Dynamics of Product Category Emergence: Social Influence within and Between Hybrid Electric Vehicles

Jeroen Struben,*) Jad Sassine
McGill University
*) Corresponding author (jeroen.struben@mcgill.ca)

Version
June 9, 2014

Dynamics of Product Category Emergence: Social Influence within and Between Hybrid Electric Vehicles

Abstract

This paper explores the dynamics of consumers' product and category consideration during category emergence. Extending ideas of two-stage adoption of category consideration and product selection this paper develops a dynamic model with a focus on the interplay between social exposure at the product and category level. The model differentiates between audience-level exposure (at the collective level) and direct exposure (at the individual level). We apply the model to the context of hybrid electric vehicles (HEVs) in the United States between 2000 and 2010. Specifically, we study spatio-temporal interactions between adoption patterns of the category's most successful product – the Toyota Prius – and other HEVs. We find that their joint adoption is marked, initially, by strong positive-cross product exposure spillovers. Yet, as the category matures positive spillovers reduce, in particular at the audience level. Further, we show, through counterfactual analysis that while social exposure spillovers were critical for the buildup of consideration, their subsequent reduction facilitated patterns of dispersed adoption, shielding the category from within category competition. Our findings provide insights into both timing and direction of category emergence. We discuss generalizations to other contexts, including to those involving emerging categories with boundaries that are fuzzier than those of HEVs.

Keywords: category emergence, valuation, social exposure, diffusion, hybrid electric vehicles

Introduction

Our society in its dependence on nonrenewable resources is unsustainable, with resulting greenhouse gas emissions rapidly changing the climate (IPCC, 2013). Yet, contrasting a strong scientific consensus on the causes, human role, and risks of climate change, altering patterns of consumption within the public has proven a major challenge (Sterman, 2008; 2011). The importance of markets as a locus for achieving more sustainable pathways has been well recognized. Concerns about sustainability have stimulated a proliferation of actions by industry players, governments, and non-governmental organizations. However, markets have shown a remarkable resistance to interventions aimed at improving the share of alternative products whether economic incentives, raising awareness about them through education, or increasing availability and appeal through innovation. For example, despite high anticipation and media attention, alternative fuel vehicles (AFVs) are slow to penetrate, and there have been many failed attempts across countries. Our lock-in into unsustainable pathways is much more a socio-behavioral-systems-problem rather than a technology problem. In short, we need to better understand how markets form and the pathways they may take. To this end, this paper focuses on one aspect of market formation: the dynamics of product category formation, highlighting coevolutionary process of public acceptance of and choice between multiple product alternatives.

Central to the dynamics of new market formation is the formation of new product categories. New markets do not emerge out of a vacuum. Rather, their pathways are shaped through joint competition and coevolution between multiple partially similar products. While research typically stylizes the new-product-formation problem as a competition between the dominant and a challenging technology, often multiple alternative products compete for acceptance and future dominance. For example, AFVs on the market include, among others, hybrid electric vehicles, pure electric vehicles, hydrogen fuel cell vehicles, diesels, and natural gas vehicles. Moreover, within each alternative multiple options are available: over 25 hybrid models are available in the US, with offerings ranging from trucks to small cars. Likewise multiple electric vehicles are currently being introduced in the market, some of them pure electric, others with “gasoline range extensions“.

Category emergence depends on the increasing presence of appealing products. On the other hand, consideration of the category’s products requires a stock of shared meaning (Hannan et al. 2007). Understandings of the process of product category emergence build on Zuckerman's (1999) two-stage category framework on how actors select between alternative options. This framework reconciles insights from behavioral- and rational-choice-based decision-making literatures. First, consumers, with limited information processing capabilities, only have attention products that have entered in their “consideration sets” (Hauser et al. 1993). The second explains how consumers select between products based on attribute-level comparison of available options (McFadden, 1978). In Zuckerman's framework consumers first perform a "social-screening" of illegitimate options, and only thereafter select "more rationally" among legitimate alternatives. Applied to the case of minivans, Rosa et al. (1999) describe in dynamic terms the process of product category legitimization, and in particular role of social influence. Social exposure, in the early stages, builds consumer familiarity not only with the product (Mahajan et al., 2000), but also for the emerging category as a whole (Navis and Glynn 2010). Struben & Sterman (2008) develop a formal dynamic model of category legitimation consistent with the above ideas, and building explicitly on behavioral- and discrete-choice theories.

However, despite growing interest in dynamics of category emergence around multiple product alternatives (Durand & Paolella, 2012; Kennedy et al., 2010), so far the processes by which consumers compare and select between multiple partially distinct alternatives during the early stage are not well understood. Addressing this gap requires attention to the mechanisms by which product category boundaries form around and between the multiple partially distinct alternatives. But this grouping is dynamic as "shifts in the category membership parameters will exclude some previously acceptable members of the category and affirm other members" (Rosa et al., 1999). Moreover, leading producers may strategically position their products more central within (Santos & Eisenhardt, 2004) or more distant from (Barney, 1991) the core of the emerging category. The dynamics of i) category legitimation and ii) within-category positioning jointly shape the direction and success of the emerging product category.

The creation of novel categories involves deliberate deviation from the norm. Here we consider the dynamics of category emergence that implies a number of tensions. We focus on the multiple interactions between alternatives from the perspective of audiences. First, producers must be willing to position their products within an illegitmate space. From within that space, new categories may begin to build legitimacy. However, legitimacy requires products to position themselves close enough to others to achieve a recognizeable set of possibilities. The problem of conformity and differentiation is parallel and there is a dynamic tension between them.

Here we work towards developing a behavioral, dynamic and empirically grounded category emergence model for multiple products. To do this we combine, iteratively, empirical analysis with initial steps in theory building. We begin describing the overall framework in causal loop diagram. Next we introduce how discrete choice theory can be applied to test basic ideas. We then apply this approach to the case of hybrid electric vehicles (HEVs), using a large spatio-temporal dataset. Testing the model empirically on the case of Hybrid Electric Vehicles in the United States (2001-2010) we find a positive feedback between adoption and consideration. We also observe solidification into a single category. In addition we find the importance of category-level consideration spillovers across alternatives. Over-time the relative strength of these spillovers decline. But, this decline is much stronger at the level of salient selection, than at the global level. In fact, at the local level we observe selection pressures based on not only present consideration but also relative presence of alternatives. These findings suggest i) importance of path-dependence and ii) the importance of category-level identity as well as differentiation in the process of emergence.

Category Emergence

Dynamics of category emergence and stabilization involves a process of producers and audiences responding to each other's actions. Within more stable categories, we understand, first, that producers tend to produce collective self-reproducing role structures as they continue conformity with recognized "schedules" of cost-quality niches (White 1981; 1993; Leifer and White 1987). In watching their competitors firms engage in a process of social conformity influence among structurally equivalent rivals (Burt 1987; White 1993). Producers benefit from cross-product comparisons that sustain the market in contributing to more recognizable prototypes.

Audiences exert a similar influence towards macro-level conformity (Zuckerman 1999). Products compete for the favor of audiences that select among alternatives. Producers positioned in illegitimate space do not receive attention – or sales. The selection pressures induce competitive tensions among focal actors and with that pressure for conformity. Contested – with ambiguous meaning and complex products - with derived value depending on others' consumption of similar resources or technology, where the need for compatibility generates "network externalities" (e.g., Farrell and Saloner 1985). In such situations critics and knowledgeable actors play a crucial role in providing guides to current and future valuation. Indeed, with this, central in product category emergence is that a high degree of similarity with a categorical prototype assists producers and their products at a collective level audiences a class about whose meaning an audience segment has reached a high level of intentional semantic consensus (Hannan et al., 2007; Hsu et al., 2009). Collective identities develop around common dimensions, and audience recognition. Such audience recognition produces collective meanings materialized by codes through exposure at the population level (Hannan et al. 2007). Thus, consideration is governed itself by the adoption history.

