedure results in construction and, finally, actual operation. Following this overview, we track the total number of fuel stations F_{vs} of type v, in region s which integrates entrance e_{vs} less exits x_{vs} :

$$\frac{dF_{vs}}{dt} = e_{vs} - x_{vs} \tag{25}$$

While the higher-order process is captured in the model, in this exposition I collapse the process of location selection, permitting, and construction into one, with aggregate entry time τ_e. Then, new-to-industry stations in region s, F_{vs}^n , enter the market as:[13]

$$e_{vs} = F_{vs}^n / \tau^e \ , \tag{26}$$

Where the indicated new-to-industry stations equal the new-to-industry capacity intended for region s, divided by the desired fuel station capacity k_{vz}^* ,

$$F_{vs}^n = K_{vs}^n / k_{vs}^* \tag{27}$$

Location s receives share σ_{vs}^k of the total new-to-industry capacity:

1949-2004 Annual Energy Review 2004 Report No. DOE/EIA-0384(2004) Accessed August 15, 2005 http://www.eia.doe.gov/emeu/aer/petro.html).

[13] The model includes the higher-order entrance process and allows for varying the extent to which the supply line is taken into account.

$$K_{vs}^{n} = \sigma_{vs}^{k} K_{v}^{n} \qquad (28)$$

High returns at the industry level lead to expansion of existing capacity, $K_{v} \equiv \sum_{s} k_{vs} F_{vs}$.

The total market for fuel v grows at rate g_{v}^{k}, which increases with industry profits:

$$K_{v}^{n} = g_{v}^{k} K_{v}$$
$$g_{v}^{k} = g_{v}^{k0} f^{g}\left(\overline{\pi}_{v}^{e}\right); f\left(\ll 0\right) = 0; f\left(0\right) = 1; f' \geq 0; f'\left(\gg 1\right) = 0; \qquad (29)$$

where $\overline{\pi}_{v}^{e}$ is the perceived returns minus the desired, normalized to the desired

$\overline{\pi}_{v}^{e} \equiv \left(\pi_{v}^{e} - \pi_{v}^{0}\right)/\pi_{v}^{0}$. The constraints imply, first, that the growth rate equals g_{v}^{k0} when

perceived returns on investment equal desired returns; second, that the growth rate

increases with return on investment, which could differ by fuel, because of potential

variation in constraints. Further, the shape is bounded, at zero, for extremely negative

profits, and, at some finite value, for extremely high returns. The most general shape that

satisfies these conditions is an S-shape (see appendix 2c for the exact functional form).

Finally, region s's share of total new capacity is a function of the expected relative return

on investment within each region, π_{vs}^{β}, compared to that of alternative regions. A logit-

expression is sensible, given the noise in the relevant information for those who have to

decide what area to locate in:

$$\sigma_{vs}^{k} = \exp\left(\beta^{k} \overline{\pi}_{vs}^{\beta}\right) / \sum_{z} \exp\left(\beta^{k} \overline{\pi}_{vs}^{\beta}\right) \qquad (30)$$

where β^{k} is the sensitivity, which depends on the accuracy of information on differences

in profitability. Expected return on investment is derived through a net present value

calculation of future profits streams π_{vs}^{β}, compared to the desired return on investment,

$\overline{\pi}_{vs}^{\beta} \equiv \left(\pi_{vs}^{\beta} - \pi_{v}^{0}\right)\big/\pi_{v}^{0}$. Entrepreneurs use heuristics to estimate how much demand would

be induced by their entrance., based on reference demand generated by existing transport

patterns (see appendix 2d).

Exits are driven by recent station performance and follow a standard hazard formulation,

where the hazard rate λ^{x} is a function of anticipated return on investment, $\overline{\pi}_{vs}^{x}$, compared

to a required profitability, π_{s}^{x} , $\overline{\pi}_{vs}^{x} \equiv \left(\pi_{vs}^{x} - \pi^{x}\right)\big/\pi^{x}$:

$$
\begin{aligned}
x_{vs} &= \lambda^{x} F_{vs}; \\
\lambda^{x} &= \lambda^{x0} f^{x}\left(\overline{\pi}_{vs}^{x}\right); f(0) = 1; f \geq 0; f' \leq 0; f'' > 0
\end{aligned}
\tag{31}
$$

where λ^{x0} is the exit rate when recent profits equal desired profits. A general shape that

satisfies these conditions is an S-shape, such as the logistic curve (see appendix 2c for the

exact functional form).

To determine their own anticipated return on investment π_{vs}^{x} , mature stations rely on

recent performance π_{vs} ; new to industry stations use their expected return on investment

figures, π_{vz}^{β} . The different emphasis is captured by the weight w_{vs}^{m} given to the recent

profits streams, which increases with the average station maturity:

$$
\pi_{vz}^{x} = w_{vs}^{m} \pi_{vs} + \left(1 - w_{vs}^{m}\right) Max\left[\pi_{vz}^{\beta}, \pi_{vs}\right]
\tag{32}
$$

where the weight increase is zero for entirely new to industry stations, and equals one for

old stations. A reasonable form is an S-shaped form, centered around the age m^{*},

$\overline{m}_{vs} \equiv m_{vs}\big/m^{*}$:

$$w_{vs}^m = f\left(\overline{m}_{vs}\right); f\left(\infty\right) = 1; f\left(1\right) = 1/2; f\left(0\right) = 0\ f' \le 0 \tag{33}$$

Station maturity is derived through a simple age co-flow function (Sterman 2000) that tracks the average age of fuel stations. Appendix 2e provides the selected functional form. Appendix 2c of Essay 1 describes the formulation of co-flow structures.

The final decisions to be described are decisions within industry to alter capacity of fuel stations. Existing stations adjust capacity, in terms of the number of pumps, to the desired level y_{vs}^* over an adjustment time τ_{vs}^k, accounting for both the time to actually learn about the optimal size, as well as the time to alter capacity, which can differ by region.

$$\frac{dy_{vs}}{dt} = \left(y_{vs}^* - y_{vs}\right)\!\Big/\tau_{vs}^k \tag{34}$$

where the desired number of pumps allows the utilization to reach its desired level:

$$y_{vs}^* = \left(\upsilon_{vs}^k / \upsilon_{vs}^{k*}\right) y_{vs} \tag{35}$$

where υ_{vz}^{k*} is desired utilization. Stations desire high utilization, as profits increase with utilization; however, very high utilization will lead to congestion at stations and customer defection. Thus, desired utilization is well below 1. A heuristic estimate, observing fuel stations gives utilization levels on the order of 0.2, that is, well below maximum utilization .[14]

---

[14] The desired utilization is therefore linked to the wait time in equation (17), with a likely optimum at the point where its slope begins to increase sharply, which is also well below full utilization. For less regular demand patterns, or fewer pumps, desired utilization would be lower. On the other hand, adjustment constraints can lead to a utilization that is higher than desired, while competition effects can render it lower.

This finalizes the model structure. Key decisions on the supply side were: market entry decisions that were based on expected NPV; exit decisions, in response to realized profits; fuel station location decisions, based on relative expected profitability between different locations; and finally, capacity adjustment, in response to utilization.

Analysis

The analysis begins with consistency tests, illustrated through a comparison with empirical data based on the state of California. Next key insights of the basic behavior of the model found by analyzing the introduction of a hypothetical AFV in California are discussed. Given these understandings, I discuss the generality of these results and explore the value of relaxing technical and behavioral assumptions. Finally, I analyze implications when technical and economic parameters are varied and discuss implications for the introduction of various types of AFVs.

Ie use the state of California as a reference region for analysis. That is, the demographic, economic, and technical parameter settings as well as the reference data, are equivalent to those typically found in California. **Table 2** provides a summary of the relevant statistics. The default parameters in the model are provided in Table 3, and are used for the simulations, unless otherwise stated. Parameter settings for particular simulations are discussed in the text for each figure. To determine behavioral parameters, such as the consumer sensitivity parameters, or those that relate to station entrance and exit, a combination of heuristics, published empirical findings, sensitivity analysis, and

calibration are used to select reasonable values. To simplify analysis and dynamics, one type of consumer is assumed: households that generate trips conforming to a frequency distribution $T_{zz'}^{\max}$ that is generated by a lognormal in distance with average trip length of 20 miles and , with a uniform distribution of the direction, subject to boundary conditions. If all trips would be made by vehicles, it generates the maximum 15,000 vehicle-miles per person per year.

Fundamental behavior

Several partial model tests, sensitivity analyses, and calibrations have been carried out to confirm behavioral consistency and heuristic parameter settings. Figure 9 shows, as an example, the results of a partial model test, which was to replicate the distribution and total number of ICE gasoline stations in California. Figure 9a shows the actual gasoline fuel station distribution in California in 2003 (N=7949) on a 625 patch grid. For the stations we used actual GIS data provided by the National Renewable Energy Laboratory. Throughout these simulations vehicle ownership was held fixed at 2003 levels (17.126e6) with a distribution identical to that of the population, with an adoption fraction equal to 0.91 throughout. For each trip destination, the desired fraction of trips to be performed with an ICE/gasoline vehicle was 0.8, which would yield the average of 12,000 miles per vehicle, if realized. Simulations began with 10% of current stations, uniformly distributed, with 8 pumps per station. Supply was subsequently allowed to adjust over time through entry, exit, and capacity adjustment. **Figure 9**b shows the simulated results based on the heuristic parameters, obtained without optimization.

Without relying on detailed data inputs regarding items such as traffic flows, the model performs quite well, though there are a few regions that over or underestimate the number of stations. For example, the model places some stations in mountainous regions, or deserts, while it ignores a disproportionately high number of stations in big transit hubs, e.g., to Las Vegas. While these deviations are small, it is easy to correct for such deviations, without much additional data being required. This is discussed later. On a final note, the model performs much better, compared with simulations where the number of pumps per station was held fixed at the average of 8, illustrating the relevance of such additional behavioral feedbacks.

Now we perform an analysis in which both supply and demand are endogenous. **Figure 10** shows the base case simulation. In the base case, the initial ICE fleet and infrastructure size and distribution are set to 2003 California values: 15.5 million vehicles and 7949 gas stations. In the base case, to emphasize the spatial co-evolution of vehicles and infrastructure, we assume full familiarity with AFVs and set AFV economic and technical parameters of merit equal to those of ICE. The simulation begins with an AFV adoption fraction of 0.1% and 200 fueling stations (these numbers approximate station values for CNG in California in 2002, including private fleets and fueling stations). We assume, optimistically, that all AFV fuel stations are accessible to the public. Initially, investors and other partners will be committed to and collaborate to achieve a successful launch and hence they attempt to keep stations open, even when making losses. We capture this by subsidizing, on average, 90% of a station's losses for the first 10 years. This scheme disproportionally favors those stations that are in more vulnerable locations and receive more support.

Figure 10 shows the alternative fuel stations and fleet. The top graph shows the simulated adoption fraction, stations, and fuel consumption, relative to normal, over time.

The bottom graph shows the geographical distribution of the adoption fraction and number of fuel stations at time t=45. Qualitatively, three important results are revealed. First, despite performance equal to ICE/gasoline and full familiarity with the AFV, overall diffusion is very low, especially in the early phase. Net fuel station growth initially lags that of the fleet.

Second, many stations are forced to exit when subsidies expire, while entry ceases somewhat earlier, as the expected value from subsidies starts to decline. Average capacity increases strongly after the shakeout (see number of pumps, right axis) because of two effects: a selection effect is that those who exit are generally the smaller stations; in addition, those that remain in business experience increased demand, which drives their capacity expansion. For the same reason average profitability increases dramatically. However, these effects have a limited effect on the overall demand growth. Eventually, with the gradual increase of demand and constraints in capacity expansion, station entrance accelerates.

Note that the growth of fuel consumption lags adoption, especially earlier in the simulation. This is because of the limit on the destinations that can be reached with the AFV because of absence of stations in rural areas and overcrowding in urban areas. Time to adopt and settle is much longer than one might expect from the time delays in vehicle replacement and station entrance only, which total up to 12 years for this simulation. This behavior is a result of closing the feedback between the interdependent relationship of vehicle demand and infrastructure development, each of which only gradually increases to an indicated level, as shaped by the other, and thereby also only slowly adjusts the goal for the other.

Third, the end state that emerges shows a spatially bi-stable equilibrium in which essentially all AFVs and fueling stations are concentrated in the major urban centers. Miles driven per year and actual consumption for the typical AFV are also far less than for ICE vehicles. Limited diffusion is a stable equilibrium in the cities, because high population density means fuel stations can profitably serve the alternative fleet, and low refueling effort induces enough people to drive the alternative vehicles. **Figure 11** shows the underlying hypothesis. Both urban and rural demand is subject to chicken-and-egg dynamics (R1, B1). For metropolitan areas, potential demand would be sufficient if all demand would be generated from within and to within; however, rural areas would never be able to generate a self-sustaining market. Though AFV fuel stations do locate in rural areas during the period they are subsidized, rural stations remain sparse, so rural residents and city dwellers needing to travel through rural areas find AFVs unattractive (demand spillover, R2). Further, urban adopters, facing low fuel availability outside the cities, use their AFVs in town, but curtail long trips (demand spread R3). Consequently, demand for alternative fuel in rural areas never develops, preventing a profitable market for fuel infrastructure from emerging, which, in turn, suppresses AFV adoption and use outside the cities.

Consideration of relaxing assumptions

The benefits and costs of expanding the model boundary are discussed in this section and provide further support for the insights. Central to the model structure is its ability to capture the dynamics of supply and demand that interacting through space. Therefore I discuss first the appropriateness of the level of spatial detail. The AFV introductory

scenario that was discussed earlier is used as the basis. Figure 12a illustrates the

sensitivity of the model behavior to changes of the patch length (the square root of the

patch area). To control for large rounding errors with very large patches, population

density is kept fixed at the California average (109 households/sqml). Tracing the

equilibrium adoption as a function of path length, the results show: no self-sustaining fuel

demand for the alternative when patch length are above 200 miles; equilibrium demand

peaking when patch length is about 100 miles; and convergence for patch length below

30 miles. The variation in the equilibrium demand for larger patches is explained as

follows: the extreme case of one single patch corresponds to assuming a uniformly

distributed population. This assumption does not bring out strategic location incentives

on the supply side, and, even under rich behavioral assumptions, will yield demand

/supply responses that correspond with the qualitative sketch in **Figure 2,** where demand

is adjusted for the number platforms. The single patch dynamics will therefore result in a

limited amount stable equilibria of which the number depends on the number of

competing platforms. In the case of two platforms, as here, there are at most three stable

equilibria. Two equilibria provide full adoption of either platform, and zero for the other.

Whether a third equilibrium allows for both platforms to be self-sustaining depends if, in

the case of equivalent platforms, if 50% of the demand yields a profitable market (see

appendix 4b that this is indeed the case). Whether such equilibrium is actually achieved,

depends on whether the subsidy schemes can bring the adoption/fuel stations past the

boundary separting the low and the 50% equilibrium. We see that in this case this did not

yield enough adoption for take-off towards the 50% adoption fraction. At more moderate

patch length of say 100 miles, some patches capture major urban clusters, but also their

hinter land. Within such patches, average potential demand is large enough to yield the penetration to 50%. Further, virtually all trips of drivers are covered within that area, resulting in more adoption and demand.

Under such assumptions of regional uniformity, expected distance to a station is identical from all locations within that region, however, the fraction of long trips fulfilled relative to short trips differs slightly. This is mostly because of the varying dependence of effort and out of fuel risk, for short and long trips. In sum, this level of granularity brings out bi-stable character, associated with adoption clustering, but not the coupling between the different regions (as indicated by the demand spread and demand spillovers loops in Figure 10).

AFV demand and fueling infrastructure supply exhibit more subtle long distance interdependency that drives dynamics. Decreasing the patch length further brings the feedbacks associated with the long-range interactions into consideration. The explicit consideration of the existence of vast rural areas outside, and between urban regions, results in a reduction of demand as compared to the coarser grid, which is also illustrated by a lower ratio of large to small trip fulfillment for these patches. Decreasing patch length from here on, allows capturing population level and demand fluctuations, but does not affect the overall patterns of demand and supply. However, as simulation run time increases exponentially with the number of patches, computational constraints become another factor of consideration. The current patch length of 18 miles falls within the region where dynamics are insensitive to a change in its length. This example further

illustrates that the current analysis not only allows exploration of behavioral explanations of why take-off might stall, or of what policies might lead to success, but uncovers fundamentally different dynamics and equilibria, compared to assumptions that ignore the spatial heterogeneity, or that only focus on local supply and demand interactions. Studies of symmetry breaking in spatially distributed systems are more and more appearing in biological research (e.g. Sayama et al. 2000). In this particular case the consumer and supplier behaviors mediated interactions that are different over short- and long-distance and drive dynamics that cannot be observed with mean-field approaches. Finally, it is expected that dynamics are not affected by disaggregating other parameters, such as consumer types, for reasons similar to those arguing against reducing the patch size.

A second assumption to explore in more depth involves that of randomly distributed trip destinations. Such an assumption greatly limits data and modeling requirements and is certainly useful for shorter trips. However, long-distance travel occurs at least partly over highways and is thus considerably more concentrated. The impact of relaxing the assumption of undirected travel for longer trips on the overall dynamics is not straightforward. Highway travel creates corridors that reduce the effective dimensionality for parts of the long-distance trips. This lowers the effective distance between stations and thus, holding actual stations constant, has a declining effect on driving effort. On the other hand, including road travel increases the typical length for the same absolute distance, and thus the required total number of stations per trip. Further, availability of

sufficient fuel stations throughout a trip is imperative for drivers' willingness to adopt

and drive, but long-distance travel is a relatively low contributor of the total demand

volume and provides limited potential for revenues, especially in the less high-volume

regions (even for gasoline competitive highway stations are frequently more than 30

miles apart).

I address these considerations in a simulation that generates different driving patterns

with a different treatment for short and long distance trips. Short trips with random

directionality, distributed following the same assumptions as in the previous analysis are

generated. The concept of gravitational models (e.g. Fotheringham 1983) is used to

generate long-distance trip destinations as a function of the population density, with

populated areas serving as the main destinations. Next, the repertoire of highly frequented

destinations is expanded by including several destination hotspots, such as Las Vegas,

Lake Tahoe, and the north east border, Crescent City.[15] In the model, the high density

traffic between cities and to hotspots form natural corridors for demand and serve as a

useful proxy for directed trips. We perform a simulation that is further, where possible

identical, to Figure 10, in terms of parameter settings, initial conditions and subsidy

scenario for the entrant equivalent to ICE. However, to conserve computational efforts, I

limit the simulated area. I choose one that includes the complete LA region, until the

Mexican border, including Lake Tahoe towards the North-East, and San Jose on the

---

[15] Even though such hotspots may lie outside the modeled grid area, their drivers destined for these locations generate demand within the modeled grid.

North-West. The average population density for the selected region is 30% higher than the California average. Details are provided in appendix 4c.

Figure 12b shows the results. We see, first, that adoption attains a low level equilibrium, only slightly higher than in Figure 10. Further, there is a strong discrepancy between urban and rural adoption. Comparing these results with a simulated equilibrium of ICE in absence of the equivalent entrant (ICE equilibrium), illustrates that a high equilibrium with stations throughout, can be achieved. Performing analysis at this more disaggregated level requires careful calibration and more work is needed to confirm these results. However, the results are strong: the assumptions for this simulation strongly favor take-off: besides the higher average population density any station that appears along the corridors is easily accessible for regional demand. This favors especially rural stations.

Adding more behavioral detail does matter. An analysis of the role of endogenous topping-off behavior illustrates this. Drivers can adjust their topping-off level, trading off the frequency of refueling for a reduction in needing to go out of the way, crowding and out-of-fuel risks, by selecting more convenient locations before the actual need appears. To test the implications, we represent the endogenous topping-off level relative to the normal toping-off buffer r_{i0}^{b}, that adjusts to the indicated level r_{iz}^{b*}, which is a function of the average utility of driving, which can be seen to represent the certainty of availability of fuel and service:

$$r_{iz}^{b*} = f\left(u_{iz}^{t}\right)r_{i0}^{b}; f' \leq 0; f\left(0\right) = r_{\max}^{b}/r_{i0}^{b}; f\left(1\right) = 1; f\left(\infty\right) = r_{\min}^{b}/r_{i0}^{b}$$

The relative top-off buffer increases with decreasing utility, but stabilizes at r_{max}^b for very low utility, as drivers will not want to be constrained by refilling on average too early.[16] Further, when drivers are fully confident, they will reduce their buffer to r_{min}^b, which can be below the indicated level by the warning sign, r_0^b. The exact form, yielding one sensitivity parameter α^f, is derived in Appendix 4d, also including a graphical representation. When the value of the sensitivity parameter α^f equals 0, the topping-off buffer remains constant for all utility parameters, when it is equal to 1, the buffer changes linearly with utility. The reference topping-off buffer is $r_{i0}^b = 40$ miles (10% of the total range), $r_{max}^b = 200$ miles (50% of the total tank range), and $r_{min}^b = 20$ miles.

Figure 13 illustrates the results. Respective simulations involve increasingly sophisticated assumptions about refueling behavior. Varying α^f, and β_{ref}^f, a measure for refueling location sensitivity to a change in the relative effort in refueling, I show 1) the case of responsive behavior, for which the topping-off buffer is held fixed and drivers are assumed to always start searching for fuel when they reach their buffer $\left(\beta_{ref}^f = 0; \alpha^f = 0\right)$. In this case, within each trip, the refueling location shares $\sigma_{i\omega_{zz}\cdot s}$ is exactly equivalent to the share of driving through the various locations; 2) balancing behavior, in which drivers hold their topping-off buffer fixed, but are allowed to select

---

[16] This level depends on the physical constraint of refueling elsewhere; see also equation (7) and Figure (5). From this behavioral reasonable parameters could be derived.

refueling sites, based on the $\left(\beta_{ref}^{f}=2;\alpha^{f}=0\right)$;[17] iii) adjusting behavior, in which drivers

endogenously adjust the effective tank range $\left(\beta_{ref}^{f}=2;\alpha^{f}=0.5\right)$. We see from the results

that endogenous topping-off does not stimulate, but hinders adoption. Ignoring other

effects, facing an increase of uncertainty of fuel availability, a driver's adjustment of its

topping-off buffer can improve her utility, from being able to locate at more favorable

locations, at the cost of a little increased frequency. However, once that happens, two

major reinforcing feedback loops become active: first, drivers contribute to an increase in

crowding, because of their lower effective range, without increasing net consumption,

this further triggers upward adjustment of buffers, leading to more crowding. Second, the

reallocation of demand for fuel implies that more fuel goes to more favorable locations.

This further reduces demand in already ill-served areas, contributing to more station

exits, increasing uncertainty and reducing further demand in those areas. This last

feedback is intrinsic to the urban-rural inequality, as well as the behavioral and

disequilibrium character of this system.

Varying AFV characteristics

How is adoption affected when AFVs differ from the incumbent technology, ICE, along

technical and economic dimensions of merit? To answer this question using simulations,

use more favorable conditions than before to generate a successful take-off in the

reference case that represents the ICE-equivalent AFV. Besides high station and vehicle

subsidies, favorable assumptions regarding vehicle/fuel performance, cost parity,

awareness and acceptance of the alternative technology - already assumed in the previous

---

[17] These parameter settings correspond with the assumptions for all other simulations

simulation – are used. In addition, lower consumer sensitivity to the additional effort/risk associated with low station coverage is included. In **Figure 14**, the blue line (highest penetration) shows the reference case, a successful penetration. The left axis shows adoption, with a value of 0.5 corresponding to a 50% share of the market, which is expected to be the maximum equilibrium situation for an ICE equivalent. Equilibrium market penetration still saturates at a level lower than that of the status quo due to a high degree of clustering near metropolitan centers.

Particular starting assumptions are relaxed step-by-step to allow for a comparison of three other fictitious AFVs that are also shown in **Figure 14**. The parameters that are varied and their values for each run are shown in Appendix 4e . To illustrate the role of increased efficiency, the red line shows the dynamics for scenario 2, representing a fuel-efficient fictitious AFV with fuel efficiency three times that of the reference case and total vehicle driving range held constant, as compared to the reference case (to achieve this, the tank size is set to equal 1/3 of the reference case's). This scenario could represent the introduction of small fuel-efficient AFV vehicles, at first sight an attractive candidate for early adopters. While adoption takes off fast, it stagnates early; surprisingly, more efficient vehicles are not necessarily more successful. **Figure 14**, right, shows that the supply collapses after the subsidies come to an end. The increased demand is not sufficient to make up for lower revenues, and no self-sustaining market emerges at low levels of penetration. Thus, this counterintuitive result illustrates a large trade-off between the end goal of increasing fuel efficiency and diffusion: on the one hand, there is efficiency, which reduces the environmental footprint (the energy dependence of

transportation), and may drive adoption; on the other hand, we see the importance of rapid supply growth to achieving successful diffusion.

Dispensing capacity is expected to be a constraint for many alternatives, especially gaseous fuels (CNG, hydrogen), and EVs. Scenario 3 (the green line in **Figure 14**) illustrates the role of dispensing rate on the dynamics. It shows the dynamics for parameters similar to scenario 2, except for an assumption of dispensing capacities being 25% as compared to the reference. Entrant technologies also have the burden of limited performance. In this case, adoption is suppressed directly as well. Due to the significant overcrowding at stations, which has a very non-linear response to station/pump utilization levels, attractiveness for potential adopters remains low. On the other hand, stations have limited incentives to expand or enter in places where utilization does not achieve very high levels. When fuel efficiency is lower, fueling frequency and crowding go up considerably. This dramatically increases the refueling time, making the effect even stronger. The final simulation represents early stage HFCVs, with DOE's 2015 targets for HFCVs as a reference for the parameters (Table 3, case 4). Importantly to stress, without sophisticated introduction policies and under the current model assumptions, these parameters result, in no take-off at all. To point of this last simulation is not to show expected failure for HFCV, but to illustrate that for different configurations, dynamics can be disproportionally influenced.

Different technologies result in different challenges. For example, introduction of hybrid vehicles, that use an infrastructure that is compatible with gasoline, and further have

lower fuel consumption, will lead to fast penetration (ignoring other feedbacks that relate to familiarity, technology learning and policies). In this model, if utility from hybrid vehicles equals that of gasoline vehicles, 25% of the market share is attained in 5.5 years, and 40% in 11 years, solely constrained by replacement dynamics of vehicles. The infrastructure can easily absorb the reduced demand, while still providing fuel throughout. However, for most of the alternative technologies, for which the infrastructure is not compatible, the dynamics as discussed above will be critical.

Bi-fuel and flex-fuel vehicles will exhibit a significantly reduced out-of-fuel risk, compared to alternative fuels, such as CNGs and HFCVs, as they can rely on pure gasoline as backup, while they can select the cheapest vehicle. But for hybrid solutions, there are inherent tradeoffs. This is illustrated by the case of CNG-gasoline vehicles. The fixed cost is higher, while vehicle performance and space are compromised. More importantly, the spatial dynamics of bi- and flex-fuel vehicles might play out quite differently than is the case for a mono-fuel: the reduced dependence of drivers on availability of remote stations reduces demand in the low-volume regions even further, which further reduces incentives for a widespread network to build up. Plug-in EVs also pose challenges. Charging at home solves part of the service time challenge of EVs. However, a side-effect is that the demand volume outside the home location is virtually non-existent, again providing little incentive for infrastructure to build up. In summary, for bi-fuel vehicles, the low-demand bi-stable equilibrium might emerge more easily and quickly, but the gap with full-scale penetration can become even larger than is the case

for mono-fuel vehicles that depend on an infrastructure that is incompatible with gasoline.

This analysis also brings to mind the coordination and standardization challenge that stakeholders, fuel suppliers, automotive manufacturers, and governments face. Similar coordination issues contributed heavily to the stalling of the EV infrastructure in the early 20th century. Not until it was too late were inventors, entrepreneurs, owners of central electricity stations, and policy makers able to coordinate on a viable infrastructure solution by providing large-scale, low-cost, off-peak refueling opportunities at central stations; sufficient coordination did not occur despite many viable ideas that were proposed early, included battery change services, leasing by central stations, and curbside pump networks stations (Schiffer 1994).

Also here, choice to seek early standardization occurs at several levels: across AFV portfolio choice, such as internal combustion hydrogen versus hydrogen fuel cells; within an AF technology, such as forms of on-board storage (comprising a variety of gaseous low- and high pressure, liquid, Nanotube solutions); with respect to individual technologies; or, regarding practices and regulations, such as on-site fuel storage modes, or the dispensing process. While technology diversity may be beneficial to the innovation rate of the technologies involved, absence of standards produces many difficulties. First, this greatly increases incompatibility for users. For example, different forms of on-board storage require different dispensing technologies. From the preceding analysis, one can readily interpret the dramatic negative impact this would have on the early market

formation. Similarly, for fuel stations absence of standards implies higher cost and increased space constraints. Further, different technologies that share much lower volumes have less learning and cost-reduction. Finally, absence of standards, rules, and legislation greatly increases permitting time for fuel stations.

Discussion and conclusion

Modern economies and settlement patterns have co-evolved around the automobile, internal combustion, and petroleum. The successful introduction and diffusion of alternative fuel vehicles is more difficult and complex than for many products. The dynamics are conditioned by a broad array of positive and negative feedbacks, including word of mouth, social exposure, marketing, scale and scope economies, learning from experience, R&D, innovation spillovers, complementary assets including fuel and service infrastructure, and interactions with fuel supply chains and other industries. A wide range of alternative vehicle technologies – hybrids, biodiesel, fuel cells – compete for dominance.

This essay focuses on only one interaction: the co-evolution between alternative fuel vehicle demand and the refueling infrastructure. I developed a dynamic behavioral model, with explicit spatial structure. The behavioral elements in the model included drivers' decisions to adopt an AFV, their trip choices, and their decisions to go out of the way to find fuel, as well as their topping-off behavior in response to the uncertainty of finding fuel. The responses to fuel availability included the effort involved in searching or getting to a station, the risk of running out of fuel, and the service time (as a function

of supply and demand), and number of service points. The supply-side decisions included station entry and location decisions, exit, and capacity adjustment.

The local scale, but long-distance correlation of interactions is paramount in this dynamic and behavioral setup. Fuel availability differs for each driver based on their location and driving patterns relative to the location of fuel stations. Often labeled as "chicken-and-egg" dynamics, these co-evolutionary dynamics are much more complex. The increasing interest for spatial symmetry breaking in biological systems (e.g. Sayama et al. 2000) is also justified for the complementary interactions between vehicle demand and its fueling infrastructure. Analysis of local adoption and stagnation provides an explanation for persistent clustering phenomena, with low levels of adoption and usage, for AFVs that are introduced in the market. For example, in Italy, with a CNG penetration of 1% in 2005, 65% of the CNG vehicles and 50% of the CNG fuel stations are concentrated in 3 of the 20 regions (Emilio-Romagna, Veneto, and Marche), together accounting for about one-sixth of the population and area (Di Pascoli et al. 2001). In Argentina, the largest bi-fuel CNG market with a penetration of 20%, 55% of the adopters live in Buenos Aires and 85% in the biggest metro poles. Similarly, in the beginning of the 20th century, EVs remained clustered in urban areas, with virtual absence of recharging locations outside urban areas (Schiffer at al. 1994). Many attempts to introduce AFVs collapsed after government support, subsidies, or tax credits were abandoned, for example with bi-fuel CNG/gasoline in Canada and New Zealand (Flynn 2002). While islands of limited diffusion might be sustained in the cities, as can be seen in Argentina, broad adoption of AFVs can easily flounder even if their performance equals that of ICE. The

acknowledgement of different relative "tipping points" for rural and urban markets and their interdependency can inform the evaluation of different hydrogen transition strategies and policies. The clustering and stagnation behavior is significantly different than the basic chicken-egg dynamics suggests, or than can be inferred from standard economic analysis of complementarities. Modeling the behavioral decision making and the spatial aspects dynamically is essential for revealing these patterns of low penetration.

This model is in the early stages of development and requires more intense calibration, validation, and extensions. Yet current analysis considerably enhances our understanding of previous alternative fuel experiences and future alternative fuel transition strategies. The tight coupling between components of the system that are physical (such as typical replacement time and the spatial characteristics), behavioral (trip choice, sensitivity to availability of fuel), or technical/economic (e.g., fuel economy, tank size, fuel price) influence the dynamics. The analysis illustrates a bi-stable equilibrium with urban adoption clusters and limited aggregate demand. This fully dynamic perspective illustrates some counterintuitive results: more efficient vehicles are not necessarily improve the transition dynamics, for the emergence of a self-sustaining market, and can in fact harm it. More generally, the analysis illustrates the trade-off between the long-term goal of low consumption and emission vehicles and the necessary market take-off.

The behavioral character of the model, within the spatial context, provides significant insights with driver behavior, for instance fuel station capacity adjustment, being

138

endogenous. For example, the number and length of trips increases as fuel availability rises, and only then demand spillovers from urban to local regions, allowing for sufficient demand for take-off in those regions. Finally, we saw that dynamics were critically impacted when we allowed topping-off levels to be endogenously adjusted. Drivers who perceive refueling effort to be high - say, because some fuel stations are distant or crowded - will seek to refuel before their tanks are near empty, balancing increased efforts from more frequent refueling stops against reduced out-of-fuel risk. However, the side effects of increased crowding, and reallocation of demand to the higher volume regions, set in work self-fulfilling prophecies of the uncertainty of supply. More generally, including these behavioral aspects highlights the distributed nature of the system. The local adjustments of supply and demand can easily be absorbed in a well established high volume system and provides increased adaptability and efficiency that can thus be expected to improve successful transitions. However early in the transition the negative side effects of such adjustments can and lead to a failed transition.

The analysis focused on the impact of supply-demand interactions relevant for aggregate diffusion dynamics. This model's finite element approach suggests several research directions. For example, one could focus on specific state-level location strategies, by reducing patch size and incorporating detailed data such as traffic flow information. However, we saw that for the transition dynamics, capturing heterogeneity at the scale below the typical trip length, in combination with the behavioral feedbacks, was critical to obtaining the results, but the high-frequency noise from smaller-scale fluctuations could be ignored. In addition, we saw that the fundamental conclusions are not changed,

when relaxing the assumptions of randomly directed trips. Assuming random directions saved scarce resources for computation and analysis, and critically reduced data requirements. Also, analysis at a higher level of aggregation allows including more behavioral feedbacks that, as we saw throughout, but in particular with the topping-off dynamics, contribute significantly to the aggregate dynamics.

Transition challenges are different for different AFVs. Successful introduction of hybrid vehicles poses much fewer and smaller challenges than achieving this for HFCVs. It is valuable to think how the dynamics observed here would interact with other elements of the socio-technical system. For example, suppressed diffusion also limits the accumulation of knowledge that is critical for improving AFV performance. Further, automotive OEMs are likely to respond to the observed demand patterns for AFVs that favor cars for city-dwellers. In response, their portfolios would come to consist mainly of small, efficient, inexpensive models, adapted for commuting but ill suited for touring. Such behavior further reduces their attractiveness in rural areas, and likely restricts adoption to affluent households who can afford an AFV for commuting and an ICE vehicle for weekend excursions. These feedbacks can further constrain diffusion.

Taking a broad system perspective allows exploring at high leverage interventions. As we discussed with hybrid vehicles, a transition is certainly possible. For example, in Essay 1 I focus on the role of social exposure dynamics: as vehicles are complex, and emotions, norms and cultural values play an important role, social exposure dynamics will have

significant influence on the transition dynamics. Combining the partially local diffusion aspects with the spatial infrastructure dynamics will provide more insights into challenges and levers for adoption. As an "inverse" analogy to ring vaccination policies (designed to contain viruses), peripheral dotting of metropolitan regions at edges between urban and rural areas might be used to bridge demand for drivers towards more remote regions, thereby lowering uncertainty in demand. This robustness of this policy can be further tested with this model.

Other policy levers lie in the collective action problem that is deeply rooted in AFV transition dynamics. Without coordination between automakers, fuel suppliers, and governments, adoption will not take off. First, there is the challenge of coordination on strategic investment. As we saw above, if AFs are initially only introduced in light, compact, efficient cars, there might be little incentive for the supply side to roll out a large infrastructure. On the other hand, if the benefits are too little from the consumer perspective, demand will not develop. This suggests high leverage can be found in coordination across stakeholders on issues such as pilot region selection, target market, vehicle portfolio selection, asymmetric incentives for urban and rural stations, other incentive packages, and standardization. Second, governments' policies need to be aligned with those of the industry: a gasoline tax alone might spur demand for other fuels, but it might take a long time before good alternatives became available. Further, as we saw, if the alternative does not provide incentives for suppliers to build fuel stations or for automakers to build alternative vehicles, impact will be small. Finally, the lack of standardization is a strong cause and effect of the coordination problem. Further

application of the present model can reveal high-leverage coordination policies between these (and other) stakeholders. Subsequent research will be targeted at such questions.

The observations in this discussion suggest that, for exploration of robust alternative fuel transition strategies, full policy analysis, and development of incentives of proper kind and duration, other feedbacks need to be included as well. Inclusion of other feedbacks -- such as scale and scope economies, R&D, learning by doing, technology spillovers, familiarity through word of mouth and driver experience, and production/distribution of fuels and other complementary assets -- are crucial for understanding the transition challenges. Initially, the technologies of AFVs will perform much worse than ICE, significantly increasing the threshold for the formation of a self-sustaining market. The strong dependency of model behavior on economic/technical characteristics suggests that full inclusion of these feedbacks is critical. Building towards this, essay 3 discusses the inclusion of learning and technology spillovers. Finally, full analysis must include various alternatives at the same time also competing with each other.

The variety of success and failures of AFV market formation in the past suggests strongly that our understanding was unguided by reliable insight. This essay demonstrates the importance of dynamic models – when they incorporate behaviorally rich detail and focus on those factors that increase the dynamic complexity – for understanding the dynamics of market formation that involves consumers, producers, regulators, and producers of supporting infrastructure.

References

Abernathy, W. and J. Utterback (1978). "Patterns of Industrial Innovation." *Technology Review* **80**(7): 40-47.

Arthur, W. (1989). "Competing Technologies, Increasing Returns, and Lock-in by Historical Events." *Economic Journal* **99**(394): 116-131.

Bass, F. (1969). "A New Product Growth Model for Consumer Durables." *Marketing Science* **15**(5): 215-227.

Ben-Akiva, M. and S. Lerman (1985). *Discrete choice analysis: theory and application to travel demand.* Cambridge, Mass., MIT Press.

Bentham, J. (2005). CEO Shell Hydrogen Testimony to the U.S. Senate Subcommitee on Energy.

Berry, S., J. Levinsohn, and A. Pakes. (2004). "Differentiated products demand systems from a combination of micro and macro data: The new car market." *Journal of Political Economy* **112**(1): 68-105.

