Agent Based Simulation for modelling the distribution of online music

Stephan Schuster and Nigel Gilbert
Centre for Social Simulation
School of Human Sciences
University of Surrey
Guildford, Surrey GU2 7XH
UK
s.schuster@soc.surrey.ac.uk and n.gilbert@soc.surrey.ac.uk

Abstract

The online content market for music is changing rapidly with the spread of technology and innovative business models. It is difficult for suppliers of online content to anticipate these developments and the effects of their businesses. The paper describes a multi-agent simulation to model possible scenarios in this market and argues that agent-based modelling can be a useful tool in thinking about future developments in these markets. It demonstrates this by applying the model to two simple scenarios of interest in the domain, the disintermediation of the value chain in the internet and the lock-in of consumers to Apple's iTunes download platform.

1. Background

Business models for online content production and distribution are still new and frequently changing.

Online music distribution takes place in a very dynamic environment where many factors are unknown. For example, it is impossible simply to transfer existing knowledge about consumer behaviour from the offline to the online domain The typical online customer is (still) not comparable with offline customers – they are usually innovators, rarely bound to specific brands names etc. and as a result, policies that used to work offline do not work online.

The introduction of new technologies, distribution and marketing strategies is therefore risky and the success of such strategies depends on external factors that are often hard to evaluate and in dynamic environments can be very short-lived.

From a modelling perspective, it would be useful to reflect the many different consequences of strategic choices. There may be many possible different outcomes of market processes, depending on the behaviour of the actors.

The work reported here was conducted as part of a larger project[1] whose main aim is to develop simulation tools that enable stakeholders in the online music and news content sectors to explore possible future developments and scenarios.

These simulation tools use agent based modelling (ABM). Very broadly speaking, agent based models simulate phenomena by specifying the actors and their behaviours in the system. The outcome emerges from the simulation due to the actions and interactions of the agents.

Compared to other modelling techniques (for example, differential equations) agent based modelling is particularly appealing for the purpose (see also [1]):

- *Bottom-up approach*: Knowledge or assumptions about the domain's driving factors are the more important input, not detailed hypotheses and equations about relationships on the aggregate level as required by other approaches (e.g. System dynamics). Instead of defining these relationships in advance, computers can be used to systematically explore and characterise these relationships.
- *Natural way of modelling*: Agent based models are described in terms of actors, behaviours and events. They allow non-experts in modelling to think about their domain in a much more realistic way.
- *User friendly output:* Simulations produce data on the individual level and these data

[1] Simweb (http://www.simdigital.com/), sponsored by the European Union, contract IST-2001-34651.

can be analysed in the same way as any other individual-based data sets.

Apart from these general benefits of ABM, it is possible to analyse results on both the individual and aggregate levels, which can be useful when looking at business strategies where one might be interested in the interdependency between the business environment and business strategy. Many concepts of strategy build on this dichotomy (e.g. [2]) and see strategy as the result of adaptation of firms to their environment. Processes in this framework are by definition non-linear and usually adaptive – individual behaviour changes as the environment changes. Often the agents initiate these changes in the system themselves. Agent based modelling can account for such complex interactions by endogenising them into the behaviours of agents.

2. A model of the online music market

The objective of the project for which this model was built was very broadly defined - to investigate online content markets, find out how they are shaped, what the driving forces in market development are, and where these developments might lead.

Qualitative research into the domain and discussions with business experts ([3]) yielded some general key features including:

- Increasing level of competition: More and more firms are entering the market, often imitating successful business models
- Reluctance of internet users to pay for downloads
- Technological changes influencing demand, e.g. the availability of new music playing devices, or the expected influence of broadband penetration on Internet purchases
- Consumer behaviour and preferences are largely unknown and unpredictable
- On a more general level, it has been argued (see, e.g. [4]) that the technological nature of the Internet may alter the traditional supply chain: the new possibilities of direct contact and negotiation between end consumer and producer could make intermediaries superfluous.

The domain is modelled as a simple market where consumer and provider agents trade products – in this case online music. Realised trades depend on certain product characteristics and consumer characteristics that are assumed to influence purchase decisions, for example (preferences for) download formats or pricing. As the distribution of these characteristics is unknown, one possibility is to model them as random; but the model also allows the parameterisation of the demand side.