On the other hand products differ from each other. For example, while the "minivan" is a category that emerged over time, with products containing similar attributes (Rosa et al. 199), not one minivan is the same. As a contrasting force, audiences induce actors to differentiate themselves from others (Hotelling 1990). Hence there is a natural tradeoff between similarity and differentiation within a product category. Together these forces explain observations of stable categories. For example, Zuckerman (1999) develops an equilibrium mode, connects sociological models of organizations and markets and the models of consumer decision-making and market structure that prevail in the marketing literature. Thus producers engage in a process of "conformity and differentiation". Players, in turn, vie with one another to promote their offers to audiences. Each player tries to differentiate its offer from those advanced by its peers and establish its relative desirability.

More operationally, audiences are making two types of decisions. First, consistent with consideration set theory (e.g., Nedungadi 1990), options that do not meet minimal criteria of acceptability and comparability are screened out. Next, consumers compare among members of their "consideration sets" and select a final choice.

Category dynamics

Situations of new category emergence require more careful scrutiny however. For new categories producers introduce potentially viable products away from a conventional category and compete for legitimacy to achieve and maintain value and meaning. Category emergence centers on the process of building legitimacy (Meyer and Rowan 1977; DiMaggio and Powell 1983). In the early stages, as no subcategory can be distinguished, market discipline faced by niche defectors, important in stable markets, cannot be a dominant force yet. First, in the very early stages audience evaluations and consumer choice are, due to lack of clear values, dominated by the screening out of illegitimate options rather than by rational attribute comparison. Thus, individual consumers will consider only a few among many options but, initially, there is no cohesion in this choice among the many consumers. In the early period of confusion, as "social screens…given in the categories that comprise market structure" (Zuckerman 1999) do not exists yet between alternative products, all that deviate from accepted categories are penalized because of limited consideration. Over time however, consideration sets begin to emerge endogenously, rather than originating in individual tastes, as publicly discussion of product categories produce social boundaries (Urban et al. 1993; Bronnenberg and Vanhonacker 1996). As consumers collectively select some products more than others, their market share feeds back to provide further exposure, public discussion, and media attention. Thus, the cumulative sales increasingly produces material for legitimacy (Nedungadi 1990). Consideration build up inducing a positive feedback of social exposure dynamics (Valente 1996; Struben and Sterman 2008).

Second, the positioning of the products within the category matters, just as it does for static processes. First, because salience and memory are important (Nedungadi 1990) in the buildup of category legitimacy, products that are positioned proximate others that are perceived as viable benefit from the spillovers. Consistent with this, the general pattern of category formation that researchers suggest is that producers initially emphasize category development in the early stages of the category–by highlighting consistency with the forming category (Navis and Glynn 2010). But, as the category begins to solidify, lead producers find it sensible differentiate themselves from the main category (Navis and Glynn 2010; Santos and Eisenhardt 2009).

The tradeoff between competition and coevolution suggests that dynamics within (and between?) categories depend on the degree to which one or the other force dominates - how similar consumers perceive products are across subcategories to be and how producers position those products. Because of these positive feedback dynamics in emergent categories. The degree of differentiation imposes a trade-off between direct and longer-term benefits to individual producers as well as to the category as a whole. This trade-off however is dynamic. more at stake in terms of positioning. Results also depends on conditions: differentiability and how to differentiate. Combination of strategy and plasticity of the product category. Similarity affects the various feedback relations in two ways. First, during category emergence, similarity across choice options strengthens the social exposure spillover between subcategories. Second, similarity increases selection strength between products. In the extreme, when two products are perceived to be identical, assuming full familiarity, consumers simply tend to select the superior product. Thus, rather than highlighting the category above the firm identity, strong identification with the emerging category contributes to the perceived visibility of the category as a whole.

Within the frame of our model this over-over time individual firm orientation means that lead-producers may seek to reduce perceived similarity as the category matures or as the market share of products within the subcategory increases. The intended rational underlying this would be that reduced similarity increases market share directly, as benefits from further building the category reduce. However, several questions remain open: we do not understand the process extend producers actually follow such a process. Further, what can we say about *when* it makes sense for producers to alter their identification with the emerging category? A third major question is the role of spatial niches. For example, spatial differentiation between subcategories may be an intermediate strategy allowing producers of subcategories to support overall category building while benefiting from high individual market shares niche within niches.

In what follows we develop a dynamic model with a focus on the interplay between social exposure at the product and category level. The model differentiates between audience-level exposure (at the collective level) and direct exposure (at the individual level). We apply the model to the context of hybrid electric vehicles (HEVs) in the United States between 2000 and 2010. Specifically, we study spatio-temporal interactions between adoption patterns of the category's most successful product – the Toyota Prius – and other HEVs. We find that their joint adoption is marked, initially, by strong positive-cross product exposure spillovers. Yet, as the category matures positive spillovers reduce. Further, while the social exposure spillovers were critical for the buildup of category consideration, reduced spillovers as the category matures directly facilitated differentiated adoption, thus shielding the category from within category competition. Our findings provide insights into pathways – timing and direction - of category emergence. We discuss generalizations to other contexts, including to those involving emerging categories with boundaries that are fuzzier than those of HEVs.

. What matters is how recognizable product categories are. We propose a model that include these trade-offs: legitimacy building across subcategories, differentiation in terms of within and between category. We suggest that social exposure dynamics are such that the positive feedback of spillover drives category formation importantly.

A Dynamic Model of Category Emergence

Overview

We develop a generic model of category emergence that highlights the over-time emergence of a category as multiple potential product members get adopted and diffuse in the market (Figure 1). The basic unit of analysis is the subcategory, representing a subset of one of multiple similar alternatives, considered, at the population level, as "similar" (Hannan et al. 2007). The installed base of a subcategory increases with sales, a function of market share and total market size. The subcategory market share depends on its market share within the emerging category. This market share depends on the intrinsic-valence of the product and consumers' consideration of the product. Consumers willingness-to-consider (WtC) a product increases as audiences – media, evaluators, potential consumers, etc.. - get exposed to the product and collectively develop shared codes and inscribe meaning (Hannan et al. 2007). As in turn product sales increase so does the material for further audience exposure (R1a, *Shared Code and Meanings*). Because products across subcategories partially overlap, from the perspective of audiences, social exposure may spillover across subcategories (R1b, *Spillovers*). Whereas the audience feedback is governed at a collective level, as consumers are directly exposed to a product – within their local context r -, they may become more knowledgeable about the product. Subsequently, WtC can increase. As in turn product sales increase and get more exposure through social exposure to the product, which in turn increases with adoption and usage of the products within the subcategory (R2, *Social Exposure*). Likewise, as products within an emerging category share similar features exposure to those related products also increases consideration (*R2b, Spillovers*). On the selection side, as the affinity of products within its constituent subcategories increase, a category's market as a whole also increases. Hence, the improvement of affinity with products of a subcategory improves also the affinity with the category as a whole, increasing the category market share and thus the total market share of the subcategory, with installed base further improving, this sets in place a second loop (R3, *Category Affinity*). Hence products between subcategories not only compete but also help each others' adoption as well as share.[2]