Brownstone, D., D. Bunch, and K. Train. (2000). "Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicles." *Transportation Research Part B-Methodological* **34**(5): 315-338.

Brownstone, D. and K. A. Small (2005). "Valuing time and reliability: assessing the evidence from road pricing demonstrations." Transportation Research Part a-Policy and Practice **39**(4): 279-293.

Centola, D. and M. W. Macy (2005). "Complex Contagion and the Weakness of Long Ties." Under review, American Journal of Sociology.

Collischonn, W. and J. V. Pilar (2000). "A direction dependent least-cost-path algorithm for roads and canals." *International Journal of Geographical Information Science* **14**(4): 397-406.

Cowan, R. and S. Hulten (1996). "Escaping lock-in: The case of the electric vehicle."

Technological Forecasting and Social Change 53(1): 61-79.

David, P. A. (1985). "Clio and the Economics of Qwerty." *American Economic Review* **75**(2): 332-337.

Di Pascoli, S., A. Femia, et al. (2001). "Natural gas, cars and the environment. A (relatively) 'clean' and cheap fuel looking for users." *Ecological Economics* **38**(2): 179-189.

Domencich, T. A., D. McFadden, et al. (1975). Urban travel demand : a behavioral analysis : a Charles River Associates research study. AmsterdamNew York, North-Holland Pub. Co. ;American Elsevier.

DOE (2006). Department of Energy Hydrogen Program - H2A Analysis

http://www.hydrogen.energy.gov/h2a_analysis.html.

Dijkstra, E. W. (1959). "A Note on Two Problems in Connexion with Graphs." *Numerische Mathematik* **1**: 269-271.

Energy Information Administration (2005). Alternative Fuel Reports, Analyses 1993-2003. Energy Information Administration Accessed 11-20-2005. http://tonto.eia.doe.gov/reports/reportsD.asp?type=Alternative%20Fuel

Farrell, A. E., D. W. Keith, et al. (2003). "A strategy for introducing hydrogen into transportation." *Energy Policy* **31**(13): 1357-1367.

Flynn, P. C. (2002). "Commercializing an alternate vehicle fuel: lessons learned from natural gas for vehicles." *Energy Policy* **30**(7): 613-619.

Fotheringham, A. S. (1983). "A New Set of Spatial-Interaction Models - the Theory of Competing Destinations." *Environment and Planning* **A** 15(1): 15-36.

Fujita, M. and P. Krugman (2004). "The new economic geography: Past, present and the future." *Papers in Regional Science* **83**(1): 139-164.

Fujita, M., P. Krugman, et al. (1999). "On the evolution of hierarchical urban systems." *European Economic Review* **43**(2): 209-251.

Geels, F. W. (2005). Technological Transitions and Systems Innovations - A Co-evolutionary and Socio-Technical Analysis. Northampton, MA, Edward Elgar.

Greene, D. (2001). TAFV: Alternative fuels and vehicle choice model documentation. Center for Transportation Analysis, Oak Ridge National Laboratory report ORNL/TM-2001/134.

Gnedenko, B. V. and I. B. N. Kovalenko (1989). Introduction to queueing theory. Boston, Birkhäuser.

Katz, M. L. and C. Shapiro (1994). "Systems Competition and Network Effects." *Journal of Economic Perspectives* **8**(2): 93-115.

Keeling, M. (1999). Spatial Models of Interacting Populations (Ch3). *Advanced ecological theory: principles and applications*. J. M. McGlade. Malden, Mass., Blackwell Science**:** xvii, 353.

Klepper, S. (1996). "Entry, exit, growth, and innovation over the product life cycle." *American Economic Review* **86**(3): 562-583.

Krugman, P. (1996). The Self-Organizing Economy. Birmingham, Blackwell Publishers.

Langer, J. S. (1980). "Instabilities and Pattern-Formation in Crystal-Growth." *Reviews of Modern Physics* **52**(1): 1-28.

Lefeber, L. (1958). Allocation in space; production, transport, and industrial location. Amsterdam, North-Holland Pub. Co.

Lösch, A. (1940). The economics of location. New Haven, Connecticut, English translation: Yale University Press.

Lovins, A. and D. Cramer (2004). "Hypercars((R)), hydrogen, and the automotive transition." *International Journal of Vehicle Design* **35**(1-2): 50-85.

Lovins, A. and B. Williams (1999). A Strategy for the Hydrogen Transition. Tenth Annual US Hydrogen Meeting.

McFadden, D. (1963). "Constant Elasticity of Substitution Production Functions." *The Review of Economic Studies* **30**(2): 73-83.

McFadden, D. (1978). Econometric Models of Probabilistic Choice. In *Structural Analysis of Discrete Data with Econometric Applications*. C. Manski and D. McFadden (eds.). Cambridge, MA, MIT Press.

McFadden, D., Ed. (1998). Measuring Willingness-to-Pay for Transportation Improvements. Theoretical Foundations of Travel Choice Modeling. Amsterdam, Elsevier Science.

National Academy of Engineering (2004). *The Hydrogen Economy: Opportunities, Costs, Barriers, and R&D Needs*. Washington D.C., National Academy Press.

Ogden, J. (2004). Where will the Hydrogen Come From? System Considerations and Hydrogen Supply. *The Hydrogen Economy Transition: Moving Toward the Post Petroleum Age in Transportation.* D. Sperling and J. Cannon (eds.) Amsterdam, Elsevier Academic Press.

Romm, J. (2004). "The hype about hydrogen." *Issues in Science and Technology* **20**(3): 74-81.

Sayama, H., L. Kaufman, and Y. Bar-Yam (2000). "Symmetry breaking and coarsening in spatially distributed evolutionary processes including sexual reproduction and disruptive selection." *Physical Review E* **62**(5): 7065-7069.

Schiffer, M. B., T. C. Butts, et al. (1994). Taking charge: the electric automobile in America. Washington, Smithsonian Institution Press.

Small, K. A. (1992). *Urban transportation economics.* Chur, Switzerland ; Reading, Harwood Academic Publishers.

Small, K. A., C. Winston, et al. (2005). "Uncovering the distribution of motorists' preferences for travel time and reliability." *Econometrica* **73**(4): 1367-1382.

Steinmetz, S. S. C. and D. Brownstone (2005). "Estimating commuters' "value of time" with noisy data: a multiple imputation approach." Transportation Research Part B-

Sterman, J. (2000). *Business dynamics: systems thinking and modeling for a complex world.* Boston, Irwin/McGraw-Hill.

Struben, J. (2005). Space matters too! Mutualistic dynamics between hydrogen fuel cell vehicle demand and fueling infrastructure. 23rd International System Dynamics Conference, Boston, MA.

Struben, J., and J. Sterman (2006) "Transition Challenges for Alternative Fuel Vehicle

and Transportation Systems". *MIT Sloan Research Paper* No. **4587-06** Available at
SSRN: http://ssrn.com/abstract=881800

Schiffer, M. B., T. C. Butts, et al. (1994). *Taking charge : the electric automobile in America*. Washington, Smithsonian Institution Press.

Train, K., and C. Winston (2005). "Vehicle Choice Behavior and the Declining Market Share of U.S. Automakers" *Working Paper,* The Brookings Institution.

Turing, A. M. (1952). "The Chemical Basis of Morphogenesis." <u>Philosophical</u> *Transactions of the Royal Society* **B237**: 37-72.

Witten, T. A. and L. M. Sander (1981). "Diffusion-Limited Aggregation, a Kinetic Critical Phenomenon." *Physical Review Letters* **47**(19): 1400-1403.

Figures
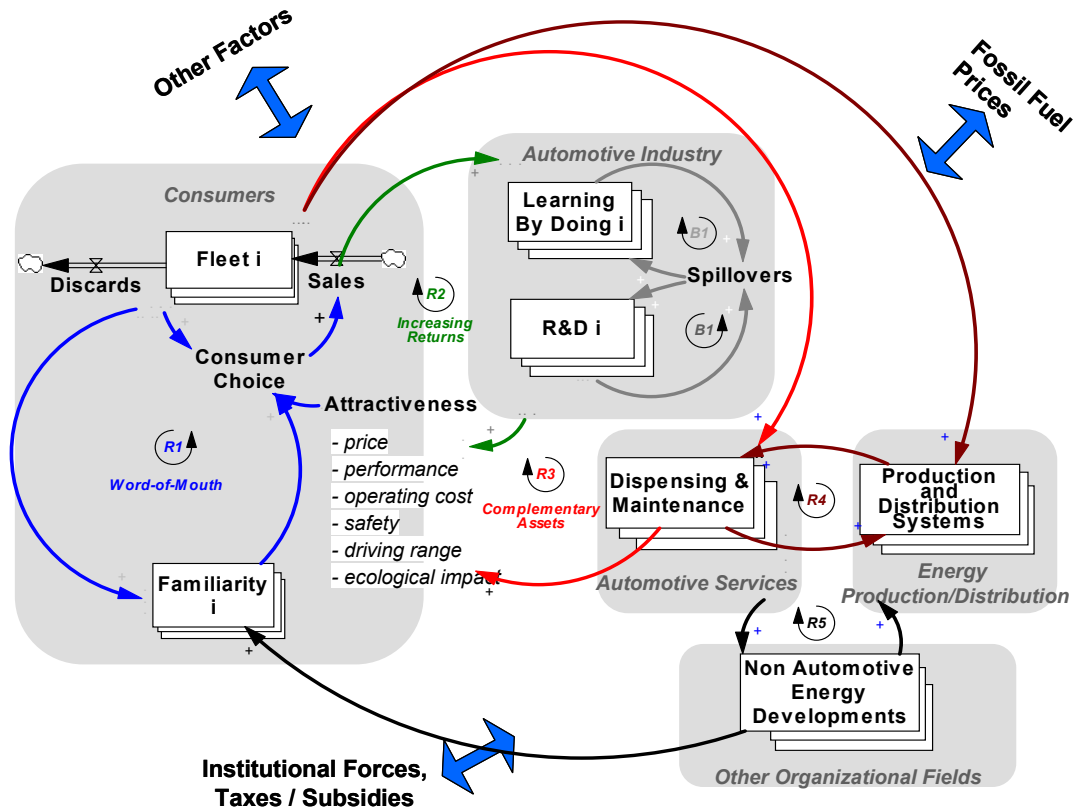
Figure **1** Full model boundary.
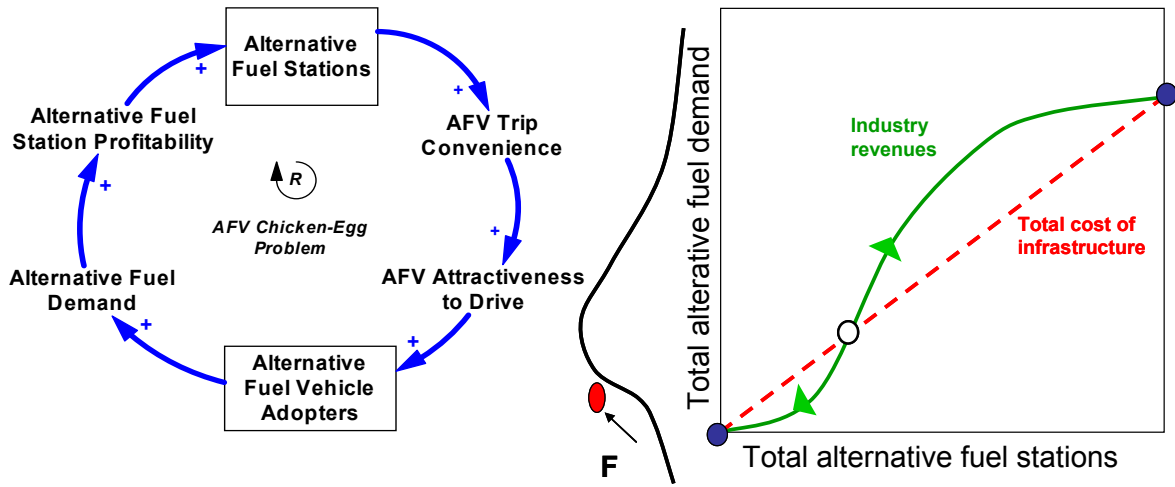
Figure 2 Rudimentary hypothesis for chicken-and-egg dynamics.

household density in z

trip generation to z to z'
{length l, direction θ}

station density in s

trips $z\,z'$ refueling in s

trip choice $z\,z'$

adoption fraction in z

Figure 3 Spatial representation of model concept: demand and supply are subject to short and long-range interactions; demand decisions involve adoption, trip-choice, and refueling locations, and topping-off behavior; supply decisions involve entrance, exit, location selection and capacity adjustments.

Figure 4 Principle feedbacks governing the co-evolutionary dynamics between vehicle fleet and fueling infrastructure.

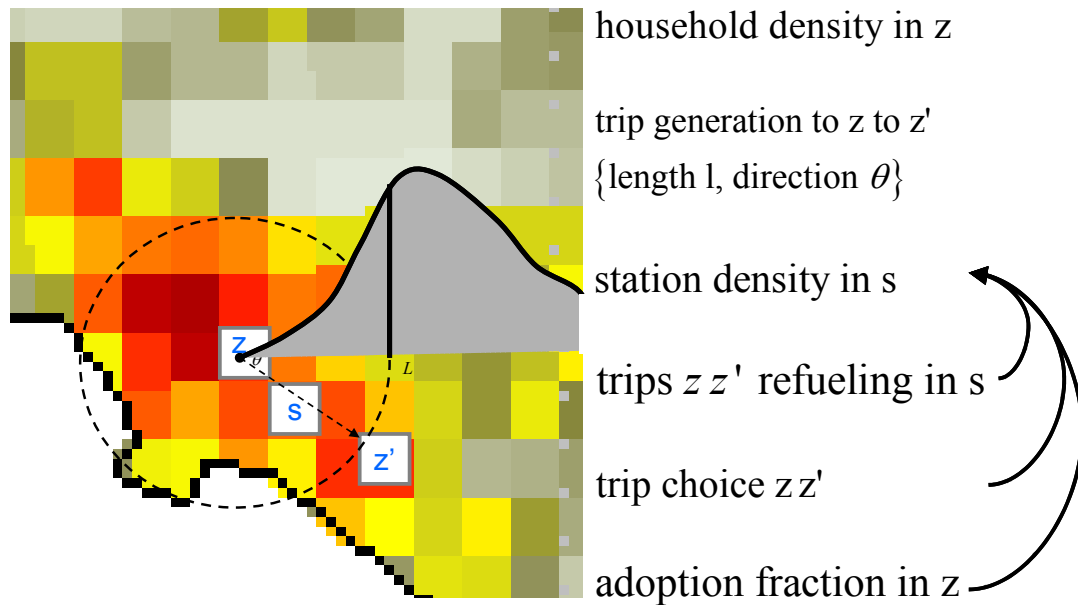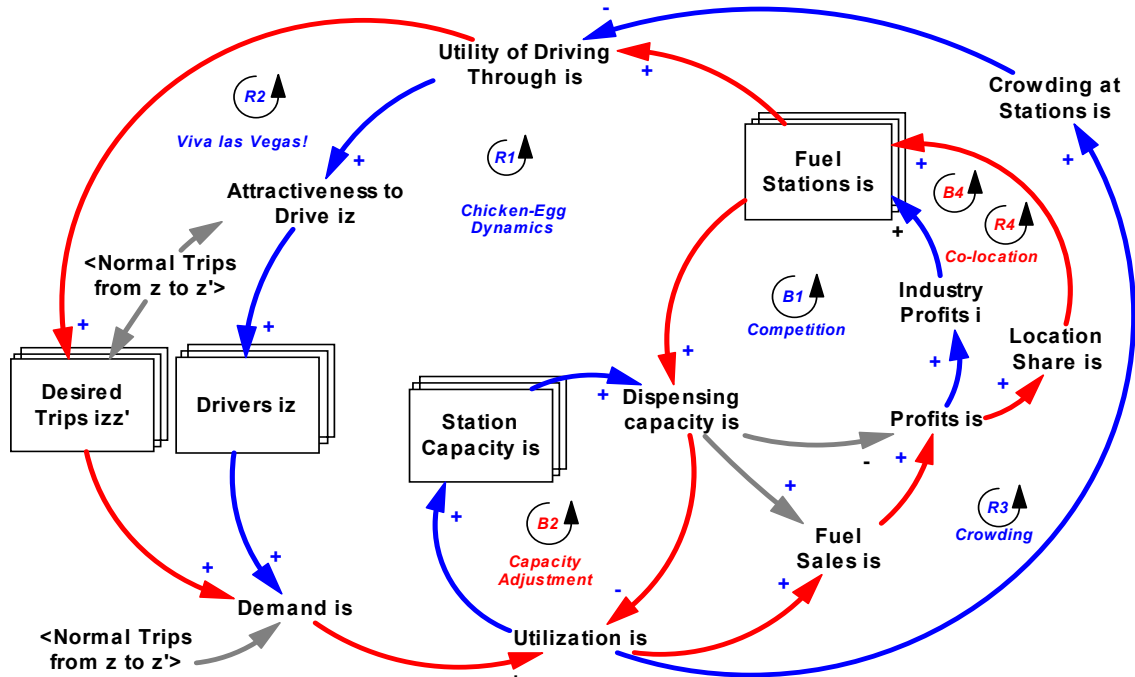| Overview of decision structure	choice function (MNL)		aggregate utility and effort function (CES)
	$\sigma_{iz} = \dfrac{u_{iz}}{\sum_j u_{jz}}$	$u_{iz} = u_{iz}^0 u_{iz}^t \, e^{\sum_l \beta_l \frac{a_{izl}}{a_l^*}}$	$u_{iz}^t = \left(\sum_{z' \in Z} w_{zz'} u_{izz'}^{t\,\mu^t} \right)^{\frac{1}{\mu^t}}$
	$\sigma_{izz'} = \dfrac{u_{izz'}^t}{u_{izz'}^t + u_{zz'}^o}$	$u_{izz'} = u_{izz'}^0 u_{izz'}^t$	$u_{izz'}^t = \left(\sum_{\omega_{zz'} \in \Omega_{zz'}} \sigma_{\omega_{zz'}} u_{i\omega_{zz'}}^{t\,\mu^\omega} \right)^{\frac{1}{\mu^\omega}}$
	$\sigma_{i\omega_{zz'}} = \dfrac{u_{i\omega_{zz'}}}{\sum_{\omega_{zz'} \in \Omega_{zz'}} u_{i\omega_{zz'}}}$	$u_{i\omega_{zz'}}^t = e^{\beta^\omega \frac{a_{i\omega_{zz'}}^t}{d_{zz'}^*}}$	$a_{i\omega_{zz'}}^t = a_{\omega_{zz'}}^{t0} + \phi_{i\omega_{zz'}} a_{i\omega_{zz'}}^f$ $a_{i\omega_{zz'}}^f = \sum_{s \in \omega_{zz'}} \sigma_{i\omega_{zz'},s} a_{is}^f$
	$\sigma_{i\omega_{zz'},s} = \dfrac{\varphi_{i\omega_{zz'},s}}{\sum_{s' \in \omega_{zz'}} \varphi_{i\omega_{zz'},s}}$	$\varphi_{i\omega_{zz'}} = r_{\omega_{zz'},s} \, e^{\beta^f \frac{a_{is}^f}{d_{zz'}^*}}$	a_{is}^f

Figure 5 Consumer choice decision tree: left, diagrammatic representation; right, functional forms used for choice structure (multinomial logit (MNL)), and utility and effort structure (non-linear weighted average (CES)).
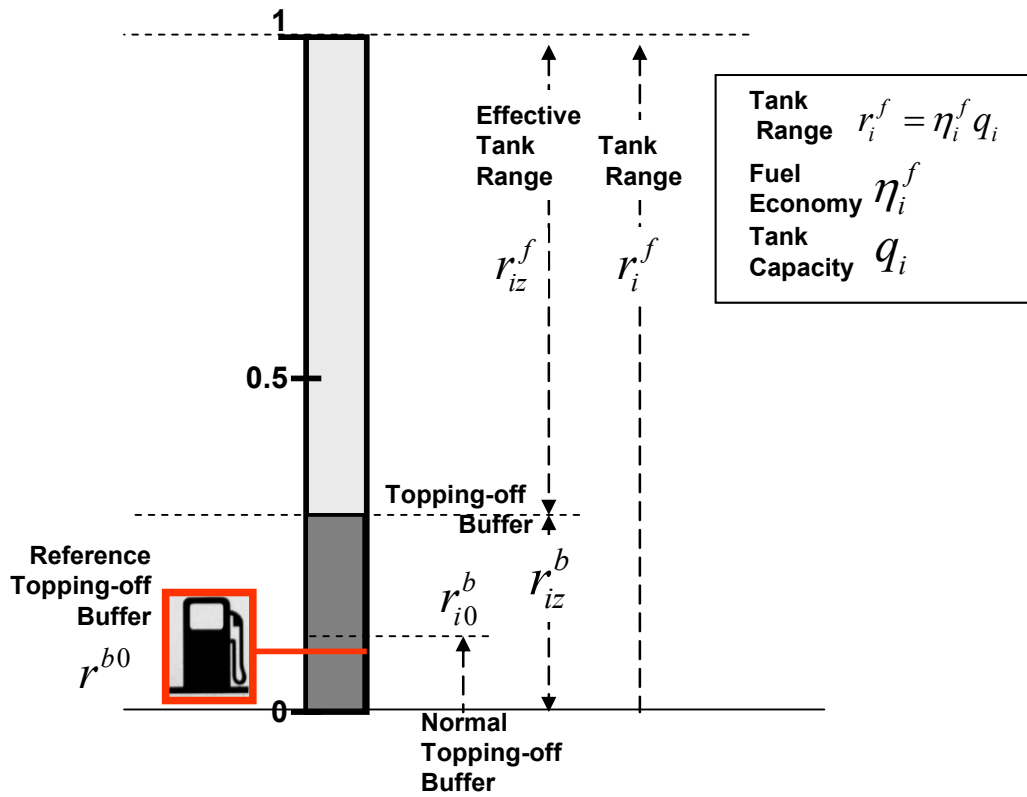
Figure 6 Tank range and topping-off parameters.
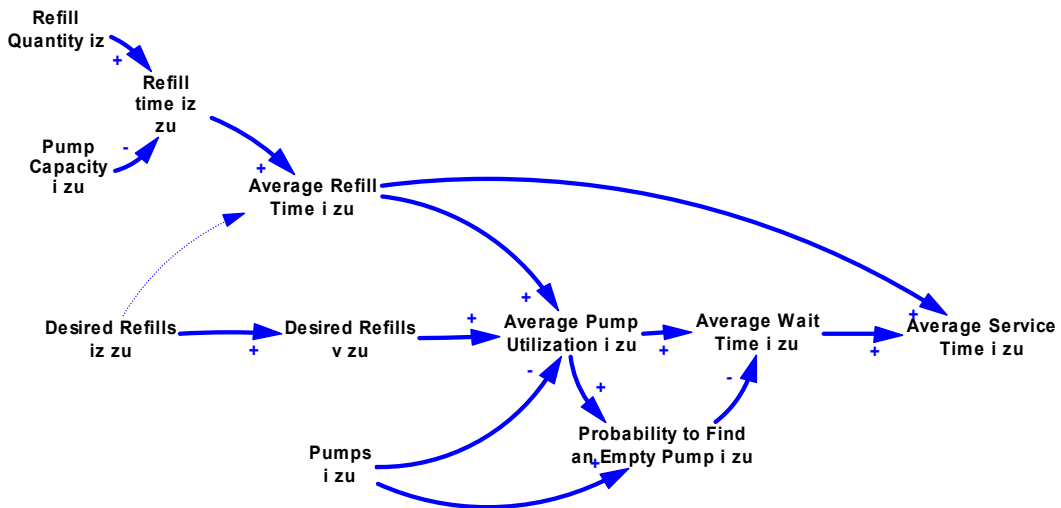
Figure 7 Station utilization and servicing time.
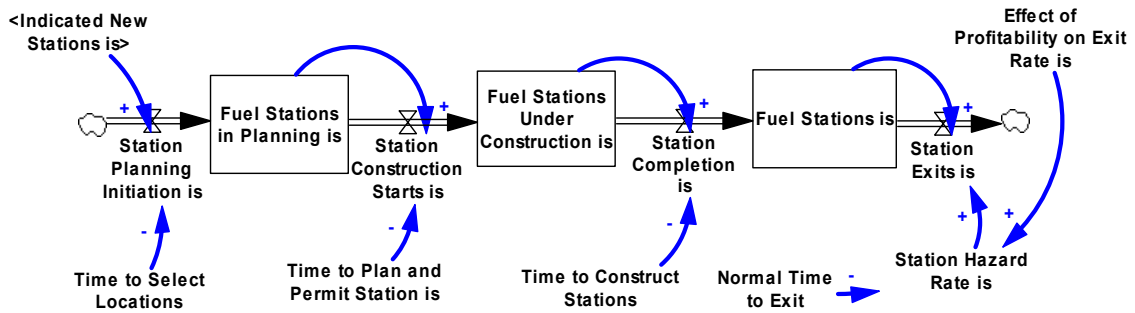
Figure 8 Fuel station entrance and exit process.
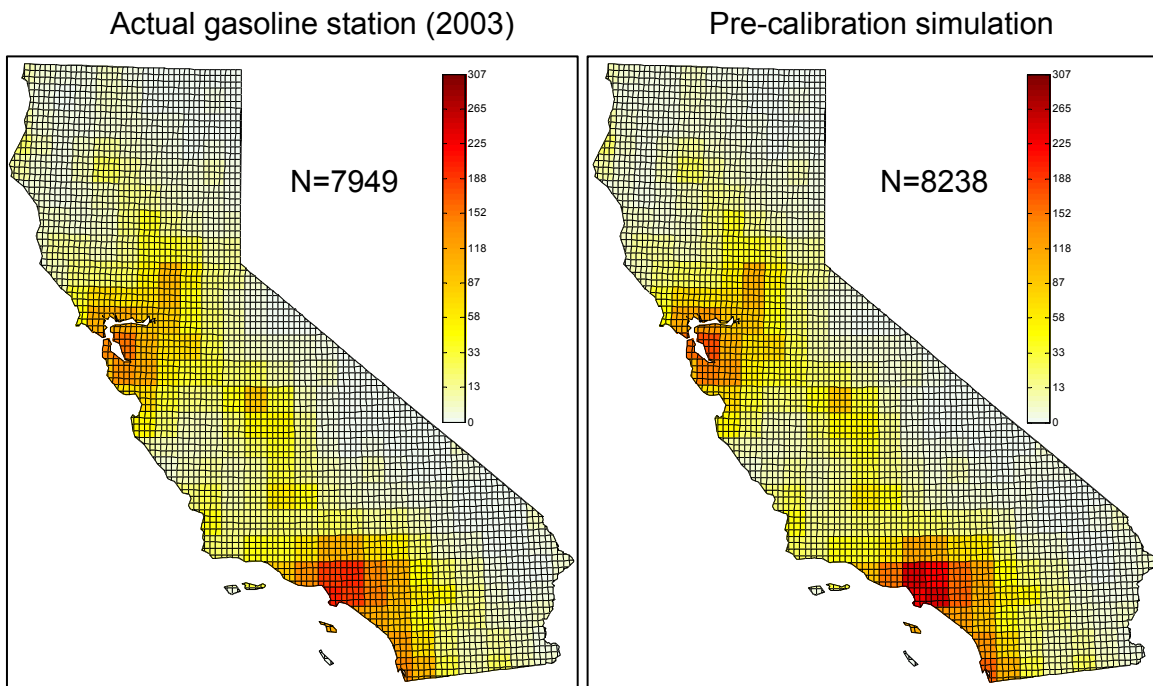
Figure 9 Pre-calibration performance test of station entrance behavior: a) Actual California station distribution; b) Simulated under fixed adoption.
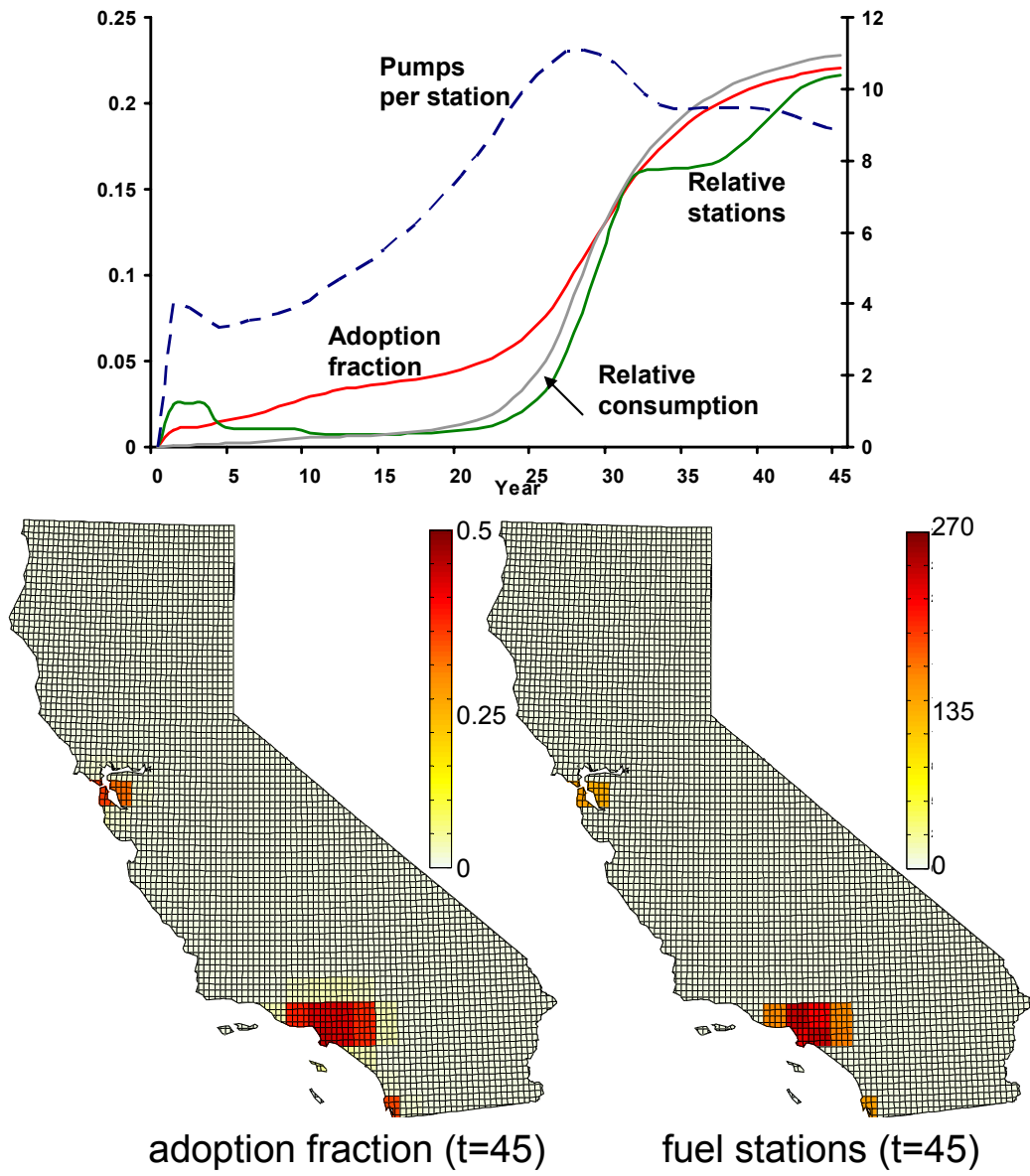
Figure 10 Behavior of spatially disaggregated model calibrated for California.
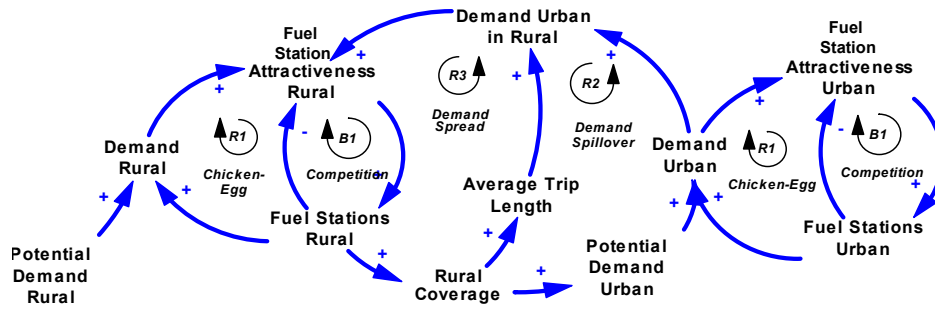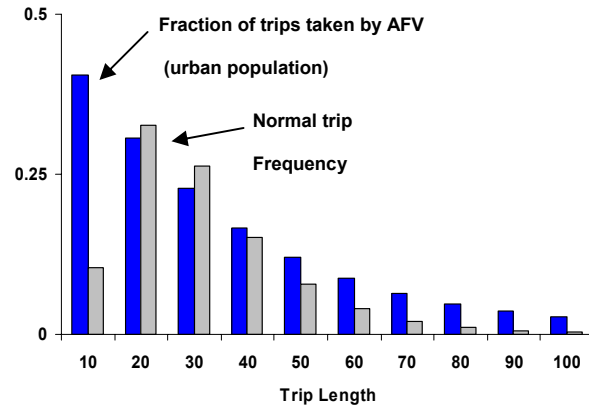
Figure 11 Hypothesis for bi-stable equilibrium with low level adoption and urban clusters.

a) Sensitivity of adoption to patch length

b) Relaxing random trip distributions
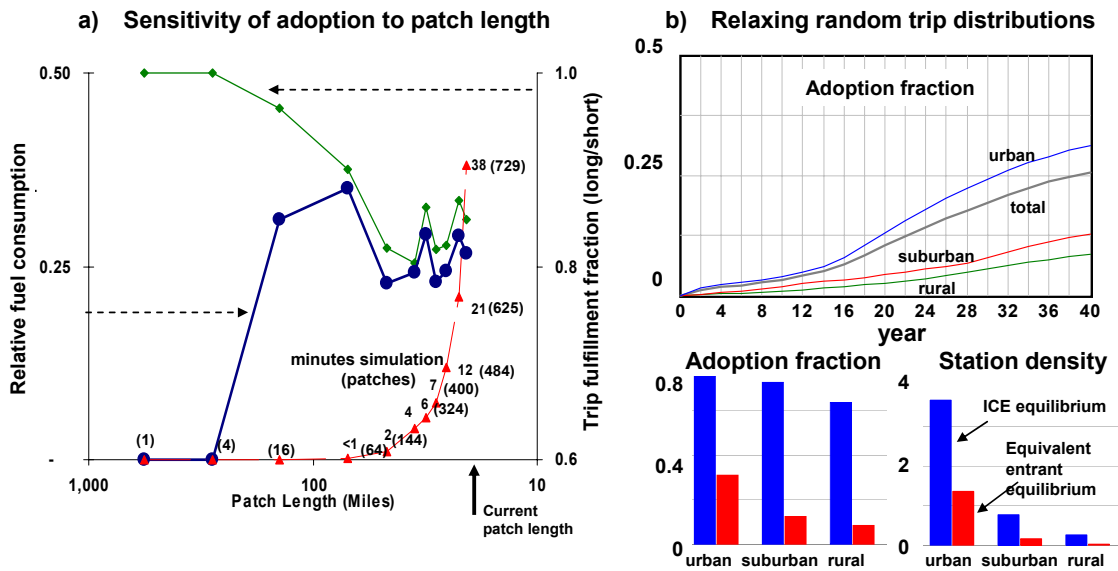
Figure 12 Model sensitivity to spatial detail: a) sensitivity of equilibrium behavior to patch length, with equilibrium fuel consumption (left axis), relative trip fulfillment short versus long trips (right), and simulation time (number of patches) as function of patch length; b) relaxing the assumption of randomly distributed long-distance trips, with adoption fraction over time (top) and the equilibrium adoption fraction for urban, suburban and rural, compared to the results for a simulation of ICE.
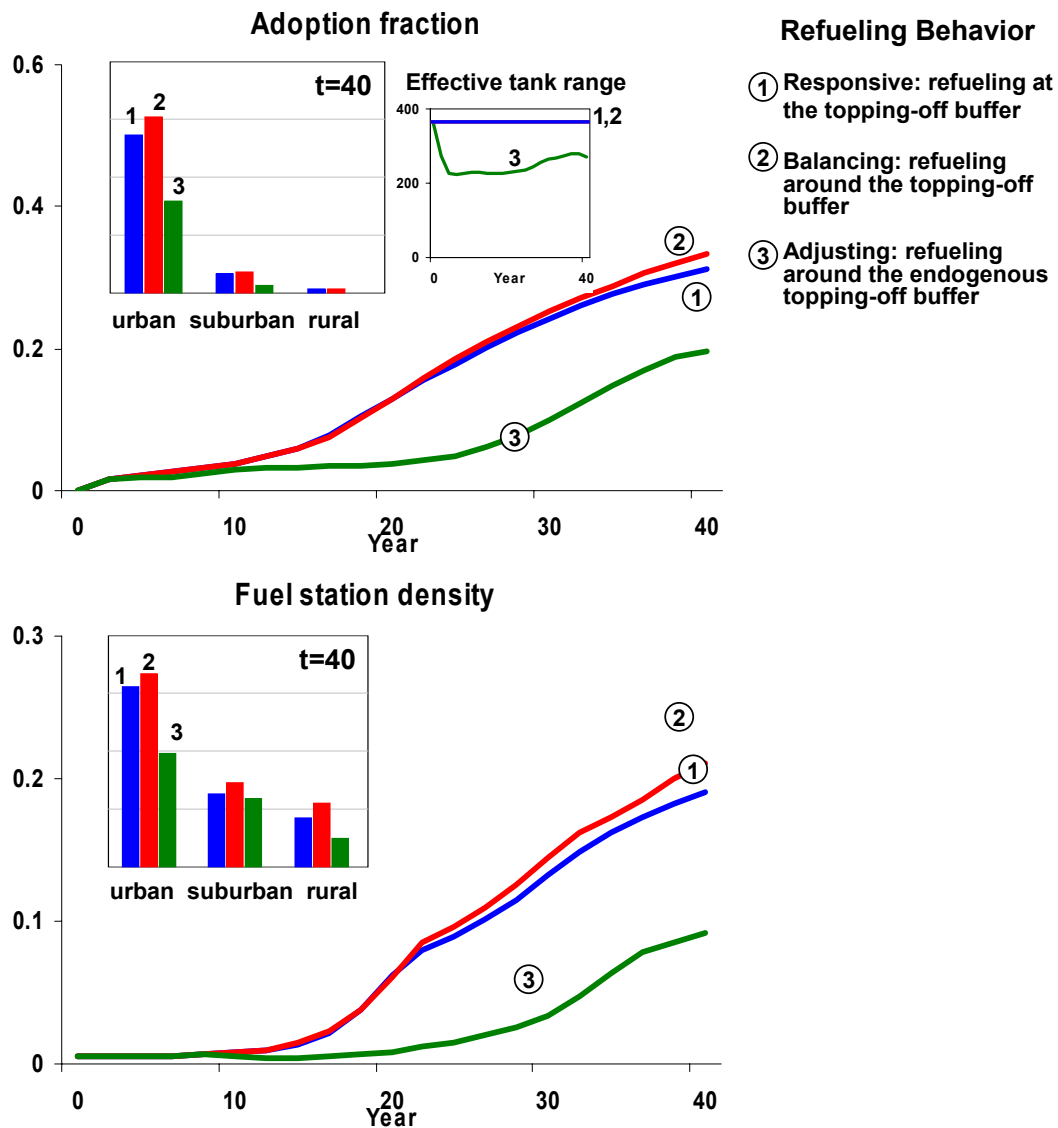
Figure 13 Sensitivity to topping-off behavior: adoption fraction (top) and fuel station density (bottom) for increasingly behavioral assumptions: 1) responsive, drivers always start searching when they reach their topping-off buffer; 2) balancing, drivers refuel on average at their topping-off buffer, allowing some flexibility to refuel at more favorable locations 3) adjustment: topping-off buffers are adjusted in response to changes in the uncertainty of availability of fuel. Left insets show the adoption fraction and fuel station density at t=40 for urban, suburban and rural populations. The right inset shows the effective tank range. For simulation 3 the effective tank range adjusts over time.