The framework used to represent the overall dynamic of the market and strategic decisions taken by companies is based on Porter's generic strategy model [6]. He distinguishes three major strategies, depending on the target scope (narrow or industry wide) and a firm's strength (either in minimising costs or in product differentiation, including features such as brand name):

1. The cost leadership strategy tries to gain advantage through lower costs at a given level of quality. Although usually aiming at a broad market, it is also possible to apply this strategy in narrow segments. The risk is that competitors may be able to find similar ways to offer their products cheaply, so that the advantage can be lost after some time.
2. The differentiation strategy offers unique product attributes that make customers perceive the product as more valuable than those of the competition. The firm may be able to charge a premium for this added value. However, firms using this strategy can be easily imitated and can quickly loose their advantage.
3. The focus strategy concentrates on a narrow market segment and tries to achieve either a cost advantage or differentiation. By concentrating on a single segment, the advantage is usually a higher customer loyalty that discourages other firms from competing directly. Changes in customer tastes may make the advantage obsolete.

The argument is that, to be successful, one of these three generic strategies has to be applied, otherwise the firm will fail to gain a clear advantage. However in practice we find many combinations.

Porter's classification was used to define a two-dimensional strategy space (building on the ideas of modelling strategies in [7]) in which consumers and firms can be located. One dimension represents the target scope. At one extreme are located firms aiming at a certain niche segment in the market and differentiating between the customers they want to attract. At the other extreme are firms trying to offer something that is similarly appealing for most consumers. This dimension can be interpreted as a continuum. For example, many companies targeting the whole market often also have certain specialisations, such as music in specific genres or from certain regions. The second dimension represents the choice between a cost efficiency based and a product differentiation based strategy. Again, this dimension can be interpreted as a continuum; for example companies may choose to offer a very cheap, but standardised product, a medium-priced product with some distinctive features, or a real premium product that might involve high cost (such as exclusive pre-releases that may require expensive licenses).

Consumers may be defined in the same space, with a different interpretation of the dimensions, but with the same logic: There may be customers looking for cheap, standardised products (bargain hunters), or customers willing to pay a premium for a product provided that it provides unique value to them. Depending on their taste for certain genres (for example classical music), they can be classified as either being in a niche segment or belonging to the 'mainstream'.

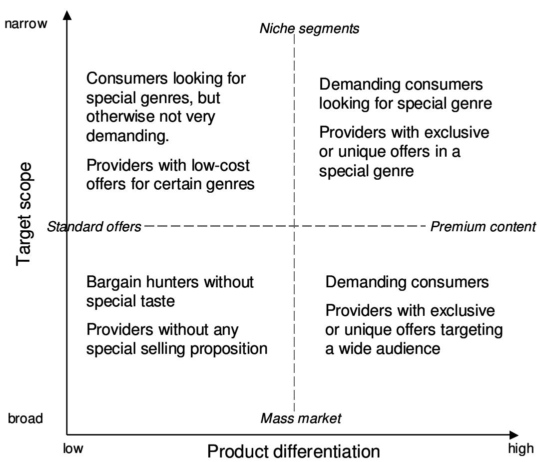

Figure 1. Strategy space in which firms and consumers interact

Although invented for offline markets, it has been argued that these dimensions capture the environment of online companies sufficiently to be useful [8], because the firms are acting in a competitive market environment. Strategies unique in the Internet environment are built in into the model, augmenting this general framework, and are strategies for cooperation [9]. Two cases can be distinguished: First, temporal cooperation among competing firms to speed up new product and services development and to share possible risks in an uncertain environment ('co-optition'). Second, often longer term cooperation between companies that stand in less competition to each other, e.g. they target a similar customer base, but provide different services, for example Internet Service Providers and content providers.

The outcome of simulations with this model (e.g. how many firms are successful in terms of market share or profits) depends mainly on the specification of the scenarios to be simulated, for example by defining how strongly the technological changes reflect upon consumer decisions, how strong the competition is or how demand varies. In short, the model sets the boundaries of the system. It defines the (possible) behaviour of agents by the modeller imposing constraints on the strategies adopted by firms and allows manipulation and tuning of the parameters that generate the result.