*** FIGURE 1 HERE ***

Category Choice

Choice for a product p is modeled as a two-component decision-making process (Zuckerman 1999). Consumers, within geographic or demographic region *r*, select between categories, yielding category share σ_{pr}, and then choose products *p* among those within the category, yielding product share $\sigma_{p_c r}$. Choice of products may differ across regions because of varying exposure to the various products and with that different consideration (Keith et al. 2014), in turn affecting category choice and emergence. Then a product *p*'s, member of category *c*, market share within region *r* is:

$$\sigma_{pr} = \sigma_{p_c r}\sigma_{cr} \tag{1}$$

Between-category selection is based on the relative affinity (or "currency", Kennedy et al. 2010) a_{cr} across categories:

[2] While naturally the valence derived from the products within the various subcategories does not tend to be stable during the period of category emergence. This however is not the focus of our current analysis. For the purpose of analytical clarity we leave any interaction of valence and adoption out of the scope of the current model. Dynamics of valence, capabilities, and consideration between categories are the topic of another paper.

$$\sigma_{cr} = \frac{a_{cr}}{\sum_{c'} a_{c'r}} \tag{2}$$

A products' share within category c, $\sigma_{p_c r}$ is based on relative affinity a_r across individual products, with product substitution parameter ρ indicating how related consumers see the products within the forming category compared to products that fall outside the boundary of the category. (This equation follows the classic nested logit-choice model, which captures interrelated choice situations (McFadden, 1978; Ben-Akiva and Lerman, 1985):

$$\sigma_{p_c} = \frac{a_p^{1/\rho}}{\sum_{p \in c_p} a_p^{1/\rho}} \tag{3}$$

A small substitution parameter ρ ($\rho > 0$) corresponds with products perceived to be more similar to each other (than across categories). Thus a small ρ indicates a strong perception of category membership between a group of products p_c with clear boundaries between those and other products; this clarity induces stronger product selection within the category. By contrast, situations of $\rho = 1$ correspond with products where consumers see the products as unrelated. (In this case equation (1-3) reduce to a symmetric choice structure between $P = \Sigma_{cp} P_c$ equally dissimilar alternatives.)

Category affinity depends on the affinity with the set of member products, as well as on how those products are related.

$$a_c = \left(\sum_{p \in c} a_p^{*1/\rho} \right)^{\rho} \tag{4}$$

For products that are perceived to be close substitutes of each other (small ρ) the category affinity is less elastic to additional product introductions. For example, adding another brand of tissues (of similar affinity as others) does not increase the overall affinity of the tissues segment. By contrast, when a category is perceived to constitute relatively unrelated products (or, when a category is not very well established), $\rho \approx 1$ when adding a new product fills previously unsatisfied demand.

Affinity with an individual product p is a multiplicative function of the WtC and the intrinsic product's valence (Struben and Sterman 2008):

$$a^{*}_{pr} = C_{pr} v_{pr} \tag{5}$$

A product's intrinsic valence captures the potential for a product to yield high affinity and thus market share. It is a function of product attributes and the characteristics of the population. While malleable by producers, valence is not directly subject to social influence. (Hence our definition of valence differs slightly from the one in Kennedy et al. (2010). By contrast, consideration is a social construct, capturing cumulative effect from audience evaluations, from collective meanings and shared codes that audiences over-time ascribe to the products, as well from the knowledge that consumers have acquired about those products. For the purpose of this paper we focus on dynamics around consideration of the product, rather than on dynamics of the valence.

Willingness-to-Consider

Consumers' consideration of a product captures the extent to which consumers are sufficiently familiar with a product and the broader product category that they are willing to include it in their consideration set. WtC is memory driven (Nedungadi 1990), building and sustained through the "…cognitive and emotional processes through which [consumers] gain enough information about, understanding of, and emotional attachment to a platform for it to enter their consideration set" (Struben and Sterman 2008). Consumer WtC accumulates the impact of various social influences e_{pr} about the category and product, including direct social exposure to related products, broader audience influences, marketing etc.. Further, as actors rely on their memory, WtC decays over tme, at rate ϕ. Then, a product's WtC is:

$$\frac{dC_{pr}}{dt} = e_{pr}\left(1 - C_{pr}\right) - \phi_{pr} C_{pr} \tag{6}$$

Social influence for product consideration e_{pr} is mediated through different channels, marketing $e_{m,pr}$, the collective production of codes and shared meanings by broader audiences $e_{a,pr}$ and the individual, direct exposure to the products $e_{d,pr}$:

$$e_{pr} = e_{m,pr} + e_{a,pr} + e_{d,pr} \tag{7}$$

Exposure at the individual and collective level each involve product-specific influences as well as spillovers from products that are perceived close enough related. That is, a high degree of similarity with a categorical prototype contributes to legitimacy of other products (Hannan et al., 2007; Hsu et al., 2009). Within-category spillover effects γ depends may differ between influence from the collective or individual level. Thus:

$$\begin{aligned} e_{a,pr} &= e_{a,pp} + \gamma_a \sum_{p' \in c_p; p' \neq p} e_{a,p'p} \\ e_{d,pr} &= e_{d,ppr} + \gamma_d \sum_{p' \in c_p; p' \neq p} e_{d,p'pr} \end{aligned} \tag{8}$$

This equation captures how categories develop around common dimensions, on the one hand as audience recognition defines common traits and categorical identity, and, on the other hand as consumers are directly exposed to and are more familiar with the collective of products with those common traits. Weather direct or audience exposure is more important depends on the characteristics of the good. Products that are more complex can be expected to require more direct exposure (Centola and Macy 2007).

Irrespective of the source, exposure increases with the presence of the products and category products. We model audience exposure as a function of total recent category sales across the market $e_{a,p} = e_{g0} f\left(\overline{\sigma_p}\right)$, with $\overline{\sigma_p} = \sum_r \overline{s_{pr}} \Big/ \sum_{p'r} \overline{s_{p'r}}$,where s_{pr} total sales, the market share times total replacements, $s_{pr} = \sum_{p'} N_{p'r} / \tau$, with τ the turnover rate of products and N_{pr} the cumulative sales of a product. Direct exposure instead is about visibility of the products in use, potential consumers talking to current users of the products, and is therefore a function of the installed base N_{pr} the consumers of segment r are exposed to:

$e_{d,pr} = e_d f\left(N_{pr}/N_r\right)$, with $N_{pr} = \sum_{p'} N_{p'r}$

Unless consumers actively pay attention to new product categories WtC erodes over time (Nedungadi 1990; Struben and Sterman 2008; Keith et al. 2014). Moreover, consideration loss rate should be highest when the consumer's exposure to a product category is low. Knowledge of the new product category is incomplete and personal commitment to it is not strong. However, once a product category becomes ubiquitous, and constant reminders of its presence exist, WtC becomes highly durable. Hence:

$$\phi_{pr} = \phi_0 f\left(C_{pr}\right); \qquad 0 \le f \le 1; \qquad f' \le 0 \tag{9}$$

Application: The emergence of the hybrid electric vehicle category

The early diffusion of Hybrid Electric Vehicles (HEVs) forms an excellent setting to examine this model. HEVs are alternative fuel vehicles combining a gasoline engine with a storage-battery-powered electric motor. In the US, during 15 year since the introduction of the first hybrid vehicle (the Honda Insight), multiple HEVs, including cars and SUVs, have been introduced with different and evolving characteristics (Figure 2).