158
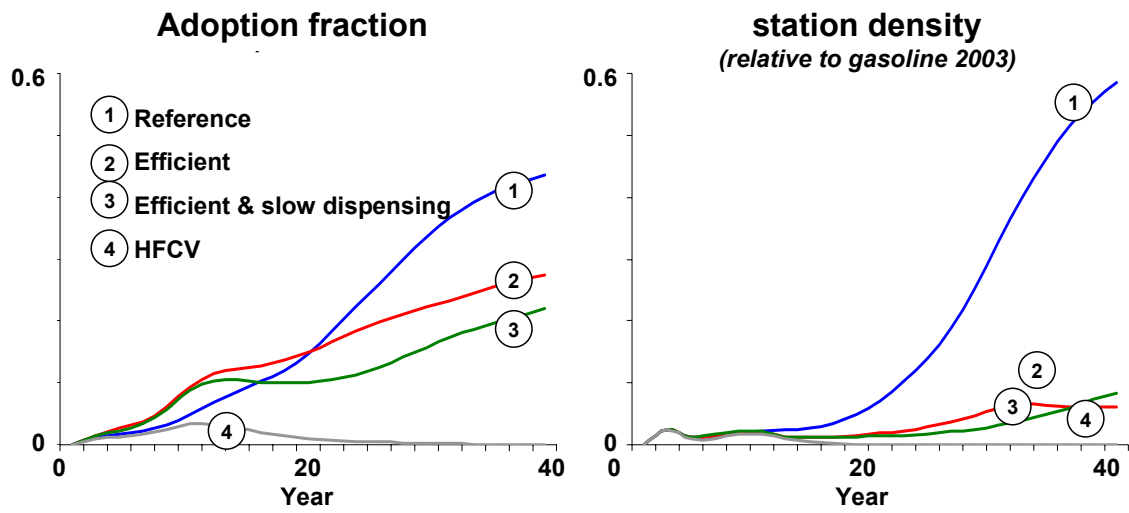
Figure 14 Introductions of hypothetical alternative fuels; details in **Table 3**. Run 2 shows a failure of a more efficient vehicle relative to the reference.

Tables

Table 1 Sources of dynamic complexity of market formation for alternative fuel vehicles.

Characteristic	AFV market formation example
Dynamics / time scale of change	Vehicle turnover; technological progress; infrastructure replacement.
Multiple stakeholders	Consumers; automotive companies; energy companies; fuel cell developers; policy makers; media.
Multiple feedbacks	learning from R&D- and user experience, and by doing; word-of-mouth, technology spillovers; complementarities (fueling infrastructure).
History dependent	Cumulative knowledge; efficacy- and safety perceptions; oil infrastructure.
Nonlinear	Effect of fuel availability on trip effort.
Spatial heterogeneity	Urban/rural asymmetries; short haul/long haul trips; station locating strategies

Table 2 Summary statistics for the state of California.

Statistic	value	Unit	Source
Population	35,537,438	People	US Census 2000
Households	50.2%	Dmnl	US Census 2000
Land area	155,959	Miles^2	US Census 2000
Fraction population metropolitan	84	Dmnl	US Census 1996
Fraction land metropolitan	0.08	Dmnl	US Census 1996
Registered automobiles	17,3e6	Vehicles	Bureau of Transportation Statistics http://www.bts.gov/publications/state_transportation_profiles/california/html/fast_facts.html
Gasoline fuel stations	7,949	Fuel stations	Provided by National Renewable Energy Lab (year = 2003)
Mean travel time to work	27.2	Minutes/ trip	US Census 2000
Annual vehicle miles	12,000	Miles/ year	Average US

Table 3 Default parameter settings; defaults not listed here have been specified in elaboration sections in Appendix 2.

Short	Description	Value	Units	Source/Motivation
Demand - Consumer Choice
τ^d	Time to discard a vehicle	8	Years	Close to Census values
u^*	Reference Utility	1	Dmnl	Free choice
$u^o_{zz'}$	Utility of alternative to drive	0.25	Dmnl	Heuristic
μ^t	Trip distribution parameter	-2	Dmnl	See discussion in text
β^t	Route choice sensitivity	∞	Dmnl	Simplifying dynamics
μ^ω	Route distribution parameter	1	Dmnl	
β^c	Elasticity of Utility to Cost	-0.5	Dmnl/ ($/trip)	Used to compare (coarsely) across elasticity
v^t	Value of Time	40	$/Hour	See research by e.g. Train (2005). Used to specify value of elasticity parameters, including refueling
v^r	Value out of Fuel	200	$/Empty Tank	Used to calculate w^r
γ^s	Relative Value of Time Service	1	Dmnl	Used to calculate w^x
γ^f	Acceptable refueling effort as fraction of trip effort	0.25	Dmnl	
v_u	Average drive speed	40	Miles/hour	
τ^s	Time to observe station density and wait time	1	Dmnl	As close as possible to 3 Months, simulation time constraints
r^{b0}	Reference Toping-off buffer	0.1	Dmnl	
Demand - Platform specific
q_i	Storage capacity per Tank	20	Gallon Equivalent	Equivalent to typical ICE
η^f_i	Vehicle fuel Efficiency	20	Miles/Gallon Equivalent	Equivalent to typical ICE
Station Economics
c^f_{vz}	Whole sale fuel price	1.65	$/gallon	Typical for US
c^o_{vz}	Non Fuel Variable Cost	0.6	$/gallon	Typical for US
f^a_{vz}	Ancillary revenues as fraction of value of 1 gasoline gallon equivalent consumed	0.2	Dmnl	Typical for US
m_{vz}	Fuel margin	0.5	dmnl	Typical for US

Short	Description	Value	Units	Source/Motivation
π_v^0	Reference Profitability	0.1	dmnl	
y_{iz_u}	Reference number of pumps per station	8	Pumps/station	Typical for US (Gasoline)
$k_{iz_u}^p$	Normal Pump Capacity	400	Gallons/hour	Average for California (Gasoline)
Station Behavior				
τ_{iz}^{ep}	Time to Permit Stations	1	Year	Part of τ^e
τ_{iz}^{el}	Time to Select Locations	1	Years	Part of τ^e
τ_{iz}^{ec}	Time to Construct Stations	2	Years	Part of τ^e
λ_{ref}^x	Normal station hazard rate (station hazard rate at zero ROI)	0.1	Dmnl/year	
β^k	Sensitivity of Entry to Local Profits	1	Dmnl	
τ_{vz}^k	Time to adjust capacity	1	Year	Though longer when population density is larger.

Technical appendix accompanying Essay 2

1 Introduction

The model described in this Essay is designed to capture the diffusion of and competition among multiple types of alternative vehicles and their fueling infrastructure. In the appendix I discuss additional components of the full model, highlighting those structures required to capture the full model. Further, this appendix provides additional information to accompany the model and the analysis of Essay 2. Each subsection is pointed to from a paragraph within the Essay. Following this introduction, subsequent sections group issues by:

2 **Elaborations on the model** that provide details on expressions that were not fully expanded due to space limitations (in particular we discuss functional forms).

3 **Derivations,** which discuss expressions that be derived through closed form derivations. These were highlighted in the paper but not fully expanded due to space limitations.

4 **Notes on simulations**, providing completions or complementary information to analysis in the paper.

5 **Stipulations:** Additional notes that provide insight in the model or analysis

6 **Model analysis and documentation:** Essay 2, in combination with the first two Appendix sections allows replicating the model. The third section allows the reader to replicate those analyses that did not provide sufficient information in the Essay to do so. Here I point to additional supporting documentation to do so.

7 **References**

2 Elaborations on the model

This section elaborates segments of the model that were highlighted in the paper but not fully expanded due to space limitations. These elaborations include in particular selected functional forms for functions that were provided in general form in the model.

a) Refueling sensitivity parameter

The refueling sensitivity parameter captures the propensity of drivers to refuel outside the location where they reach their normal refueling buffer r_{iz}^b (Equation 7 of the Essay).

The two functions g and h determine how this propensity depends on the effective buffer and driving range, relative to the trip length. The functions should increase in both inputs, but are bounded, as the relevant area of sensitivity is at the order of the trip length (if one has a top-off buffer of 100 miles and the trip is 20 miles, we can typically refuel anywhere we like along the way). We use:

$$g\left(r_{iz}^b/r_{i\omega_{zz'}}^t\right) = \min\left[1, r_{iz}^b/r_{i\omega_{zz'}}^t\right]^{\eta^b}$$

$$h\left(\left[r_{iz}^f - r_{iz}^b\right]/r_{i\omega_{zz'}}^t\right) = \min\left[1, \left[r_{iz}^f - r_{iz}^b\right]/r_{i\omega_{zz'}}^t\right]^{\eta^f}$$

The default setting in the model of both η^b and η^f are set equal 1.

The reference sensitivity β_{ref}^f determines how the refueling location sensitivity is constraint by a combination of behavioral and physical constraints of refueling on a different location. Formally, it gives the elasticity of refueling shares to a change in the utility of a refueling at a location, when the refueling buffer equals the trip length

determined by the physical and behavioral factors. Here we assume this to be fixed (see Table 3). Note that if drivers always intend to top-off at their topping-off buffer, this parameter is zero.

b) Scale economies for fuel stations

Fixed costs are defined in Equation 23 of the Essay as:

$$c_{vs}^{k} = c_{vs}^{k,ref} f^{k}\left(y_{vs}/y^{ref}\right); f(0) > 0; f(1) = 1; f' > 0; f'' < 0$$

Economies of scale in fuel station cost follow the standard diminishing returns to scale function:

$$f^{k}\left(y_{vs}/y^{ref}\right) = \left(y_{vs}/y^{ref}\right)^{\eta^{k}}$$

Further, station cost may, labor, and land may differ per region. In particular fuel stations in urban areas have a totally different costs than those in rural areas. Higher cost in urban areas wil suppress expansion and entrance. Population is a good proxy for consistent variation between them. Thus, I include a population dependent factor:

$$c_{vs}^{k,ref} = c_{vs}^{k,o} + c_{vs}^{k,h}\left(h_{s}/h^{avg}\right)^{\eta^{h}}$$

In the simulation I use the following parameters: $\eta^{k} = 0.25$, and

$\eta^{h} = 0.25; c_{vz}^{k,o} = 250,000; c_{vz}^{k,h} = 250,000$

Note that these parameter settings disvafovor adoption in urban areas relative to rural areas. Effects of excluding this have limited impact on the dynamics.

A note on explicit representation of multi-fuel stations:

Assuming that scale economies do not hold across technologies, it is reasonable to exclude the role of multi-fuel stations. Under such conditions, we can see a multifuel station as two neighbouring monofuel stations. This is assumption is reasonable when involving entirely different fuels such as natural gas and gasoline. Flexfuels are more likely to be substitutes from the station's perspective and require some more complicated scaling. In that case multi-fuels might offer lower barriers than specialized stations. In the current simulations I exclude the explicit representation of hybrid fuel stations.

c) Entry and exit sensitivity to profits

Equation 30 in the Essay defines the industry growth rate as

$$K_v^n = g_v^k K_v$$

$$g_v^k = g_v^{k0} f^g \left(\overline{\pi}_v^e \right); f(\ll 0) = 0; f(0) = 1; f' \geq 0; f'(\gg 1) = 0;$$

The constraints imply, first, that the growth rate equals g_v^{k0} when perceived returns on investment equal desired returns; second, that the growth rate increases with return on investment, which could differ by fuel, because of potential variation in constraints. Further, the shape is bounded, at zero, for extremely negative profits, and, at some finite value, for extremely high returns. The most general shape that satisfies these conditions is an S-shape. The logistic curve is used here, with the following parameter settings:

$$f^e \left(\overline{\pi}_v \right) = f^{LG} \left(\overline{\pi}_v ; \pi_v^0 ; 1; 0; 10; \alpha^{\max} \right).$$

See this Appendix, section 2f for a detailed specification of the functional form and interpretation of the parameter entries, but in short, the elasticity of industry growth to market profits equals 1 at the normal profits, and the growth entrance rate is smoothly bounded by 0 and 10 times the normal growth rate.

Exits, specified in Equation 32 of the Essay also follow an S shape, but have a negative elasticity of 1:

$$f^x\left(\pi_{vz}^x\right)=f^{LG}\left(\pi_{vz}^x;\pi_v^0;-1;0;10;\alpha^{\max}\right)$$

See this Appendix, section 2f for a detailed specification of the functional form and interpretation of the parameter entries, but in short, the elasticity of exits to profits equals -1 at the normal profits, and the growth entrance rate is smoothly bounded by 10 10 times the normal exit rate (at large losses) and 0 (large profits) .

d) Expected return on investment

Entrepreneurs derive the perceived net present value of operation over planning horizon τ^p and continuous time discount rate β. Expected retuns on investment $\pi_{vz\beta}^e$ are determined by the net present value of expectected revenues, minus net present value of costs, divided by net present value of cost:

$$\pi_{vz\beta}^e=\left(r_{vz\beta}^e-c_{vz\beta}^e\right)\!\big/c_{vz\beta}^e$$

The net present value of a constant stream s (income or expense) is represented by:

$$s_\beta=\int_0^{\tau^p}se^{-\beta t}=s/\beta\left[1-e^{-\beta\tau^p}\right]=v_\beta s \tag{A36}$$

This formulation is a good representation for expected net present value of, say, cost of capacity, or price. However, other values, in particular sales, adjust gradually over time. For, instance, the expiration of subsidies can be anticipated, which results in a gradual reduction of entrance in the last years of such a program. Similarly, placement of 5 stations in a periphery around, say, Sacramento can considerably increase the

attractiveness of AFVs, driving up sales of vehicles of that platform, followed by increased fuel consumption, but the impact on the return on investment depends strongly on the adjustment time (in relation to discount rate). The more general representation of an expected net present value of a variable stream s_β that adjusts with adjustment rate λ to its indicated value s^* equals:

$$s_\beta^e = \int_0^{\tau^p} \left[s^* - \left(s^* - s\right)e^{-\lambda t} \right] e^{-\beta t} = s_\beta + \Delta s_\beta - \Delta s_{\beta'} \tag{A37}$$

Where $\Delta s_\beta = s_\beta^* - s_\beta$ is the net present value of the goal (structurally defined in equation (A36)) and, with and $\beta' = \beta + \lambda$, $\Delta s_{\beta'}$ is the correction for the time needed to adjust to it. Note that if $\lambda \to \infty$, the third term drops out, and net present value equals that of a constant s^* stream. Further, the net present value of the sum of two variables is additive, while the net present value of the product of two variables is found through additivity in the adjustment rate and we can also write for therefore we can also write for equation (A37): $s_\beta^e = s_\beta + \Delta s_\beta^e$ and:

$$\Delta s_\beta^e = \Delta s_\beta - \Delta s_{\beta'} = \left(v_\beta - v_{\beta'} \right) \Delta s$$

The main challenge for stations is estimation of future sales $s_{vz\beta}^{fe}$ at entrance, which feeds into revenues $\left(r_{vz\beta}^e = p_{vz} s_{vz\beta}^e \right)$ and variable cost $\left(c_{vz\beta}^a = c_{vz\beta}^v + c_{vz\beta}^k ; c_{vz\beta}^v = c_{vz}^v s_{vz\beta}^e \right)$. We will discuss this here. Expected present value of revenues $s_{vz\beta}^e$ are those of current fuel sales plus the adjustment for growth $\Delta s_{vz\beta}^e$ induced by entrance, but can not exceed one's planned capacity:

$$s_{vz\beta}^e = \min\left[k_{vz}^*, \frac{F_{vz}}{(F_{vz}+1)} s_{vz\beta} + \frac{1}{(F_{vz}+1)} {}_\Delta S_{vz\beta}^e \right]$$

The first term on the right hand side of the min function equals current demand patterns at stations, adjusted for sharing of sales by an increased base of fuel stations. The second term captures the share of (net present value of) an anticipated increase of sales due to increased coverage, going to a new station. This increase in sales comprises four components: i) closing the gap between demand and sales, in the case of full utilization ii) increased share of current driver's refuelings in that area, iii) an increase of trips by adopters iv) an increase in adoption.

In an earlier version this has been derived and implemented, using the actual demand elasticities. An alternative, simpler, approach that is used in this and is now discussed. Potential entrants expect that demand in a zone can grow more the wider the gap is between perceived potential demand and perceived actual demand. Perceived potential demand s_{vs}^m, equals the total current demand in that region, of which the potential for the entrant is corrected by α_{vs}, that captures fuel specific factors (e.g. higher fuel efficiencies result in less potential demand), and contextual factors (the aggregate of factors discussed above):

$$\Delta s_{vs}^* = f_{vs}^* \max\left[0, \left(\alpha_{vs} s_{vs}^m - s_{vs} \right) \right]$$

The effectiveness to attract more demand, f_{vs}^*, depends on how the infrastructure coverage is changed as a function of entrance. The heuristic follows the one that let us draw the demand curve in Figure 2 of the Essay. At zero existing stations, the responsiveness will be very low, similarly when infrastructure is already very abundant.

170

However, when stations are reasonably spars, say halfway the normal demand, entrance responsiveness is expected to be high:

$$f_{vs}^* = f_{vs}^0 g\left(s_{vs}/s^{ref}\right) h\left(s_{vs}/s^{ref}\right) \begin{cases} g(0)=0; g(1)=1; g'\ge 0; g''(0)>0; \\ h(0)=1; h(1)=0; h'\le 0; h''(0)<0; \end{cases}$$

We use two standard symmetric, bounded at 0 and 1 logistic curve functions (see Appendix 2f), with sensitivity parameters of respectively g and h being 2 and -2.

e) Exits: weight of expected profits for mature stations

Stations enter based on expected return on investment, and during a honeymoon period, losses may well be anticipated. Equation 34 in the paper captures the different behavior for mature and new to industry stations, through the weight of the relevant expected profits. As we study early transition dynamics, it is important to capture the reality that new stations can stay in business, holding on to their business case, eventhough no profits are being made. New stations therefore base their exit rate on adjusted expected profits.

For the weight function given to recent profit streams increases with the average maturity of the stations we use the logistic curve:

$$f^L\left(L_{vz}\right) = f^{LG}\left(L_{vz}; 4; 0.25; 0; 1; 1\right)$$

See this Appendix, section 2f for a detailed specification of the functional form and interpretation of the parameter entries, but in short, the elasticity of growth to profits equals 4 at the normal profits, and the growth entrance rate is smoothly bounded by 0 and 1.

f) General forms, the logistic curve

In the Essay several functional forms were specified in general terms, including boundary constraints as normal values, extreme conditions, first and second derivatives. For several of them a general S-shape curve is a natural form. For those we specify here the exact expressions used in the simulation. While many forms are available, I use the Logistic Curve. I will present this here in a more general form, and specify parameters where applied:

$$f^{LG}\left(x; x_0; \beta; \max; \min; \alpha\right) \equiv \min + \frac{\alpha\left(\max - \min\right)}{\alpha + \exp\left[-\beta'\left(\dfrac{x - x_0}{x_{ref}}\right)\right]} \tag{A38}$$

With min and max as specfified, output at x=0 equal to:

The inflection point at x0 has value:

$$f^{LG}\left(x_0\right) = \frac{\alpha \max + \min}{\alpha + 1} \Rightarrow$$

$$\alpha = \frac{f^{LG}\left(x_0\right) - \min}{\max - f^{LG}\left(x_0\right)} = \frac{y^{LG}\left(\max - \min\right) + \min - \min}{\max - \left(y^{LG}\left(\max - \min\right) + \min\right)} = \frac{y^{LG}}{\left(1 - y^{LG}\right)}$$

Where y^{LG} is the locus of the inflection point as fraction between the max and the min. If we want to set $f^{LG}\left(x_0\right)$ to 1, provided min<1:

$$y^{LG} = \frac{1 - \min}{\Delta}; \Delta = \left(\max - \min\right)$$

Next, if

$$\beta' = \frac{\min + y^{LG}\Delta}{f_0^x y^{LG}\left(1 - y^{LG}\right)\Delta}\beta; f_0^x = \frac{x_0}{x_{ref}}$$

Then the elasticity of output to the input at the inflection point equals:

172

$$\varepsilon_{LG/x} = \frac{x}{f^{LG}} \frac{df^{LG}}{dx} \Rightarrow \varepsilon_{LG/x}\big|_{x_0} = \beta$$

Note further that the symmetric configuration, $\alpha = 1$, renders the standard logistic curve.

However, with minimum at min, maximum at max, elasticity at the inflection point specificed and the output equal to 1 at x_0, we have:

$\alpha = \dfrac{1}{(\text{max}-1)} \equiv \alpha^{\text{max}}$. Note that with max to infinity we get the exponential function:

$$f^{LG}\left(x; x_0; \beta; \infty; 0; 0\right) = \exp\left[\beta\left(\frac{x - x_0}{x_0}\right)\right]$$

3 Derivations

In this section I derive analytical expressions, including the average route effort (discussed with Figure 5), refills per trip, and the trip effort inputs average refueling distance, out of fuel risk and service time.

a) Notes on derivation of trip effort components

In the following treatment we assume that all searches for fuel occur within the zone (used interchangeably with "patch") s in which refueling is desired. This is justified as in the current analysis the zones are naturally chosen large enough such that in search for fuel within a zone, and small enough to allow capturing the effects of heterogeneous population concentrations. For deriving the average risk of running out of fuel, the average refueling effort, I use a discrete grid, with patches defined at a much smaller

scale than that of the patches *s* or *z*. Where preferred I will resort to polar coordinates, using (l, θ).

b) Route effort and probability of a refill

Here I explain how I derive at the route effort expression in Figure 5 (row 3 of the aggregate utility and effort column). For the average trip effort I use an approximation of the expected trip effort, which aggregates over the probability of refueling n times $p_{i\omega_{zz'},n}$, multiplied with the corresponding effort $a^t_{i\omega_{zz'},n}$. Assuming that various refueling events are uncorrelated, which holds true when averaging over a large population, this equals the effort of not having to refill, plus the summation (to infinity) over the number of rfill events n , of n multiplied with the refueling probability and the net effort of refueling 1 time, $a^{f1}_{i\omega_{zz'}} - a^{f0}_{i\omega_{zz'}}$

$$a^t_{i\omega_{zz'}} = \sum_n p_{i\omega_{zz'},n} a^t_{i\omega_{zz'},n} \approx a^{t0}_{\omega_{zz'}} + \sum_n n p_{i\omega_{zz'},n} \left(a^{f1}_{i\omega_{zz'}} - a^{f0}_{i\omega_{zz'}} \right)$$

The product $n p_{i\omega_{zz'},n}$ is the only part that is a function of n. This summation equals the expectd refills per trip, $\phi_{i\omega_{zz'}}$, and the previous expression can be further simplified to

$$a^t_{i\omega_{zz'}} = a^{t0}_{\omega_{zz'}} + \phi_{i\omega_{zz'}} a^f_{i\omega_{zz'}}$$

Where $a^f_{i\omega_{zz'}} \equiv \left(a^{f1}_{i\omega_{zz'}} - a^{f0}_{i\omega_{zz'}} \right)$. For each individual refueling location this equals the average effort of refueling $a^f_{i\omega_{zz'}}$.

c) Refills per trip

The refills per trip can be found by solving from Equations (6) and (8), using that the refills per effective tank range r_{iz}^f equals 1. Then:

$$\phi_{i\omega_{zz'}s}^* = r_{\omega_{zz'}}^t \bigg/ \left(r_{iz}^f - \sum_{s \in \omega_{zz'}} \sigma_{i\omega_{zz'}s}^f r_s^f \right).$$

The denominator provides a corrected effective tank range that is reduced because of the search for fuel. We see that if the expected distance to obtain fuel approaches the actual tank range, this term diverges. This is the situation corresponds with the situation that there is not enough fuel to be found along the whole trip, to bring us home. The divergence is physically sound. Note for instance that at this point the utility for making the trip approaches zero (see Figure 5). However, the negative constraint is not. To deal with this in a consistent manner, I assume that the range to find fuel in each location is bounded by the actual tank range, representing an option to call a service to fill you up. However, the cost in time and money is very large, thus at this time the effect of refueling effort on utility is at this point already reduce it to zero, consistent with this, thus:

$$\phi_{i\omega_{zz'}} = r_{\omega_{zz'}}^t \bigg/ \left(\max\left[0_+, r_{iz}^f - \sum_{s \in \omega_{zz'}} \sigma_{i\omega_{zz'}s}^f r_s^f \right] \right)$$

d) Average refueling distance

The average distance of a refueling point to the desired refueling location, $\langle r_{is}^d \rangle$ is found by summing over the probability that the nearest station is at a distance r_l from the desired refueling point in s, p_{isl}^* multiplied by the distance:

$$\left\langle r_{is}^d \right\rangle = \sum_{l \geq 0} \frac{n!}{l!(n-l)!} r_l p_{isl}^*$$

which equals the probability that at least one station exists at a ring with radius r_l and with dl, minus the probability that a station within that ring within r_l of s, p_{isl-1}:

$$p_{isl}^* = p_{isl} - p_{isl-1}$$

The probability of finding a station until l, equals 1 minus the probability of finding no station:

$$p_{isl} = 1 - p_{isl}^0$$

Given the poisson characteristics, p_{isl}^0 this equals (e.g. Pielou 1977):

$$p_{isl}^0 = \exp\left(-F_s A_l / A_s\right) \tag{A39}$$

Where $A_l = 2\pi r_l dl$.

Below I plot, for reference the relative effective trip duration, as a function of the effective tank range (wich determines the trip frequency), and thus the effective search time, and the trip length.
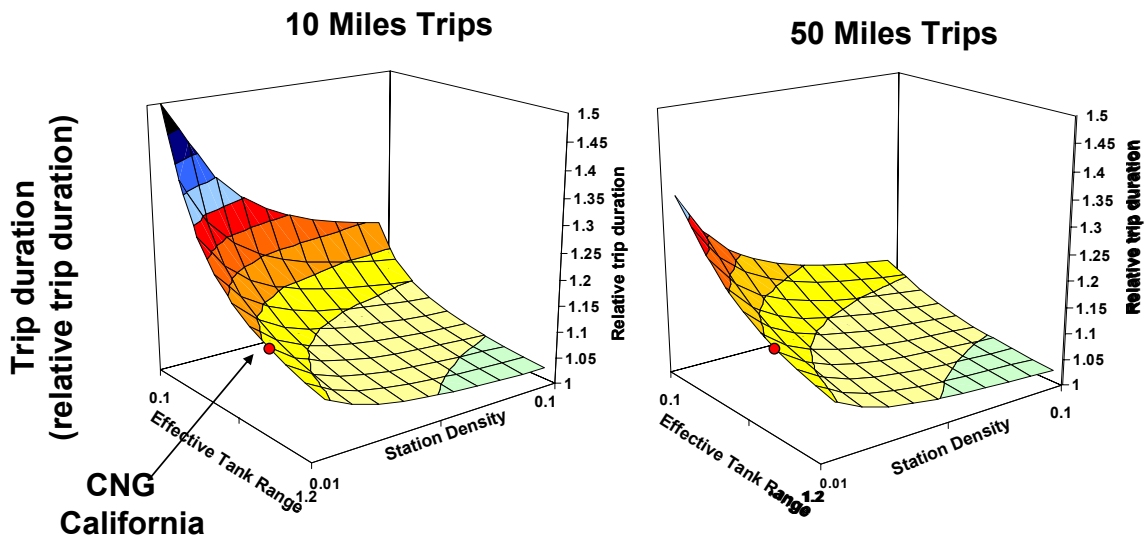
Figure A1 – trip duration

No parameters are required to derive this function. Trip duration is especially sensitive to station density for short trips.

e) Out of fuel risk

The expected risk of running out of fuel is derived by summing over probabilities of running out of fuel at distance r_l from the desired refueling location, given topping-off buffer r_{iz}^b. Such a probability requires not having encountered a station within one's driving radius r_l, p_{ls}^{0-}, times the probability of running out of fuel in location p_{izsl}^o, conditional upon not having run out of fuel before:

$$\langle o_{izs} \rangle = \sum_{l>1} p_{izsl}^{o1} p_{izsl}^{o0-}$$

Where an out of fuel in location l equals to the probability of getting out of fuel at a distance r_l from the desired refueling point p_{izsl}^o, conditional upon not having been out of fuel before:

$$p_{izsl}^{o1} = p_{izsl}^o c_{izsl-1}^{-o} \tag{A40}$$

The cumulative out of fuel probability is a function of the tank range r_i^f and the buffer r_{iz}^b:

$$c_{izl}^o \quad = f\left(r_l / r_i^f\right); f(0) = 0; f\left(r_{iz}^b / r_i^f\right) = 0.5; f(1) = 1; ; f' \geq 0$$

The probability of not having been out of fuel can be derived from this expression, when using simple exponential expressions for p_{izl}^o, otherwise it can be approximated through:

$$c_{izl-1}^{-o} = \left(1 - c_{izl-1}^o\right) \tag{A41}$$

While

$$p_{izl}^{o} = \left(c_{izl}^{o} - c_{izl-1}^{o} \right) \qquad \text{(A42)}$$

Figure A2 shows a graphical representation for relatively short (10 miles) and longer trips (50 miles).

Figure A2– fuel reliability

Comparing this with the general results for search efforts, we note that the driving effort component is most easily affected during short trips, while the out of fuel risk grows faster in larger trips.

f) Mean waiting time for service

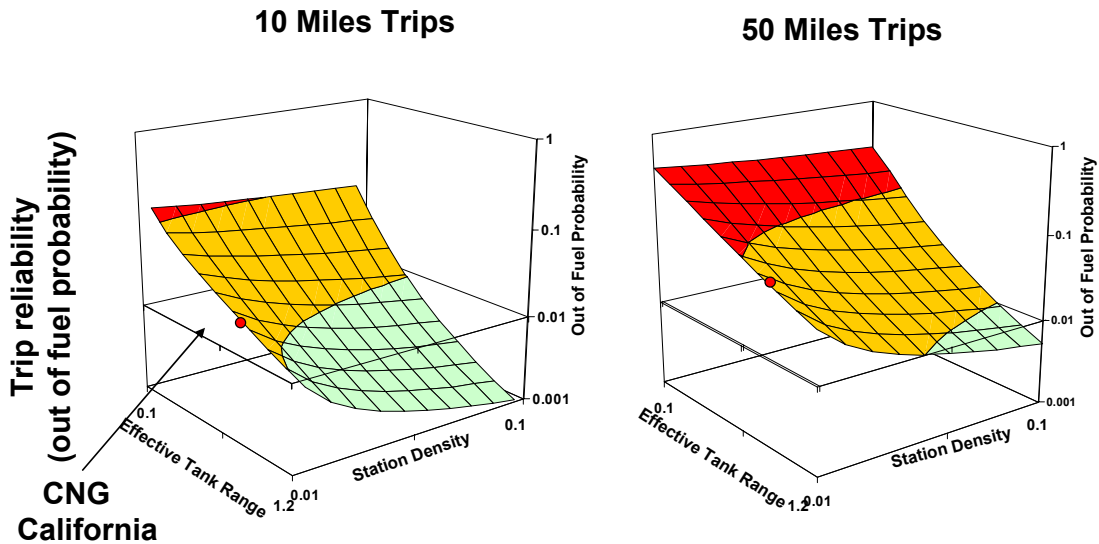Disequilbirum between supply and demand are very critically felt at the pump. In Argentina and New Zealand that have experienced a take-off of CNG, waiting times have

been found to be in the order of 2 hours.[18] The mean waiting time at a station is derived

through stationary solutions of basic queuing theory concepts. This provides insights on

the average wait time as a function of average utilization, number of pumps and pump

capacity. The assumptions we make are simplified, but provide excellent insights on the

strong non-linearities involved. We assume customer arrival rate at stations in s, for

drivers of platform i to be uncorrelated and Poisson distributed. Assuming more complex

demand patterns, such as peak behavior would yield average wait times that are even

larger. The average arrival rate per station is the sum over arrival rates from all regions

z, $\lambda_{is}^s = \left(\sum_z \lambda_{vzs} \right) \Big/ F_{vs}$. The arrival rate for refills in region s for platform v from region

s, λ_{vzs} , is given by the average refills during trips between z and z', the actual trip

distribution between zz' and the number of adopters in z:

$$\lambda_{vzs}^s = \sum_{i \in v, z'} \phi_{izz's} T_{izz'} V_{iz} \qquad (A43)$$

With refills per trip from location z to z', with underway s:

$$\phi_{izz's} = \sum_w \sigma_{izz'sw}^f \sigma_{izz'w}^t \phi_{izz'w} \qquad (A44)$$

A second requirement for the (basic) queuing concepts is to assume the servicing time at

the pump to be exponential. The strong assumption can be easily relaxed, for instance by

assuming more sophisticated Erlang distributions, but this is sufficient to surface the

strong non-linearity and analytically convenient. Other second order effects derive from,

for instance, the number of stations in an area.

---

[18] Jeffrey Seissler, Executive Director of the European Natural Gas Vehicle Association (ENGVA) –
personal communication July 2006.

The average duration τ_{is}^{sf}, is found by averaging over the service duration from all

customers:

$$\tau_{is}^{sf} = \sum_z \lambda_{izs}^s \tau_{izs}^{sf} \Big/ \sum_z \lambda_{izs}^s \tag{A45}$$

The steady state wait time, that is, the waiting time for $t \to \infty$ is derived from the

constraint that the sum of over all probabilities of finding k customers in the system

should be equal to 1. The derivations are done for total demand being smaller than supply

(the interesting area). First we define the average station load factor, $\rho_{is} = \lambda_{is}^s \tau_{is}^{sf}$, which is

by the foregoing assumption smaller than the number of stations, and the average stations

available $\alpha_{is} = (y_{is} - \rho_{is})$. Then, the probability that k customers demand fuel, P_k,

expressed in the probability that no customers demand fuel P^0 (for derivation see e.g.

Gnedenko and Kovalenko (1989)), omitting subscripts for clarity:

$$P_k = \begin{cases} \dfrac{\rho^k}{k!} P^0 & 0 \le k \le y \\[2ex] \left(\dfrac{\rho}{y}\right)^{k-y} P_y & k > y \end{cases}$$

Then, with $(y-\rho)/y \sum_{k>y}^{\infty} (\rho/y)^k = (\rho/y)^{y+1} \Rightarrow y^y \sum_{k>y}^{\infty} (\rho/y)^k = \rho^{y+1}/(y(1-\upsilon))$ the

probability of having no customers waiting equals:

$$P^0 = \left[\sum_{k=0}^{y} \frac{\rho^k}{k!} + \frac{1}{(1-\upsilon)} \frac{\rho^{y+1}}{y+1!} \right]^{-1} \tag{A46}$$

And the probability that all pumps are busy equals

$$P^q = \sum_{k=y}^{\infty} P_k = \frac{1}{(1-\upsilon)} P_y = \frac{1}{(1-\upsilon)} \frac{(y\upsilon)^y}{y!} P^0 \tag{A47}$$

In the case of one pump per station, this equals the average utilization of a station υ, according to intuition.

This intermediate outcome is important: it shows that when the average station utilization becomes high $(\upsilon \to 1)$, the probability that someone finds all pumps increases dramatically. Further, this probability is also highly dependent on the number of busy pumps, even when the total utilization is constant. Finally the mean waiting time can be shown to be:

$$\left\langle \tau_{is}^{sw} \right\rangle = \frac{P_{is}^{q}}{y_{is}\left(1 - \upsilon_{is}\right)} \tau_{is}^{sf} \tag{A48}$$

Figure 3 shows, for one set of parameters, technological parity with gasoline stations and ICE vehicles (e.g. pump capacity, vehicle driving range), utilization and relative service time for increasing demand supply imbalance and increasing number of pumps per station. I use an estimated 8 pumps per station. Of interest is the steep non-linearity of service time for low utilization, especially for fewer pumps per station).
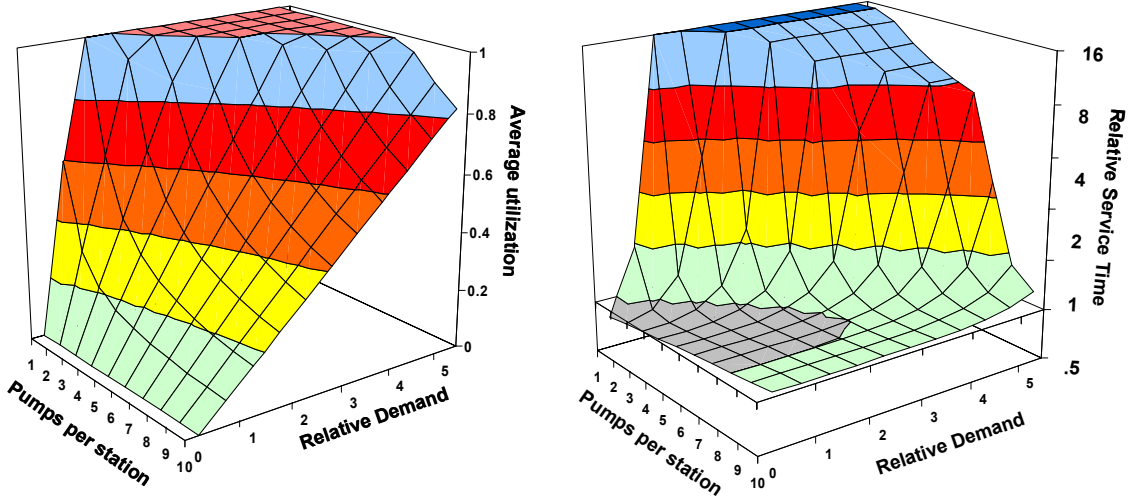
Figure A3– Utilization and Relative Service Time for increasing pumps per station and relative demand (Technological Parity with Gasoline stations and Drive Range).