3. Model implementation

Consumer agents have a set of preferences. A preference is defined as an ideal product a consumer wants to buy. Each product is represented as a bundle of attributes of the product itself and the provider it is bought from, and each attribute is weighed according to the importance it has to a customer as a feature of the product. Consumers are fully and perfectly informed about the market and search it for platforms maximising their utility.

The utility of an agent depends on several factors that are formalised in two satisfaction functions based on the following assumptions. First, it is assumed that consuming music is a need whose satisfaction competes with other needs. How much music an agent wants to consume depends therefore on the relative importance of this need compared to other needs – music fans will certainly spend more time listening to music than the average consumer. Furthermore, it is assumed that the perceived quality influences the quantity demanded: The higher the quality the quicker a consumer reaches a high satisfaction level. These facts are captured in the two satisfactions,

$$\text{satisfaction}_{music} = 1 - e^{-\alpha*\gamma*q}$$

and

$$\text{satisfaction}_{other} = e^{-\left(\frac{1-\sigma}{c}*\frac{1}{\beta}\right)*q} .$$

Both functions have a concave shape as common in consumer theory (e.g. [5]) – satisfaction increases stronger with the first units consumed, and reaches later a saturation level at which additional units contribute only marginally to the satisfaction level. In the first function the satisfaction level is dependent upon the intrinsic value α an agent has for downloading and listening to music, the perceived quality γ and the quantity q. α and γ vary between 0 and 1. The larger the product of α and γ, the steeper the curve and the less quantity is needed to become satisfied. Whereas α is a fixed consumer attribute, γ is calculated as the degree to which the offers of a platform match a consumer's preferences[2]. The larger this value, the better the match and consequently, for a given α, higher satisfaction with fewer songs is possible. In the second function the competing set of needs are summarised, where the (fixed) parameter β is the valuation of doing things other than consuming music, and q the respective quantity (q could also be interpreted as time units). It follows the opposite direction of the first function as higher satisfaction on one dimension is traded off against lower satisfaction on the other. The term $(1-\sigma)/c$ is a means to vary the relation between the two curves as a function of a parameter σ, $< 0\ \sigma < 1$, and a constant c, which is used to calibrate the final individual demand. σ can be used to represent exogenous factors that reduce (as $\sigma \rightarrow 1$) or increase (as $\sigma \rightarrow 0$) the tension between downloading music and other activities. For example, a faster internet connection allows consuming the same music in less time. Assuming that this does not reduce the total time spent on downloading music, higher satisfaction can be achieved without compromising on other activities - the relative importance of competing needs decreases, and this is translated into a flatter curve for these needs.

Individual demand is calculated in two steps. First, the optimal quantity a consumer wants to download, D_{opt}, can be computed as a function of the two satisfaction functions by solving for q. In this model, it is the approximate intersection of the two curves: Consumers aim at reaching similar satisfaction levels in all their consumption activities (of course there are many other ways like maximising the product of both). In a second step, the actual demanded quantity D_{act} is calculated by reducing D_{opt} until its price fits the customer's budget. Using this information, the agent calculates the actual satisfaction that can be achieved at the various platforms, ranks them and chooses the provider at the top of this list. The act of downloading is implemented by sending a message to the respective platform containing D_{act}.

Consumers in the different segments of Figure 1 can now be created by assigning agents different preferences, budgets and taste parameters. Preference attributes represent the two dimensions of the strategy space, like certain music genres in the target scope dimension, and additional features on the differentiation dimension. γ is dependent on the type of offers during the simulation, i.e. the initial set-up of the suppliers and the strategies these apply during the simulation. Thus, if products of different suppliers are similar, customers decide only on the basis of price, independent of their preferences. However some consumers - provided they have a high enough budget - may find themselves better off by buying fewer, more expensive songs from suppliers who can offer them higher subjective value (as $\gamma \rightarrow 1$ and β is constant, the intersection of the two functions shifts to a lower q). For example, a heavy downloading low budget customer could be defined by a low budget, a low value for α, a high value for β, random preferences for different genres and low preference attribute values and weights for additional features, which places him/her into the bargain hunter segment. Such agents will choose their supplier mainly based on price, and are likely to switch quickly when cheaper offers appear in the market. On the other hand, consumers in the other segments (high preference values and weights for added-value features and/or rarer genres) might consider paying a premium for high-quality offerings and/or become loyal to specialised suppliers.