Among HEVs the Toyota Prius has been a market leader ever since its introduction. Except for the two-door Honda Insight, the Toyota Prius was, in the years 2000-2001, the only alternative fuel vehicle competing in the US market. In 2002 came the Honda Civic, with market share relatively close to that of the Prius. After the introduction of the 2nd generation Prius (2004), Toyota Prius sales doubled those of the Civic (Polk 2010). From 2003 onward, other hybrid vehicles were introduced, including the Ford Escape, the Lexus RX300 or the Mercury Mariner. Figure 3 summarizes the evolution of HEV in the US market, showing total HEV market share, Prius market share, best follower sales, and the number of HEV models on the market.

While HEVs were perceived as a category, the Prius has been its dominant product, or "industry referent" (Santos and Eisenhardt 2004). The Toyota Prius was introduced in Japan in 1997 and in the United States in 2000. It was the first hybrid four-door sedan available in the US (Hybridcars 2011). While competition increased, the Toyota Prius remained the leading alternative fuel vehicle with more than 50% market share. The Toyota Prius is on the one hand a reference for the 'Hybrid' category but has been positioned much differently from other HEVs (Garland et al. 2013). Figure x shows Prius advertisement and a comparable.

*** FIGURE 2 HERE ***

*** FIGURE 3 HERE ***

The emergence of the Hybrid category forms an excellent case to study the social influence dynamics also because its underlying technology fairly stable. For example, compared to new personal computers, the improvement of hybrid technology is fairly stable. Other factors such as varying technology, while a potential important extension, does not importantly affect the dynamics discussed here. Further, HEVs is a very successful diffusion case.

Hypotheses

From our generic model on category emergence case we develop three hypotheses about category emergence of HEVs. We develop our hypotheses in the context of HEVs because the model parameters and the dynamic implication are context-dependent. Testing these hypotheses form nevertheless an important first test of the general usefulness of the model. In addition, they provide insights into the specific case of HEV emergence.

We expect the HEV to be relatively well defined as a category, right from the launch of introduction. While currently alternatives increasingly enter the market that are pure electric vehicles, battery electric vehicles mixes thereof (such as the plug in hybrid electric vehicle), in the first decade since the introduction of the Honda Insight, the HEV category was unambiguously defined by cars having an extra capability that can captures energy while driving, store this in the battery, and deliver energy to the motor when needed. Those characteristics are consistent, for example, with a major HEV patent on Hyperdrive power-amplified internal combustion engine power train in 1994 (Patent Grant US5343970 A, http://www.google.com/patents/US5343970, by Alexei Severinsky). These ideas have remained fairly stable. Thus, we expect that products that having this capability is perceived as closer to each other than to the conventional products.

Hypothesis 1: Substitutability between HEVs models is smaller than 1.

While category membership is expected to be well defined, initial consideration of products part of the category for consumption will be low. Many consumers, while potentially aware of the existence of the category are unfamiliar with its products and have low knowledge about them. From a consumer point of view products are complex and different from existing models. Investments in vehicles are large so they cannot easily be tried out. Thus, as emergence of the category requires buildup of consideration we expect that the social exposure feedback is an important process governing the diffusion of the category. The requisite growth of consideration at the category level means that spillovers across products within the HEV category.

Hypothesis 2: In the early stages of HEV category formation, social influence spillovers among its category members is strong.

Consideration is governed by both local and audience category influence. Such influence is important in the early stages of category formation. However, as the category matures and becomes legitimated, consideration saturates and producers, audiences and consumer choice begin to focus on product selection.

Hypothesis 3: In the later stages of diffusion, products differentiate themselves from other hybrids, reducing consideration spillover between the Prius and non-Prius HEVs.

Model Estimation

Data

To test our hypotheses we draw upon a unique spatio-temporal dataset (by zip code and by quarter) on vehicle adoptions in the United States between 2001-2010. The data set contains all-US new registrations by zip code / quarter between 2001-2009, for individual HEV models, total industry sales (cars and trucks separately), obtained from Polk. The data set was collected jointly with other co-authors (David Keith and John Sterman).[3] Our main goal is to understand social exposure dynamics at the category and product level. Because , first, dynamic estimation of multiple HEVS (up to 23 models) is computationally intensive,

[3] The full data set contains incentives, model-specific attribute details (including fuel economy, vehicle size, and vehicle price), state-level fuel price, state and federal policies, zip-code-level demographics including median household income, education-level, party affiliation and voting and news articles on climate change (Keith et al. 2014).

we produce results using the simplest possible representation, and, second, we are primarily interested in interactions within HEV adoption.

First, to simplify the estimation across multiple hybrid electric vehicles we distinguish the Prius (P-HEV), a leading product in the emergence of the category. Estimation requires a dynamic model. To reduce computational intensity, we aggregate other non-Prius hybrids (NP-HEV) and sales by conventional cars (CONV). Second, with low overall HEV adoption, especially for NP-HEVs, to assure sufficient adoption events we bound our estimation here to a region with successful al-round HEV adoption. We exclude truck sales out of the hybrid and conventional data set (subsequent analysis can include this). We selected the bay area of the state of California (counties: Alameda, Contra Costa, Marin, Napa, San Francisco, San Mateo, Santa Clara, Solano and Sonoma), a region with the highest HEV adoption rate in the country. This subset contains in total 428 zip codes and 40 quarters.

Social exposure variables

We represent audience-level collective exposure, $e_{a,pr}$, through the total sales in the Bay area. Direct social exposure, $e_{d,pr}$, is captured through the exposure to the within zipcode installed base. The zip code forms a reasonable proxy for a region for representing consumers' exposure to vehicles on the road or to being informed about them by people from within their community (Keith et al. 2014). To simplify the vehicle fleet model we assume that individuals do not replace their newly purchased vehicles for at least 10 years. The installed base share of a hybrid car within a zip code is therefore computed as the cumulative product sales divided by the stock of all cars within the zip code. The stock of all cars is calculated as the 10-year average sales within the zipcode. As we aggregate all other hybrids into one choice, we correct the exposure effects so that only the fraction $w=1/(m_h-2)$ of the NP-HEV installed base share contributes to direct exposure, where m_h is the number of HEV models on the market (including the Prius). *(1- w)* of the PN-HEV installed base share contributes to within-category spillovers.

While we possess marketing data on the Prius, we do not have this for each of the individual models. To nevertheless account for the marketing effect, estimate the marketing term component e_m as a constant, consistent with the Bass Model (Bass et al. 1994; Mahajan et al. 2000; Struben and Sterman 2008).

Choice structure

The aim of this estimation is to identify the relative importance of the various social exposure variables. However, we need to control for alternative-, environment-, and individual-specific factors that may influence adoption. Therefore, in the intrinsic valence component of the choice structure we include three factors to be estimated. First, we identify an unobservable alternative-specific variable β^h for any choice option that is part of the hybrid choice set. Second, *m(t)* captures the number of HEV models in the market (Figure 2, 3). Third we capture a fixed effect that provides an alternative explanation of unobserved locational specific effects, k_{pr} to the direct and collective social influence effects. Then, with $v_{pz} = e^{u_{pz}}$

$$u_{pr} = \begin{cases} \beta^h + k_{pr} & p = PRIUS \\ \beta^h + \beta^m(m-1) + k_{pr} & p = NP-HEV \\ 0 & p = CONV \end{cases} \tag{10}$$

The factor k_{pr} allows controlling for alternative explanations for the pattern of adoption, tied to the local region (zip codes) environmental factors such as education, variation in regional fuel prices (see Keith 2012). A related explanation that may lead to confounded results is the presence of homophily. Homophily is the principle that a contact between similar people occurs at a higher rate than among dissimilar people (McPherson et al.