4 Notes on simulations

a) Trip generation and trip relevance

For the simulation we derived the aggregate for each driver d from a two-parameter lognormal distribution:

$$f_n^t = f^{TOT'}\left(r^{t,mean}/r^t\right)\Big/\left(\sigma^t\sqrt{2\pi}\right)\exp\left[-\left(\frac{\sigma_t}{2\sqrt{2}}+\frac{\ln\left(r^t/r^{t,mean}\right)}{\sqrt{2}\sigma^t}\right)^2+\frac{\sigma_t^2}{2}\right]$$

with σ_t the standard deviation and mean distance, $r^{t,mean} \equiv r^{t,ref}e^{0.5*\sigma_t^2}$; $f^{TOT'}$ equals the total annual trip frequency $f^{TOT'}$, times the cumulative distribution until a maximum range r^{max}. Specific data can be derived from trip-tables (e.g Domencich et al. 1975), but here we assumed identical average trip behavior across the regions:

$r^{t,mean} = 27, r^{max} = 120$ miles per trip and $\sigma^t = 0.5$, yielding the total f^{TOT} of 300 trips per

year (f^{TOT} derived through integration over the populations corresponds with average

annual miles $m = h \int_0^{r^{\max}} \mathrm{r}^{\mathrm{t}} f_n^t$).

The total vehicle miles for a drivers of platform i equal:

$$m^{v,\max} = \sum_{z'} r_{zz'} T_{zz'}^{\max}$$

And are set to 15,000 miles per person per year. Subsequently, $T_{zz'}^{\max}$ was derived by dividing of trips between region z and z' assuming uniform distribution in radius .

Trips between regions are weighted by desired frequency and distance, thus, with

$$m^{v,\max} = \sum_{z'} r_{zz'}^{\eta^w} T_{zz'}^{\max}$$, we have $w_{zz'} = m_{zz'}^{v,\max} \Big/ \sum_i m_{zz'}^{v,\max}$.

In the simulations $\eta^w = 2$. The combination of trip weight and frequency render the following distributions:

Trip Frequency and Weigth
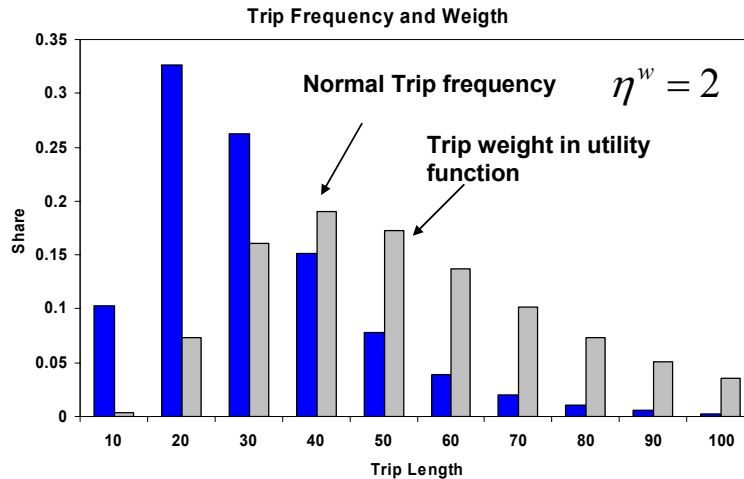
Figure A4 Normal Trip Frequency, and trip weight for the deteriming average trip attractiveness (see Figure5 in the Essay). To speed up the computation, throughout the analysis, drivers only select the direct route $\beta^t \to \infty$ (see Table 3).

b) Figure 12a – tipping point for a one patch model

Figure 12a in the Essay discusses the equilibrium dynamics as a function of the patch length. Under the assumption of one patch, the model structure corresponds with assumptions of uniform population distribution. This assumption does not bring out strategic location incentives on the supply side, and, even under rich behavioral assumptions we can plot a unique adoption curve as a function of the number of fuel stations, that will yield demand /supply responses that correspond with the qualitative sketch in **Figure 2.** This graph shows the equilibrium adoption fraction under the default

simulation assumptions:
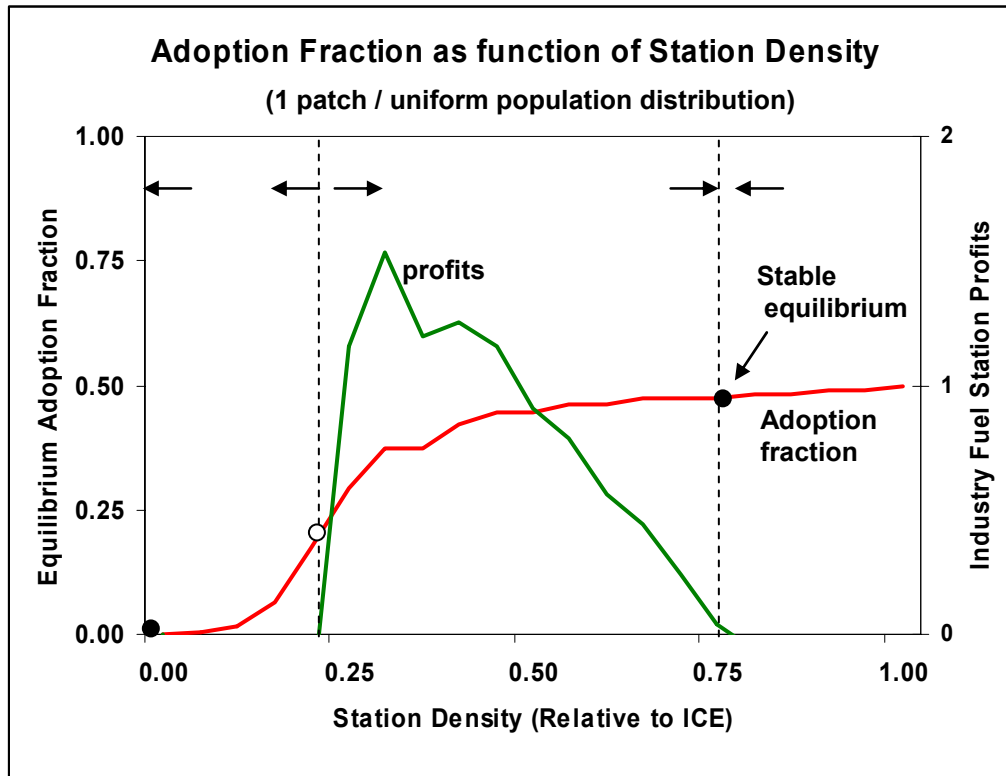
Figure A5 Adoption fraction and fuel station profits for a 1 patch simulation

Where profits equal zero (or for 0 fuel stations), we can expect an equilibrium. We see that under these assumptions 2 platforms can be supported. But getting towards that requires significant investment. While useful as a starting point, the assumptions for the uniform distribution ignore many feedbacks that involve critical dynamics.

c) Figure 12b – directed trip simulation

The full Los Angeles region is included, including San Diego. Further, to the north, Fresno, San Jose, and Sacramento. Figure A6 right provides summary statistics of population distribution for each landuse, as well as the average trip frequency and miles

(short and long trips). Minimum pop density indicates the selection criterium for each landuse type and is measured against the average population of the region. The region contains 83 % of the California populion and 58% of the land area. (Figure A6).

Region Parameters	Urban	SubUrban	Rural	Total
Minium Pop Density	10	0.75	0	
Area Fraction	0.09	0.29	0.62	1.00
Population Fraction	0.70	0.27	0.01	
Population (M)	19.8	7.6	0.33	28
Short Trip Frequency	97	89	86	94
Long Trip Frequency	24	11	7	20
Total Trip Frequency	122	100	93	114
Short Trip Distance	7,453	5,380	5,155	6,767
Long Trip Distance	6,131	2,870	1,992	5,120
Total Trip Distance	13,584	8,250	7,147	11,887

California Parameters	Urban	SubUrban	Rural	Region / California
Fraction of Total Area	0.05	0.19	0.75	0.58
Fraction of Population	0.63	0.32	0.05	0.83

1. Shasta Cascade
2. North Coast
3. Central Valley
4. Gold Country
5. High Sierra
6. San Francisco Bay Area
7. Central Coast
8. Desert
9. Los Angeles Count
10. Inland Empir
11. Orange C
12. San

Figure A6 Selected region for directed trip simulation (left) and summary statistics of geographical and the generated desired driving behavior (right).

Short-distance trips follow the same distribution as those in the base model (with random direction). However, to conserve computation time, trips of that category were cutoff beyond a 50 miles radius for urban and suburban population and a 30 miles radius for suburban and rural population. Long-distance trip destinations drawn from each location (23 for urban, 10 for suburban and 8 for rural), to a limited set of destinations as well,

weighted by population and distance from home to final destination: Population weights

increased linearly with density: $w^p = \left(h_{z'}/h_{ref} \right)$ if $h_{z'} \geq h_{ref}$, 0 otherwise; distance between

zones, with $w^d = \max\left[1, \left(d_{ref}/d_{zz'} \right) \right]$, with $d_{ref} = 100\, miles$. Thus, a long distance

destination 200 miles away from one's home location was twice as likely to be drawn

than a trip 400 miles away from z. Finally 6 hot-spot areas were handpicked: Las Vegas,

Lake Tahoo, Crescent City, Alturas, Mamoth Lakes, with weight being set equal to a

region with 10-20 times the average population. If destination fell outside the boundary

of the region, the nearest point to the region was selected. (. Figure A7 shows the demand

profile that is generated by total population and its trip profile in Mgallons/year/ zone

(top). Also in Figure A7 (bottom) the actual gasoline stations. Average demand per

station is 1.1 Mgallons/year. Therefore, the demand/ gasoline ratio in each region is a

good indicator for proficiency of the generated trips.

Figure A7 Comparing generated demand with actual supply with in selected region: top shows generated gasoline demand (Million Gallons per year per zone; D=6,751e6);

bottom shows the distribution of gasoline stations (Top, N=6,499). White areas <5 units; green <20; yellow < 50; orange ≥50.

d) Figure 13 – endogenous topping off

Figure 13 in the essay simulates the endogenous topping-off buffer. The general functional form was provided to be:

$$r_{iz}^b = f\left(u_{iz}^t\right)r_{i0}^b; f' \leq 0; f\left(0\right) = r_{max}^b / r_{i0}^b; f\left(1\right) = 1; f\left(\infty\right) = r_{min}^b / r_{i0}^b$$

The relative top-off buffer increases with decreasing utility, but stabilizes at r_{max}^b for very low utility, as drivers will not want to be constrained by refilling on average too early.[19] Further, when drivers are fully confident, they will reduce their buffer to r_{min}^b, which can be below the indicated level by the warning sign, r_0^b. To satisfy these conditions we use a simple one parameter form for f:

$$r_{iz}^b = \max\left[r_{min}^b, r_{max}^b\left(1/\left(1 + xu_{iz}^t\right)\right)^{\alpha^f}\right]; x = \left(r_{max}^b / r_{i0}^b\right)^{1/\alpha^f} - 1$$

which yields, for the selected parameters:

_____

[19] This level depends on the physical constraint of refueling elsewhere; see also equation (7) and Figure (5). From this behavioral reasonable parameters could be derived.

189
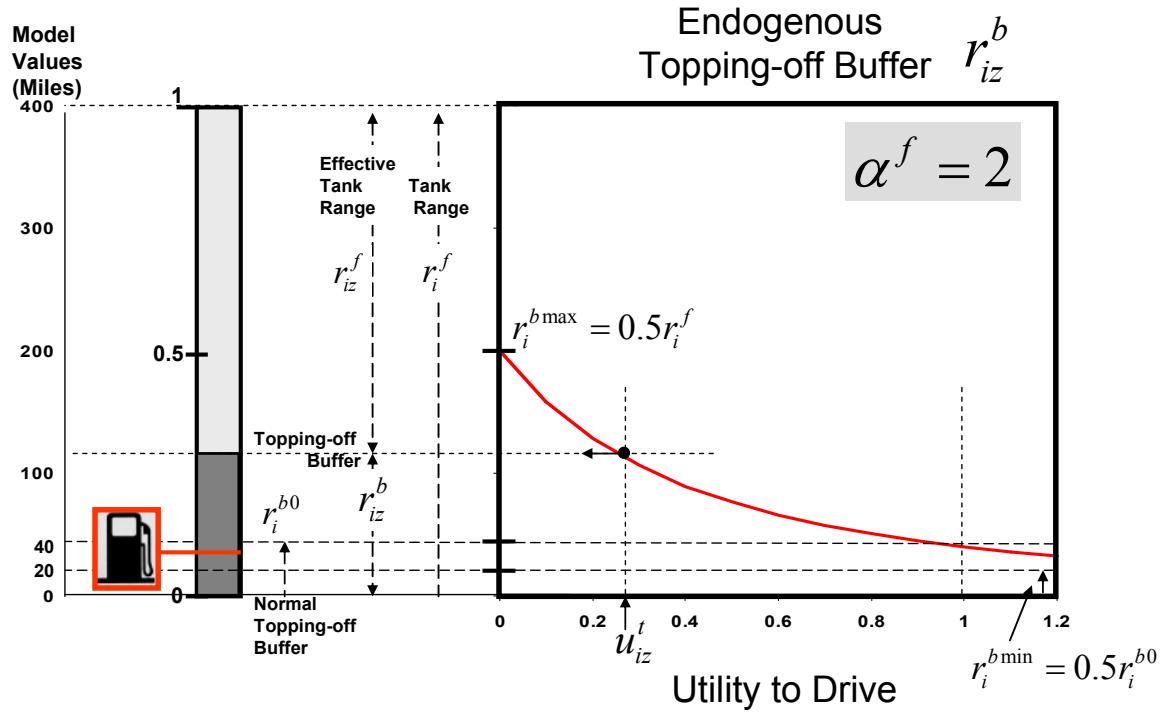
Figure A8 Endogenous topping-off buffer as a function of utility.

The reference topping-off buffer is $r_{i0}^b = 40$ miles (10% of the total range),

$r_{max}^b = 200$ miles (50% of the total tank range), and $r_{min}^b = 20$ miles. We see that, for

instance, the utility equals 0.27, the topping-off buffer becomes equal to 100 miles

(corresponding, under current assumptions, but with uniform population and fuel station

distribution, and in absence of crowing, with a station density of about 19% of the normal

density).

e) Figure 14 – table for technology parameters

Table 1 Parameters for the 3 scenarios:

Parameter	Reference	Relative to reference		
		Scenario 2	Scenario 3	Scenario 4
Relative fuel efficiency	20 miles/gallon	3	3	3
Relative tank size	20 gallons/tank	0.33	0.33	0.25
Relative pump refill rate	400 gallons/hour	1	0.25	0.25
Relative fixed cost (at four pumps)	50,000 $/refill point	1	0.25	0.40
Relative wholesale fuel cost	1.65 gallons/hour	1	1	1.25

5 Stipulations

This section provides additional comments and clarifications on assumptions, or on
connections.

a) Sensitivity parameters for trip efforts

The Essay describes how several attributes are brought together to detremine the trip
one's utility to drive u_{iz}, (Figure 5). Below I discuss how the relative weights can be
interpreted and validated.

The elasticity of one's utility to drive u_{iz}, to a change effort component c, with $c=\{$drive
time, out of fuel risk, refueling service$\}$, when all attributes $a^f_{zz'c}$ are at their reference
level $a^{f*}_{zz'c}$ equals:

$$\varepsilon_{u-a^{fc}_{zz'}} = \frac{a^{f*}_{zz'c}}{u}\frac{du}{da^f_{zz'c}}\bigg|_{a_{zz'c}=a_{zz'c}} = \beta^t\frac{\phi_{izz'}w_c a^{f*}_{zz'c}}{a^{t*}_{zz'}} \tag{A49}$$

Where $a_{zz'}^{t0}$ is the shortest trip effort between z and z' and $\phi_{izz'}$ is the normal refueling

frequency. Further, The reference effort, $a_{zz'}^{t*}$, equals the reference trip time, plus the

frequency of refueling multiplied with the reference levels for each attribute:

$$a_{zz'}^{t*} = a_{zz'}^{t0} + \phi_{izz'} \sum_c w_c a_c^{f0} \qquad (A50)$$

With a_c^{f0} being the acceptable level (e.g. $a_2^{f0}=0$, we don't accept out of fuels). For

example, ignoring the role of out of fuel and refueling time, we see that the actual

elasticity of utility to drive depends on the search time for fuel, relative to the normal

travel time for a trip, times the refueling frequency, times weight of finding fuel, and the

elasticity to trip effort:

$$\varepsilon_{u/a_{zz'}^{fd}} = \beta^t \phi_{izz'} w^d \frac{\tau_{zz'}^{fd*}}{\tau_{zz'}^{d*}}$$

Note further that the elasticity of utility to refueling in total equals:

$$\varepsilon_{u-a_{zz'}^f} = \frac{a_{zz'}^f}{u} \frac{du}{da_{zz'}^f}\bigg|_{a_{zz'}^f=a_{zz'}^{f*}} = \beta^t \frac{\phi_{izz'} a_{zz'}^{f*}}{a_{zz'}^{t0*} + \phi_{izz'} a_{zz'}^{f*}} \left(= \beta^t \sum_c \varepsilon_{u-a_{zz'}^{fc}}\right) \qquad (A51)$$

and the elasticity of utility to a change of a component, at the normal level, relative to the

elasticity of utility to a change of another component is a direct measure of their relative

weight:

$$\varepsilon_{u-a_{zz'}^{fx}} \Big/ \varepsilon_{u-a_{zz'}^{fx}} = w^x a_{zz'}^{f*x} \Big/ \sum_x w^x a_{zz'}^{f*x} .$$

Together this gives an interpretation of the relative importance of the attributes, with

respect to each other and compared to the trip effort as a whole, determined at some

useful reference point, e.g. at 30% station density of current. At that level, the out of fuel

risk might be very low, say 1%, but its weight can be very large.

We now set the weight for searching for fuel equal to the value of time v^t, divided by the a parameter that measures how time of getting fuel is weighted against spending effort /time driving towards a destination, $\gamma^f : w^d = v^t / \gamma^f$. Similarly, for out of fuel risk: $w^r = v^r / \gamma^f$, while the weight for service time is equal to that of searching fuel, corrected for a parameter that measures the weight of time waiting for fuel, with that of $w^d = \gamma^s w^d$.

6 Model and analysis documentation

The model and analyses can be replicated from the information provided in the Essay and the first two sections in the Appendix. However, analysis involved several steps and different tools. For instance, the population distribution for the proper gridsize was derived in Excel, while the static trip distributions (trip generation) were calculated in Matlab, using also the population information. Next each was uploaded in Vensim for simulation. Model source code and instruction for replication of the analysis can be downloaded from

http://web.mit.edu/jjrs/www/Thesis Documentation.htm

7 References

Domencich, T. A., D. McFadden, et al. (1975). *Urban travel demand : a behavioral analysis : a Charles River Associates research study.* Amsterdam New York, North-Holland Pub. Co. ;
American Elsevier.
Gnedenko, B. V. and I. B. N. Kovalenko (1989). *Introduction to queueing theory.* Boston, Birkhäuser.

Essay 3

***Alternative fuel vehicles turning the corner?:* A product lifecycle model with heterogeneous technologies**

Abstract

The automotive industry may be on the verge of a technological disruption as different alternative fuel vehicles are expected to enter the market. Industry evolution theories are not unified in suggesting the conditions under which different types of entrant technologies can be successful. In particular, the competitive dynamics among a variety of technologies with varying potential for spillovers are not well understood. This essay introduces a product life cycle model used to analyze the competitive dynamics among alternative fuel vehicles, with explicit and endogenous product innovation, learning-by-doing, and spillovers across the technologies. The model enables in particular the exploration of the spillover dynamics between technologies that are heterogeneous. I explore how interaction among learning and spillovers, scale economies, and consumer choice behavior impacts technology trajectories of competing incumbents, hybrids, and radical entrants. I find that the existence of learning and spillover dynamics greatly increases path dependence. Superior radical technologies may fail, even when introduced simultaneously with inferior hybrid technologies. I discuss the implications for the prospective transition to alternative fuels in transportation. While the dynamics are discussed in relation to the automobile industry, the model is general in the sense that it can be calibrated for different industries with specific market, technology, and organizational characteristics.

Introduction

Mounting economic, environmental, and security-related concerns put long-term pressure on a largely oil-based transportation system. In response, automakers are developing alternative technologies, such as hydrogen fuel cell vehicles (HFCVs), to transition away from the petroleum-guzzling internal combustion engine (ICE) vehicle fleet. A central and hotly debated issue among stakeholders is the feasibility of various transition paths

towards a vehicle fleet powered by renewable energy. For instance, according to some, HFCVs are a radical innovation with long-term socio-economic advantages and are therefore bound to replace current automobiles (Lovins and Williams 1999). On the other hand, current cost and performance factors disadvantage hydrogen relative to the established ICE-gasoline system, creating large barriers to entry (Romm 2004).

Adding to the complication is the plurality and diversity of other alternatives being considered. Besides leapfrogging to HFCVs or electric vehicles (EVs), some automakers are focused on increasing the efficiency of the current ICE technology. Others emphasize shifting to alternative fuels, such as compressed natural gas or blends of bio- and fossil fuels or are exploring various combinations of these alternative technologies, such as ICE-electric hybrids (ICE-HEVs), diesel-electric hybrids, or hydrogen-ICE (MacLean and Lave 2003). Beyond the fact that each technology trajectory involves large upfront investments, an alternative fuel transport system will drastically transform the social, economic, and organizational landscapes, with implications well beyond the automotive industry. With so much at stake, a thorough understanding of the transition dynamics is crucial.

How do different technologies come to be, gain traction, and sustain themselves? The general pattern dominating the post-industrial perspective regarding technological innovation is the S-shaped diffusion path of superior or novel technologies (e.g., Griliches 1957). This diffusion pattern is currently considered a stylized fact (Jovanovic and Lach 1989), with numerous documented examples including: end products such as

motor cars (Nakicenovic 1986) and laser printers (Christensen 2000); process

technologies (Karshenas and Stoneman 1993); enabling products such as turbo jet

engines (Mowery and Rosenberg 1981) and mini mills (Tushman and Anderson 1986);

ideas and forms of social organization (Strang and Soule 1998). While a powerful for ex-

post finding, this transition concept is useful for the dynamics of prospective transitions if

we have a thorough and detailed understanding of the mechanisms underlying the

outcomes[20].

Examination of the mechanisms underlying transitions is required, first, because several

hypotheses about the mechanisms underlying the S-curve pattern co-exist (Geroski

2000). For example, while the role of word-of-mouth is emphasized in diffusion models

(Bass 1969), game-theoretic models emphasize the process of learning-by-doing and

spillovers as fundamental (Jovanovic and Lach 1989). Furthermore, many diffusion

patterns deviate from the typical S-shape. Henderson (1995) records unexpectedly long

lifecycles for lithographical technologies while other technologies, such as

supercomputers and nuclear energy, have saturated at low levels. Also, as Homer

showed, diffusion is often much more complex, with a boom-bust-recovery being

common (Homer 1987, Homer 1983). In line with this, the empirical literature

increasingly identifies cases of diffusion challenges for new technologies across a wide

range of complex environments, such as medical applications (Gelijns et al. 2001),

---

[20] The S-curve literature is guilty of selection bias: successful technologies are the focus of explanation. Yet failures (instant or fizzle) are surely numerous.

renewable energy (Kemp 2001; Garud and Karnoe 2001), or automotive industry (Geels 2005).

The reason for such a high degree of heterogeneity in hypotheses and outcome is due in part to the differences in potential performance and productivity of individual technologies across cases. Further, the literatures emphasize different drivers of diffusion. The marketing literature emphasizes social dynamics and consumer choice, while the literature on industry dynamics emphasizes the technological S-curve. In each system, both are present, but their influence differs across cases. In several cases it is justified to filter out the most dominant mechanisms; however, this is not always true. However, other critical factors can make similar, or stronger, contributions to the dynamics: a technology transition includes network effects, scale economies and other increasing returns to scale, co-evolution with complementary systems, consumer behavior and learning, public rules and regulations, and competing technologies.

 It is such interplay within and with its context that makes a technological trajectory path-dependent. Such path dependency is a particularly important consideration for the evolution of the automotive industry. Figure 1 illustrates the evolution of the installed base of various fuel technologies between 1880 and 2005. ICE vehicles displaced the horse-drawn carriage as the dominant mode of transport through a very rich set of interactions that included the competitive development of various types of platforms (that is, vehicles defined by the technology but also their complimentary and institutional elements) with technological innovations for each that partly spilled over between them,

but also competitive and synergistic interactions with other emerging modes of transportation, such as trolleys and railways. Furthermore, co-evolution of fueling and maintenance infrastructure, roads, and driving habits played a large role in the adoption dynamics (e.g., Geels 2005). In the first decades there was little agreement on what the outcome of the transition would be. For example, around 1900 EVs were very much in competition with steam and internal combustion engines (ICE): they held the world speed record of 61 mph in 1899 (Flink 1988); their performance was superior in many other key attributes (e.g., simplicity, cleanliness, noise); they had strong support from leaders in industry, including Thomas Edison. However, soon after, sales of automobiles powered by ICE surpassed electrics and ICE became the dominant design (see Essay 1 for a more detailed discussion).

With the prospective transition challenges within the automobile industry in mind, we develop a model that captures a broad scope. In the other Essays, the role of feedbacks related to consumer familiarity (Essay 1) and to infrastructure complementarities (Essay 2) are analyzed in depth. This essay focuses on the mechanisms that involve technological innovation, learning, standardization, and spillovers among various technologies. Technology spillovers are a central contributor to advancement of technology throughout industries (Jovanovic and Lach1989). For example, a critical invention for the advancement of ICE vehicles was the electric starter. Its idea, built on the use of a battery and dynamo, was derived from the EV. The experience with the EVs was fundamental to its successful implementation in ICE vehicles the dynamo, wiring,

non-standardized batteries, and starter system all needed to be adjusted properly to each other.

The power of spillover is also illustrated by the emergence of the wind-power industry. In the early 1980s, two drastically different approaches competed with each other. First, a US-based approach was founded on superior and top-down design, based on aerospace fundamentals, and backed by fundamental R&D. In contrast, the Danish wind industry supported development of diverse alternatives, by individual entrepreneurs, and was geared to stimulate spillovers among them. It was the low-investment, large-spillover approach that out-competed the superior designs (e.g., Karnoe 1999).

One key question to understand in relation to such technology competition is, Under what conditions is leapfrogging, rather than gradual change, more likely to lead to success? A related question is whether broad deployment of competing alternatives constrains or enables a transition. Radically different technologies will experience limited exchange of knowledge with incumbents. For example, HFCVs can share part of the gains in body weight with ICE/gasoline vehicles, and vice versa, but their fuel-cell stacks and electric motors will not benefit from the 100 years of experience with ICE. On the other hand, contemporary HEVs can learn from experience with both ICE and HFCVs.

While strategic and policy implications are enormous, the concept of spillovers has been treated explicitly in only a few models (notable exceptions are Klepper 1996, Jovanovic and Macdonald 1994, Cohen and Levinthal 1989). Here I introduce and explore a model

with endogenous innovation, learning and spillover, and resource allocation. This model contrasts with the traditional models regarding three critical aspects. First, this model explicitly captures the notion of variation in the substitutability of knowledge across platforms. Second, advances within an entrant technology can spill over to the market leader. That is, market leading and technology advances are decoupled. Third, the model includes scale effects that are external to the technology and analyzed in interaction with the spillover dynamics.

These differences will permit focus on the specific challenges related to technology transitions. The first two distinctions imply relaxing the implicit assumption of technology convergence to one standard. The third will be shown to have significant implications for the dynamics, even when weak in isolation. Further, we can examine the competitive dynamics between entrants, hybrids, and more radical technologies.

I begin with a short discussion of the literature on technological change patterns. Next I will provide an overview of the model. Thereafter I present the model structure. In the analysis I demonstrate the possibility of superior technologies failing in competition with inferior ones. In addition, while the isolated effects of spillovers and scale effects can be limited, their interaction can dramatically influence the dynamics and reduce the take-off opportunities for more radical technologies. I also point to the path dependency of multiplatform competition. In the final section, I state conclusions and discuss implications for the AFV transitions.

Modeling competitive dynamics between heterogeneous technologies

This section provides an overview of the central factors affecting "technology trajectories" and next describes the model boundary and scope.

The literature on technological change patterns

" In product life cycle (PLC) theories, radically different technologies start with an initial low level of agreement about the key dimensions of merit on the producer side, along with limited attention to the technology from consumers (Abernathy and Utterback 1978). A subsequent rise of entrants with different ideas drives up product innovations. As industry and average firm size grow, and an increase in capital intensity forms barriers to entry, benefits from engaging in process innovations increase, which lowers cost. A shakeout results in a reduction of variety and total product innovation, stabilizing the standard product (Klepper 1996), or, alternatively, a dominant design results in stabilization and shakeout through subsequent process improvement (Abernathy and Utterback 1978). Ultimately, market shares of firms' products stabilize, indicating the final stage of the PLC. Table 1 presents an impressionistic overview of the evolution of the automobile industry, novel in 1890, infant around 1910, and mature by 1960, corresponding with the general PLC observations. The industry is currently experiencing a period of change.

Disruptive innovations are hard to establish in a mature and oligopolistic market. Barriers to change are formed: first, because incumbents can deter entry through preemptive patenting out of fears of cannibalization of existing market share (Gilbert and Newbery

1982, Arrow 1962); and, second, because of the existence of various increasing returns to adoption economies (Arthur 1988). Others describe conditions under which disruption is possible, for example, under sufficient uncertainty of the timing and impact of the innovation (Reinganum 1983).

Addressing the issues of barriers from increasing returns, the literature builds on Dosi (1982), who distinguishes market-performance attributes, organizations' value networks, and technology cost structures. For example, Tushman and Anderson (1986) distinguish capability-enhancing and capability-destroying disruptions: that is, cumulative experience and scale can either help or hinder incumbents producing the old technologies, but not entrants. This asymmetry allows barriers for development of a new technology to be broken down either because incumbents have an incentive to rely on scale economies and experience or because the entrants are not locked-in to the sunk cost and experience of the old technology. Incumbents have inertia because of cost in adjusting their channels (Henderson and Clark 1990) or because of cognitive biases (March 1991; Tripsas and Gavetti 2002). Christensen (1997) notes that disruptive technologies can emerge in a neighboring market and compete on dimensions of merit previously ignored. For the incumbent it is not attractive to invest in a small infant market product, but they can fend off threats by shifting upward in the market. However, as the experience of the entrant grows, its superior performance in the new attributes allows the entrant to outplay the incumbent.

While the unit of analysis of these studies is the firm, when the focus shifts to technology entrant and incumbent, the conclusions are similar. Firm capabilities are built up around particular technologies. Learning and accumulation of experience are central in the study of technological change. Four types of channels are usually distinguished: product innovation through R&D, learning by doing (often equated with process innovation) (Arrow 1962; Zangwill and Kantor 1998), learning by using (Mowery and Rosenberg 1989), and spillovers (e.g., Cohen and Levinthal 1989). Developments in each channel can be tightly interdependent. For example, tasks (processes) depend on design (product). To what extent this is the case depends on technology design characteristics, such as its complexity and modularity (Clark 1985; Sanchez and Mahoney 1996; Baldwin and Clark 2000) and its vertical integration (Henderson and Clark 1990; Christensen and Rosenbloom 1995; Ulrich 1995; Fine 1998).

The window of opportunity for a disruption is discussed by Tushman and Rosenkopf (1992). They expand the "dominant design" model to incorporate the social dynamics by which networks of power rearrange during the ferment period, subsequentially changing the institutional structures and driving the next process towards standardization. Holling (2001) provides a similar ecological view of succession. On the other hand, Basalla (1988) describes a much more evolutionary process of change. Finally, the invention and progress rate depends also on potential rates of discovery (Aghion and Howitt 1992; Aghion et al. 2001), technological characteristics (Iansiti 1995), firm goals and perceptions of the technology potential (Henderson 1995). The relevance of these

different observations depends on industry specific parameters and the stage of the industry.

Technological innovations spill over between technologies. The effect increases with the gap between laggards and leaders (Jovanovic and Macdonald 1994; Aghion et al. 2001), and with the capability to extract knowledge from the outside (Cohen and Levinthal 1989). At the industry level, competence building is a social, distributed process of bricolage (Garud and Karnoe 2003). This view emphasizes the value of technological diversity as was discussed for the emergence of wind energy by (Karnoe 1999; Kemp 2001; Garud and Karnoe 2003). Whether innovations of a potential entrant will generally trigger increased R&D activity and performance increases of incumbents, the so-called sailing-ship effect (Rosenberg 1976), has also been observed in the automobile industry (Snow 2004). It is these combinations of interactions that suggest that hybrid technologies can serve as temporal intermediate bridges between an incumbent and a radical innovation (Utterback 1996).

Other dynamic factors are early uncertainty about the efficacy and safety of new technology, the role of complementary assets, economies of scale, scope, and other market externalities. They drive increasing returns to scale (Arthur 1989) and network externalities (Katz and Shapiro 1985) that play a central role in the emergence of a standard designs (David 1985; Sterman 2000; Klepper 1996).

Model Boundary and scope

The model represents the evolution of an industry's technology over time and is in spirit similar to the product life cycle model of Klepper (1996) that is based on the concepts of industry evolution (Nelson and Winter 1982). The current formulation captures the new and replacement sales of semi-durable goods. The model is discussed in the light of the vehicle market. Klepper focuses on interactions between market structure (patterns of firm entry, exit, and concentration) and innovation, with heterogeneity in capabilities of firms as a main driver of dynamics. In this paper the unit of analysis is the technology rather than the firm. Figure shows the boundary of the model. Layers indicate different platforms. Further, as with other PLC models, this model captures learning-by doing and R&D, and endogenous allocation of resources that are adjusted with the relative productivity of the production inputs. Technological diversity evolve over time and substitutability between variants explicitly and endogenously. However, central to this analysis is the assumption that technology is inherently multidimensional. This means, first, that spillovers can also flow to the market leader, as platforms lead at some aspects of technology, but lag at others. Second, technologies can be non-uniform across platforms. Finally, to explore the dynamics, the model allows examining the interaction with other scale effects, external to the technology.

Figure 3 shows the principal feedbacks that drive technological change. Sufficient attractiveness of a product increases its market share and sales and allows for allocation of resources for R&D that in turn improves the knowledge and technology, and subsequently the product attractiveness. This further increases market share (R1, *learning*

206

by R&D), as well as *learning-by-doing* (R2) through accumulation of production

experience. The first results in product improvement, the second mostly in process

improvement. Improvements occur with *diminishing returns* (B1). On the other hand,

resources can be allocated to absorb *knowledge spillovers* (B3) from other platforms.

Resources are allocated to those activities with the highest perceived *productivity* (R3).

While not shown explicitly in this high-level overview product and process improvement

is separately represented in the model. Also not shown, but included, are several

increasing returns to scale. Without a priori assumptions that impose conversion of

technologies to one standard, we can explore here under what conditions these dynamics

benefit or harm different technologies.

The model

For platform economies I use a simple model of cost, volume and profits. Aggregate

profits earned by producers of platform type *j*, $j = \{1,...,n\}$, depend on the net profits π_j^n

minus capital cost, C_j^k, and investments in R&D, C_j^{RD}:

$$\pi_j = \pi_j^n - C_j^k - C_j^{RD} \tag{1}$$

The price equals unit cost plus markup $p_j = (1 + m_j)c_j$. Then, net profits equal the

markup multiplied by unit cost c_j and total sales s_j,

$$\pi_j^n = (p_j - c_j)s_j = m_j c_j s_j \tag{2}$$

A key structure in the model is how experience and revenues feedback to improve knowledge, technology and then consumer choice and sales. Figure 4 shows the modeled chain of operations that connects the producers' resource allocation decisions to the consumers' purchase decisions, through knowledge accumulation and technological improvement. The chain is comprised of three main segments: consumer choice, effective technology and knowledge accumulation, and resource allocation. The consumer's choice of platform j, $j \in \{1,..., N\}$, depends on the utility u_j that consumers derive from platform j, and is determined using a multinomial logit function. Utility is derived from two attributes a_l, $l \in \{\text{performance,price}\}$ that are a function of the state of the effective technology associated respectively with cost and technology performance. There are two types of activity, $w \in \{\text{product,process}\}$, that each determine the state of technology. To simplify the analysis, I assume that the state of technology associated with product improvement yields performance improvements and those with process improvements yield solely cost improvements. The technology frontier moves with an increase in the effective knowledge, with diminishing returns. Effective knowledge aggregates knowledge from all sources i that contribute to the state of the technology and that are associated with activity w, this is done through a constant elasticity of substitution (CES) function. Knowledge of platform j accumulates, through internal learning-by-doing and product improvement $(i = j)$, or through spillovers $(i \neq j)$. The third section comprises resource allocation decisions made to maximize marginal returns.

This structure rests upon several significant simplifications. While the key arguments of this paper do not rest on the current level of detail, a more detailed transition exploration

208

of the transition dynamics would benefit from relaxing some assumptions. Four are especially important to highlight at this point. First, I collapsed several consumer choice attributes into two that map on to cost and performance. However, consumers base their choice on a series of attributes (price, operating cost, convenience, reliability, driving range, power, etc…). Capturing these details can be important, for example because complementarities from fueling infrastructure affect attractiveness at this level, but can differ by platform. Second, I map cost and performance one on one onto respectively process and product innovations. In reality both process and product innovations contribute to both performance and cost. Third, vehicles comprise different modules (powertrain, body, brake-system, electrics). It is at this level that spillovers and improvements occur, and the degree of this depends very much on the specific module. Thus, an analysis of transition dynamics for specific AFVs should rest on a structure at the module level. Fourth, product and process improvements are tightly coupled due to the design/task interdependencies of complex products. For example, the unit production cost of technologies may increase temporarily after a product innovation cycle. This is because product innovations partly render previous process improvements obsolete. Appendix 3a describes the generalization of the model that includes these more general formulations. This expanded model allows testing of the extent to which the key dynamics hold when the boundary is expanded. It also allows for the exploration of dynamics within a larger set of environments.

I proceed here with an exposition of the core model. In the next section I provide the functional relationships for the central parts of the model: technology, and knowledge

accumulation. Thereafter I discuss the resource allocation process. I end the exposition

with notes on consumer choice and accounting that includes the elasticity of substitution

between the various sources of knowledge, effective technology, and the input factors.