The adaptive nature of firm behaviour described in the conceptual model is implemented using a reinforcement learning approach (e.g. [11]). In reinforcement learning, agents choose exactly one of several different actions according to a stochastic rule. After executing an action, an agent receives a feedback from its environment which it uses to update the probability with which an action is selected in the following time steps. This idea has already been applied in economic contexts, where choice can be seen as an action selection problem in repeated situations (e.g. [12]), and has a clear application in the business strategy domain: Firms continuously have to choose among different operations over the course of time in order to adjust to a complex and changing environment. Here, the environment of the firm is composed of a consumer population whose demand functions are unknown and context-dependent, and the actions of other companies. In the reinforcement learning context, strategy can then be seen as a process that combines a subset of successful actions out of a large space of possibilities over time, dependent on the state of the environment.

[2] Using a fuzzy matching algorithm provided by iSOCO, one of the project partners. It calculates scores based on the weighted difference between preference and product attributes.

More specifically, firm agents dispose over a set of actions {0...i…n} that modify their position in the strategy space. Analogous to consumers, this position is computed by the attribute values of their product offerings that make up their position in the target scope and product differentiation dimensions. These attributes have numeric value ranges, and one action can change the concrete value of exactly one attribute by 0.1.
The probability of choosing an action i is determined by the selection rule

$$p(i) = \frac{e^{s(i)/\alpha}}{\sum_j e^{s(j)/\alpha}},$$

where s(i) is called the strength of action i, and α is a learning parameter. The strength can be calculated in various ways. In this model, it is the normalised change in profits that occurred after the action was applied,

$$s(i) = \frac{profit_t - profit_{t-1}}{profit_{t-1}}.$$

This selection rule thus assigns a probability to each action based on the payoffs it achieved in the past; the higher the payoff was, the higher the selection probability (using an exponential form has the more general advantage that it allows negative rewards and assigns positive probabilities also to actions not yet applied). The learning parameter α, $0 < \alpha < 1$, controls the speed of learning by weighting differences in the respective action strengths; the smaller this value, the stronger the weight of small profit changes. α may be set to larger values to represent errors in strategic choices or encourage the experimentation with different actions. In certain conditions larger values may be the better option, for example where the environment changes frequently (e.g. a high rate of market entries), and sticking too long to formerly successful actions can reduce competitive advantage. However, in the scenarios described below, α was constantly set to 0.1, which has proven to find a reasonably efficient combination between exploration and exploitation (see, e.g., [11]). An agent can also choose the special action of 'doing nothing' if it does not want to modify its position anymore. If this action is becoming the prevalent action during the simulation, the model can be said to be converging to an equilibrium state.

Porter's strategy ideal types can now be defined. Analogous to the consumer population, the attributes of their initial product offerings place agents into the strategy space described above. Then, different actions can be assigned to company agents, operationalising different strategies. For example, the low-cost strategy could imply only actions focusing on price changes, whereas a differentiation strategy could be restricted to actions modifying only on special feature attributes, possibly treating price as constant. Or, as the first scenario below demonstrates, all firms of a certain type start with the same actions, but may develop different strategies over time or under different conditions.

4. Two specific Scenarios

The following two scenarios illustrate how the agent-based model described above may generate very different results if applied to specific simulations. Also, due to the bottom-up approach of agent-based modelling they also show the possibility to analyse these results on different levels of aggregation.

The two scenarios have been chosen on the basis of existing cases to illustrate how agent-based models can formalise and reproduce existing social or economic phenomena that are often hard to formalise with mathematical or statistical methods, and how it is possible to use these scenarios as a starting point for what-if analyses that can explain the singular observation they were derived from.

4.1. Disintermediation

For the first scenario, one question of interest was how the nature of the Internet may or may not affect the traditional supply chain of companies active in e-business in general.