2001). For example people with similar interest, taste, education, and values tend to live together (Lazarsfeld and Merton 1954). Under stable patterns of social relations within communities (whatever the over time mechanisms that led to this) those living nearby may therefore be more likely to share values, attitudes, and beliefs. Hence, socio-demographic clusters may stratify not only acquired characteristics including behavior patterns. Thus, homophily creates correlated outcome patterns among "nearby actors" that merely mimic viral contagions without direct causal influence (Aral 2010). Homophily would predict outcomes conform the final adoption patterns. However, whatever the effect, influence effects would be constant rather than propelling a positive feedback (Aral 2010). We capture this by specifying k_{pr} to predicts adoption conform final adoption. Hence we specify k_{pr} = α.fix*ln($\sigma_{T,pr}$)/sum($k_{pr} = \alpha_{fix} \ln\left(\sigma_{T,pr}\right) - \mu_h$ where $\mu_h = \sum_{p \in h,r} \ln\left(s_r \sigma_{T,pr}\right) \Big/ \sum_r \ln\left(s_r\right)$ is the weighted average of the installed base share at the final time.[6]

Over-time structural changes

We test for structural changes in the process of category formation over time. For example, Hypothesis 3 suggests that the role of spillovers reduce as the category matures. To examine this we include tests that allow detection of change in the estimated parameter strength of social influence and spillover strength. In specific models, we estimate a multiplier after a selected "change" year (2005, and 2006).

Estimation Procedure

Discrete choice theory is an empirically robust framework aimed at understanding why individuals make mutually exclusive, discrete choices. We use the method of Maximum Likelihood Estimation (MLE) is used to find the parameters that maximize the probability of observing the data. The likelihood function can be specified that s consistent with the logit-

[6] In a static model version, which permits more easily estimation of a large parameter set, we tested the role of other controls such as education, operating costs, and truck share, presentation voters green (See also Keith 2012; Keith et al. 2014). While some effects such as operating costs are significant, the variables have little effect on the social exposure effects.

nested-logit choice (Train 2003). The models assume iid extreme valu distribution of the error terms. Although dynamic models rarely obey these assumptions, they can practically effectively be estimated using these models (Train 2003; Struben et al. 2014). In the context of logit, estimating the logit model using the method of maximum likelihood estimation (MLE) means finding the vector $\hat{\theta}$ that maximizes the probability that each individual choses the alternative he was observe to choose. We developed the model and estimate this using the open source program R (R-Core Team 2012).

Results

We first discuss the general results and then interpret those findings in relation to the hypotheses. Table 1 summarizes the estimation results. First, in model 1 ("product diffusion") we assume only audience-level exposure to products. In this model, products within the HEV category are not considered significantly different from the conventional products: the substitution parameter $\rho \approx 0.99$ and not significantly different from 1 (at 95% confidence levels, at asymptotic confidence intervals[7]). Adding category effects at the audience level (Model 2) shows an important role of spillovers between products ($\gamma_a \approx 0.57$, significantly >0 and <1). In addition, the overall role of social exposure increases compared to model 1 (e_a alone more than doubles). We also note a small but significant effect of the within-category substitutability ($\rho \approx 0.99$, significantly <1). Finally, the category effect slightly suppresses the importance of fixed effects (α_{fix}) and of HEV model introductions.

We can reject the simpler model 1 (LL=685254 vs LL=68474). The test statistic (being twice the difference between the unconstrained and the constrained model) is distributed chi-squared. Here, the constrained model is model 1, the unconstrained model is the model 2 and the difference is 1 degrees of freedom. The critical value for a 99%

[7] We also derived univariate confidence intervals with mostly consistent outcomes. However, the univariate intervals tended to be narrower hence we work with the asymptotic intervals

confidence interval being 13.27, we can reject model 1 with a high degree of confidence. The same holds true for subsequent models - as we move to the next model the simpler model is rejected (except when comparing model 2 and 3, as for the latter we did not find full convergence).

Turning to model 3, which adds direct social influence effects (at the zip code level) we see, first, the importance of direct social exposure ($e_d \approx 0.024$, significantly >0) as well as of spillovers. In fact, the direct spillover effect *($\gamma_d \approx 1.25$)* is estimated to be slightly stronger than the within effect. Finally, models 4a-c include sensitivity to over-time changes in the exposure parameters: t_change = 2005 for model 4a and t_change = 2006 for models 4b and 4c, with the period before t_change representing early category formation. First, including sensitivity to change in social influence parameters, maintains the main effects of early category formation. While the importance of product exposure increases as the category matures, spillovers decrease after t_change (ga change and gd change < 1). We suggest that as overall legitimation increases, the role of shared meaning formation becomes saturated. Finally, model 4c adds the influence of an industry referent. We examined the spillover influence of the on NP-HEVs separately. The Prius – which can be seen as an industry referent - has a nearly 50% stronger influence compared to the average non-Prius HEV. This finding is consistent with those from others who suggest that a categorical prototype may contribute more to legitimacy of other products (Hannan et al., 2007; Hsu et al., 2009).

*** TABLE 1 HERE ***

Hypotheses tests

Hypothesis 1 states that HEV models would be considered a category. Indeed we find for any of the model (except the simplistic diffusion model 1). Therefore, we can reject the null hypothesis that HEV models are perceived to be within the same category as conventional models. We also tested whether ρ changes as the category matures but did not find this. Hypothesis 2 states that spillovers are important in the early stages of category emergence. We find this indeed to be the case. We found that spillovers are important both at the audience and at the direct exposure level. Hypothesis 3 states that the category influence decreases as the category emerges. We find this particular to be true for the audience effect. Nevertheless, direct influence across products remains strong throughout. We interpret that as installed base and WtC increases, shared meanings about the HEV have been well-developed and the collective idea of the HEV category is well established and legitimated. However for actual adoption decisions, consumers are still not sufficiently comfortable with the idea of an HEV. Consumers remain to have a need to become more knowledgeable. We believe that this interpretation is also consistent with a two–stage decision-model on product-category adoption (Zuckerman 1999).

Further Interpretation of Results

Figure 5 provides an overview of the estimation of model 4c. The two top rows show the actual and simulated market shares for PRIUS and NP-HEV, respectively for the Bay Area and for three selected zip codes (those include largest market share for HEVs, the Prius, and for NP-HEV). The bottom row showed simulated consideration (left) and the relative effect of direct versus audience and within versus spillover influence, averaged across zipcodes.[8] (Note that category influence is larger than just the total spillover influence because part of the category influence derives from the within effect.) The simulated results further suggest that category formation does not solely rely on higher-level meaning formation but also on specific local exposure. While early audience exposure about the

[8] Since the fixed hybrid effect and the buildup of WtC involve colinearity, it is hard to detect the specific level of consideration. We believe that the estimated level of consideration is too high. However the variation of effects between different channels of influences is not affected by different combinations of fixed hybrid effects WtC.

product is stronger than direct exposure about the product, the direct category effects are important. Moreover, they remain important as the category matures.

*** FIGURE 5 HERE ***

Category effects in social influence – spillovers across products – form an important driver of WTC. However they affect adoption dynamics in more subtle ways. Category influence increases the tendency for adoption of different HEV models (Figure 2) to occur more than otherwise in similar regions. This correlated adoption among HEV helps build up WtC among clusters of potential adopters and in turn contributes to audience-level category effects. Counterfactual analysis on the important social influence parameters illustrates this effect. Table 2 measures the correlation in adoption between PRIUS and NP-HEVs, across zip codes. Actual data and simulated results (Model 4c) shows high correlation (c=0.927 and 0.937 respectively). Reducing direct (or any) spillovers, while naturally suppressing diffusion, also has the effect of reducing correlation. This means that HEVs adoption is more dispersed and WtC is slower to build up. By contrast, joint-adoption of different HEV models within a demographic region contributes to exposure, WtC and subsequently increased audience exposure .