Cost have a fixed component c^f and a variable component that decreases with the

advance of relative process technology θ_{j2} (index $w=2$, $\theta_{jw} \equiv T_{jw}/T_w^0$). The variable costs

are equal to c^v when relative technology is equal to the reference technology T_2^0 :

$$c_j = c^f + c^v / \theta_{j2} \tag{3}$$

Technology, T_{jw}, adjusts to its indicated level T_{jw}^* with adjustment time τ^t, while

technology exhibits diminishing returns in accumulation of effective knowledge K_{jw}^e.

$$T_{jw}^* = T_{jw}^0 \left(K_{jw}^e / K_w^0 \right)^{\eta_w^k} \tag{4}$$

where T_{jw}^0 represents the quality of a platform, or its technology potential. The state of

technology adjusts to T_{jw}^0 when internal knowledge equals the mature knowledge K_w^0. η_w^k

is the diminishing returns parameter, $0 \le \eta_w^k \le 1$.

Much of the knowledge that is accumulated within one platform can spill over to others.

One firm and platform may lead on certain aspects of technology and lag on others,

simultaneously being both the source and beneficiary of spillovers. To allow for varying

substitution possibilities, the knowledge base for each platform is a constant elasticity of

substitution (CES) function of the platform's own knowledge K_{jjw}, and the knowledge,

spilled over from other platforms, K_{ijw}, depending on the spillover effectiveness κ_{ijw}:[21]

$$K_{jw}^e = \left[\kappa_{jjw} \left(K_{jjw} / K_w^0 \right)^{-\rho_{jw}^k} + \sum_{i \neq j} \kappa_{ijw} \left(K_{ijw} / K_w^0 \right)^{-\rho_{jw}^k} \right]^{-1/\rho_{jw}^k} \tag{5}$$

I separate the contribution from internal knowledge to emphasize the different process

(see below). The spillover effectiveness is not identical across technologies. For instance,

the fraction of the knowledge of a HEV powertrain that is relevant to ICE vehicles differs

from the fraction relevant from a biodiesel powertrain. Parameters will depend on

differences in the technologies. For example, ICE experience is relevant to biodiesel

vehicles, but less relevant to General Motors' HyWire HFCV, which radically alters most

design elements. We specify this spillover potential between two technologies, with

respect to activity w as κ_{ijw}, $0 \leq \kappa_{ijw} \leq 1$ and, by definition, for internal knowledge there is

full spillover (carry over) potential, $\kappa_{jjw} = 1$.

Further, $\rho_{jw}^k = \left(1 - \varsigma_{jw}^k\right) / \varsigma_{jw}^k$ is defined as the substitution parameter, with its transformed

value ς_{jw}^k being a measure of the elasticity of substitution between the various knowledge

---

[21] This expression is a natural generalization of McFadden's (1963) multiple input CES function. This
significantly increases the production possibilities. For instance the elasticity of substitution does not have
to be identical for all inputs (see also Solow 1967). See the analysis for an explanation of how this function
behaves naturally with accumulation of knowledge.

sources for platform j.[22] For such technologies $1 < \varsigma_{jw}^{k} < \infty$. Further, we see that one way

for the effective knowledge to be equal to the normal knowledge is when internal

knowledge equals the mature knowledge K_{w}^{0} in absence of any spillover knowledge.

Accumulation of knowledge

Knowledge accumulates through four distinct processes: product improvement through

R&D, process improvement through learning-by-doing, and spillovers of both product

and process knowledge. Knowledge production occurs through directed search (trials)

(Simon 1969) and following standard search models, actors take random draws from a

large pool of potential ideas (Levinthal and March 1981). Product improvement trials can

be undertaken with increased R&D. Process improvements accumulate through learning-

by-doing, increasing with production rates and investment (Arrow 1962; Zangwill and

Kantor 1998). Knowledge production grows with diminishing returns in the number of

resources, reflecting the several organizational and time constraints in doing more trials.

More formally, product innovation and process improvement knowledge accumulate at a

rate Γ_{w} when resources are equal a normal value R_{0}. The accumulation rate increases

with allocation of resources, an endogenous productivity effect ε_{jw}^{i}, and relative resource

allocation:

---

[22] In a two platform context, ς_{jw}^{k} would measure exactly the elasticity of substitution between spillover

knowledge and internal knowledge. In a multiple platform situation the definition of elasticity of

substitution is not well defined.

$$\frac{dK_{jjw}}{dt} = \varepsilon_{jw}^i \left(R_{jw} / R_0 \right)^{\eta_w^i} \Gamma_w \tag{6}$$

Benefits to resource allocation exhibit diminishing returns: $0 \leq \eta_w^i \leq 1$.

For product improvement the productivity effect is constant, $\varepsilon_{j1}^i = 1$. Process

improvement is subject to learning-by-doing effects and the effectiveness is a concave

function of the relative resources per volume produced: [23]

$$\varepsilon_{j2}^i = \left(s_j / s_0 \right)^{\eta^s} \tag{7}$$

with $0 \leq \eta_j^s \leq 1$. The unit of analysis is the platform. Capturing learning-by-doing at this

level is justified for that knowledge that can flow easily between firms with similar

technologies are fast relative to the industry evolution time scale). However this is

certainly not true for all knowledge. As the typical number of firms that are active in an

industry can change significantly over time, this also means the learning-by-doing

effectiveness can do so. This is discussed in Appendix 2d.

Process knowledge and the knowledge embedded in the product can spill over to other

technologies. Imitation, reverse engineering, hiring from competitors and other processes

---

[23] We can arrive at the combined effect of equations (6) and (7) following a different train of thought: process knowledge grows linear with sales, holding resources per unit produced equal to its reference value, while reference resources per unit produced increase with sales (as it is harder to capture all the benefits); and finally, the productivity of resources per unit produced has diminishing returns Thus:

$dK_{jj2}/dt = s_j \left[\left(s_j / R_j \right) / \left(s_0 / R^s \right) \right]^{-\eta_2^i} \Gamma_1; R^s = \left(s_j / s_0 \right)^{\eta^s} R_0$, with constraints:

$-1 \leq -\eta_2^r - \eta^s \leq 0$ guaranteeing diminishing returns in sales following this expression, and

$0 \leq \eta_2^r, \eta^s \leq 1$, because of the interpretation in the main text.

that enhance spillovers take time and resources. Further, spillovers close the gap between the *perceived* knowledge of platform i as perceived by platform j, $K^{\sim}_{ijw}$, and the knowledge that has already spilled over K_{ijw}. Further, spillover increases with resource allocation, and fractional growth rate g^o_w:

$$\frac{dK_{ijw}}{dt} = g^o_w \left(K^{\sim}_{ijw} - K_{ijw} \right) \left(R_{jw}/R_0 \right)^{\eta^o_w} \tag{8}$$

Note that the model exhibits diminishing returns in the accumulation of technology, in relation to effective knowledge, but that there are constant returns to the accumulation of knowledge itself. In real life, the exact locus of diminishing returns is not always easy to measure. For instance whether aggregate diminishing returns are the result of constraints at knowledge collection, effectiveness of knowledge, or transforming knowledge into technology is not easily to observe. Moreover, all will be true in reality, in the long run. In appendix 3b I show that we can be indifferent to where we impose diminishing returns, as they are mathematically interchangeable. Therefore I collapse all sources of diminishing returns into one parameter. I further discuss how the current formulation relates to standard learning curves.

Supply decisions

Here I describe how the resource allocation process is captured. Upfront investment in R&D can increase total profits in the long run, either by improving performance or by lowering costs (and subsequently price). Both have a positive effect on attractiveness and sales. Actual resource allocation decisions then depend on expected demand elasticity

under the existing market structure, and effectiveness in improving platform

performance, as compared to reducing its cost.

Decision makers within organizations are bounded rational (Cyert and March 1963;

Forrester 1975; Morecroft 1985). They learn about relevant knowledge and productivity

over time and resources are allocated based on the relative perceived marginal returns

(Nelson and Winter 1982). Further, decisions are made locally. Managers push projects

by pushing those allocations that are perceived most beneficial, modules that are

outsourced are optimized at the module level. This concept is used here for the resource

allocation decision. While the key findings of this paper do not rest on the concept of

local decision making, it is robust as compared to globally optimal decision making, but

also mathematically convenient, for the same reason that actual decision making is local.

Resource allocation decisions include: i) allocation of a share of total revenues going to

R&D, σ_j^r ; ii) the share of total R&D resources of platform j that the chief engineers

dedicates to process or product improvement, $\sigma_{jw}^r, \sum_w \sigma_{jw}^r = 1$; iii) the share of total R&D

resources of platform j activity w, that managers dedicate to internal knowledge

accumulation, σ_{jjw}^r, as opposed to spillovers $\sigma_{\sim jjw}^r = 1 - \sigma_{jjmw}^r$; and finally, iv) the share of

total R&D spillover resources of platform j, activity w , that engineers dedicate to

extracting knowledge from platform $i \neq j$, $\sigma_{ijw}^r, \sum_{i \neq j} \sigma_{ijw}^r = 1$.

We will discuss one resource allocation decision here, others follow the identical

structure. Resources that are dedicated by platform j to spillovers, $R_{\sim jjw}$, need to be

distributed to capture spillovers from the various platforms. The distribution results in

resources $R_{ijw} = \sigma^r_{ijw} R_{\sim jjw}$, going to platform i, with σ^r_{ijw} being the share of the total budget

going to i. The share adjusts over resource adjustment time τ^r to the desired share for

platform i, σ^{r*}_{ijw}, which equals desired resources R^*_{ijw} divided by the resources others

bargain for:

$$\sigma^{r*}_{ijw} = R^*_{ijw} \Big/ \sum_{i' \notin j} R^*_{i'jw} \tag{9}$$

Desired resources for platform i increase with expected return on effort ς^{r*}_{ijw} relative to the

reference returns ς^k in knowledge generation.

$$R^*_{ijw} = f\left(\varsigma^{r*}_{ijw} \big/ \varsigma^k\right) R^r_{ijw}; f' \geq 0; f \geq 0; f(1) = 1 \tag{10}$$

Returns are measured in relation to the relevant lowest level performance indicator that is

perceived to be fully influenced by the decision, capturing the essence of local decision

making. The planning horizon over which the expected performance is estimated is τ^p.

In the case of resources for spillovers across platforms, the reference indicator is total

spillover knowledge, $K_{\sim jw}$, with $K_{\sim jw} \equiv \left[\sum_{i \neq j} \kappa_{ijw} \left(K_{ijw} \big/ K^0_w \right)^{-\rho^k_{jw}} \right]^{-1/\rho^k_{jw}}$, which follows

from equation (4).

In Appendix 3c show that when the expected returns on effort, ς^{r*}_{ijw}, equal marginal

returns on effort, the resource allocation is locally optimal. Here I assume, optimistically,

216

that decision makers understand the structure that drives marginal returns on effort and that they can learn this, with perception delays, under the local conditions of holding current resources and all outside conditions constant (see Appendix 2b for a detailed motivation and example).

A final set of decisions involve entry and exit. Entry decisions are conditional on realization of discovery of a particular technology. Entry depends further on expected return on investment (ROI), which follows similar heuristics as outlined here for the resource allocation process. Expected ROI depends on the spillover effectiveness with incumbents, on the current state of the industry, the initial experience that platforms are endowed with, the initial state of their technology, and on the size and duration of seed funding. Platforms exit when profits fall below a reference value. This will be discussed more in the analysis

Platform sales

The total number of vehicles for each platform $j = \{1,...,n\}$, V_j, accumulates new vehicle sales, s_j, less discards, d_j:[24]

$$\frac{dV_j}{dt} = s_j - d_j \qquad (11)$$

---

[24] I ignore the age-dependent character of discards in this discussion (see for this Appendix 2a in Essay1).

Total potential sales going to platform j equal considered sales from non drivers adopting at rate s^n and all discards from all platforms, multiplied by the share going to platform j, σ_j:

$$s_j = \sigma_j \left(s^n + \sum_i d_i \right) \tag{12}$$

The replacement decision involves a choice of whether to adopt or not, and conditional upon adoption, platform selection. This is captured through a nested logit-model (Ben-Akiva 1973). Further, Essay 1 discusses the social factors influencing utility such as familiarity and experience from driving, as well as perceptions of attributes' state as input. Appendix 2e provides the detailed nested-logit formulation, and how familiarity and perceived utility are integrated in the nested-logit formulation. In the model exposition here we proceed with an extreme case of the nested form: the normal multinomial form in which all alternatives are compared at par:

$$\sigma_j = u_j \Big/ \left(u^o + \sum_{j'} u_{j'} \right) \tag{13}$$

For non-drivers, the total purchase rate, in the absence of capacity constraints, equals:

$$s^n = N \big/ \tau^a \tag{14}$$

Where, $N = H - V; V = \sum V_j$ are the non-drivers, with H being the total number of households, while τ^a is the average time between two adoption considerations.[25]

---

[25] The proper interpretation of a "share" that is allocated based on relative utility is thus defined as individuals' allocation between two alternatives at a decision point, rather than a fixed fraction of the population adopting or not. The steady state total adoption fraction depends thus on the consideration time. For instance, if $u_j = u^* \forall j$ and non-drivers, the total adoption fraction equals $\tau^a \big/ \left(\tau^a + \tau^d \right)$, and is therefore not necessarily 50%.

The perceived utility of a platform captures the aggregate of experience across various dimensions of merit. Ignoring variation in perceptions for drivers of different platforms, we can write $u_{ij} = u_i \ \forall i$. Further, with utility being equal to the reference value u^* all attributes equal their reference value, we have:

$$u_j = u^* \exp\left[\sum_l \beta_l \left(a_{jl}/a^* - 1\right)\right] \qquad (15)$$

where β_l is the sensitivity of utility to a change in the attribute $l \in \{1, 2\}$. The first attribute captures the performance, and thus state of the production technology, $a_{j1} = \theta_{j1}$. The second attribute captures price $a_{j2} = p_j$, where price is an indirect function of the state of the process technology, θ_{j2}, discussed above.

This concludes the fundamental structure of the model, relevant and sufficient for explaining the key insights of this essay. The model has been subjected to its robustness by testing the role of other factors. None of them have critical impact on key insights of this paper, however, those that I include in Appendix 4 do allow studying a richer variety of contexts and also serve for detailed testing of the conditions under which the key insights hold. Besides the expanded structure regarding technology accumulation, discussed above, additional boundary conditioning structures that I subjected the model to are: i) endogenous elasticity of substitution, which allows capturing consistently spillover dynamics of multiple endogenously platforms over long time horizons; ii) interaction effects between different activities, which traces the effective technology more closely; iii) spillover potential ; iv) endogenous capacity adjustment, constraining

the sales growth rate after sudden technology shocks. So far we ignored that demand and actual sales can become decoupled through capacity constraints, accumulating backlogs. v) backlogs and churn, which properly deals with demand responses to supply shortages; vi) adjustment of markups, which allows one to proxy different market structure and competitive effects; vii) scale economies within a platform, which allows to distinguish these effects, that are not prone to spillovers, from learning by doing.

Analysis

We will first explore the basic behavior by testing basic PLC dynamics for two extreme cases: i) a single platform, without spillovers; ii) multiple platforms that enter endogenously, and are subject to complex spillover interactions. Next we analyze the spillover mechanisms in detail by examining the isolated case of two competitive dynamics between two platforms. To understand how these mechanisms play out in a richer context, we also explore the role of scale effects. With the insights from these analyses, we will study implications for AFV transitions and focus in particular on competitive interactions between three heterogeneous platform.

Testing basic model behavior

I first test whether the model is able to generate the stylized patterns of behavior we should expect from a PLC model. Figure 5 shows the product lifecycle dynamics generated by the model, representing the introduction of a new technology in isolation, such as the basic technology related dynamics of the emergence of the automobile

industry. Parameter settings for this and other simulations are provided in **Table** .

Discovery probabilities for all but one technology are set to zero, while this technology is

introduced at t=0. The installed base reaches 90% of the potential market over time

(utility of not adopting, u^o equals 0.1). The improvement rate of product technology

precedes that of process technology. Vehicle performance improves initially very steeply,

while costs rise after initialization, because of the inexperience with the new products.

After year 5 costs start to decline rapidly as well due to the rapid increase in scale,

spurring learning-by-doing effects. After year 13 the improvement rate of process

technology dominates benefits from the increased scale. From then on, costs fall over 50

percent, while vehicle performance improves marginally. Investment in R&D increases

rapidly, due to considerable returns on investment and larger scale, but decays gradually

subsequently, as ROI evaporates when a reasonable large market share is reached.[26]

However, ultimately, rapid experience overcomes this. At the same time, the scale is

large enough that net cost reductions remain positive. Clearly, other modes of behavior

can be generated depending on these assumptions and on the initial experience of product

and process innovation. However, by using typical parameter settings, the fundamental

PLC patterns are well represented by the model.

---

[26] In these simulations we have ignored the number of firms within a platforms and their effect of the

market concentration on scale economies (see Klepper 1996). This will be treated in later versions. See

also Appendix 2c.

The PLC scenario in **Figure** 5 represents the aggregate behavior of a market that in reality is comprised of multiple technologies that compete, enter, and exit with various degrees of spillover among them. As the goal is to understand inter platform competition, it is imperative that this model can also reproduce such dynamics deriving from a lower level of disaggregating, in which entrance is endogenous. I analyze here if and how competitive and multiple platform dynamics lead to stabilized market concentration and performance. To do this I explore simulations in which platform entrance is a stochastic process. I first discuss the setup for these simulations and then discuss typical results. The results comply with robustness requirements of the model. In the subsequent section I explore the underlying drivers for spillovers dynamics in depth.

The expected entrance rate for a platform depends on the expected returns and on the normal entrance rate, which can be seen to represent the aggregate barriers to entry due to various factors such as technological complexity, economic barriers, rules and regulations. Expected returns $\varsigma_i^{\pi *}$ are compared to the required returns ς_{ref}^{π} :

$$\langle e_i \rangle = f\left(\varsigma_i^{\pi} / \varsigma_{ref}^{\pi}\right) / \tau^e \; ; f\left(1\right) = 1; f' \geq 0; f \geq 0; f\left(\infty\right) = f^{\max}; \tag{16}$$

Expected returns depend on the type of current platforms in the market, their market shares and the distribution of knowledge across the various platforms. Expected returns will vary by the technology potential as perceived by those who consider to enter, In this simulation I assume that potential entrants have the same information about the market as actual entrants. Potential entrants are endowed with, and take into consideration, additional seed funding of 5 years of 1.5 Billion $ (equal to 1% of normal industry revenues). Expected entrance increases with expected profits, but saturates for large

profits. I use the logistic curve for this, with sensitivity parameter $\beta^e = 1$. To represent the distribution of technologies available for the market, I vary the distribution of spillover effectiveness between technologies, κ_{ij1} and the technology potential, T_{j1}^0. For spillover potential I define $\kappa_{ij1} = \alpha |i - j|^{(\upsilon-1)/\upsilon}$, for $i \neq j$ $(\kappa_{ii1} = 1 \forall i)$, where α is a scaling parameter for spillover strength, and υ is the uniformity index for the available technologies in the market, with $0 \leq \upsilon \leq 1$. When υ is close to zero, the spillover potential between technologies approaches zero very fast over different platforms, representing a more heterogeneous market. While υ equal to 1 implies that spillover across platforms is equal to the maximum α for all platforms. The technology potential is varied randomly across platforms, with an average of 1 and standard deviation of 0.5.

I am interested in the competitive dynamics between the various platforms over time, and the market behavior with respect to knowledge accumulation and performance. To analyze the competition over time, I use the Herfindahl index, which measures the market concentration and is defined as:

$$H = \sum_{i=1}^{N} \sigma_i^2$$

The Herfindahl Index (H) has a value that is always smaller than one. A small index indicates a competitive industry with no dominant platforms. If all platforms have an equal share the reciprocal of the index shows the number of platforms in the industry. When platforms have unequal shares, the reciprocal of the index indicates the "equivalent" number of platforms in the industry. Generally an H index below 0.1 indicates an unconcentrated market (market shares are distributed equally across technologies). An H index between 0.1 to 0.18 indicates moderate concentration, An H

index above 0.18 indicates high concentration (most of the market share is held by one or two platforms).

Figure 6 shows representative results. Different simulations each start with 16 potential entrants. Across simulations I vary the technology heterogeneity, with $\alpha = 0.75$ for each, and for simulation $s \in \{1,...,7\}$, $\upsilon \in \{0.91, 0.83, 0.67, 0.5, 0.33.25, 0.1\}$. Figure 6a shows the average spillover potential κ across platforms, weighted by market share in equilibrium ($t=100$). Technology heterogeneity in equilibrium corresponds with the distribution of technologies available in the market. Further, an increase in the spillover potential also results in increased resources being allocated to spillovers. Aggregate behavior of all simulations is consistent (Figure 6b). Figure 6c shows the Herfindahl index over time. First, we see, that for these simulations the market can only support a limited amount of platforms (in equilibrium, $H_{\min} \sim 0.15$, or 5-6 firms). This is in absence of any scale effects that are not related to R&D and learning. We also see that concentration increases with the uniformity of the technologies. Absent any potential for spillovers, entrants can partly catch up, despite initial experience deficit. This holds especially true for those platforms that have superior technology potential. The spillover dynamics work in favor of more superior technologies that have for example, more resources available, providing scale economies associated with learning by doing. [27] Note that these dynamics do not reflect the concept of niche formation, as performance is a scalar. Including additional increasing returns to scale will reinforce this significantly.

---

[27] An additional analysis to separate micro effects from macro effects would be to look at the seniority of those who have an advantage.

Figure 6d shows that the increased spillovers also lead to a greater attractiveness of the average platform in the market (weighted by market share), in the capacitated market. Attractiveness behaves properly, with diminishing returns. The aggregate market dynamics are robust and intuitive.

Reducing the barriers to entry for new platforms, which can be emulated by lowering τ^e, results in an increase in the number of entrant attempts throughout. The result is that increased spillovers compete with a more intense competition, but before hand it is not clear which effects are stronger. Doubling the normal entrance rate for these simulations has no significant effects on the Herfindahl, and on average a 5-10% increase in the market attractiveness. Increasing the barrier to entry leads to a 5-20% increase in the Herfindahl and a 10-25% decrease in market attractiveness, in all cases with diminishing returns. The results of endogenous entry dynamics illustrate the consistency and robustness of the model behavior over a wide range of contexts. However, a deeper understanding of the dynamics should come from various levels of analysis. I now concentrate on a deeper understanding of the spillover dynamics.

Analysis of spillover dynamics

To understand the basic spillover dynamics, I analyze the competition between the incumbent I_1 and one alternative entrant platform E_2. Figure 7 shows simulated adoption over time for cases with varying, but symmetric, spillover effectiveness across platforms, $\kappa_\Delta \equiv \kappa_{i,i+1} \in [0, 0.1, ..., 1]$. Technology potential is identical. The adoption rate for the entrant and its equilibrium adoption fraction increase with spillover effectiveness: when

225

all platform technology of each platform is fully appropriable, $\kappa_\Delta = 0$, the entrant reaches

about 10% of the installed base. However, the entrant can catch up fully, reaching 50%

of the market, when spillover effectiveness equals 1. However, note that it takes 40 years

to reach the equilibrium, even under maximum spillover effectiveness, while the

technology replacement time is 10 years. Figure 7b) and 7c) show the allocation of

resources to R&D for two cases of very low, and very high spillover effectiveness,

$\kappa_\Delta = \{0.1, 0.9\}$. Figure 7b) shows total resources that are allocated to R&D. The entrant

technology, being less mature and having a lower market share, invests heavily as it can

capture significant returns on its R&D, especially in the high spillover case. Note that

returns and thus investment in R&D would be considerably suppressed in the presence of

scale effects. The incumbent experiences several effects. A first order effect is that

reduced revenues also lower R&D spending. However, other effects lead to an increase in

spending: irrespective of any spillover, demand elasticity to innovation increases when

market share is reduced. This effect is however stronger for the high spillover cases, as

these are the scenarios under which the entrant captures a larger market share. This effect

is combined with an effect that is directly a function of spillover strength: as the entrant

develops its technology, so does the spillover potential for the incumbent. These two

second order effects lead to an increase in R&D investment by the incumbent and are

different manifestations of the sailing-ship effect (Rosenberg 1976; Snow 2004). Further

(Figure 7c), after entrance, both parties dedicate indeed the largest portion of their

resources to spillovers. Once the core technology has been established it becomes much

more beneficial for the entrant to improve technology through internal R&D.

In summary, dynamics between various technologies in a market unfold with three competing effects at work: first, there are competition effects that distribute the market shares; second, there are the learning and R&D feedbacks at work (as well as external increasing returns); finally, there are spillover effects between the technologies. Competition effects pressure established technologies' installed base through the balancing feedback of reallocation of vehicle discards according to platforms' relative attractiveness. Those that receive a larger market share than their installed base share, will grow until they match. Attractiveness depends on each platform's technology potential, the current relative state of their technology. Learning-by-doing and -R&D, allow improving the technology performance through internal processes that further build attractiveness which can drive up sales, feeding back to investment in those processes. Generally these are subject to diminishing returns, and therefore, when presented in isolation, they will allow laggards to catch up (see Essay 1). Finally, the spillover effects derive from interaction between competitors' relative performance that borrow ideas from each other. The net spillover effect involves a flow towards the entrant, and the magnitude depends on their amount of internally produced knowledge.

Equilibrium is established when the forces from these three interactions offset each other, balancing market share, relative resources, relative flows of internal and spillover flows. We saw that with for two platforms an increase in spillovers benefits the entrant. However, for differences in technology potential, for multiple technologies, or when other scale effects are included, one can see the existence of different conditions for

equilibrium, or multiple equilibria and strong path-dependency, based on the specific interdependencies between platforms. This is will be analyzed next.

Analysis of AFV competition: spillovers, scale effects and multiple entrants

Having an increased understanding of the general dynamics generated by the model, I now analyze how different technologies fare in a multi-platform race, focusing on the role of spillovers and learning, on their interaction with scale effects and with the effect of differences in the technology potential of the platforms. I specify an incumbent, I_1, with a large and saturated installed base, analogous to ICE in 2000. I first analyze the dynamics when entrance is limited to one platform only.

The model captures internal economies of scale that represent, for instance, reduced production cost when production plants are scaled up or economies of scope. However, platforms are also subject to increasing returns to adoption related to external factors, such as complementarities or other (network) externalities that affect the perceived consumer utility in one way or the other. In particular, the co-evolution of demand for alternative fuel vehicles and infrastructure is an important feedback for many technologies, especially hydrogen, but also to some extend CNG, flex-fuels, EVs, and plug-in hybrids. Further, as is discussed in Essay 1, the requirement of building up familiarity greatly. Other increasing returns result from economies of scope such as increased sales and experience, the number of models offered (which will greatly enhance demand, as vehicles have limited substitutability). Expanding the product

portfolio also results in a wider experience, both in using (users will drive the vehicles in different climates, or environments), and in production (the variety of trials available for innovation is wider). These increasing returns to adoption can be a function of cumulative adoption, the current installed base, or the current sales rate. To test how the learning and spillover effects I have analyzed so far interact with such external scale effects. I introduce as the third attribute, one aggregate scale effect as a function of installed base share $\sigma_j^v = V_j / V^T$:

$$a_{j3} \equiv \varepsilon_j^s = f\left(\sigma_j^v\right); f' \geq 0; f\left(\infty\right) = 1; f\left(\sigma_{ref}^v\right) = \varepsilon_{ref}^s \qquad (17)$$

Appendix 2f discusses the functional form used, but Figure 8 shows the shape of the function and the parameters. At the reference installed base share σ_{ref}^v , the scale effect on attractiveness relative to the case of full penetration equals ε_{ref}^s . The scale factor, defined as the inverse of the relative scale effect, $f_j^s \equiv 1/\varepsilon_j^s$, serves as a measure of the strength of the scale. The scale factor gives the relative attractiveness of an entrant when its installed base share equals the reference installed base share, compared to when it is fully penetrated. At full penetration all scale effects work maximally to its advantage. For the reference I use an installed base 5% of the fleet and sensitivity parameter $\beta^s = 1$, which measures the slope at the reference installed base share.

Figure 9 a) shows the sensitivity of the entrant's equilibrium installed base share to scale effects (technology potential is equal that of the incumbent, $T_\Delta^0 = 1$; the same holds true for other parameters). **Table** 3 lists parameter manipulations for all the following analyses. The equilibrium installed base is very sensitive to scale effects. For any scale

factor f^s larger than 4, equilibrium penetration remains below 0.1 (that is, all results fall below the iso-installed base line of 0.1). Increasing the spillover effectiveness improves the range of scale factors that result in take off. However, the entrant, otherwise equivalent to the incumbent, approaches 50% of the market only in absence of scale factors. Thus, while the scale effects have no effects when learning is ignore, and limited effects when spillover potential is large, the interaction of this feedback with those from learning lead to strong strong barriers to entry, when spillover effects become smaller.

The installed base for different values of technology potential ($T_\Delta^0 \equiv T_{21}^0 / T_{11}^0$, see equation(13)), and spillover potential κ_Δ is illustrated in Figure 9b. Absent any spillover and learning, the predicted share of the entrant is equal to:

$$\sigma_2^s = \frac{\varepsilon^s \left(\sigma_2^v \right) u_{21}^\Delta}{\left(\varepsilon^s \left(\sigma_2^v \right) u_{21}^\Delta + \varepsilon^s \left(1 - \sigma_2^v \right) \right)}$$

with $u_{21}^\Delta = \exp\left[-\beta^\theta \theta^I T_\Delta^0 \right]$, where β^T is the aggregate sensitivity of adoption to technological advance and θ^I is the status of the incumbent technology, relative to the reference.[28] In our case, $\beta^T \approx 0.9$ and $\theta^I = 1$. Equilibrium requires that the sales share equals the installed base share, $\sigma_2^s = \sigma_2^v$. This equilibrium is indicated Figure 9b. We see here that, when learning dynamics are included, a superior entrant technology reaches a larger share in equilibrium, provided presence of limited spillovers (above the dotted

---

[28] The technology state parameter of the incumbent makes explicit that MNL models predict that, holding the relative difference between two technologies constant, the gap between the relative shares that technologies receive increases with the advancement of the technology (as the effect of the unobserved characteristics remains constant).

line). An entrant with technology potential that is equivalent to the incumbent (T_Δ^0 close

to 1) achieves equal shares only when spillover effects are very strong. A weak

technology never approximates its potential.

Under what conditions can superior or equivalent entrant technologies catch up with

incumbents? The process of learning and spillover determine the technology trajectory.

This, however, is very much a function of the mix, diversity, and quantity of alternative

technologies available in the market. The analysis illustrates that scale effects create a

barrier to entry, as can be seen in the low spillover case. Beyond that, they allow for

spillovers to flow to the incumbent, before the entrant catches up. This was the situation

for example in the case for EVs in the early 20th century. They diffused slowly with

limited progress in critical aspects such as battery life, recharging speed, and availability

of recharging points, due to limited penetration and limited standardization of electricity

systems at that time. Gradually, the batteries and dynamo system improved and around

1910 they experienced a second wind. However, this also provided spillovers to the more

established ICE platform, and led in particular to the commercialization of the electric

self-starter by Kettering, a critical device that was implemented in ICE vehicles as of

1911 (Schiffer et al. 1994). Ultimately, more and more ICE vehicles were able to gain

market share in areas that were previously considered EV niches. This supports the

notion that neither learning and spillover dynamics, nor scale effects must be explored in

isolation. They interact tightly with each other and also with others such as vehicle

placement and consumer choice dynamics. Together they determine the transition

trajectories and potential for different technologies. It is for this reason that we need to explore dynamics of multiple platforms in depth.

In reality competition plays out not between one incumbent and one entrant, but between a mix of platforms, as was illustrated by Figure 1. Further, such platforms are different from each other across different attributes. For instance, where ICE and HEVs share an engine, HEVs and HFCFs share an electric motor system. Advances in ICE experience, with respect to the engine, are thus relevant to great extent to HEVs, but not so to General Motors' HyWire HFCV, which radically alters most design elements (Burns et al. 2002). On the other hand advances in some elements, such as body weight, are relevant to great extent across all platforms. Many more of such cases can be found considering the enormous set of combinations of mono-, bi-, flex-fuel vehicles, or the consideration of gaseous versus liquid fuels. This context of multiple, heterogeneous platforms greatly limit our ability to intuitively grasp the dynamic implications of the basic interactions discussed above.

 I study the fundamental dynamics of such a situation, by analyzing the case in which one hybrid platform (E_2) that has reasonably large overlap with the incumbent (I_1), and a radically different platform (E_3), with technology that has little in common with the incumbent, but significant overlap with the hybrid. To do so I define the spillover effectiveness between the ith and the $i+1^{th}$ as $\kappa_{i,i\pm1} \equiv \kappa_\Delta$, representing the spillover effectiveness between the incumbent and the hybrid, but also between the hybrid and the radical. In addition I also define the spillover effectiveness between the ith and $i+2^{nd}$

232

platform as $\kappa_{i\pm2} \equiv \kappa_{\Delta2} \le \kappa_\Delta$, setting spillover effectiveness between two platform pairs equal. Thus, $\kappa_{\Delta2}$ represents the spillover effectiveness between the incumbent and the radical. Figure 10a) shows the simulated trajectories of the installed base shares of both entrant technologies, for four different spillover configurations: symmetric and asymmetric situations between {S,A}, for which respectively $\kappa_{2\Delta} = \{\kappa_\Delta, 0.4\kappa_\Delta\}$ and high and low spillover effectiveness {H,L}, for which respectively $\kappa_\Delta = \{0.75, 0.25\}$. See also **Table** 3.

Technology potential and scale factors are equal to one. The dotted line along the 45-degree line show the trajectory for the symmetric, high spillover effectiveness scenario {S,H}. Dots represent samples with a 2.5 year interval. The three other trajectories with dots show trajectories of the asymmetric, high spillover effectiveness scenario, in which the hybrid and radical technology are introduced, simultaneously $\left(\tau_2^i = \tau_3^i \right)$ and with 15 years between them. Both trajectories appear to yield the same equilibrium. In fact, the case where the radical technology is introduced later, results in the highest market share. This is because the hybrid technology matures before being able to capture some benefits from the HFCV. Along the axes we can observe the trajectories where only one entrant is introduced ($\tau_{-i}^I \to \infty$).The equilibrium installed base shares for these cases are equal to those with corresponding parameters in Figure 9a, where the scale factor 1, and spillover potential is 0.25 (E3 in this analysis) and 0.75 (E2 in this analysis).

While the combined market share is considerably higher than for the individual introductions, the individual shares of the entrant platforms are lower than in the case when they are introduced individually. That is, under current assumptions, the competition effects limit market share and dominate the spillover effects. For instance, the hybrid technology learns much from the incumbent. This, however, is of limited value to the radical technology. Further, the incumbent also learns and, while attractiveness of the platforms is higher than is the case with individual introductions, this is also the case for the incumbent. Also shown is the equilibrium installed base share for the symmetric and asymmetric, low spillover effect case {S,L}, {A,L}. Figure 10b) shows the evolution of installed base share for the {A,H} trajectory with late introduction of the radical technology against time.

This simulation reveals that the radical technology does not reach as much of its potential as the hybrid does. In equilibrium, all market shares remain constant while for each platform internal knowledge as well as spillover knowledge can be different. However, the growth rate of total knowledge must be identical across each. Three competing effects are at work to contribute to knowledge. First, there are competition effects that distribute the instantaneous market shares based on platforms' relative attractiveness. Second, there are internal learning and innovation feedbacks at work as production and sales proceed, allowing for improved attractiveness and that further build production and sales. Finally, there are spillover effects between the technologies. Initially the radical can catch up with the hybrid, through spillover. However, it will also build up knowledge itself, through learning-by-doing and R&D investment. However, that is partly available to the hybrid.

234

The net spillover effect to the radical captures the flow towards the radical (from mainly the hybrid), less those towards the incumbent (from the mainly the hybrid), and the hybrid (from both other players), each closing the gap with the other's learning. However at the same time there is also intensive interaction between the hybrid and the incumbent. This additional feedback, results in a steady state advantage for the hybrid.

Generally the technology potential is not identical across platforms. For example, hybrid vehicles will have to sacrifice space and weight to offer multiple propulsion technologies. Vehicles that propel on gaseous fuels have lower energy density, in volume, compared to those that drive on liquid fuels and thus generally lower tank ranges. Radically different designs, such as HFCVs could offer more space, and more features than others due to their inherently electric system, which also requires few moving parts. Figure11 adds this dimension to the analysis, showing scenarios as before, for varying technology potential, while we explore with it the role of scale effects. Figure 11a) shows the equilibrium penetration levels for the high, symmetric spillover effectiveness scenario, in the absence of scale effects. I show the equilibrium installed base share for E2 and E3, as a function of the technology potential of E3, relative to the incumbent, keeping the product of the hybrid and the radical identical to that of the incumbent: $T_{\Delta}^{0} = T_{3}^{0}/T_{1}^{0} = T_{1}^{0}/T_{2}^{0}$ (thus values $T_{\Delta}^{0} > 1$, corresponds with the technology potential for the radical being larger than that of the incumbent, while that of the incumbent is larger than that of the hybrid). The hatched line shows the analytically derived equilibrium for when all technologies are equal to their potential value. We see that including dynamic effects of learning and spillover reinforces the effects of a difference in technology potential on the installed

base shares. Figure 11b) shows the same scenarios, except that now we also apply a weak scale factor of value 3. This scale effect is considered weak as for this value no effect can be detected for the equilibrium value in the static case (dotted lines are identical to those in Figure 11a). In the dynamic case, we now see a tipping point: only one entrant will survive – the most superior.

Figure 11c) and d) show the same scenarios as in Figure 11a) and b), but for asymmetric spillover effectiveness, representing the true situation of a hybrid and a radical entrant. In absence of scale effects, the point where the hybrid and radical have identical market share is shifted to the right - the situation where the radical is superior and the hybrid is inferior to the incumbent. The case where all technologies are identical corresponds with the equilibrium of simulation (1) in **Figure** 10b, which was identical to the case of simultaneous introduction). Figure 11d shows again the weak scale effect scenario, now under asymmetrical spillover effects. In this case there is again a tipping point, allowing for only one entrant to succeed. This graph reveals how the superior technology can fail dramatically. In fact, successful penetration occurs for the radical only under extreme conditions. The weak scale effect imposed was sufficient to greatly reinforce the effect already apparent without any such effects. The radical succeeds only when it is significantly superior to the incumbent and the hybrid. For asymmetric spillover potential, the hybrid can accumulate its technology much faster than the radical, diffuses and sustains successfully for intermediate scale factors as well, while the radical fails for a larger range of scale factors. Under these conditions, hybrids can benefit enough from the spillover dynamics, improve their technology, and offset limitations from scale

effects. The more radical technology does not improve its technology fast enough to overcome the initial barriers. The hybrid survives under more adverse conditions, in the presence of a weaker alternative.