Arguments from current literature are twofold: a common assumption was that a business model combining content production and distribution within a single firm would be the most profitable – the so-called disintermediation thesis. With the advent of the internet, the distributor in middle is no longer necessary and the savings can be partitioned between the producers and/or consumers. More recent evidence (e.g. [10]) suggests (re-) intermediation. Due to increasing competitiveness in the end-consumer market, more and more resources have to be spent on attracting new and keeping existing customers (which may be further exacerbated by the relatively low degree of loyalty of Internet users). It can be argued (e.g. [8], [10]) that handling both aspects of selling online content will lead to inefficiencies and that producers should concentrate on their strengths, i.e. providing content, and look for profitable partnerships to reach end-consumers indirectly.

This scenario can be modelled by formalising the choices the different stakeholders, i.e. content producers and intermediaries (content distributors), face, and the factors by which these choices are influenced. For content producers it is a decision between increased spending in establishing their own distribution channels or finding a partner to sell their products, and this choice depends on the expected returns. If distributing by oneself becomes too expensive, a search for partners should be observable. For distributors, the model assumes that they can only buy from producers and sell to end-consumers and that they have no option to alter the way they do business.

Three types of agents are defined to set up the scenario: end-consumers; intermediaries selling to end-consumers and buying from producers; and producers selling to intermediaries and end-consumers. Producers can choose between concentrating on selling their content to intermediaries or selling to end-consumers directly. Profit is generated from end-consumers who buy content products; depending on the decision of producers this profit is accrues to producers and/or intermediaries.

In their role as sellers to end-consumers, content producers differ from intermediaries to reflect the fact that intermediaries are specialised in distributing and marketing, whereas producers are specialised in content production. This difference is implemented by a higher 'presence' of resellers in the end-consumer market, e.g. because they run expensive awareness campaigns; specialise in certain areas or offer additional services that attract more Internet users.

The process determining how the market is shared depends on the calculations of producer agents who follow the logic described above. By observing the success of their strategy they learn over time which direction is the better one. Thus, what is a more feasible strategy depends mainly on what can be seen as external conditions to the agents – how many customers they can reach and how many intermediaries exist in the market. For example, if every producer cooperates with a distributor, it can become unprofitable for a single producer to invest in its own distribution, as the reach of specialised distributors in the end-market limits the possibilities to gain a substantial share.

Thus, the result depends on the environment content producers find themselves in. To keep the analysis of this example scenario simple, simulations are run holding the providers' behaviour constant. A measure of competitive pressure is defined as the ratio of firms per customer in the market (the actual magnitude is of no concern here as the measure is only of interest to compare settings as more or less competitive). Simulations in environments of different competitive pressures are reported below.

Running the simulation at one of two extreme settings – either very competitive or with very low competition – can serve as a convenient illustration of how different outcomes can develop from the same set of initial strategies:

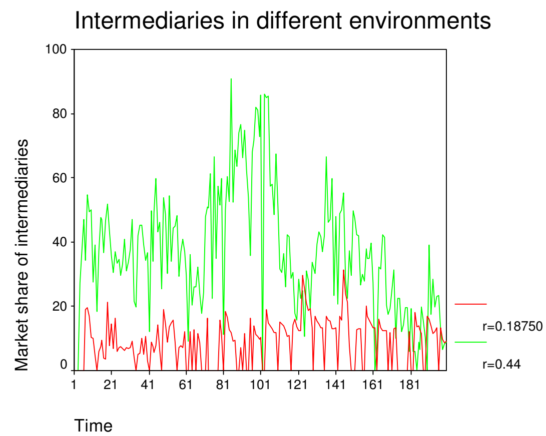

Figure 2. Market share of intermediaries in the end-consumer markets as a function of competitive pressure. The parameter r is the ratio of firms per consumer. Low r means low pressure

Figure 2 shows the results of running the simulation twice: once with a low competitive pressure (the low line) and once with high competitive pressure. The figure plots the mean market share of intermediary firms over 200 time steps (the correspondence between model and real time is arbitrary, but can be thought of being calibrated in weeks). It can be seen that while the mean share of intermediaries fluctuates considerable, as the firms react and counter-react to the changing market environment in which they find themselves, with low competitive pressure, the intermediaries almost always have a lower market share than they do in a high competitive pressure situation.