However, as the category matures, we see that spillovers weaken also locally (contrasting the increase of local direct product exposure). As a consequence correlation between HEV models in adoption reduces. However, as the category is more mature, this dispersion allows the HEVs as a category to better compete with the conventional vehicles.

*** TABLE 2 HERE ***

In this particular case the product category boundary was relatively well-defined from the beginning. The estimated ρ smaller than 1, but not changing over time is consistent with this observation. However, r is only moderately smaller than 1. More distinct and well bounded categories may lead to smaller values. Counterfactual analysis with different values of ρ suggest a strong reduction in within-category correlation of product adoption (and, not shown, on overall diffusion). Sharper boundaries increase competition within and therefore lead to dispersed adoption of alternatives. On the other hand categories with sharp boundaries – by contrast to those with fuzzy boundaries (Hannan et al. 2007) may be subject to stronger social influence effects in the build up of WtC because it is easier to provide meaning. Subsequent work can focus on empirical cases on these trade-offs involving fuzzy boundaries and membership (as we now see around the multiple types of electric vehicles)

Discussion and Conclusion

We began this paper with the specification of a generic model of category formation. Product category formation is an critical process to understand the multiple pathway markets may take. The key mechanisms we focused on were the competition and cross-fertilization between multiple subcategories as the category emerges. Examining this empirically in the context of HEVs in the (2003-2009), we found that early on (2003-2006), the within-category cross-product effects on consideration are important. This supports our hypothesis that, in the early periods of adoption, non-Prius hybrids and the Prius are considered by consumers to be members of the same category. But because they are considered to be members of the same category, they draw *more* from each other than they draw from the conventional cars. In the later period of adoption, the category/spillover effects reduce.

This paper is a beginning to the development of a more dynamic model about how categories emerge and coevolve. First, we note that different contexts involve quite different parameterizations. Examining such differences will provide deeper understanding into the pathways of emerging categories and the conditions under which they are more or less stable.

Further, building on the findings here one can expand the model by (Struben and Sterman, 2008), incorporating the multiple sub-category interactions (Ben-Akiva et al., 2002). Models of category emergence need to address patterns of success, failure, and distinct pathways. In particular we must explore under what condition two competing effects dominate dynamics. Our empirical and theoretical analysis suggests additional questions regarding a tradeoff between the effect of social exposure spillovers between subcategories and the effect of differentiation. Differentiation, by associating consumers with their products more directly, may offer producers additional market share. However, because differentiation reduces benefits from social exposure spillovers, there is a tradeoff. Differentiation may help only the individual subcategories that have sufficient legitimacy. While our findings are consistent with existing research on how categories form insofar as they suggest that social exposure dominates in earlier stages while differentiation is more important in later stages, we do not know yet under what conditions this may be effective. One important factor interacting with product-related differentiation is geographic-market differentiation. Subsequent analysis can deepen insights into when and under what conditions these various forms of differentiation interact and are important. Finally, the model may be used to explore differentiation strategies and their timing and how this may help or hurt individual models and the category as a whole.

References

Barney, J. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management*, *17*(1), pp.99–120.

Bass, F. M., T.V. Krishnan, and Dipak C. Jain. "Why the Bass model fits without decision variables." Marketing science 13.3 (1994): 203-223.

Ben-Akiva, M. and S. Lerman. (1985). Discrete Choice Analysis. Cambridge:MIT Press, 1985.

Burt, R. (1987). Social contagion and innovation: Cohesion versus structural equivalence. American Journal of Sociology.

Centola, D., & Macy, M. (2007). Complex contagions and the weakness of long ties. American Journal of Sociology, 113(3), 702–734.

Davis, J. P., Eisenhardt, K. M., & Bingham, C. B. (2007). Developing theory through simulation methods. *Academy of Management Review*, *32*(2), 480–499.

Durand, R., & Paolella, L. (2012). Category Stretching: Reorienting Research on Categories in Strategy, Entrepreneurship, and Organization Theory. *Journal of Management Studies*.

Hannan, M. T., Polos, L. and Carroll, G. R. (2007). Logics of Organization Theory: Audiences, Codes, and Ecologies. Princeton, NJ: Princeton University Press.

Hauser, J. R., Urban, G. L., & Weinberg, B. D. (1993). How consumers allocate their time when searching for information. Journal of Marketing Research.

Hotelling, Harold. Stability in competition. Springer New York, 1990.

Hsu, G., Kocak, O., & Hannan, M. T. (2009). Multiple Category Memberships in Markets: An Integrative Theory and Two Empirical Tests. American Sociological Review, 74(1), 150–169.

IPCC (2013) Climate Change 2013: the physical science basis. Cambridge University Press, Cambridge UK; available at www.ipcc.ch

Keith, D. (2012). *Essays on the Dynamics of Alternative Fuel Vehicle Adoption: Insights from the Market for Hybrid-Electric Vehicles in the United States*. Massachusetts Institute of Technology. PhD Thesis.

Keith, D. , J. Struben, and J. Sterman. (2014). Spatio-temporal Diffusion of Durable Goods – The Case of the Toyota Prius Hybrid-Electric Vehicle. (Working Paper, May 2014)

Kennedy, M., Lo, J., & Lounsbury, M. (2010). Category currency: The changing value of conformity as a function of the ongoing construction of meaning. Categories and markets. *Research in Sociology of Organizations*, *31*.

Lazarsfeld P., R. Merton (1954). Friendships as a social process: a substantive and methodological analysis In: Freedom and Control In Modern Society. Eds M. Berger pp. 18-66. NewYork: Van Nostrand

Mahajan, V., E. Muller, Y. Wind (2000) New-Product Diffusion Models (Kluwer Academic, Boston,MA)

McFadden, D.(1978) "Modelling the Choice of Residential Location" in A. Karquist et al., eds., Spatial Interaction Theory and Planning Models, Amsterdam: North-Holland.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. Annual Review of Sociology, 27, 415–444.

Navis, C., & Glynn, M. (2010). How New Market Categories Emerge: Temporal Dynamics of Legitimacy, Identity, and Entrepreneurship in Satellite Radio, 1990–2005. *Administrative Science Quarterly*.

Nedungadi, P. (1990). Recall and Consumer Consideration Sets - Influencing Choice Without Altering Brand Evaluations. *Journal of Consumer Research*, 17(3), 263–276.

R Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.Rproject. org/.

Rosa, J., J. Porac, J. Runser-Spanjol, & M. Saxon (1999). Sociocognitive dynamics in a product market. *Journal of Marketing*, *63*, 64–77.

Santos, F. M., & K.M. Eisenhardt (2004). Constructing markets and organizing boundaries: Entrepreneurial action in nascent fields. *Academy of Management Journal*

Sterman, J. D. (2008). Economics - Risk communication on climate: Mental models and mass balance. *Science*, *322*(5901), 532–533.

Sterman, J. D. (2011). Communicating climate change risks in a skeptical world. *Climatic Change*, *108*(4), 811–826.

Sterman, J. D. (2012). *Sustainability Science*. (M. P. Weinstein & R. E. Turner, Eds.) (pp. 21–58). New York, NY: Springer New York. doi:10.1007/978-1-4614-3188-6_2

Struben, J., & Sterman, J. D. (2008). Transition challenges for alternative fuel vehicle and transportation systems. *Environment and Planning B-Planning & Design*, *35*(6), 1070–1097.