The mechanisms that were discussed to be at work in **Figure** 10, are drastically reinforced under the scale effects: while initially the system might get close to equilibrium, the scale advantage of hybrids widens the gap between the hybrid and the radical. Importantly: as the hybrid benefits, by definition, much more from the mature technology, the incumbent will generally lag, which makes the relevance of a installed base gap larger.

While the scale effects have little impact in isolation and the asymmetric spillover effects alone do not lead to the dramatic tipping, their interaction results in the real dramatic failure. With understanding from the preceding analyses it may seem likely that there are a large number of combinations of contexts that can generate conditions that result in failed diffusion of superior radical technologies. However, these conditions, when examined in isolation, do not have any significant impact. For instance, alternative fuels are introduced in the market at different times, after much of the competitive landscape has changed, they rely upon different fueling, distribution, and production infrastructures, parts of which may be compatible with those of other AFVs. I address this in the concluding analysis with three scenarios that capture different, small dissimilarities. Figure 12, left columns (a-c 1), show successful transitions towards the radical entrant. The right columns (a-c 2) show the failed transitions for the radical platform, achieved

by one parameter departure from the corresponding scenario on the left. Detailed parameter settings are provided in **Table** . The scenarios show for the failed cases: a) a further reduced spillover effectiveness between the incumbent and the radical, in absence of scale effects; b) less scale effects for the hybrid compared to the radical, in the case of more superior radical. This may be the case, for instance because the hybrid depends on an infrastructure that is compatible with that of the incumbent, which is the case for gasoline ICE-HEVs; and c) a lagged introduction of the radical with respect to the hybrid, which is a natural situation. In this case the combination of an (already improved) incumbent and maturing hybrid, the performance gap is too big to be overcome through spillovers.

Discussion and conclusion

The early decades of the transition to the horseless carriage in the late 19th and early 20th century constituted a period of excitement, but also a period of great uncertainty about which technology would prevail. The technology of steamers, EVs and the eventually prevailing ICE vehicles all changed dramatically during those periods. Technological change was particularly large when the industry became more organized and sales increased. Also, there were large spillovers between the various technologies within and outside the infustry. As Flink (1988) argues, critical to further development of the automobile was the development of the bicycle around 1890. Key elements of the automotive technology that were first employed in the bicycle industry included product innovations such as steel-tube framing, pneumatics, ball bearings, chain drive, and differential gearing, as well as process innovations, such as quantity production, utilizing

special machine tools and electric resistance welding. Importantly, not all vehicles benefited in the same way from this. For instance the differential gears contributed to those of ICE and steamers, while steel-tube frames were particularly beneficial to EVs, making them significantly lighter, providing a larger action radius (McShane 1994; Schiffer 1994).

Another types of interaction involved induced research intensity in response to upcoming threats. For instance, the light two cylinder cycle car stormed the market in the 1910s, responding to increasing congestion in the urban streets. But it did not take long before genuine vehicles became smaller in response to this threat, soon after which the cycle cars disappeared from the landscape, not being able to keep up with their limited experience. The prospective transition in the automobile industry, this time away from the fossil fuel burning ICE vehicles with many alternatives enter the market is subject to similar complex dynamics.

In this paper I emphasized the dynamics of and interaction between technology trajectories. This analysis was supported by a dynamic model that included explicit and endogenous consideration product innovation, learning-by-doing, investment decisions, and spillovers between the technologies. In contrast to other treatments of technology spillovers (e.g. Cohen and Levinthal 1989; Jovanovic and Macdonald 1994; Klepper 1996), spillovers, in this paper, are a function of the relative similarity between heterogeneous technologies. Further, in this setting, leading technologies may also learn from laggards, capturing various forms of sailing-ship effects.

To provide sufficient but controlled variation of relevant interactions, the analysis focused on the competitive dynamics of up to three players, one incumbent, one hybrid, and one radical platform. The competitive landscape under which the alternatives are introduced matters enormously for their likelihood of success. I analyzed in detail the dynamics resulting from three competing effects at work: competition effects that distribute the market shares, internal learning and R&D feedbacks, and spillover effects between the technologies. I found plausible conditions under which a superior technology may fail, competing against inferior entrants.

As expected, an entrant with a radically different technology, say the HFCV, may benefit from the existence of a hybrid technology, such as HEVs, when its technology potential is significantly higher than that of the hybrid. Alternatively, various alternative technologies may co-exist in equilibrium. The net spillover effect to the radical captures the flow towards the radical (from mainly the hybrid), less those towards the incumbent (from the mainly the hybrid), and the hybrid (from both other players). However, to illustrate the dynamic complexity, at the same time there is also intensive interaction between the hybrid and the incumbent. This is why a radical platform, occupying the margin within the space of spillover can be suppressed, even when equivalent or even superior to its competitors in terms of technology potential.

The automobile industry is subject to various forms of scale effects. The challenges for policy and strategy makers become apparent in when these are included in the analysis.

Successful diffusion and sustenance of AFVs are dramatically affected when spillover dynamics are allowed to interact with scale effects. Scale effects are important in the automotive industry. New platforms, consumer and investor familiarity needs to build up before they are considered on equal par (see Essay 1). Similarly, complementarities, such as fueling infrastructure need to build up with the vehicle fleet (see Essay 2). The analysis in this paper illustrates how such scale effects, modeled in reduced form, can have drastic effects on the technology trajectory and adoption dynamics, even when the effects in isolation are moderate. In particular technologies that develop slower, for instance those on the outside of a spillover landscape, are negatively affected.

On top of that, HFCVs will be introduced later and their scale effects are much stronger. Such a situation was the case with the transition towards the horseless carriage, with EVs having the burden of a slow developing support infrastructure, and steamers experiencing a liability of public acceptance from earlier times. This allowed ICE vehicles to gain market share, build experience and innovate more, and keep learning from its slower developing competitors. Similarly, in the modern transition, the various hybrid technologies might be well positioned. However, for a full policy analysis, an integrated model is needed that explicitly captures the various feedbacks of infrastructure, consumer acceptance, and fuel production and distribution dynamics, that all act differently for the various alternatives. The model must be subjected to more empirical cases and in more depth analyzed. A particular enrichment will be to study introductions that had variations of success.

With respect to the model structure, for the purpose of analytical clarity, I have allowed several simplifications. For instance individual firms were not modeled explicitly. Doing so will allow for a more elaborate capturing of industry level effects from the bottom up, such as learning-curves. Further, some firms will produce multiple platforms, thus yielding a richer distribution of spillover rates. Facing the transition challenges, several consortia emerge, but also partial collaborations across them. For instance GM, BMW and Toyota collaborate on hybrid technology, but not on their HFCV related R&D. Capturing such firm detail will also allow exploration of firm specific strategies. However, I do not expect that the central conclusions of this paper will be affected.

Another potential area of expansion is the consumer choice structure. While the technology heterogeneity was captured carefully, from the demand side substitutability among platforms differs as well. For instance the total portfolio of gaseous fuel vehicles might be treated by consumers as one "nest" of partially substitutable choices. Advances and increased demand for one platform of such a nest can have a positive effect on market shares of others that are also considered part of that nest. For instance, once familiarity of one type of gaseous fuels grows, others also benefit from this. Beyond our research focus, transition dynamics in the automobile industry, the PLC model can find a broader application in various new and mature markets, especially those that involve more complex products, with large diversity and large volumes, such as the upstream-high tech sector (e.g. semiconductor), as well as downstream high-tech sector (computers, PDA, cameras, mobile phones), energy (wind-energy), and aircrafts.

Besides opportunities for further work, the findings illustrate already the enormous challenges for policy and strategy makers. There are a wide range of patterns of behavior possible, including early success and failure, even of superior technologies. Small differences that have limited significance in isolation may have dramatic impact. Strategy and policy makers that support technology neutral incentives, such as fuel taxes, to stimulate AFVs may see unexpected side-effects through the co-development of the various other AFVs and incumbents that compete at the same time. On the other hand focused support of a single technology such as E85 or HFCVs is likely to stall when interdependencies between the technologies are not well understood. Further many other non-technology related dynamics, including those related to consumer acceptance and learning (as discussed in Essay 1), to infrastructure complementarities (Essay 3), or to product portfolios will dramatically alter strategies and policies of preference. However the research also suggests that there are opportunities for management at the level of technology portfolios. With the tools that are geared to support analysis of the dynamic complexity, the challenges to the transition can be understood, allowing for high leverage policies to be identified.

References

Abernathy, W. J. and J. M. Utterback (1978). "Patterns of Industrial Innovation." Technology Review **80**(7): 40-47.

Aghion, P., C. Harris, et al. (2001). "Competition, imitation and growth with step-by-step innovation." Review of Economic Studies **68**(3): 467-492.

Aghion, P. and P. Howitt (1992). "A Model of Growth through Creative Destruction." Econometrica **60**(2): 323-351.

Arrow, K. (1962). Economic Welfare and Allocation of Resources. Rate and Direction of Resources. R. Nelson.

Arrow, K. J., H. B. Chenery, et al. (1961). "Capital-Labor Substitution and Economic Efficiency." Review of Economics and Statistics **43**: 225-50.

Arthur, 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: 590-607.

Arthur, W. B. (1989). "Competing Technologies, Increasing Returns, and Lock-in by Historical Events." Economic Journal **99**(394): 116-131.

Baldwin, C. Y. and K. B. Clark (2000). Design rules the power of modularity, volume 1. MIT Press.

Basalla, G. (1988). The evolution of technology. Cambridge Cambridgeshire ; New York, Cambridge University Press.

Bass, F. M. (1969). "A new Product Growth Model for Consumer Durables." Marketing Science **15**(5): 215-227.

Ben-Akiva, M. E. (1973). Structure of passenger travel demand models. Cambridge, Massachusetts Institute of Technology.

Burns, L. D., J. B. McCormick, et al. (2002). "Vehicle of change." Scientific American **287**(4): 64-73.

Christensen, C. M. (1997). The innovator's dilemma : when new technologies cause great firms to fail. Boston, Mass., Harvard Business School Press.

Christensen, C. M. (2000). The Innovator's Dilemma, HarperBusiness.

Christensen, C. M. and R. S. Rosenbloom (1995). "Explaining the Attackers Advantage - Technological Paradigms, Organizational Dynamics, and the Value Network." Research Policy **24**(2): 233-257.

Clark, K. B. (1985). "The Interaction of Design Hierarchies and Market Concepts in Technological Evolution." Research Policy **14**(5): 235-251.

Cohen, W. M. and D. A. Levinthal (1989). "Innovation and Learning - the 2 Faces of R-and-D." Economic Journal **99**(397): 569-596.

Cyert, R. M. and J. G. March (1963). A behavioral theory of the firm. Englewood Cliffs, N.J.,, Prentice-Hall.

David, P. A. (1985). "Clio and the Economics of Qwerty." American Economic Review **75**(2): 332-337.

Dosi, G. (1982). "Technoogical Paradigms and Technoogical Trajectories: A suggested interpretation of the determinants and directions of technical change." Research Policy **11**: 147-162.

Energy Information Administration. (2005). Alternative Fuel Reports, Analyses 1993-2003. Energy Information Administration Accessed 11-20-2005. http://tonto.eia.doe.gov/reports/reportsD.asp?type=Alternative%20Fuel

Fine, C. H. (1998). Clockspeed : winning industry control in the age of temporary advantage. Reading, Mass., Perseus Books.

Flink, J. J. (1988). The automobile age. Cambridge, Mass., MIT Press.

Forrester, J. W. (1975). Collected papers of Jay W. Forrester. Cambridge, Mass., Wright-Allen Press.

Garud, R. and P. Karnoe (2003). "Bricolage versus breakthrough: distributed and embedded agency in technology entrepreneurship." Research Policy **32**(2): 277-300.

Geels, F. W. (2005). Technological Transitions and Systems Innovations - A Co-evolutionary and Socio-Technical Analysis. Northampton, MA, Edward Elgar.

Gelijns, A. C., J. G. Zivin, et al. (2001). "Uncertainty and technological change in medicine." Journal of Health Politics Policy and Law **26**(5): 913-924.

Geroski, P. A. (2000). "Models of technology diffusion." <u>Research Policy</u> **29**(4-5): 603-625.

Gilbert, R. J. and D. M. G. Newbery (1982). "Preemptive Patenting and the Persistence of Monopoly." <u>American Economic Review</u> **72**(3): 514-526.

Griliches, Z. (1957). "Hybrid Corn: an exploration in the economics of technological change." <u>Econometrica</u> **25**(4): 501.

Henderson, R. (1995). "Of Life-Cycles Real and Imaginary - the Unexpectedly Long Old-Age of Optical Lithography." <u>Research Policy</u> **24**(4): 631-643.

Henderson, R. M. and K. B. Clark (1990). "Architectural Innovation - the Reconfiguration of Existing Product Technologies and the Failure of Established Firms." <u>Administrative Science Quarterly</u> **35**(1): 9-30.

Holling, C. S. (2001). "Understanding the complexity of economic, ecological, and social systems." <u>Ecosystems</u> **4**(5): 390-405.

Homer, J. B. (1983). A Dynamic Model for Analyzing the Emergence of New Medical Technologies. Alfred P. Sloan School of Management. Cambridge, Massachusetts, Massachusetts Institute of Technology.

Homer, J. B. (1987). "A Diffusion-Model with Application to Evolving Medical Technologies." <u>Technological Forecasting and Social Change</u> **31**(3): 197-218.

Iansiti, M. (1997). "From technological potential to product performance: an empirical analysis." <u>Research Policy</u> **26**(3): 345-365.

Jovanovic, B. and S. Lach (1989). "Entry, Exit, and Diffusion with Learning by Doing." <u>American Economic Review</u> **79**(4): 690-699.

Jovanovic, B. and G. M. Macdonald (1994). "Competitive Diffusion." <u>Journal of Political Economy</u> **102**(1): 24-52.

Karnoe, P. (1999). "Wind energy in America: A history." <u>Technology and Culture</u> **40**(2): 417-419.

Karshenas, M. and P. L. Stoneman (1993). "Rank, Stock, Order, and Epidemic Effects in the Diffusion of New Process Technologies - an Empirical-Model." <u>Rand Journal of Economics</u> **24**(4): 503-528.

Katz, M. L. and C. Shapiro (1985). "Network Externalities, Competition, and Compatibility." <u>American Economic Review</u> **75**(3): 424-440.

Kemp, R. (2001). Path Dependence and Creation - Wind Energy. <u>Path Dependence and Creation</u>. R. Garud and P. Karnoe, Lawrence Erlbaum Associates, Publishers. http://tonto.eia.doe.gov/reports/reportsD.asp?type=Alternative%20Fuel

Kimes, B. R. and C. H. A. Jr. (1996). Standard Catalog of American Cars 1805-1942. Iola, WI USA, Krause Publications.

Klepper, S. (1996). "Entry, exit, growth, and innovation over the product life cycle." <u>American Economic Review</u> **86**(3): 562-583.

Levinthal, D. and J. G. March (1981). "A Model of Adaptive Organizational Search." <u>Journal of Economic Behavior & Organization</u> **2**(4): 307-333.

Lovins, A. B. and B. D. Williams (1999). <u>A Strategy for the Hydrogen Transition</u>. Tenth Annual US Hydrogen Meeting.

MacLean, H. L. and L. B. Lave (2003). "Evaluating automobile fuel/propulsion system technologies." <u>Progress in Energy and Combustion Science</u> **29**(1): 1-69.

March, J. G. (1991). "Exploration and exploitation in organization learning." <u>Organization Science</u> **2**: 71-87.

McFadden, D. (1963). "Constant Elasticity of Substitution Production Functions." <u>The Review of Economic Studies</u> **30**(2): 73-83.

McShane, C. (1994). <u>Down the asphalt path : the automobile and the American city.</u> New York, Columbia University Press.

Morecroft, J. D. W. (1985). "Rationality in the Analysis of Behavioral Simulation-Models." <u>Management Science</u> **31**(7): 900-916.

Mowery, D. C. and N. Rosenberg (1981). "Technical Change in the Commercial Aircraft Industry, 1925-1975." <u>Technological Forecasting and Social Change</u> **20**(4): 347-358.

Mowery, D. C. and N. Rosenberg (1989). <u>Technology and the pursuit of economic growth</u>. Cambridge, Cambridge University Press.

Nelson, R. R. and S. G. Winter (1982). <u>An evolutionary theory of economic change</u>. Cambridge, Mass., Belknap Press of Harvard University Press.

Nakicenovic, N. (1986). "The automobile road to technological change. Diffusion of the automobile as a process of technological substitution." <u>Technological Forecasting of Social Change</u> **29**(4): 309-340.

Reinganum, J. F. (1983). "Uncertain Innovation and the Persistence of Monopoly." American Economic Review **73**(4): 741-748.

Romm, J. J. (2004). "The hype about hydrogen." Issues in Science and Technology **20**(3): 74-81.

Rosenberg, N. (1976). Perspectives on technology. New York, Cambridge University Press.

Sanchez, R. and J. T. Mahoney (1996). "Modularity, flexibility, and knowledge management in product and organization design." Strategic Management Journal **17**: 63-76.

Schiffer, M. B., T. C. Butts, et al. (1994). Taking charge : the electric automobile in America. Washington, Smithsonian Institution Press.

Simon, H. A. (1969). The sciences of the artificial. Cambridge, M.I.T. Press.

Snow, D. (2004). "Extraordinary Efficiency Growth in Response to New Technology Entries: The Carburetor's 'Last Gasp'." Academy of Management conference Series.

Solow, R. (1967). Some Recent Developments in the Theory of Production. The Theory and empirical analysis of production. M. Brown and National Bureau of Economic Research. New York,, National Bureau of Economic Research; distributed by Columbia University Press**:** p26-48.

Sterman, J. (2000). Business dynamics : systems thinking and modeling for a complex world. Boston, Irwin/McGraw-Hill.

Strang, D. and S. A. Soule (1998). "Diffusion in organizations and social movements: From hybrid corn to poison pills." Annual Review of Sociology **24**: 265-290.

Tripsas, M. and G. Gavetti (2002). "Capabilities, Cognition and Inertia: Evidence from Digital Imaging." Strategic Management Journal **To be published**.

Tushman, M. L. and P. Anderson (1986). "Technological Discontinuities and Organizational Environments." Administrative Science Quarterly **31**(3): 439-465.

Tushman, M. L. and L. Rosenkopf (1992). "Organizational Determinants of Technological-Change - toward a Sociology of Technological Evolution." Research in Organizational Behavior **14**: 311-347.

Ulrich, K. (1995). "The Role of Product Architecture in the Manufacturing Firm."
Research Policy **24**(3): 419-440.

Utterback, J. M. (1996). Mastering the dynamics of innovation. Boston,

Harvard Business School Press.

Zangwill, W. I. and P. B. Kantor (1998). "Toward a theory of continuous improvement
and the learning curve." Management Science **44**(7): 910-920.

Figures

Alternative powered vehicles in the US 1880-2004
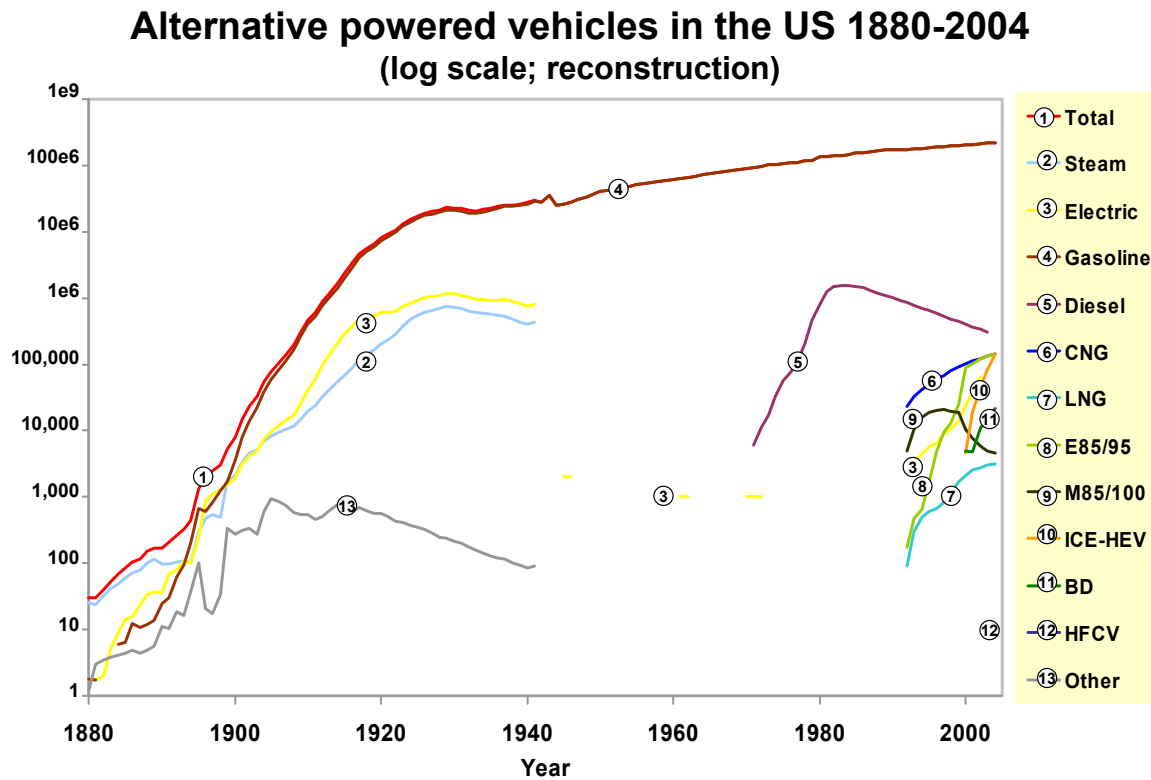### (log scale; reconstruction)

Figure 1 Early diffusion and preparation for substitution; reconstructed by author for qualitative illustrative purposes. Abbreviations of: LNG - liquid natural gas; M85 - Blend of 85% Methanol and 15% gasoline; BD - Biodiesel. Main sources: Energy Information Administration 2005, Kimes and Clark 1996).
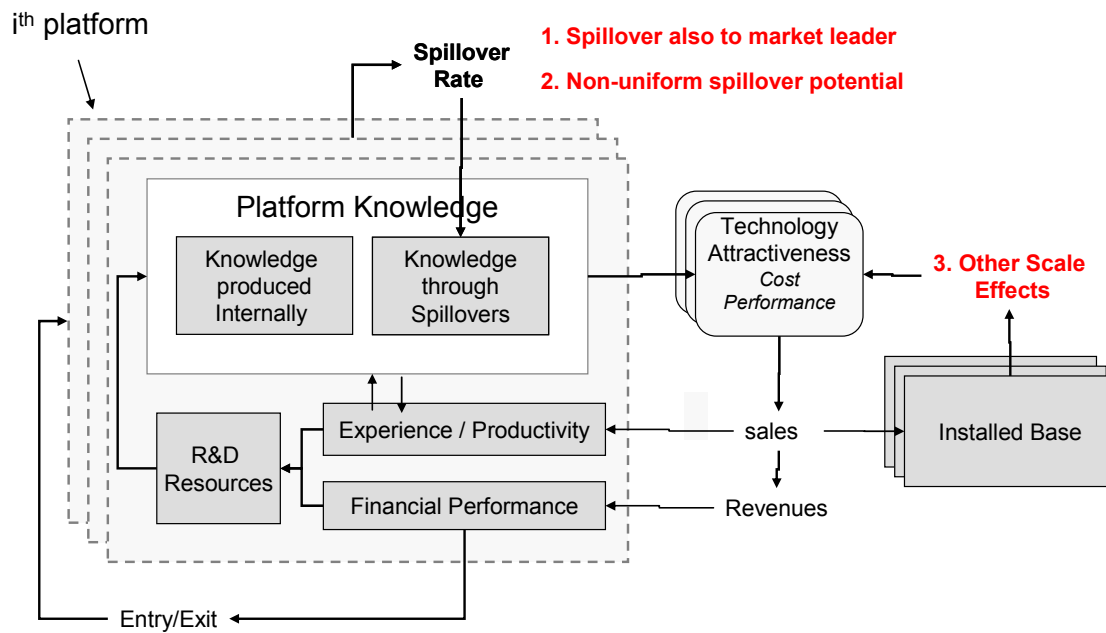
Figure 2 Model boundary. The model corresponds in many ways with the mainstream PLC models. Differences are: the unit of information and resource collection and allocation is the platform; spillovers flow between heterogeneous technologies; dynamics are explored in combination with non-technology related scale effects.
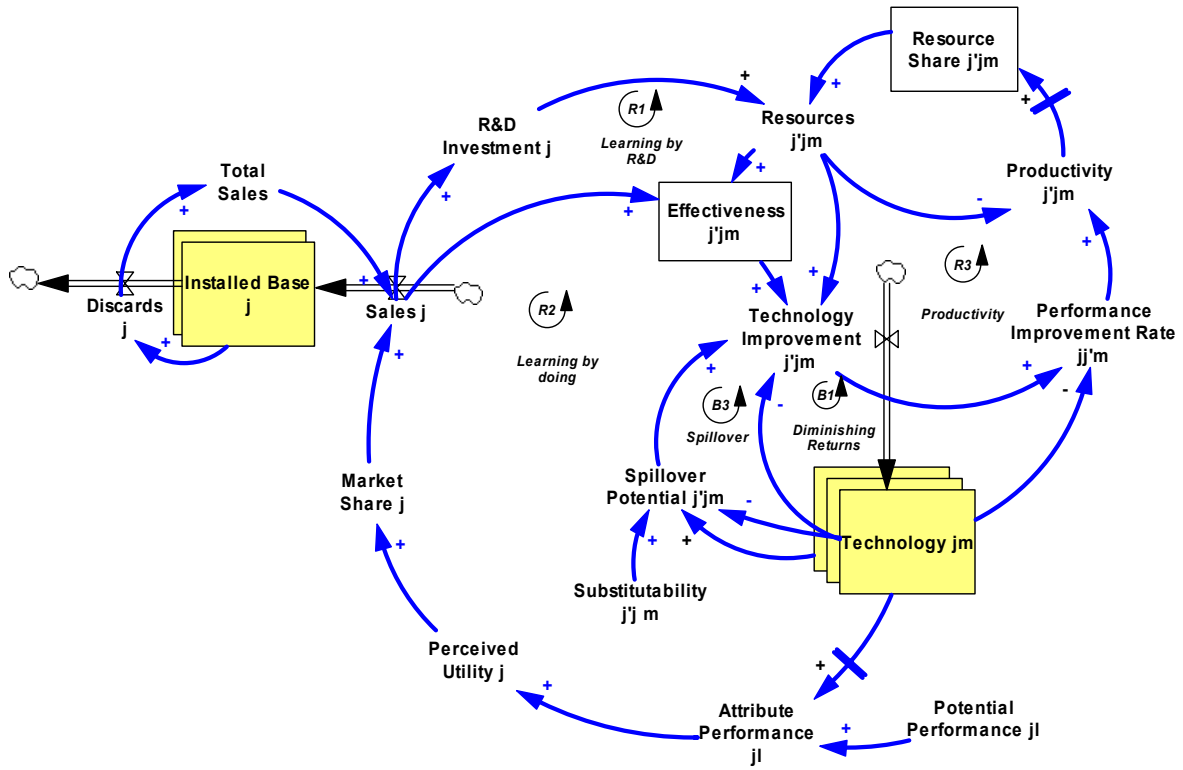
Figure 3 Principal feedbacks in the model.

segment	variable		indices	operation
Consumer choice	σ_j	Market share	Platform j Attribute l	Multinomial Logit Model
	u_j	Utility	Platform j	
	a_{jl}	Attribute	Platform j Attribute l	Mapping of technology on Attributes: {w} → {l}
Technological Performance	θ_{jw}	Relative technology	Platform j Activity type w	Normalization of technology
	T_{jw}	Technology	Platform j Activity type w	Diminishing returns
Knowledge Accumulation	K_{jw}^{e}	Effective knowledge	Platform j Activity type w	CES function
	K_{ijw}	Knowledge Input	Source platform i Target platform j Activity type w	Learning-by-doing, R&D, and spillovers;
Resource Allocation	R_{ijw}	Resources	Source platform i Target platform j Activity type w	Improve marginal return on effort

Figure 4 Diagramed representation of chain of operations from between resource allocation by producers to market share to vehicle consumer choice by consumers.
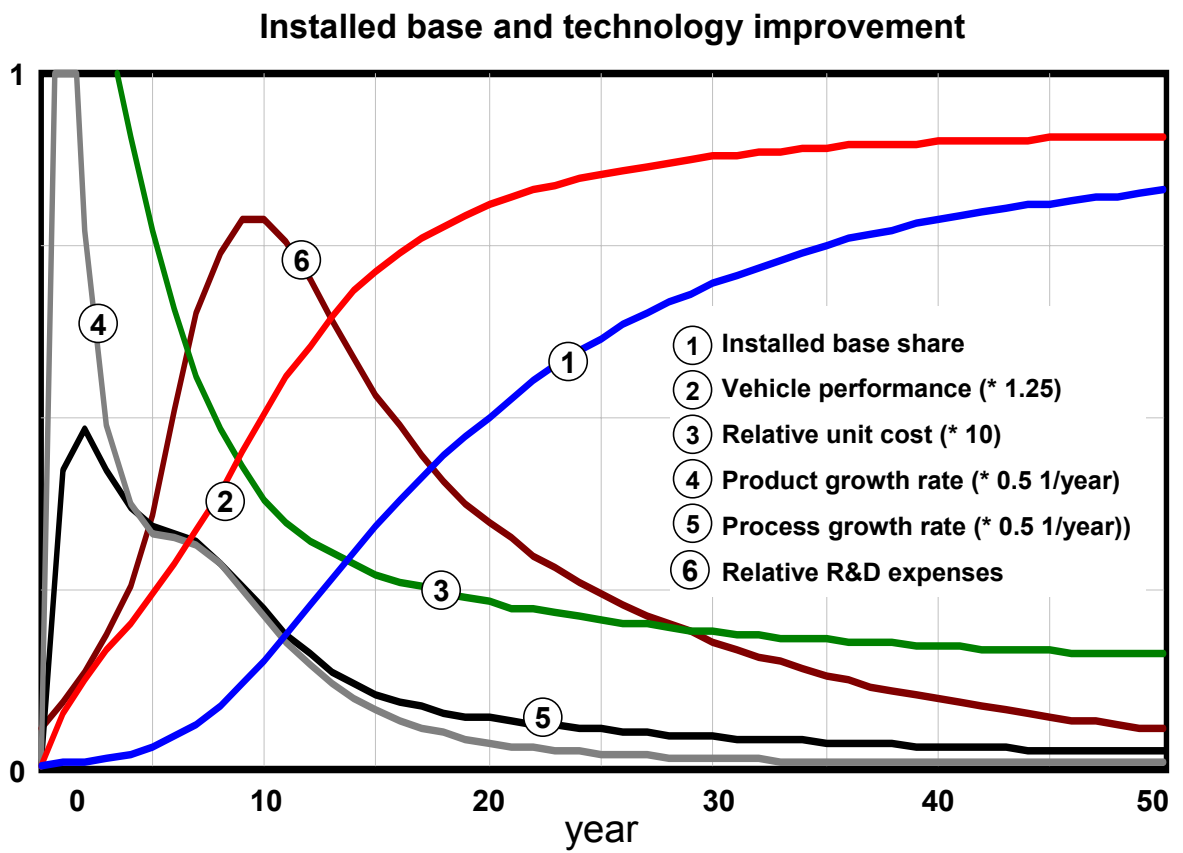
Figure 5 Simulation of PLC trajectory for single platform.
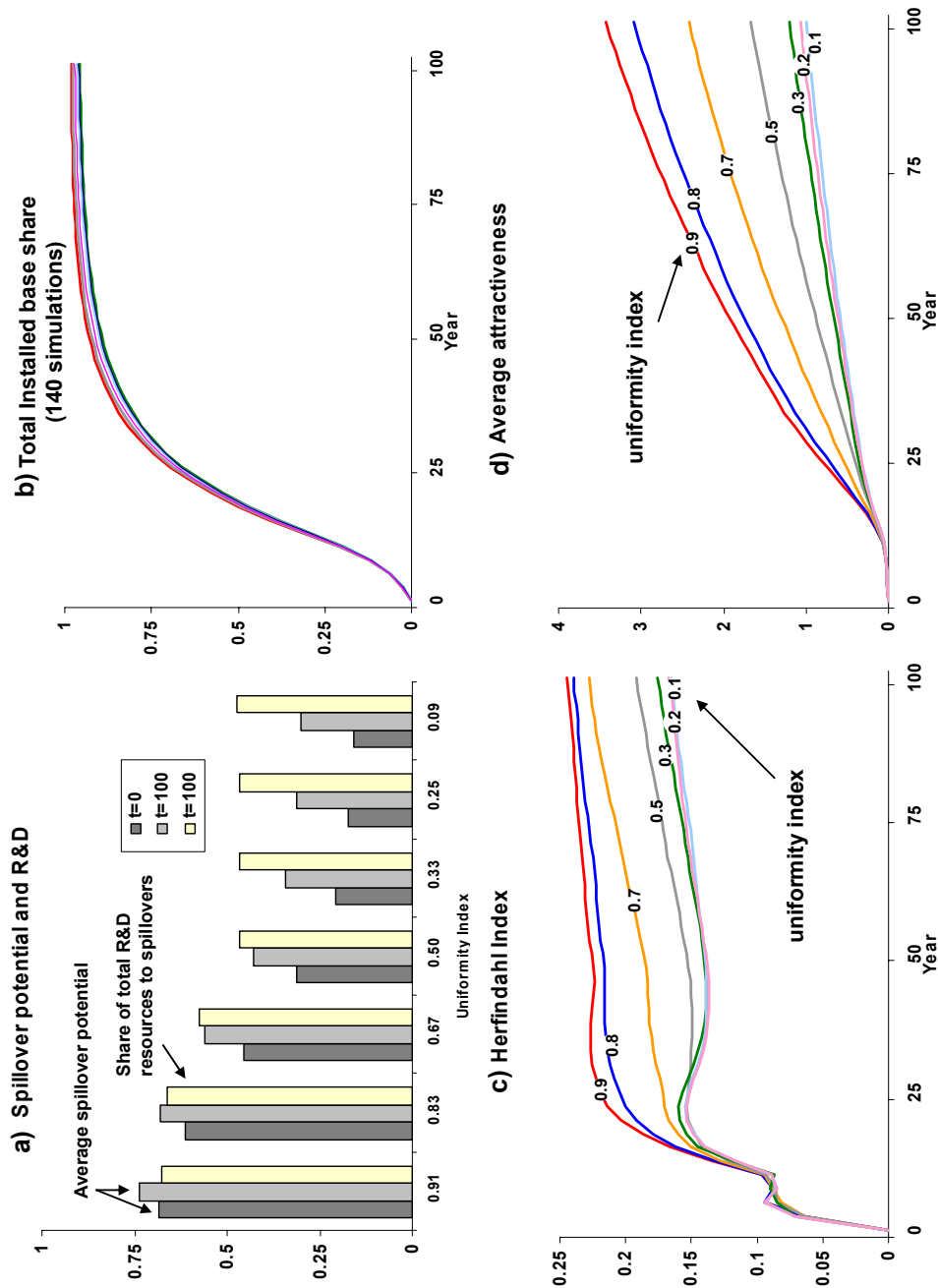
Figure 6 Endogenous platform entry. a) spillover potential, and R&D; b) total installed base for all simulations; c) Herfindahl over time for different levels of technology uniformity; c) share of total R&D resources allocated to internal R&D; b) Attractiveness of technologies in the market for different levels of technology uniformity.
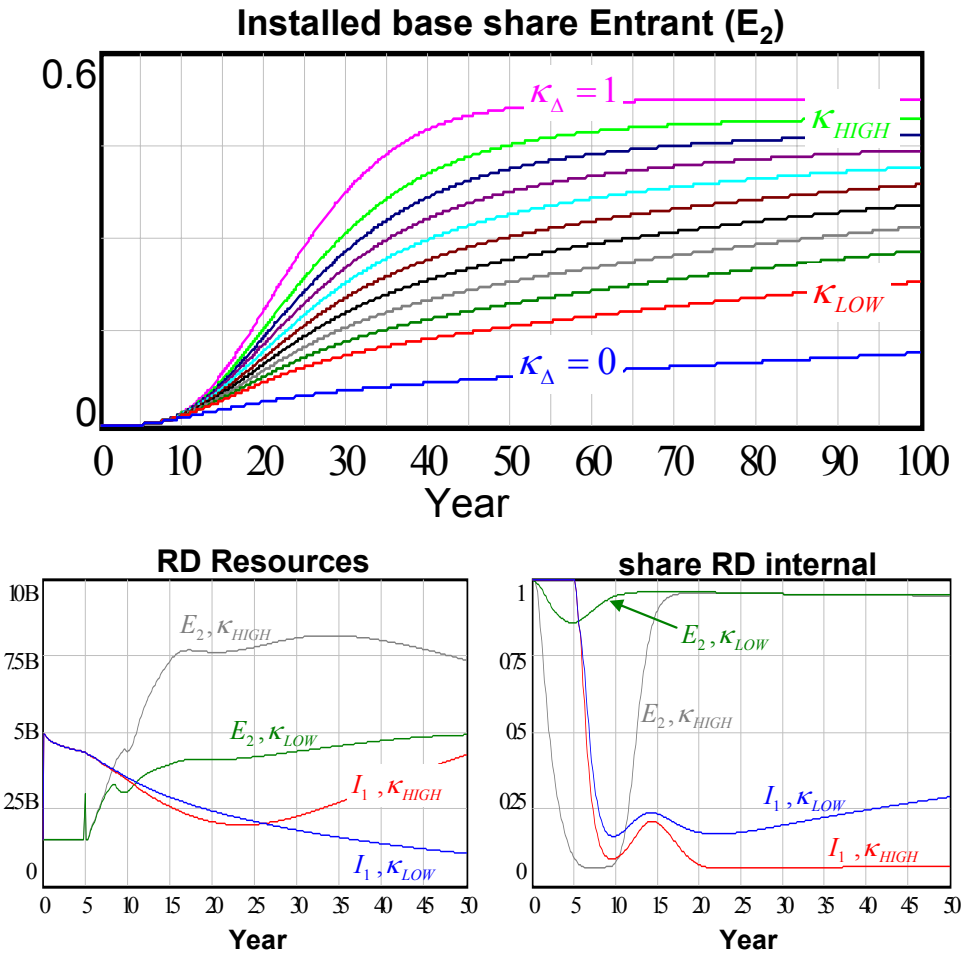
Figure 7 Base run dynamics for one incumbent and one entrant: a) entrant installed base share for various spillover potential factors; b) RD resource allocated over time for the low and high case of spillover potential factor. The low/high spillover potential case correspond each with one simulation for which both entrant and incumbent resources allocation are traced; c) share of resources allocated to internal R&D, further as in b).