As the process depends on the decision of producer, it can be concluded that with few competitors (or the potential customer base large enough), end producers have no incentive to share the end market with intermediaries, despite possible efficiency gains or consumer benefit (e.g. by more differentiated products); conversely, tougher conditions support intermediation.

Another way of looking at the model is to analyse individual behaviour and strategies. The following scenario looks closer at end-consumer market and asks how the success of firms following different strategies may depend on the interplay of these strategies over time. It was intended to reproduce the surprisingly quick success of Apple's low pricing model, but it may also serve to illustrate some wider implications of the process behind the phenomenon.

4.2. Lock-in

The case motivating the implementation of this scenario is very simple: it was commonly assumed in the online music community that paid-for downloads would not be popular. Before Apple introduced its iTunes download platform, few commercial offers existed, and these were not very successful. Whereas the music industry at that time blamed piracy as the only reason for the problem, it turned out with Apple's pricing model that consumers are willing to pay, but not at the price level the industry was wishing for. But there are other, more subtle aspects connected with this business model: Because of fixed license fees per song, margins are very low. The main advantage for Apple of offering the download platform seems to be that it is a proprietary format that is usually played with the iPod, the player manufactured by Apple. In other words, the cheap music that customers want to download can only be played using the proper hardware.

The second scenario is a less complex account of this phenomenon. It is assumed that there are consumers who are prepared to pay and some providers. However, there is a large difference in the price that the customers are willing to pay and the price at which the providers are willing to offer downloads. Nevertheless, consumers are willing to pay a premium at download platforms that offer them additional value, for example by providing extras like background information, or specialising in certain music genres. This is implemented by placing them in different regions of the strategy space (Figure 1). Providers are placed in this space in an analogous way, but in some distance to consumers. As described above, consumers and providers are matched according to their distances in that space.

The resulting sales made by providers depend on how well they manage to 'convince' consumers to pay a relatively high price for their offerings. To do this, providers can adapt their offerings over time by applying the generic strategies in a pure or mixed way. To allow for heterogeneity in the search for an optimal strategy, three different types of providers are defined. One type follows the niche marketing strategy, the two others a differentiation strategy, i.e. there is initially stronger competition between two providers aiming at the same target segment of the market than between them and the niche player.

What is investigated is how the situation changes with the introduction of a new competitor emulating Apple's low-price strategy. The download products are offered as the selling vehicle for the hardware on which the music is to b played, and customers have to buy this device in order to benefit from the cheap downloads. Buying this device is expensive, so that there is a tendency for consumers to become locked in; once they have decided to buy music in a proprietary format, the difference in expected utility between giving it up or choosing yet another proprietary format has to be large enough to persuade them to abandon the device they hold.

Running two simulations, one with an early and one with a late market entry, the model reproduces the success of the low-price business model.

The first simulation is run with the new competitor entering the market at its formative stage. In this case the early market entry targets consumer concerns – realistic prices – successfully. The market share of the new provider rises quickly and stabilises due to the lock-in effects (see Figure 3).

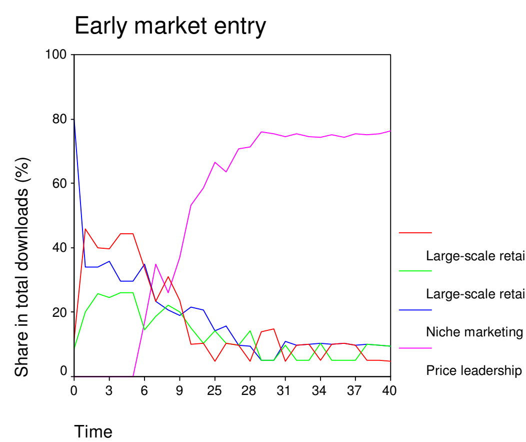

Figure 3. Share of downloads for different strategies, with the new competitor entering the market at an early stage (time step 5)

Given the assumptions of the model, the second simulation provides an example of how things could have developed if existing providers had given more attention to their potential customers' needs. In this simulation, competitors were given more time to explore the market (using Porter's recommendations). Although the effect remains large, it is well below that in the first scenario, and the shares of the market are more evenly distributed (see Figure 4). Over the course of time, the activities of the existing providers change the entry condition for the new provider – product uniqueness and niche marketing can stabilise their positions in the market in the long run.