Struben, Jeroen, John Sterman, and David Keith, (2013) "Parameter and Confidence Interval Estimation in Dynamic Models: Maximum Likelihood and Bootstrapping Methods" in: Oliva, R, N. Osgood, and H. Rahmandad (Eds) Dynamic Modeling Analytical Handbook, MIT Press, Cambridge MA (Forthcoming, Summer 2014)

Train, K. (2003). Discrete Choice Methods and Simulation.

White, H. (1981). Where do Markets Come From?. American Journal of Sociology, 87(3), 517–547.

White, H. (1993). "Markets in production networks." Explorations in economic sociology: 161-75.

Leifer, E.M. and H. White. "A structural approach to markets." (1987): 85-108. In Mizruchi and Schwartz (Eds) Intercorporate Relations: The Structural analysis of business (1987), pp. 85-108

Zuckerman, E. (1999). The categorical imperative: Securities analysts and the illegitimacy discount. *American Journal of Sociology*, *104*(5), 1398–1438.

Secondary references

Automotive Intelligence (2000) Innovative Hybrid Marketing Campaign Launches Toyota's Clean Car for the Future." Automotive Intelligence. July 25, 2000. http://www.autointell-news.com/news-2000/July-2000/July-25-00-p5.htm

"Saatchi and Saatchi, Oasis Agencies Team Up for Toyota Prius Ad Campaign ", Automotive Intelligence. Feb. 22, 2000. http://www.autointell-news.com/news-2000/February-2000/February-22-00-p6.htm

2003 Honda Civic Hybrid – Introduction", Honda Press Release, Feb. 6, 2002.

"Hybrid engine cars do better with hybrid marketing tactics", Margaret Littman, Marketing News, Chicago, Sep 25, 2000.