Effect of scale on Attractiveness

Scale factor:

$$f^s \equiv 1\Big/ \varepsilon^s_{ref}$$

β^s

ε^s_{ref}

σ^v_{ref} **Installed Base Share** σ^v_j
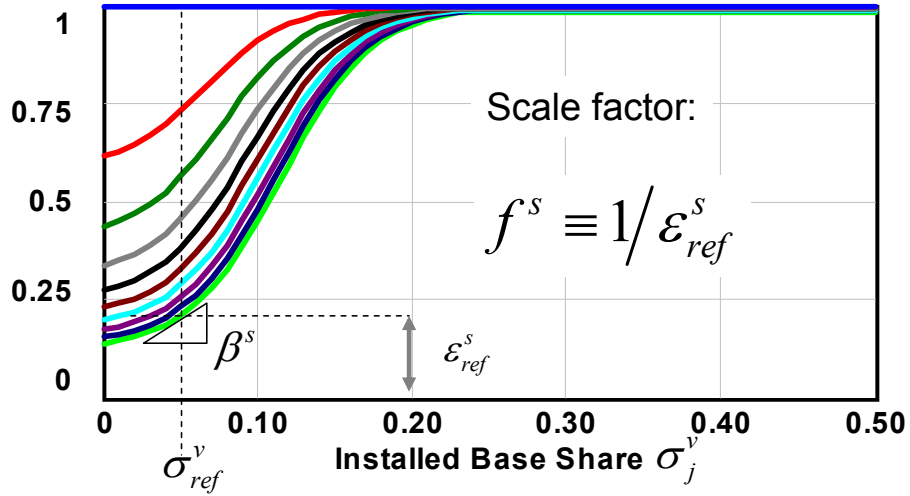
Figure 8 Incorporating complementarities and other scale effects. We vary the scale factor f^s in later analysis.

Equilibrium installed base share of the entrant

a) Varying spillover and scale

b) Varying spillover and technology
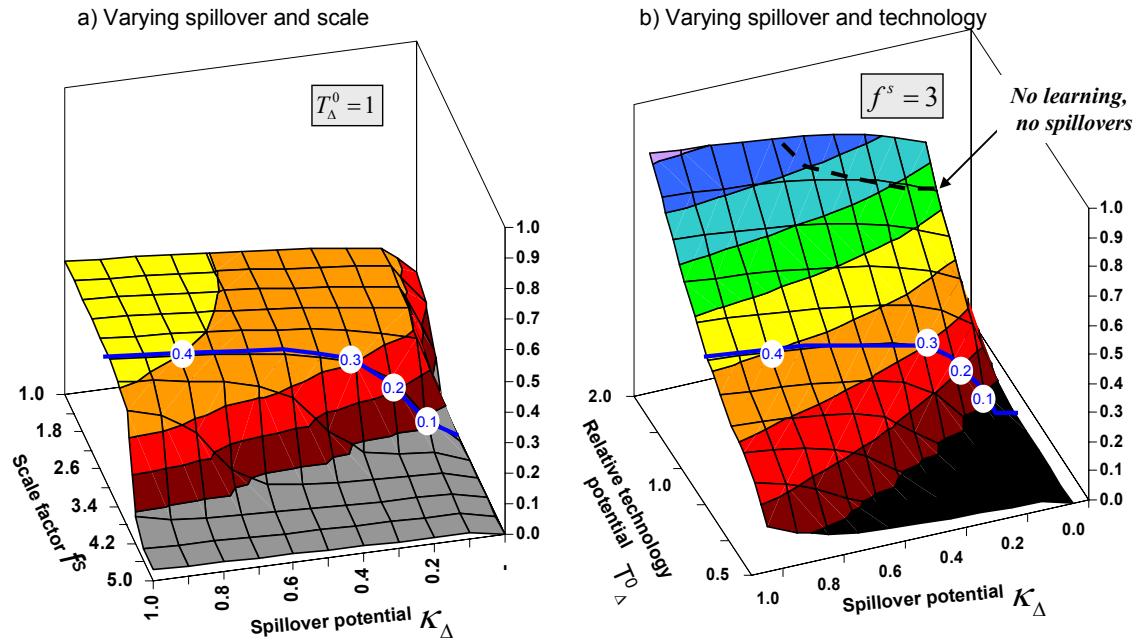
Figure 9 Entrant equilibrium adoption fraction as a function of spillover potential between the entrant and the incumbent and a) scale factor, b) relative technology potential. Thick lines correspond with identical.
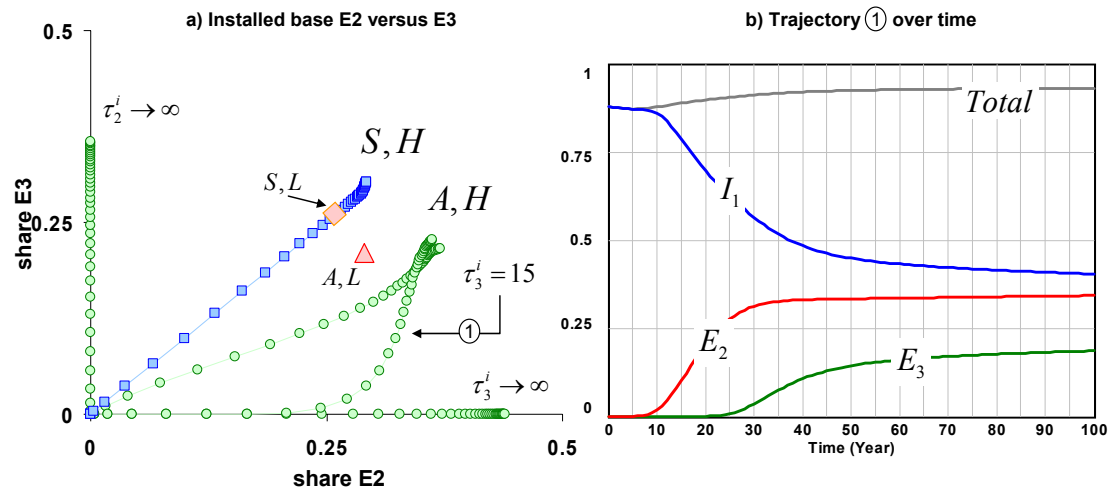
Figure 10 Technology trajectories, for incumbent and 2 entrant competition – base runs including spillovers.
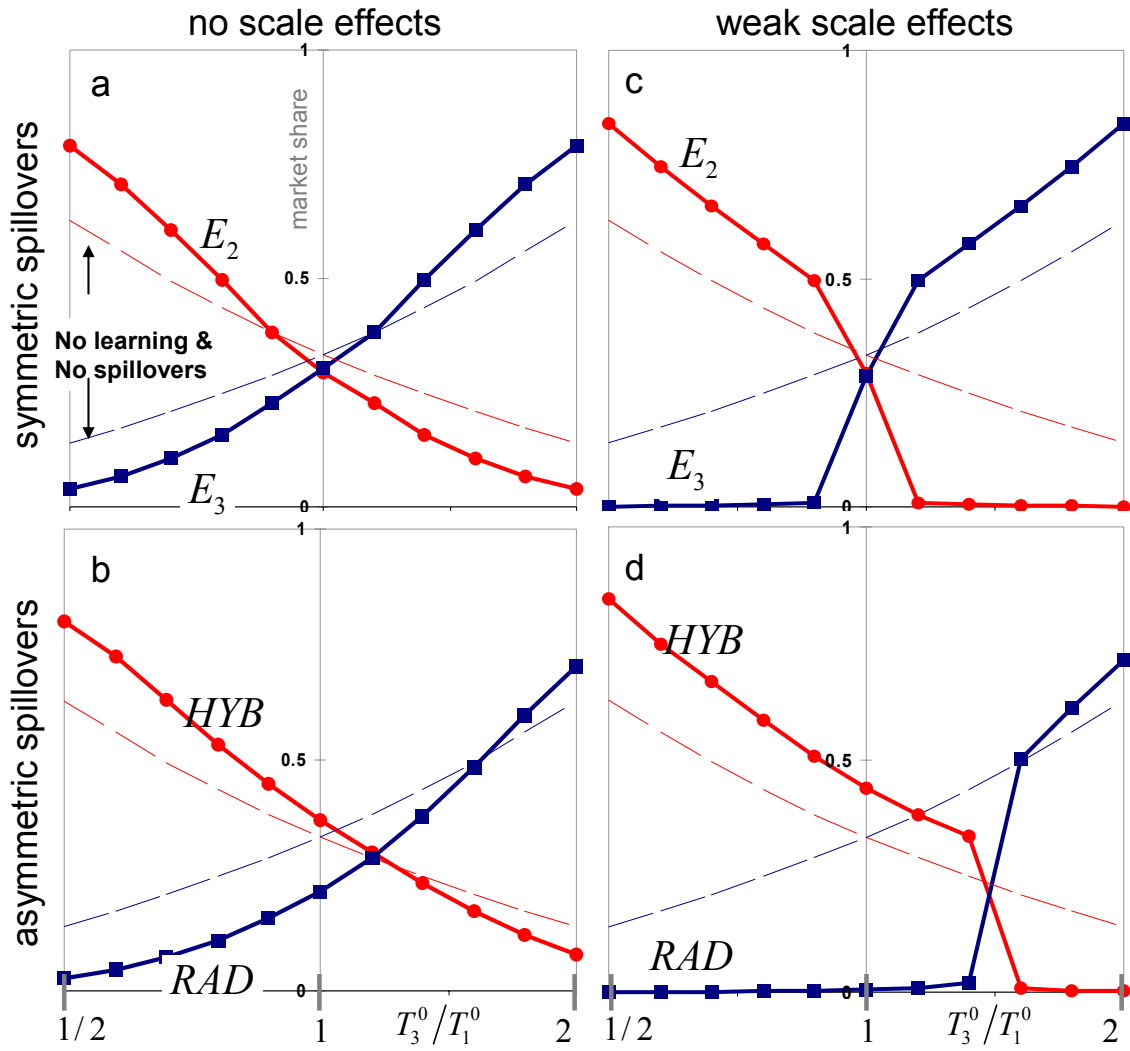
Figure 11 Scale effects and technology potential interacting with spillovers.

Installed base shares of incumbent and entrants
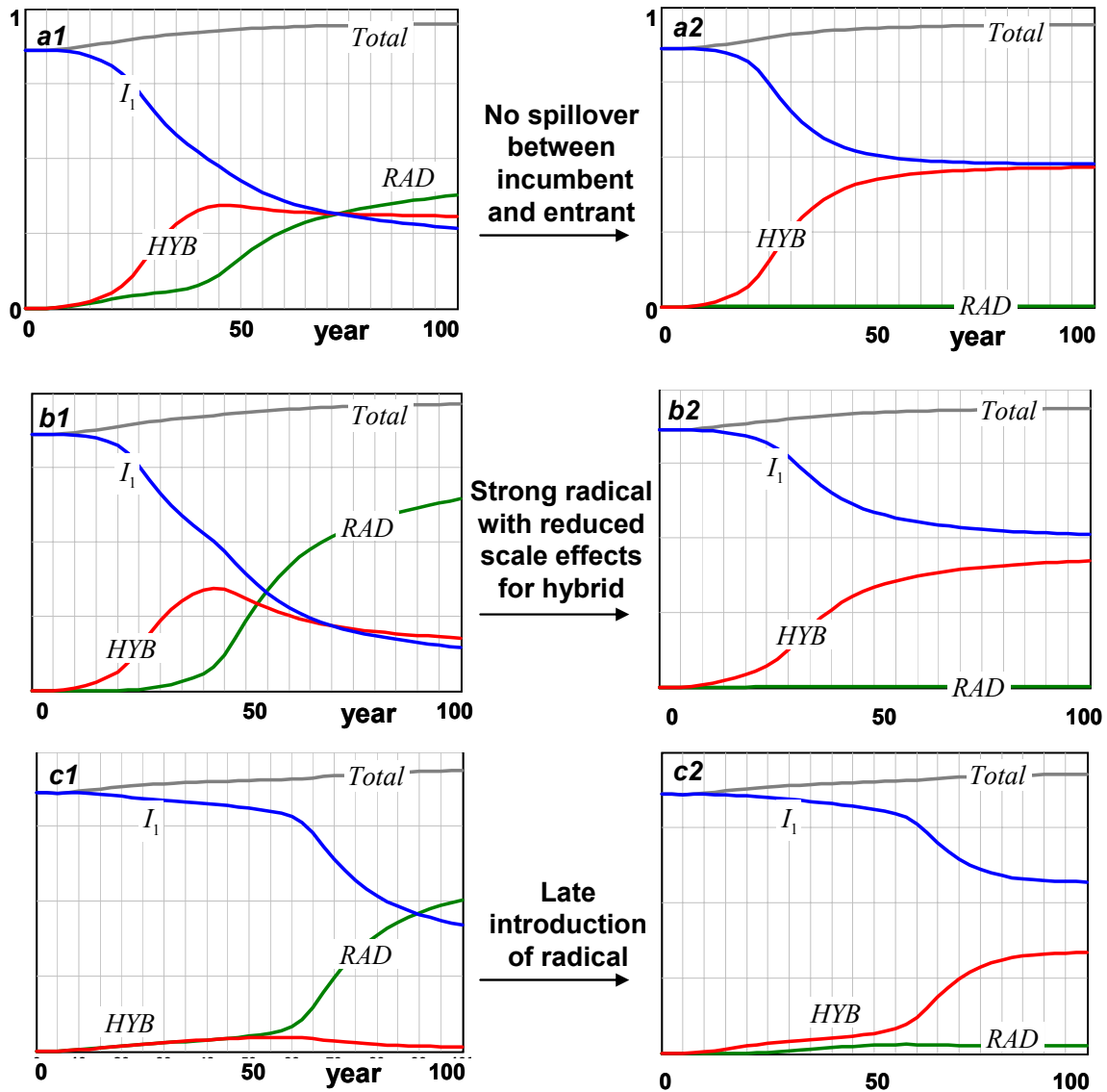
Figure 12 Transition trajectories for hybrids and radical platforms under asymmetrical configurations. The top row shows successful transitions to the radical, the bottom shows failures, as a function of: a) varying spillover potential; b) varying scale effects; c) varying introduction timing.

Tables

Table 1 Evolution of the automobile industry: users, technology, firms. Source: compiled by author.

Area	ICE 1890	ICE 1910	ICE 1960
Users	almost none	few	millions
User familiarity	almost none	moderate	high
User experience	almost none	small	large
Firms/entrepreneurs of main product	many	many	few
Firms across value chain	few	moderate	many
Performance of technology	low, growing rapidly	medium, growing	high, stable
Variety of technology	large	moderate	small
Cost of production	high, stable	medium, fluctuating	stable
Experience (cumulative vehicles)	~hundreds	~million	~billion
Diversity of Experience	large	moderate	small
Sources of innovation	many	moderate	few
Complementarities developed	few	rising	many

Table 2 Parameter settings for simulations, unless otherwise stated. All reference parameters that are not mentioned are set equal to 1.

Short	Description	Value	Units	
General				
H	Total households	100e6	people	
Firm Structure				
m_j	Markup	0.2	dmnl	
C_j^k	Capital cost	0	\$/year	
c^f	Unit production cost not subject to learning	3,000	\$/vehicle	
c^v	Unit production cost variable at normal technology	12,000	\$/vehicle	
Technology and Knowledge				
τ^t	Time to realize technology frontier	2	years	
η_w^k	Technology learning curve exponent to knowledge accumulation	0.3	dmnl	
K_w^0	Reference Knowledge	50	Knowledge units	
ς_{jw}^k	Elasticity of Substitution Parameter	1.5	Dmnl	
Γ_w	Normal knowledge growth rate	1	Knowledge units/year	
$\eta_1^i, \eta_2^i, \eta^s$	returns to resource allocation	1,0.2,0.8	Dmnl	
s_0	Reference sales for normal production	4e6	Vehicles/ year	
η_w^o	returns to resource allocation spillover	0.3	Dmnl	
g_w^o	Normal spillover knowledge growth fraction	10	Dmnl/year	
R^0	reference resources for total R&D	1.5e9	\$/year	
τ^r	Time to adjust resources	1	year	
τ^p	Planning horizon for resource allocation	5	Years	
Consumer Choice				
τ^d	Time to discard a vehicle	10	Years	
τ^a	Time between adoption decisions for non-drivers	10	Years	
β_1	Sensitivity of utility to vehicle performance	0.6	Dmnl	
β_2	Sensitivity of utility to vehicle price	-0.3	Dmnl	
K_{11w}^0	Knowledge of incumbent at introduction	1	Knowledge units	
$K_{jjw}^0 \; j \neq 1$	Knowledge of entrant at introduction	0.1	Knowledge units	
K_{ijw}^0	Spillover knowledge at introduction	0	Knowledge units	

Table 3 Parameters manipulated for graphs 8-11

Graph	Scenario	$\dfrac{T_2^0}{T_1^0}$	$\dfrac{T_3^0}{T_1^0}$	κ_Δ	$\dfrac{\kappa_{2\Delta}}{\kappa_\Delta}$	f_2^s	f_3^s	$\tau_3^i - \tau_2^i$
9a	Variable spillover potential & scale factor	VAR	-	VAR	-	1	-	-
9b	Variable spillover & technology potential	1	-	VAR	-	VAR	-	-
10a	Symmetric/Strong spillover (SS)	1	1	0.75	1	1	1	{0,15}
10a	Symmetric/Weak spillover (SW)	1	1	0.25	1	1	1	0
10a	Asymmetric/Strong spillover (AS)	1	1	0.75	0.25	1	1	0
10a	Asymmetric/Weak spillover (AW)	1	1	0.25	0.25	1	1	0
11a	Symmetric spillover, No Scale	T_1^0/T_3^0 VAR		0.75	1	1	1	0
11b	Asymmetric spillover, No Scale	T_1^0/T_3^0 VAR		0.75	0.25	1	1	0
11c	Symmetric spillover, Weak Scale	T_1^0/T_3^0 VAR		0.75	1	3	3	0
11d	Asymmetric spillover, Weak Scale	T_1^0/T_3^0 VAR		0.75	0.25	3	3	0
12a1	Base	0.75	1.33	0.75	0.25	2	2	0
12a2	RAD Fail	0.75	1.33	0.75	0	2	2	0
12b1	Base	1	2	0.75	0.25	3	3	15
12b2	RAD Fail	1	2	0.75	0.25	2	3	15
12c1	Base	0.75	1.33	0.75	0.25	3	3	0
12c2	RAD Fail	0.75	1.33	0.75	0.25	3	3	15

Technical appendix accompanying Essay 3

1 Introduction

The model described in this Essay is designed to capture the technology trajectory of and competition among multiple types of alternative vehicles, along with the evolution of the ICE fleet. For example, the model can be configured to represent ICE and alternatives such as ICE-electric hybrid, CNG, HFCV, biodiesel, E85 flexfuel, and electric vehicles. However, the Essay makes a number of simplifying assumptions that allow us to explore the global dynamics of the system. In this appendix I discuss additional components of the full model, highlighting those structures required to capture the competition among multiple alternative platforms, with their more particular context. This appendix provides also additional information to accompany the model and the analysis of Essay 3. Each subsection is pointed to from a paragraph within the Essay.

Sections group issues by:

2 **Elaborations on the model** that provide details on expressions that were not fully expanded due to space limitations (in particular we discuss functional forms).

2 **Stipulations** that provided notes on, additional motivations for, or insight into the model or analysis.

3 **Boundary constraints considered**, providing information about tests that were done by including additional behavioral and physical constraints. They partially reinforce, or otherwise dampen the dynamics, without affecting the fundamental insights of the model.

4 **Model and analysis documentation.** Essay 3, in combination with the section that

elaborates on the model provides sufficient information to replicate the model. The

Essay provides sufficient information to replicate the analysis.

5 **References**

2 Elaborations on the model

This section elaborates segments of the model that were highlighted in the paper but not

fully expanded due to space limitations. These elaborations include in particular selected

functional forms for functions that were provided in general form in the model, or more

detailed decision structures

a) Resource allocation

The process of resource allocation was discussed in the model (Equation 9 and 10, and

discussion). Here we first provide the same process in more detail and for more. **Figure

12** illustrates this process in more detail, in particular how in the model, at different

levels, competition for limited resources plays out. Constrained by allocations at more

aggregate levels we derive: i) a share of total revenues going to R&D, σ_j^{rd}; ii) the share of

total R&D resources of platform j, that managers or system integrators dedicate to the

various modules $\sigma_{m,j}^{rd}, \sum_m \sigma_{m,j}^{rd} = 1$; iii) the share of total R&D resources of platform j,

module m that chief engineers dedicates to activity w $\sigma_{jmw}^{rd}, \sum_w \sigma_{jmw}^{rd} = 1$; iv) the share of

total R&D resources of platform j, within module m, process w, that managers dedicate to

internal knowledge accumulation, $\sigma^{rd}_{j,jmw}$, as opposed to spillovers $\sigma^{rd}_{j,jmw} = 1 - \sigma^{rd}_{-j,jmw}$;

and finally, v) the share of total R&D spillover resources of platform j, R&D process w,

within module m, that engineers dedicate to extracting knowledge from

platform $i \neq j$, $\sigma^{rd}_{i,jmw}$, $\sum_{i \neq j} \sigma^{rd}_{i,jmw} = 1$.
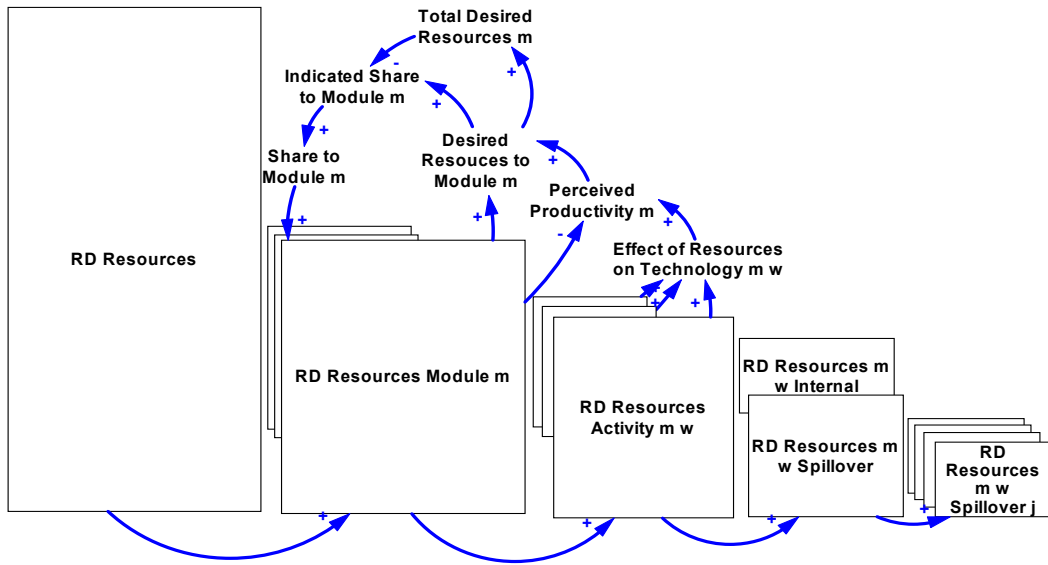
Figure 12 R&D Resource Allocation (boxes) throughout the decision making chain,

performance metric for each decision (top) and decision making process in detail for

resource allocation to module m.

Structurally each decision process is identical. Figure 4 illustrates this resource allocation

process, and in detail for one point in the hierarchy. We follow this here. We label a

decision point d, assigning lower numbers to points up the hierarchy. With the set of

potential allocations at x_d being $\{x\}_d$, the share that x_{d+1} receives from source x_d is

$\sigma^{rd}_{x_d, x_{d+1}}$. This share adjusts to its indicated share $\sigma^{rd*}_{x_d}$ over adjustment time τ^{rd}_d:

$$\frac{d\sigma^{rd}_{x_d,x_{d+1}}}{dt} = \left(\sigma^{rd*}_{x_d,x_{d+1}} - \sigma^{rd}_{x_d,x_{d+1}}\right) \Big/ \tau^{rd}_d \tag{A.1}$$

In the case of the example of **Figure 12**, this is the share of total RD resources goes to each module *m*. The adjustment time is the result of bureaucratic - and information gathering delays, depends thus on complexity, and can be different at different decision points.

The indicated share $\sigma^{rd*}_{x_d,x_{d+1}}$ is the outcome of the continuous bargaining for, given, scarce resources R_{x_d} at each decision point *d*, and equals desired resources $R^{rd\circ}_{x_{d+1}}$ divided by the resources others credibly bargain for:

$$\sigma^{rd*}_{x_d,x_{d+1}} = R^{rd\circ}_{x_{d+1}} \Big/ \sum_{\{x\}_{d+1}} R^{rd\circ}_{x_{d+1}} \tag{A.2}$$

Stakeholders at x_d can credibly bargain for more resources when expected returns $\varsigma^{rd\sim}_{x_d}$ exceed the reference value at this decision point, $\varsigma^{rd\sim}_{\{x\}_d}$:

$$R^{rd\circ}_{x_d} = f\left(\varsigma^{rd\sim}_{x_d} \Big/ \varsigma^{rd\sim}_{ref,\{x\}_d}\right) R^{rd}_{x_d}; f' \geq 0; f \geq 0; f(1) = 1 \tag{A.3}$$

Note that $R_{x_d} = \sigma^{rd}_{x_{d-1},x_d} R_{x_{d-1}}$.

b) Expected return of effort

An R&D resource allocation task involves by nature attempts to explore using some form of forward looking. I assume that for the assessment of the return on investment involves decision makers attempt to understand, at least locally, the structural factors that influence their improvement efforts. For example, one can be interested in the returns in her platform *j*'s knowledge base for activity w K_{jw} , deriving from resources dedicated to

extracting knowledge from platform i,. She will seek information about the constituents

of knowledge accumulation: i) the relevance of the source to total knowledge κ^*_{ijw} ; ii) the

perceived potential accumulating rate of spillover knowledge γ^*_{ijw} ; and, iv) the

productivity of resources r^*_{ijw} . This would imply for the expected returns on effort:

$$\varsigma^*_{ijw} = \gamma^*_{ijw} k^*_{ijw} r^*_{ijw} \qquad\qquad (A.4)$$

The assessment comprises activities as market research, interpretation of reports, study of

patents, evaluation of historic results, study of journals, and information exchanged over

coffee, in seminars, and during golf matches. Such assessments do not yield perfect

information about all factors, takes time, and are subject to information processing

constraints. Therefore we assume that decision makers: a) understand the correct

structural factors, but simplify their world by assuming that during their assessment that

the environment remains constant; b) it takes time to learn about the state of the

environment and the parameters. Thus, to assess the return on effort for collecting

knowledge from another platform, one has a perception of the knowledge of the other

platform that one assumes to remain constant during the planning horizon. Further, one

updates the perception of the knowledge base, but this takes time.

Assumptions about decision makers' available information

We now will illustrate what decision makers more generally need to know under to

allocate their resources equal to the marginal return on effort, at least for some bounding

set of assumptions. For the purpose of capturing the decisions related to resource

allocation, there are two types of activities. Some resources are typically adjusted

reactively, for instance, reallocation of resources for spillovers between source platforms.

Others involve longer term anticipation, such as adjustment of resources across modules.

In the first case improving perceived returns on effort can be captured by assuming that

the rate of accumulation of performance indicator P will be adjusted. In the second case

the NPV of P over a planning horizon τ_p will be improved.

We now will use specific examples to illustrate what decision makers need to know in

order to have their target return on effort equal the marginal return on effort.

Example: process improvement and product innovation

With the change rate of a performance indicator being $\dot{P}_i = \dfrac{dP_i}{dX_i}\dfrac{dX_i}{dt}$, the direct

adjustment of resources would require:

$$\frac{d\dot{P}_i}{dR_i} = \frac{dP_i}{dX_i}\frac{d\dot{X}_i}{dR_i}$$

We first determine the marginal return of knowledge accumulation, thus $P \equiv K_{jw}$, in

terms of resource allocation to activity w, thus R_{jw}, which implies:

$$\frac{d\dot{K}_{jw}}{dR_{jjw}} = \frac{dK_{jw}}{dK_{jjw}}\frac{d\dot{K}_{jjw}}{dR_{jjw}}$$

And with Equation (6) and the CES relation in the Essay:

$$\dot{K}_{jjw} = \varepsilon^i_{jw}\left(R_{jw}/R_0\right)^{\eta^i_w}\Gamma_w$$

$$K^e_{jw} = \left[\kappa_{jjw}\left(K_{jjw}/K^0_w\right)^{-\rho^k_{jw}} + \left(K_{\sim jw}/K^0_w\right)^{-\rho^k_{jw}}\right]^{-1/\rho^k_{jw}}$$

where $K_{\sim jw}$ is the total spillover knowledge, $K_{\sim jw} \equiv K_w^0 \left[\sum_{i \neq j} \kappa_{ijw} \left(K_{ijw} / K_w^0 \right)^{-\rho_{jw}^k} \right]^{-1/\rho_{jw}^k}$

We get:

$$\frac{d \dot{K}_{jw}}{dR_{jjw}} = \eta_w^i \kappa_{jjw} \left(\frac{K_{jw}^e}{K_{jjw}} \right)^{1+\rho} \frac{\dot{K}_{jjw}}{R_{jw}}$$

And attaining the optimal NPV, holding the environment constant would require:

$$\left. \frac{dP_i}{dR_i} \right|_{\tau_p} = \left(\left. \frac{dP_i}{dK_i} \frac{dK_i}{dR_i} \right|_{\tau_p} \right) \tag{A.5}$$

Maximization of Net Present Value yields, in this case the same type of relationship, because there is constant returns in accumulation of internal knowledge K_{jjw}.

In this case, the true values of these factors that determine the marginal return on effort are, for the productivity $\gamma_{jjw} = \varepsilon_{jw}^i \Gamma_w$, which is simply the ease at which they accumulate more knowledge, $r_{jjw} = \eta_w^i \left(R_{jw} / R_0 \right)^{\eta_w^i - 1}$, which gives the slope return on adding more resources. For instance, if the budget is a factor above than what it should be for all to be effectively spend, r_{jjw} does not increase anymore. The last term, the relevance of knowledge equals $k_{jjw} = \kappa_{jjw} \left(K_{jw}^e / K_{jjw} \right)^{1+\rho}$. A higher factor share indicates more relevance. Acknowledging this will correspond with the notion that producers of HFCVs will expect more from observing EVs, than from biodiesels. Further, if the elasticity parameter is infinite, the distribution parameter equals 0 and the whole expression is identical to its factor share. That is, when knowledge is additive, we always look towards those sources that have larger factor contributions. However, when substitutability of

knowledge is less than perfect, we increasingly expect to benefit from other sources as well.

An important point of this exposition is that decision makers will use heuristics that correspond with chunks of a structure that provides a locally optimal solution, and as a whole allows them to get reasonably close to that exact contributions.

Formulating decision making chunks at each stage, whether based on the current rates or NPV, yields similar types of structures at each level. This is how the decisions have been formulated in the model. There are important conditions, one of them is that many environmental factors are held constant in the resource allocation decisions, and thus the definition of "local" is quite narrow (and defined for this purpose above). Further, there are significant time delays in learning them. These two conditions assure that the decision structure for resource allocation conforms to the perspectives of bounded rationality.

c) Learning about productivity

Relevant knowledge, input factors, elasticity of substitution, and platform specific quality are learned over time. Biases by investors towards tested technology have important dynamics implications for the transitions. We capture this in the expanded model by allowing, for instance perceived knowledge of others, $K_{\tilde{j},ijmw}$ adjusts to the indicated level over time τ^K :

$$dK_{\tilde{ijmw}}\big/dt = \left(K_{ijmw} - K_{\tilde{ijmw}} \right)\big/\tau^K$$

A more sophisticated formulation would have the learning rate depend on attention as a function of exposure and interest, but for assessing its basic impact, this will suffice.

d) Platform- versus firm-level learning-by-doing

Equation 7 in the Essay discusses the learning-by-doing effect on process technology. Learning-by-doing can be seen as partly occurring at platform level (fast spillovers between firms for the same platform) and partly at the firm level. The number of firms changes considerably over the lifecycle of a technology, in particular, after a swift ramp-up, the number of firms tend to peak, followed by a shakeout. This would imply that the effective learning, at platform level, on is much slower early on.

It is useful to be able to capture this. I capture this by introducing the effective sales s_j^n for learning in the equation:

$$\varepsilon_{j2}^i = \left(s_j^n / s_0\right)^{\eta^s}$$

Where the effective sales s_j^n is a function of the total platform sales and the maturity of the platform and is represented by the share of the platform sales divided by the effective number of producers, for learning, n_j;

$$s_j^n = \left(s_j^n / n_j\right)$$

The effective producers assumed to decline with the maturity of the industry:

$$n_j = w_j^n n^{new} + \left(1 - w_j^n\right) n^{mat}$$

The weight of n^{new} declines with total platform sales:

$$w_j^n = f^n\left(s_j/s_0\right); f(0) = 0; f' \geq 0; f(1) = 0.5; f(\gg 1) = 1;$$

The weight starts at 1, and increases with total sales (ignoring the first ramp-up), saturating at 0. A sensible shape is the S curve. Here we use the standard logistic curve, with the inflection point at reference sales for learning s_0^n:

$$w_j^n = \exp\left(\beta^n\left[\left(s_j/s_0^n\right) - 1/2\right]\right) \big/ 1 + \exp\left(\beta^n\left[\left(s_j/s_0^n\right) - 1/2\right]\right)$$

This structure is switched off for the AFV analysis, assuming that for these dynamics very few firms will be in competition ($\beta^n = 0 \rightarrow n_j = 1$). This is valid for the basic analysis. However, in general analysis includes simulations with new industry formation. For those cases I set n^{mat}=2, n^{new}=20, $\alpha^n = -1$ and set s_0^n equal s_0, or 20% of the total potential market. The effect of this is that learning by doing is moderately slowed down in the first decade or two in the industry.

e) Vehicle choice and nesting

This focus of this Essay is on multi platform competition. Endogenous platform entrance is one of the dynamics that are captured. In the basic formulation of consumer choice between the available platforms (equation 13), when new vehicle types or a set of platforms are introduced, demand elasticity to the number of platforms is constant. This assumes the existence of a powerful feedback loop that is in reality much weaker: an increase in the variety and number of models, does not necessarily increase aggregate utility of "the vehicle" proportionally. That is, total increase in demand is generated depends on the correlation (or substitutability) of preference across a range of products in

the choice outcome. Capturing is important to generate consistent dynamics, for example in the case of endogenous platform entrance.

Here I describe the formulation of this. I also include how to incorporate familiarity with a platform in this formulation (a consumer's familiarity with a platform, through processes of social exposure is discussed in Essay 1). The basis is a nested logit formulation. The share that drivers of platform i replacing their vehicle allocate to platform j, σ_{ij}^{d}, involves a nested decision process (Ben-Akiva 1973). A share of the discarded vehicles from platform i is replaced by j, σ_{ij}^{r}, conditional upon an earlier choice of replacing the vehicle at all σ_{i}^{r} :

$$\sigma_{ij}^{d} = \sigma_{i}^{r}\sigma_{ij}^{r} \qquad \text{(A.6)}$$

For a replacement decision, all vehicle platforms form a "nest" whose utility is compared to an unspecified alternative:

$$\sigma_{i}^{r} = \frac{u_{i}^{ve}}{u_{i}^{ve} + u^{oe}} \qquad \text{(A.7)}$$

An increase in the variety of models does not necessarily increase aggregate utility of "the vehicle nest" proportionally. That is, utility of the nest depends on the correlation (or substitutability) of preference across a range of products in the choice outcome (not necessarily in direct relation to the different platforms). To capture this we introduce a scaled parameter $\mu \equiv 1/(1-\chi)$ with χ, $0 \leq \chi \leq 1$, being the correlation parameter for consumer choice with respect to the platforms within the nests (further intuition is provided following equation (A.9), the nest utility is:

$$u_i^{ve} = \left[\sum_j u_{ij}^{ve} \right]^{1/\mu} \tag{A.8}$$

While the effective utilities for the various platforms u_{ij}^{ve} are the perceived utility with each platform u_{ij}^{v} adjusted for their correlation, multiplied with familiarity F_{ij} of the population with the various choices:

$$u_{ij}^{ve} = F_{ij} \left(u_{ij}^{v} \right)^{\mu} \tag{A.9}$$

Utility, u_{ij}^{v}, depends on vehicle attributes for platform j, as perceived by driver i. For an aggregate population average familiarity F_{ij} varies over the interval [0, 1].

The correlation parameter can now be interpreted as follows, with $\chi \to 0$, the case of no correlation, platforms are perceived by the consumers as fully distinct and overall "vehicle utility" rises linearly with number of platforms. For $\chi \to 1$, full correlation, vehicle platforms are perceived to be identical, and the perceived utility equals that of the most superior. For instance, in the case of n identical products, with only different prices, all demand goes to the cheapest product. Lowering price for a more expensive product, while still being above the most affordable, has no effect on market shares, nor on the overall demand. Neither extreme is behaviorally appropriate. Further, dynamically, χ controls a potentially very strong feedback, between demand and the introduction of new platforms (with maximum strength at the default, no correlation, case $\chi = 1$). In addition, χ is arguably a function of the technological heterogeneity of products on the market. That is however not the point we want to make here. In this paper we assume that

the consumer only cares about performance, not so much about distinctiveness between

them. Thus, in this model, χ is constant between 0 and 1.

The formulation of equation (A.6)-(A.9) is equivalent to the compact general nested

formulations (Ben-Akiva and Lerman (1985), Ben-Akiva 1973), frequently used in

transportation decision making models (e.g. Brownstone and Small (1989)), industrial

organization literatures (e.g. Anderson and Palma (1992) regarding multi product firms,

Berry et al. (1995) regarding the automobile industry). We can write σ_{ij}^{d} as:

$$\sigma_{ij}^{d} = \sigma_{ij}^{r}\sigma_{i}^{r} = \frac{u_i^{ve}}{u^{oe}+u_i^{ve}}\frac{u_{ij}^{ve}}{u_i^{ve}} = \frac{F_{ij}\left(u_{ij}^{ve}\right)^{\rho}}{\left[\sum_j F_{ij}\left(u_{ij}^{ve}\right)^{\rho}\right]^{1/\rho}+u^{oe}}$$

In the nested logit model, $1 \le \mu \le \infty$ is the scale parameter for the MNL associated with

choice between alternatives within the nest (in our case the vehicles). For $\mu \to 1,$

corresponding to $\chi \to 0,$ the function converges to a standard MNL, while for $\mu \to \infty$,or

$\chi \to 1$, the model is a perfect nest.