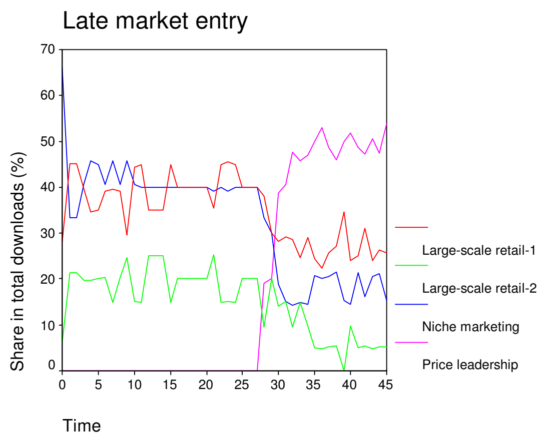

Figure 4. Share of downloads for different strategy types as in figure 3; but the new competitor enters the market later (time step 27)

5. Conclusion

These examples have been selected to illustrate the use of agent-based systems in modelling complex business environments. It has been shown that they can be analysed on different levels of aggregation and using more or less elaborated specification, depending on the question of interest.

A more accurate representation of the actual market would combine the approaches of the first and second scenario to account fully for the relationships between individual behaviour, exogenous conditions and endogenous dynamics shaping the market environment. However, this would make exploration and understanding of outcomes of the model more complicated. It is possible to simultaneously vary the environment (customer behaviour, number of firms in the market etc.) as well as the strategies. The first scenario, for example, assumes a general strategic 'short-sightedness'. If agents do not experience a rise in profits quickly enough, they will regard the strategy as not successful and explore alternatives. The outcome would certainly be different were we to vary the time-horizon of agents, i.e. modify the assumptions about their behaviour. It is however not difficult to imagine how these seemingly simple combinations and extensions will increase the difficulty of analysing the output.

The project has demonstrated how agent based models can capture complex phenomena that are difficult to model in other frameworks, especially when considering the mutual dependency and adaptation of actors as a result of the interdependencies between their different strategies, and how these models can be applied to real-world business concerns. In this, they may prove useful as a 'computer-aided thinking tool' to anticipate possible future worlds.

6. References

[1] E. Bonabeau and C. Meyer. "Swarm Intelligence: A Whole New Way to Think About Business", *Harvard Business Review*, Vol. 79, No. 5, 2001, pp. 107-114.

[2] E.E. Chaffee, "Three Models of Strategy", *Academy of Management Review,* 10 (1), 1985, pp. 89-98

[3] C. Krüger, P.M.C. Swatman, K. van der Beek, "Online News and Music Sectors in Ten European Countries", University of Koblenz-Landau, Institute for Management, January 2005

[4] J. Strauss and R.Frost, *E-Marketing*, 2001

[5] W. Jager: "Modelling consumer Behaviour", Ph.D thesis Rijkuniversiteit Groningen, 2000

[6] M.E. Porter, *Competitive Advantage*, The Free Press, New York, 1985

[7] D.A. Robertson, "Agent-Based models of a Banking Network as an Example of a Turbulent Environment: The Deliberate vs. Emergent Strategy Debate Revisited", *Emergence*, 5(2), 2001, pp. 56-71

[8] M.E. Porter, "Strategy and the Internet", *Harvard Business review*, March 2001, pp. 62-78

[9] R. Amit and C. Zott, "Value creation in e-business", *Strategic Management Journal*, vol. 22, 2001, pp. 493-520

[10] C. Krüger, P.M.C. Swatman, and K. van der Beek. 2003. "Interim Sector Reports (M 2.1) of the European Online Music and Online News Market". University of Koblenz-Landau, Institute for Management, June 2003

[11] R.S. Sutton and A.G. Barto, *Reinforcement Learning: An Introduction*, Cambridge, MA, 1998

[12] B.W. Arthur: "On designing economic agents that behave like human agents", *Journal of Evolutionary Economics* (3), pp. 1-22, 1993