Tables and Figures

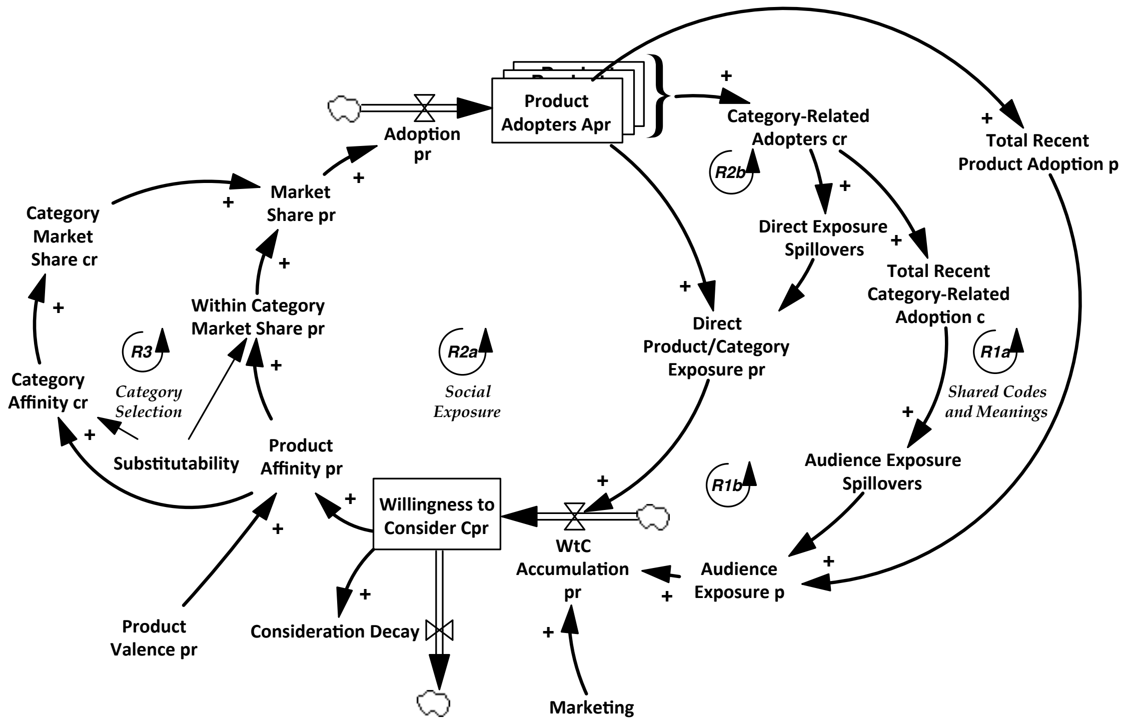

Figure 1. Model Overview

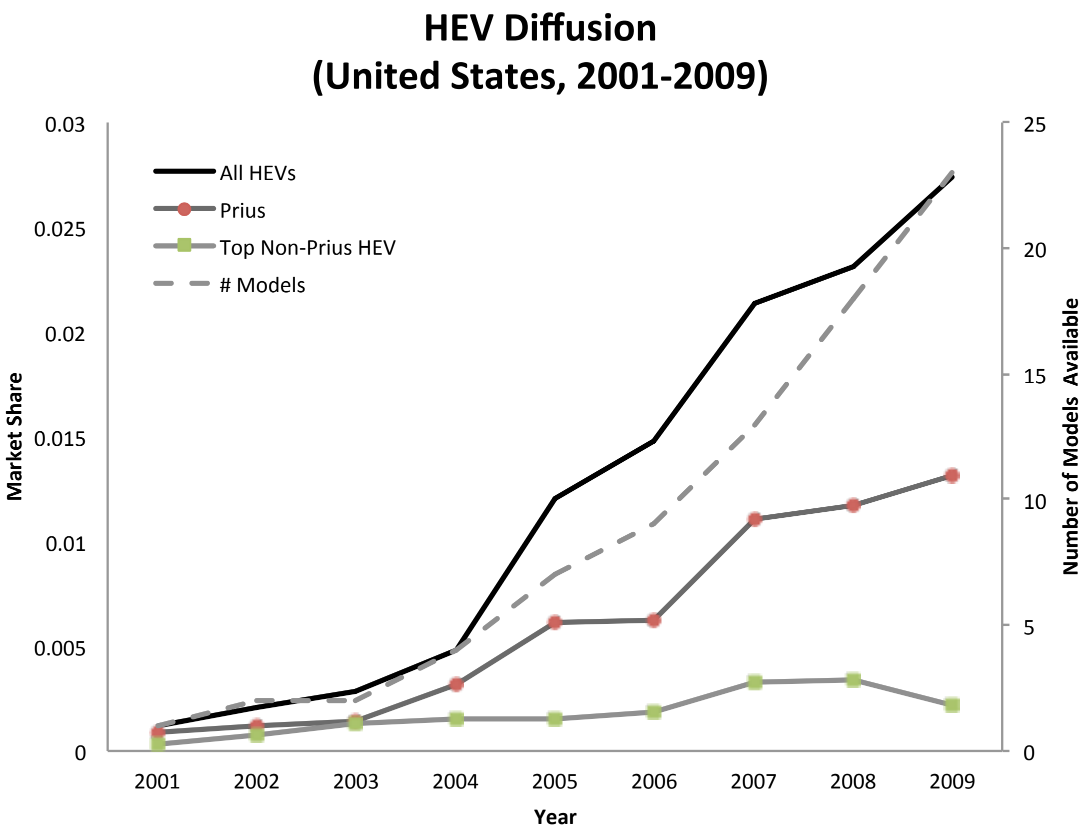

Figure 2. HEV sales

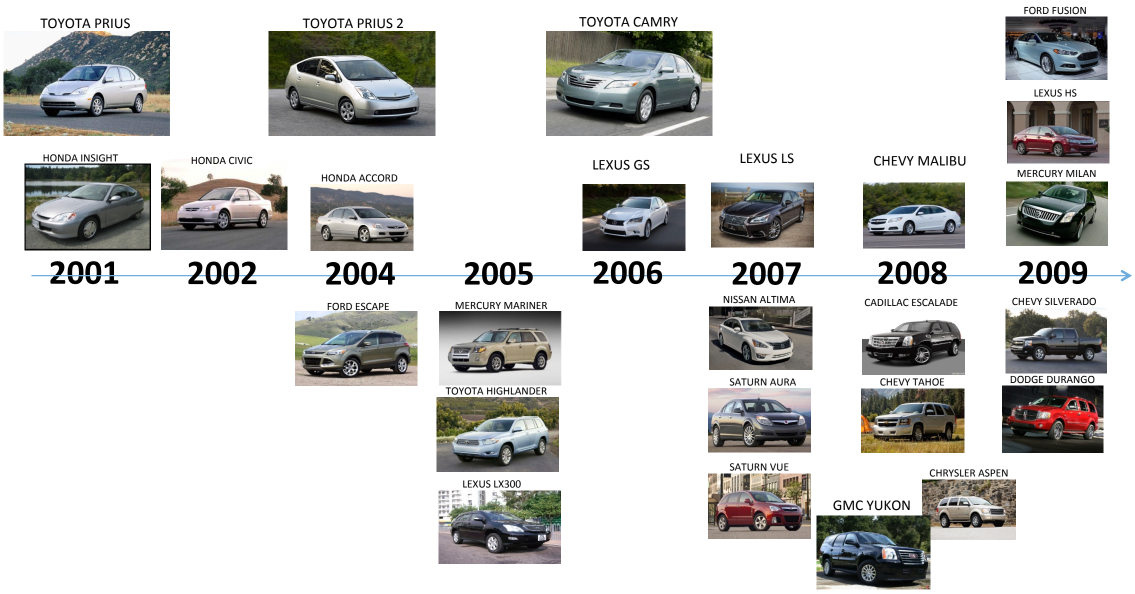

Figure 3. HEV model introductions over time. Cars above the line. Trucks/SUVs (not included in the estimated set) below the line. Larger images correspond with larger total market share.

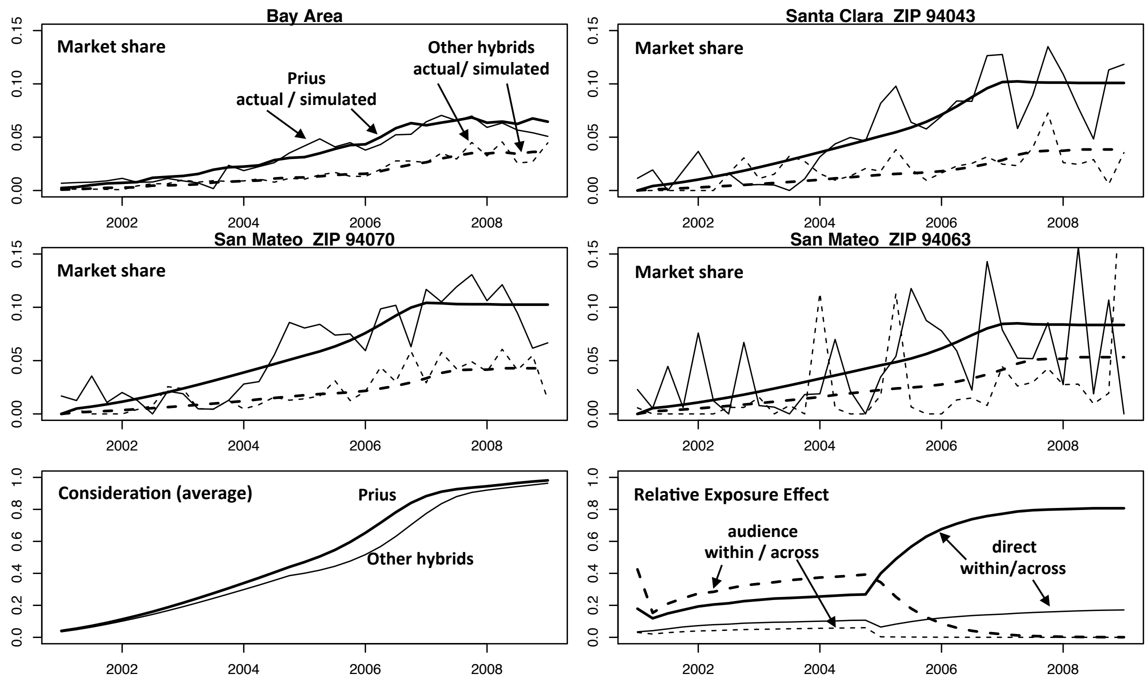

Figure 4. HEV in Bay area and within selected zips, actual and simulated; simulated consideration and social influence effects (4c)

Table 1. Estimation results

	Model 1 product diffusion	Model 2* category diffusion	Model 3 direct social exposure	Model 4a change in 2005	Model 4b change in 2006	Model 4c change in 2006; Referent effect
α_{fix}	1.023 (1.011,1.029)	0.879 (0.871,0.887)	0.856 (0.848,0.865)	0.816 (0.809,0.823)	0.830 (0.823,0.838)	0.840 (0.833, 0.847)
m	0.230 (0.227,0.242)	0.086 (0.077,0.094)	0.063 (0.057,0.070)	0.040 (0.032,0.047)	0.018 (0.009,0.027)	0.027 0.018,0.035)
ρ	0.997 (0.976,1.025)	0.812 (0.784,0.841)	0.772 (0.744,0.800)	0.749 (0.725,0.773)	0.796 (0.766,0.826)	0.767 (0.739,0.795)
e_m	0.010 (0.010, 0.011)	0.005 (0.004,0.006)	0.000 (0.000,0.000)	0.009 (0.008,0.010)	0.000 (0.000,0.000)	0.000 (0.000,0.000)
e_a	0.043 (0.041,0.044)	0.096 (0.092,0.100)	0.090 (0.087, 0.092)	0.046 (0.040,0.051)	0.089 (0.086,0.092)	0.096 (0.092,0.100)
γ_a		0.575 (0.513,0.636)	0.511 (0.471,0.550)	0.400 (0.284,0.515)	0.487 (0.433,0.540)	0.332 (0.273, 0.390)
e_d			0.023 (0.022,0.025)	0.029 (0.027,0.031)	0.028 (0.026,0.031)	0.025 (0.023, 0.028)
γ_d			1.248 (1.059,1.437)	1.017 (0.888,1.145)	1.217 (1.042,1.393)	1.174 (0.994,1.355)
e_a change				0.745 (0.643,0.847)	0.919 (*,*)	0.873 (*,*)
γ_a change				0.061 (-1.022,1.145)	0.075 (*,*)	0.086 (*,*)
e_d change				1.415 (1.399,1.432)	1.836 (1.782,1.890)	1.880 (1.852,1.908)
γ_d change				0.374 (0.304,0.444)	0.460 (0.307,0.612)	0.526 (0.414,0.637)
Referent effect (Prius)						1.479 (1.304,1.653)
logLikelihood	686254	686474.4	685080.7	684695.1	684549.2	684455

*) no convergence reached

Table 2. Within-zip code correlations between Prius and other hybrid adoption shares with counterfactual tests

Test	Correlation
Actual	0.927
Model 4c	0.937
No change	0.950
No direct spillovers	0.882
No audience spillovers	0.953
No spillovers	0.780
No within	0.943

Test	Correlation
Actual	0.927
Model 4c	0.937
Rho = 1	0.961
Rho = 0.5	0.867
Rho = 0.25	0.610