The formulation of platform utility is also consistent with general Constant Elasticity of

Substitution Production Function (CES-PF) (McFadden 1963), the functional form used

elsewhere in the paper:

$$K_i = \left[\sum_j \kappa_{ij}\left(K_{ij}\right)^{\rho}\right]^{1/\rho}$$

In this expression $\rho = 1 - \varsigma/\varsigma$, with ς the elasticity of substitution between products.

In this formulation $\mu = -\rho$. Note that the range specified for vehicle choices implies an elasticity between $-\infty < \varsigma < 0$, while elasticities on the supply side specify the positive range (complementary goods imply ς between 0 and 1, and substitutes $1 \leq \varsigma < \infty$).

In the model χ is set to 0.5 throughout.

f) External scale effects

The general formulation of the aggregate scale effect, as a function of the installed base share $\sigma_j^v = V_j / V^T$, is:

$$a_{j3} \equiv \varepsilon_j^s = f\left(\sigma_j^v\right); f' \geq 0; f\left(\infty\right) = 1; f\left(\sigma_{ref}^v\right) = \varepsilon_{ref}^s$$

We use the three parameter logistic curve to generate the patterns of Figure 7. To do so we control the value of at the inflection point, and set the fixed the rest to the selected slope at that point. This results in:

$$\varepsilon_j^s \equiv \min^s + \frac{\left(1 - \min^s\right)}{1 + \left(f_j^s - \alpha^s\right) / \alpha^s \exp\left[-\beta^s v^s \left(\frac{\sigma_j^v - \sigma_{ref}^v}{\sigma_{ref}^v}\right)\right]} \tag{10}$$

For this curve, which is we set the scale factor, as a measure for the scale effect, and fix the slope at the reference point. v^s is a scaling parameter to equalize the slope at the reference point and equals $v^s = \left(f_j^s / \alpha^s\right) \min^s$, and $\min^s = \left(f_j^s \left(1 - \alpha^s\right) + \alpha^s\right) / f_j^{s2}$, α^s is an offset parameter to determine the minimum, that is, where σ_j^v. At full penetration all scale effects work maximally to its advantage. For the default settings I use an installed

278

base 5% of the fleet, $\sigma_{ref}^v = 0.05$ and sensitivity parameter $\beta^s = 1$, which measures the slope at the reference installed base share; $\alpha^s = 0.8$. See also Appendix 2f of Essay 2 for a discussion on generating logistic curves.

3 Stipulations

a) Generalization to multiple attributes and modules

This section discusses the more general structure of which this model is a special case. In particular detailed vehicle attributes, . In the analysis these structures where all switched off, but for more detailed analysis and insights they, or parts of them, can be switched on.

Overview of expanded chain

The model as specified in the paper provides all the structure necessary to generate the key insights derived in this paper. However, in order to consistently simulate a wide range of behaviors, and more intuitive patterns we include additional structure. Figure A1 shows the chain of decisions and technological chain, between resource allocation for R&D and consumer choice regarding the technologies.

segment	variable	indices	operation
Consumer choice	σ_j Market share	Platform j Attribute l	Nested Logit Model
	u_j Utility	Platform j	
	a_{jl} Attribute	Platform j Attribute l	Mapping of module technology on attributes: {m,x} → {l}
Technological Performance	θ_{jmx} Relative technology	Platform j Function x	Mapping of activity type on Performance/cost: {w} → {x} Normalization of technology
	T^e_{jmw} Effective technology	Platform j Module m Activity type w	Complementarity across Activities {w'} → {w}
	T_{jmw} Technology	Platform j Module m Activity type w	Diminishing returns
Knowledge Accumulation	K^e_{jmw} Effective knowledge	Platform j Module m Activity type w	CES function
	K_{ijmw} Knowledge Input	Source platform i Target platform j Module m Activity type w	Learning-by-doing, R&D, and spillovers;
Resource Allocation	R_{ijmw} Resources	Source platform i Target platform j Module m Activity type w	Improve marginal return on effort

Figure A1 Diagrammed representation of chain of decisions and technological change for expanded model.

Consumer choice contains a nested logit model capturing the notion of substitutability of choice across platforms. Figure A2 shows the indices used in the expanded model. They include, from bottom to top: platforms *j*; modules *m*, the most important level at which technological change and spillovers occur; activity type *w*, that specifies whether technology advances derive from product- or process improvements; function *x* that allows to differentiate relevance of product and process improvements to either cost or performance; and attribute *l* that captures dimensions of merit from the perspective of a consumer.

attribute	$l \in \{1,...,L\}$	e.g. Price, Power, Operating Cost, Reliability.
function	$x \in \{\text{cost,performance}\}$	
activity type	$w \in \{\text{process,product}\}$	
module	$m \in \{1,...,M\}$	e.g. Body, Brakesystem, Powertrain, Wiring.
platform	$i, j \in \{1,...,N\}$	e.g. ICE, HFCV, HEV, EV, CNG.

Figure A2 Indices used in the expanded model.

Attributes derive their state from the technology performance at the module level (Figure A1), while the current unit costs at each module determine the price attribute of the vehicle. Further down the chain, another distinction is that the effective technology captures notion of complementarity between activities: advances at the product level will make process advances obsolete. We will now describe these adjustments.

Multiple dimensionality of choice attributes

The perceived utility of a platform captures the aggregate of perceived attractiveness of a platform across various dimensions of merit, for which we define the attribute set that includes price, vehicle range, power etc... With a_{ijl} being the state of the lth attribute of platform j as perceived by drivers of platform i, its perceived utility from that platform equals:

$$u_{ij} = u^* \exp\left[\sum_l \beta_l \left(a_{ijl} / a_l^* - 1 \right) \right] \tag{A.11}$$

where β_l is the sensitivity of utility to a change in the attribute. Struben (2006a) discusses the various channels through which consumers learn about and experience performance.

The performance of a user-attribute is established through technological advances with producers of platform j. Further, a technology is multidimensional. Vehicles comprise of modules that include for instance the powertrain, suspension, controls and the body. This means for different attributes, different modules m are determinants of the performance. We follow a two stage production function (McFadden 1963) to specify the attribute performance in relation to the knowledge produced at module level. We first describe in general formulations how we capture the dependence of attribute performance on knowledge, and follow that up with an example.

An attribute's performance comprises a fixed component, and one that depends on the current state of the technology.

$$a_{jl} = a_{jl}^0 + a_{jl}^v \tag{A.12}$$

a_{jl}^0 is the initial attribute independent of module level improvements through R&D, and other endogenous processes. This is the performance level that is attained at start-up depends for instance on the state of the complementary technologies, and can therefore differ per platform.

We assume that substitutability between modules' technology is maintained, independent of the rate of progress. We can thus use the standard constant elasticity of substitution (CES) function (Arrow et al. 1961), with multi inputs (McFadden 1963).[29] The CES approach to the multi input substitution problem is convenient, leads to simple estimation

---

[29] For multi inputs, the exact elasticity of substitution between two inputs is not easily to categorize, and several definitions exist. This is not a problem for our purpose.

methods and is widely used (see also Solow 1967). The behavioral characteristics are discussed further in the analysis section.

Then, performance of attribute l is a function of the relative technology of module m, and function x, θ_{jmx}. The index x represents either performance, or cost. How a technology θ_{jmx} impacts an attribute, depends on its factor contributions to, or relevance for, attribute l, κ_{jmlx}. Then:

$$a_{jl}^v = a_{jl}^1 \left(\sum_{m,x} \kappa_{jmxl} \left(\theta_{jmx} \right)^{-\rho_j^a} \right)^{-1/\rho_j^a} \tag{A.13}$$

where $\rho_j^a = (1 - \zeta_j^a)/\zeta_j^a$ is the substitution parameter and ζ_j^a, the elasticity of substitution between the effectiveness of the technology between different modules for attribute j. The attribute associated with vehicle price map strictly on the cost index of x, while all others strictly map on performance. As technologies are substitutes, the range of the elasticity is confined to $1 < \varsigma_j^a < \infty$. Further, a_{jl}^1 is the scale, or efficiency parameter for the attribute and is scaled such that $\sum_{m,x} \kappa_{jmlx} = 1$. The distribution parameters κ_{jmlx} define the relative importance of each module m to attribute j. Then, by construction, when all module technologies' effectiveness equal unity, the fixed share in the total attribute state equals $a_{jl}^0 / \left(a_{jl}^0 + a_{jl}^1 \right)$, which provides an interpretation for a_{jl}^0. Finally, we have constrained the aggregate performance to constant returns to scale with respect to the total effective technology. That is, the function has an implicit degree of homogeneity parameter that is set to 1.

Finally, the performance and cost are found from the technology as produced through process and product improvement. Index w represents different activities that allow improving a technology, such as product innovation, process improvement. Here we define strictly $w=\{$product innovation, process innovation$\}$:

$$\theta_{jmx} = \sum_w \alpha_{jmwx} \theta_{jmw}; \sum_x \alpha_{jmwx} = 1 \qquad (A.14)$$

Where α_{jmwx} represents the share of the improvements in the technology of module m, through activity w (product or process), contributing to function x (price or performance).

Illustration: from knowledge, to cost, to vehicle price

Figure A1 illustrated the how chain of decisions and technological relations connect the state of an attribute to knowledge at the module/activity type level K_{jmw}. In the exposition we just went through, we saw that this chain can be compactly captured formally, through a two-stage CES PF. The following example serves to illustrate this more clearly. The chain of connections comprises, first, the effect of the various modules on the attribute state and, second, the effect of the various sources of knowledge to improving the technology at the modular level. I use the vehicle price attribute as an example, with $a_{j1} \equiv p_j$. I select vehicle price deliberately. We have an intuition how price is connected to cost that improves especially through learning (and scale economies). By showing that also this set of relations fits in this structure, I hope to improve our intuition of it.

Following the CES expression, using index 1 of x for cost, we must get to:

284

$$p_j = p_j^1 \left(\sum_{m1} \kappa_{jm1} \theta_{jm1}^{e\,-\rho_{jm}} \right)^{-1/\rho_{jm}} + p_j^0; \tag{A.15}$$

The interpretation of the variables must become clear in the end. Further, from the bottom-up, unit costs are the sum of the system level unit cost c_j^{u0} and the cost incurred for producing modules, c_{jm}^u :

$$c_j^u = \left(\sum_m c_{jm}^u \right) + c_j^{u0} \tag{A.16}$$

Unit costs are prone to learning by doing and/or scale economy effects, ε_{jm}^c , thus: [30]

$$c_{jm}^u = c_{jm}^{u0} \varepsilon_{jm}^c \tag{A.17}$$

But only a fraction f_{jm}^s is variable with respect to scale, ε_{jm}^s , and f_{jm}^s is subject to experience through learning-by-doing ε_{jm}^e :

$$\varepsilon_{jlm}^c = \left(1 - f_{jm}^s \right)\left(1 - f_{jm}^e \right) + f_{jm}^s \varepsilon_{jm}^s + f_{jm}^e \varepsilon_{jm}^e \tag{A.18}$$

For simplification we ignore from here any internal scale economies. With $p_j = \left(1 + m_j \right) c_j$ and the derivation of the unit cost above, we now rewrite the price and derive, and interpret, the two components in (A.15):

$$p_j^1 = \left(1 + m_j \right) c_{0j}^{uve} = c_{0j}^{uve} \sum_m f_{jm}^e c_{0jm}; \sigma_{jm} = \frac{f_{jm}^e c_{0jm}}{c_{0j}^{uve}};$$

$$p_j^0 = \left(1 + m_j \right) c_{0j}^e; c_{0j}^e = \sum_m \left(1 - f_{jm}^e \right) c_{0jm} + c_{0j}$$

With c_{0j}^{uve} being the part of the total unit cost subject to learning, when inputs for learning are equal to their normal levels. We see that the interpretation of the fixed component corresponds with the one provided in equation (A.12) and (A.13). p_j^1, the efficiency or

---

[30] Implicitly, for purpose of analytical clarity, we assume here that system level costs are not subject to learning/innovation improvement. This can easily be relaxed.

scale parameter, is the variable price, when all are equal to their normal values, and

$p_j^0 / (p_j^0 + p_j^1)$ is the fixed share when all are equal to their normal values.

Further, $\theta_{jmx} = \sum_w \alpha_{jmwx} f\left(\varepsilon_{jmw}^e\right)$; with typically α_{jm21} large (most improvements from

process improvement lead to cost improvements) and $\alpha_{jm21} > \alpha_{jm11}$ (most, but certainly

not all, cost improvements come from learning by doing).

Also from bottom-up, the learning-by-doing relation also gives us:

$\varepsilon_{jmw}^e = \sigma_{jmw} \left(\dfrac{K_{jmw}}{K_0}\right)^{\lambda_{jmw}^e}$ and as cost decrease at diminishing rate with embedded

knowledge, thus, $0 < \lambda_{jmw}^e < 1$. Thus in order for the attribute vehicle price expression to

hold, $\theta_{jmw} = \varepsilon_{jmw}^e{}^{-1}$, or $-1 < \lambda_{jmw} < 0$ in equation (13) and substitution parameter $\rho = -1$.

This corresponds with the elasticity of substitution being infinite. This is intuitive: we are

indifferent to the sources of cost reductions. Further, in the case of vehicle price, an input

factor share must be interpreted as the relative contribution of each module in terms of

variable unit cost when technology is equal to normal values.

We end this exposition with the following question (and examination of it): Are

technology level returns to scale are independent of the number of modules? That is, how

can we avoid that dynamics are affected when we aggregate or disaggregate?

The dynamics are not affected. This follows directly from equation (A.13). First, a

hypothetical case: splitting the drive train into two modules into n parts that are in fact

independent, implies that the distribution parameter for each sub-module is smaller. In

the case of two equal sub modules the distribution parameters are 50% of the distribution

parameter of the whole module $\kappa_{jm'wl} = 0.5\kappa_{jmwl}$. The CES function is indifferent to this

reconfiguration. However, in this case, the state of the technology for each must now

increase at the same rate as the whole, with half the resources required. This implies that

the reference resources for the sub modules are equal to half of that of the full module:

$R_{m'}^0 = 0.5R_m^0$. The same explanation holds when we generalize to a larger number of sub

modules that have varying contribution. This also implies that, if we are interested in

more basic dynamics, we can aggregate multiple modules into one, following the same

procedure, without impacting the fundamental dynamics.

Effective technology

The complementarity between activities is captured in the net progress rate of effective

technology T_{jw}^e that depends on the progress rate of the total technology of all

activities w'. For instance, complementing a radically new body will make previous

process technology obsolete. Capturing this is important when we examine the interaction

between novel and mature technologies. For instance, mature platforms can be expected

to be conservative with innovating.

Growth of the effective technology follows that of the cumulative technology, but is

adjusted for the obsolescence rate that results from other activities. With Γ_{jW}^e being the

vector of the growth rate of the effective technology, with its wth element defined as, and Γ_{jW} being a similar growth rate vector for the technology T_{jw} : [31]

$$\frac{dT_{jw}^e}{dt} = \sum_{w'} \varepsilon_{jww'}^t g_{jw'} T_{jw'}$$

The growth rate $g_{jw'} \equiv \left(dT_{jw'}/dt\right)/T_{jw'}$. By definition, the diagonal terms are unitary.

Representing the usually negative effect of improvements in w' on activity w, the lower

triangular terms are bounded by $-1 \leq \varepsilon_{jww'}^t \leq 0$, while the upper triangular terms are zero.

Thus, with product and process innovation:

$$\mathbf{E}_j^t = \begin{bmatrix} 1 & 0 \\ \varepsilon_{j21}^t & 1 \end{bmatrix} \tag{A.19}$$

In the analysis of the Essay, I ignore any overlap and thus, $\varepsilon_{j21}^t = 0$.

Spillover potential

The process knowledge related factor shares are by construction equal to unity for

internal knowledge accumulation. However, for spillovers the factor contribution

depends on the amount of technology of j that is currently embodied in the technology i,

thus, for $i \neq j$:

$$\frac{d\theta_{ijw}^\kappa}{dt} = \sum \varepsilon_{jww'}^t f\left(K_{ijw'}/K_{w'}^0\right) g_{iw} \theta_{jw}$$
$$f(0) = 0, f(1) = \kappa_{w'}^0; f \leq 1; f' > 0 \tag{A.20}$$
$$\kappa_{ijw} = \sum_{w' \neq w} \theta_{ijw'}^\kappa / \theta_{jw}^e + \left(1 - \sum_{w' \neq w} \varepsilon_{jww'}^t\right) \kappa_{ijw}^0$$

---

[31] This could also be represented by a co-flow structure, but in this case this construction seems more intuitive.

In the analysis, the factor overlap is equal to zero, therefore the spillover potential for process improvement is independent of the product technology, and $\kappa_{ijw}^0 = \kappa_{ij}^0 \ \forall w$.

In the analysis of the Essay, I ignore any overlap and thus, $\varepsilon_{j21}^t = 0$.

b) The locus of diminishing returns

The process of accumulation of knowledge, improving technology, performance and increasing attractiveness is subject to increasing returns. In the model we limited diminishing returns to technology in an increase of total knowledge, while the attribute state has constant returns to technological change, and total knowledge has constant returns to knowledge accumulation. However, in real life it is hard to distinguish between them. Here we show that the return to scale parameter can be transferred among these three, without affecting the main dynamics. First, note that returns to scale is maintained across a constant returns function. For instance between attribute and technology, ignoring the function and activity indices, we have:

$$a_l = a^0 \left(\sum_{m=1}^{M} \kappa_m \left(\left[\frac{T_m}{T^0} \right]^{\eta} \right)^{-\rho} \right)^{-1/\rho} \approx a^0 \kappa_m^{-(1-\eta)/\rho} \left(\sum_{m=1}^{M} \kappa_m \left(\frac{T_m}{T^0} \right)^{-\rho} \right)^{-\eta/\rho} \equiv a_l^e.$$

Subsequently, we can say:

$$a_0^{\ e} = \left(a_0 \ \kappa_m^{-1/\rho} \right)^{1-\eta} \left(\frac{a_l^*}{a_0} \right)^{\eta},$$

289

where $a_i^* = a^0 \left(\sum_{m=1}^{M} \kappa_m \left(\dfrac{T_m}{T^0} \right)^{-\rho} \right)^{-1/\rho}$ is the constant returns equivalent of technology. Thus,

by approximation we can shift the constant returns parameter, only having to correct by a constant. The approximation is exact, when κ is identical for all m.

More important, diminishing returns to knowledge accumulation for source i

implies: $dK_i/dt = \varepsilon_i^k \gamma_i ; \varepsilon_i^k = \left(K_i / K^0 \right)^{\eta^k}$, with η^k . We can rewrite this, through an

intermediate variable $K_i' \equiv \left(K_i / K^0 \right)^{1/\nu}$, such that $K_i = K^0 \left(\dfrac{K_i'}{K^0} \right)^{\nu}$ and

$\dfrac{dK_i'}{dt} = \gamma_i \left(\dfrac{R_i}{R^0} \right)^{\eta^R} K^0 ; \nu = 1/1 - \eta^K$. Further, total knowledge accumulation is also a CES

function of all the various sources. Thus, we can convert diminishing returns to

knowledge accumulation for each individual to diminishing returns at the level of the

technology, by letting $\eta^K = \left(\eta^T - 1 \right)/\eta^T$ (thus, for diminishing returns, η^K is negative,

which makes sense: the accumulation of knowledge decreases with an increase of

knowledge).

This exposition can further be expanded to include consumer choice sensitivity to a

change in the attribute, and the overall elasticity of demand (including market saturation

effects). Doing this will give insights under what condition the technological progress as

a whole (temporarily) exhibits increasing, constant, or diminishing returns to scale.

c) Optimal Resource Allocation

Proposition 1 Resource allocation decisions are asymptotically optimal within the planning horizon, holding the environment constant, including the technology of other platforms.

Proof: Appendix 2A showed the decision structure and also that the perceived return on effort, $\varsigma_{x_d}^{rd\sim}$, was plausibly set equal to the marginal return on effort. Further, in equilibrium, indicated shares are equal to actual. Then (Appendix 2a), $R_{x_d} = \sigma_{x_{d-1},x_d}^{rd} R_{x_{d-1}}$, and $\varsigma_{x_d}^{rd\sim} = f\left(dP_{x_d}^{\tau_d^p} / dR_{x_d} \right)$ yield:

$$\sigma_{x_d,x_{d+1}}^{rd} = f\left(\varsigma_{x_d}^{rd\sim} / \varsigma_{ref,\{x\}_d}^{rd\sim} \right) R_{x_d}^{rd} \Big/ \sum_{\{x\}_{d+1}} R_{x_{d+1}}^{rd\circ} = f\left(dP_{x_{d+1}}^{\tau_d^p} / dR_{x_d} / \varsigma_{ref,x_d}^{rd\sim} \right) \sigma_{x_d,x_{d+1}}^{rd} R_{\{x\}_{d+1}}^{rd} \Big/ \sum_{\{x\}_{d+1}} R_{x_{d+1}}^{rd\circ}$$

And with f smooth and non-decreasing, we get:

$$\frac{\sigma_{x_d,x_{d+1}}^{rd}}{\sigma_{x'_d,x_{d+1}}^{rd}} = \frac{f\left(dP_{x_{d+1}}^{\tau_d^p} / dR_{x_d} / \varsigma_{\{x\}_d}^{rd\sim} \right) \sigma_{x_d,x_{d+1}}^{rd} R_{\{x\}_{d+1}}^{rd}}{f\left(dP_{x_{d+1}}^{\tau_d^p} / dR_{x_d} / \varsigma_{\{x\}_d}^{rd\sim} \right) \sigma_{x'_d,x_{d+1}}^{rd} R_{\{x\}_{d+1}}^{rd}} \Rightarrow \left(dP_{x_d}^{\tau_d^p} / dR_{x_d} \right) = \left(dP_{x'_d}^{\tau_d^p} / dR_{x'_d} \right)$$

Thus, in equilibrium, the marginal returns on effort of all allocations are identical. Since the costs of resources are identical across resources, this implies optimal allocation of resources.

A more formal derivation

Here we derive more formally that the preceding statement is valid. Assume a production function that improves performance indicator P, with various forms of Inputs K_i, with cost $C_i = c_i R_i$. Then, maximizing returns yields:

$$P\left(\overrightarrow{K(R)}\right) - \vec{C} \tag{A.21}$$

This implies that allocation of resources is optimal if:[32]

$$\frac{dP}{dR_i} = \frac{c_i}{c_j} \frac{dP}{dR_j} \forall i, j \tag{A.22}$$

Where $dP/dR_i = (dP/dK_i)(dK_i/dR_i)$, if the marginal productivity in resources is

multiplicatively separable in those resources, $dP/dR_i = p_i f_i(R_i)$, then the optimal

resource allocation equals:

$$\frac{dP/dR_i}{dP/dR_j} = \frac{p_i f_i(R_i)}{p_j f_j(R_j)} = \frac{c_i}{c_j} \Rightarrow R_j^* = f_j^{-1}\left(\frac{p_i c_j}{p_j c_i} f_i(R_i^*)\right) \tag{A.23}$$

And

$$\sigma_i = \frac{R_i}{\sum_j R_j} \Rightarrow \sigma_i^* = \frac{1}{1 + \frac{1}{R_i^*} \sum_{j \neq i} f_j^{-1}\left(\frac{p_i c_j}{p_j c_i} f_i(R_i^*)\right)} \tag{A.24}$$

When the functional forms are identical this simplifies to:

$$\sigma_i^* = \frac{1}{\sum_j f^{-1}\left(\frac{p_j}{c_j} \frac{c_i}{p_i}\right)} = \frac{f^{-1}(c_i/p_i)}{\sum_j f^{-1}(c_j/p_j)} \tag{A.25}$$

Note that this implies, as expected, when the production function is linear in R, it is

optimal to allocate all resources to the one with the highest marginal productivity.

Further, in equilibrium, the desired share equals the desired resources, $\sigma_i^d = \sigma_i^*$, with

---

[32] We take the Paretian profit-maximization hypothesis in which only prices are fixed and conditional on diminishing marginal productivities. This is a not unlimitedly strong but general assumption.

$$\sigma_i^d = \frac{R_i^d}{\sum_j R_j^d} \qquad \text{(A.26)}$$

and $R_i^d = R_i\, f(x_i)$.

In equilibrium, $\sum_j R_j^d = f(x_i) \sum_j R_j \Rightarrow f(x_i) = f(x_j); \forall i, j$

Thus, when $x_i \propto$ the marginal return on effort, dP/dR_i, we reach the equilibrium where shares are optimal.

As an example, assume the following multi input CES production function, as is specified for knowledge accumulation:

$$P = P^0 \left(\sum_{i=1}^{I} \kappa_i \left(\frac{K_i}{K^0} \right)^{-\rho} \right)^{-\eta^S/\rho} \qquad \text{(A.27)}$$

In this expression $\rho = 1 - \varsigma/\varsigma$, with ς the elasticity of substitution between products, and $p^0 \equiv P^0/K^0$ is the price of P. Then,

$$\frac{dP}{dK_i} = \eta^S \frac{\kappa_i \left(\dfrac{K_i}{K^0} \right)^{-\rho}}{\sum\limits_{i=1}^{I} \kappa_i \left(\dfrac{K_i}{K^0} \right)^{-\rho}} \frac{P}{K_i}; \varepsilon_{P/K_i} = \eta^S \frac{\kappa_i \left(\dfrac{K_i}{K^0} \right)^{-\rho}}{\sum\limits_{i=1}^{I} \kappa_i \left(\dfrac{K_i}{K^0} \right)^{-\rho}} \qquad \text{(A.28)}$$

Assume now the following relationship $R_i = K_i$, (e.g. forms of labor productive capital vs output). The optimal share equals:

$$\frac{dP}{dK_i} \Big/ \frac{dP}{dK_j} = \frac{\kappa_i}{\kappa_j} \left(\frac{K_i}{K_j} \right)^{-(\rho+1)} = \frac{c_i}{c_j} \Rightarrow \frac{K_i}{K_j} = \left(\frac{c_j \kappa_i}{c_i \kappa_j} \right)^{1/(\rho+1)} \Rightarrow \sigma_i^* = \frac{\left(\kappa_i/c_i \right)^{1/(\rho+1)}}{\sum_j \left(\kappa_j/c_j \right)^{1/(\rho+1)}} \quad \text{(A.29)}$$

Which is identical to equation (A.25).

Again, when the PF grows linearly with resources $(\rho = -1)$, the marginal productivity between allocation to i and the others is a fixed ratio, say $(1+\alpha)$ and in equilibrium, all shares will go to the one with largest marginal productivity:

$$\sigma_i^* = \frac{(1+\alpha)\sigma_i}{(1+\alpha)\sigma_i + \sigma_{-i}} \Rightarrow \sigma_i^{eq} = 1$$

4 Boundary constraints considered

These involved boundary constraints to which the model is tested against. I will discuss briefly what role they play in the analysis and where they influenced dynamics, sometimes in a significant way.

a) Capacity adjustment, backlogs and churn

Capacity adjustment assures robust dynamics during strong demand growth. Further, capacity adjustment is another balancing constraint on growth, relevant to many technologies. Japanese automakers face significant delays in meeting demands for their hybrids. Further, significant backlogs can have more side-effects involves churn and suppression of potential demand, as those who consider such a platform, will now abstain form selecting it. As the social behavior regarding this is hard to assess, a mismatch of supply and demand can severely hurt transition dynamics.

To simplify analysis, there are no adjustment costs, but it does take adjustment time τ^c to reach desired capacity C_j^*. Desired capacity equals current, adjusted for signals from demand:

$$C_j^* = \varepsilon_j^c C_j \qquad (A.30)$$

Where and ε_j^c is the effect of utilization on capacity adjustments:

$$\varepsilon_{js}^c = f^c \left(\left[\tau_j^d - \tau^{d*} \right] / \tau^{d*} \right); f^c{}' > 0; f^c(0) = 1; \qquad (A.31)$$

τ^{d*} and τ_j^d are the desired and current delivery time for platform j. The current delivery time is given by Little's law by backlog and capacity:

$$\tau_j^d = B_j / C_j \qquad (A.32)$$

The backlog structure is modeled explicitly to retain dynamic consistency when demand and supply are in significant imbalance, but also to allow for churning dynamics. Backlogs grow with initial purchase decisions s_j^{k*} and churn from others b_j^{ci}, and decline with actual sales, at delivery, s_j^k, and total churn to other platforms b_j^{co}:

$$\frac{dB_j}{dt} = s_j^{k*} - s_j^k + b_j^{ci} - b_j^{co} \qquad (A.33)$$

The indicated sales, under capacity constraint s_j^{k*} is equal to the sales rate discussed in the paper. However, in perceived utility for each platform is adjusted, to include an extra attribute that captures the effect of perceived wait time on attractiveness to buy, $a_j^b = \tau_j^w$ and $a^{b*} = \tau_{ref}^w$. Actual sales under capacity constraints result from deliveries to those in the backlogs at delivery rate τ_j^d, $s_j^k = B / \tau_j^d$. Further, those who are in the backlog churn

when their experienced wait time τ_j^w is much larger than their expected wait time, when

they decided to purchase τ_j^{w*}:

$$b_j^{co} = \lambda_{ref}^c f\left(\tau_j^w / \tau_j^{w*}\right); f' \geq 0; f(0) = 0; f(1) = 1 \tag{A.34}$$

Expected weight time of those who are in the backlog, τ_j^{w*} and their experienced wait

time, τ_j^w, are traced through a co-flow structure (Sterman 2000; see also Appendix ** of

Essay 1). Finally, the perceived backlog, also feeds into the initial purchase decision, thus

backlog is another attribute at purchase, with a negative elasticity.

b) Endogenous elasticity of substitution

It depends on the state of the technology how newly acquired knowledge contributes to

the total technology. In the early stages a technology trajectory is malleable. Alternative

solutions can easily be incorporated, while substitutability is low. As technology

accumulates, standards emerge, flexibility decreases, which means that substitutability

increases. Including this formulations allows exploring the fundamental dynamics

consistently over a rich set of relevant environments. For instance, the competition

between that include. For instance, incorporating this effect amplifies the fundamental

dynamics.

The functional form of Equation (5) in the Essay connects to this through the substitution

parameter: a low substitution parameter, say -1, implies that it is optimal to allocate all

resources to those knowledge sources with the highest factor shares. A substitution

parameter of 0, implies that it is optimal to build up knowledge proportional to the factor shares, and the effective knowledge is very sensitive to an increase in knowledge sources. While the substitution parameter is intimately linked with the elasticity parameter, via $\rho_{jw}^{k} = \left(1 - \varsigma_{jw}\right) / \varsigma_{jw}$, the elasticity parameter does not necessarily yield the elasticity of substitution between the knowledge between platform, when there are more than two platforms. However, following the reasoning above, the decrease of the substitution parameter is a good representation of platform maturity.

Capturing this formally through the elasticity parameter of knowledge exchange between that of platform j and that i, $\varsigma_{jw} = \varepsilon_{jw}^{\varsigma} \varsigma_{w}^{\min}$, we get.

$$\varepsilon_{jw}^{\varsigma} = f\left(\theta_{jw}\right); f\left(0\right) > 0; f\left(1\right) = 1; f' \geq 0 \qquad \text{(A.35)}$$

We imposed these conditions for the logical arguments of maturation of the technology. However, under these conditions of a non-decreasing elasticity parameter ς_{jw}, we can arrive at the same intuition more formally:

Proposition 2: For novel technologies, the effective technology exhibits increasing returns in the number of platforms. However in the long-run equilibrium, technologies exhibit neutral returns to the number platforms, even under infinite market and constant entrance probabilities.

Intuition: spillovers are a central mechanism for growth of knowledge, especially in the early stages of a product lifecycle. For a novel technology, knowledge is incomplete and

thus has complementarities with that within other technologies. As a technology matures, knowledge is increasingly complex, providing more incentive to exploit the most productive aspects.

Proof: Assume the a fortiori case in which spillover rate among platforms is infinite and free, so all resources are allocated to internal knowledge development, and we don't have to impose optimal allocation of resources.

Starting with $K_{jm}^{n} = K_{0jm} \left(\sum_{i=1}^{M} \kappa_{ijm} \left(k_{ijm} \right)^{-\rho_{ijm}^{M}} \right)^{-1/\rho_{jm}^{M}}$. Then, introducing a new platform *n+1*,

with instantaneous spillover to platform *j* implies:

$$K_{jm}^{n+1} = K_{0jm}^{\{n\}} \left(\sum_{i=1}^{n} \kappa_{ijm} \left(k_{ijm} \right)^{-\rho_{ijm}^{M}} + \kappa_{n+1,jm} \left(k_{n+1,jm} \right)^{-\rho_{ijm}^{M}} \right)^{-1/\rho_{jm}^{M}} \quad \text{(A.36)}$$

The main insights are derived when we rewrite this, so that the first order effect gets captured in the scale parameter K_{0jm} . Hereto we defining a set of distribution

parameters $\kappa'_{i,jm}$, such that $\sum_{i=1}^{n+1} \kappa'_{ijm} = \sum_{i=1}^{n} \kappa_{ijm}$. Then we can rewrite (A.36) to

$K_{jm} = K_{0jm}^{\{n+1\}} \left(\sum_{i=1}^{n+1} \kappa_{i,jm} ' \left(k_{ijm} \right)^{-\rho_{jm}^{M}} \right)^{-1/\rho_{jm}^{M}}$. With:

$$\frac{K_{0jm}^{\{n+1\}}}{K_{0jm}^{\{n\}}} = \left(\sum_{i=1}^{n+1} \kappa'_{i,jm} \middle/ \sum_{i=1}^{n} \kappa_{ijm} \right)^{-1/\rho_{jm}^{M}} \quad \text{(A.37)}$$

For immature technologies and very small number of technologies, or, elasticity of substation close to 1, $\rho_{jm}^{M} \to 0$, and *n* small, the increasing returns are very large. However, once the entrants increase, *n* large, and the technology matures elasticity $\rho_{jm}^{M} \to -1$, the

effect diminishes fully. Also, effect of entrance on knowledge growth raises with the distribution parameters. New entrants have less knowledge, so (A.37) bounds the direct knowledge. Other second order effects are also balancing, for instance, the market share goes down, reducing revenues and profits, and resource allocation for all the incumbent platforms. Generally the overlap (or quality) becomes cannot be maintained, corresponding with on average smaller spillover factors κ, further reducing the returns to the number of entrants, for incumbents. This is in particular the case for more mature technologies, for which marginal benefit of an increase in knowledge increases linearly with the input factor.

Thus, spillover is a central mechanism for growth of knowledge, especially in the early stages of a product lifecycle. For novel technologies, knowledge is incomplete and thus has complementarities with others. As technology matures, knowledge is increasingly substitutable, providing more incentive to exploit the most productive aspects. As a corollary to this, platforms with more mature knowledge fixate on fewer candidates for sources of spillover. Another interpretation of this is that with reduction of uncertainty the knowledge allocation is more accurate – closer to the concept of Jovanovic (1982) that companies only borrow from the leader.

c) Product experience

Product improvement productivity increases with effective experience in R&D. This captures an additional feedback loop that will extend the time for new technologies to

catch up. This is included by making the productivity of product innovations endogenous, tracing experience E_{j1} that accumulates historic resources allocation:

$$\varepsilon_{j1}^{r} = \left(E_{j1}/E_{0}\right)^{\eta^{e}}$$
$$\frac{dE_{j1}}{dt} = R_{j1}$$

(A.38)

where E_{0} is the reference experience at which relative productivity is equal to 1.

d) Markups

Markups adjust to desired levels m_{j}^{*} over adjustment time τ^{m}. Desired markups are equal to a reference markup, adjusted through pressure from market level prices ε_{j}^{m}

$$m_{j}^{*} = \varepsilon_{j}^{m} m_{ref}$$

(A.39)

Pressure to decrease (increase) markups result from a discrepancy between price $p_{j} = \left(1 + m_{j}\right)c_{j}$ and the market level, relevant for platform j, p_{j}^{m}:

$$\varepsilon_{j}^{m} = f\left(p_{j}/p_{j}^{m}\right); f' < 0; f(0) = 0; f(1) = 1; f(\infty) = \varepsilon_{max}^{m}$$

(A.40)

The perceived relevant market price adjusts to the actual relevant market price p_{j}^{*m} with adjustment time τ^{p}. This model ignores potential product differentiation with respect to consumer choice; therefore the indicated relevant market price is the price of all platforms weighted by their market shares:

$$p_{j}^{*m} = \sum_{j'} \sigma_{j'} p_{j'}$$

(A.41)

Thus, in the long run the market tends to produce at unit costs of the cheapest producing platform.

Throughout the analysis I hold markups fixed at 0.2.

e) Scale economies within a platform

The role of scale economies are important to consider. First they

We include two forms of scale economies. First, important economies of scale, internal to the production, and act at the modular level, ε_{jm}^s. Other scale economies are aggregated and modeled as a function of the existing installed base and introduced in the analysis section. Here we specify the internal scale economies:

$$\varepsilon_{jm}^s = f\left(s_j / s^0\right); f' \le 0; f(\infty) = 1; f(1) = 1 \qquad \text{(A.42)}$$

The selected function is a standard power law, where cost improves as $f(x) = x^{\gamma_m^s}$. The scale exponent γ_m^s is calculated from the assumed fractional cost improvement per doubling of sales, $(1 + \Delta) = (2s_0/s_0)^\gamma$, or $\gamma = \ln(1+ \Delta)/\ln(2)$. For analysis a 30% scale curve, $\Delta = 0.3$, is the default, corresponding with the scale effect parameter $\gamma^s = 0.379$. The section that discusses c) Optimal Resource Allocation shows how scale economies feed into the cost equation.

In the analyses of the Essay the scale effect parameter is set to 0.

5 Model and analysis documentation

The model and analyses can be replicated from the information provided in the Essay and the first section in the Appendix. In addition model source code and analysis documentation can be downloaded from

http://web.mit.edu/jjrs/www/Thesis Documentation.htm

6 References

Ben-Akiva, M. E. and S. R. Lerman (1985). Discrete choice analysis : theory and application to travel demand. Cambridge, Mass., MIT Press.

Berry, S., J. Levinsohn, et al. (1995). "Automobile Prices in Market Equilibrium." Econometrica **63**(4): 841-890.

Brownstone, D. and K. A. Small (1989). "Efficient Estimation of Nested Logit-Models." Journal of Business & Economic Statistics **7**(1): 67-74.