Burkhard Freiherr von Wangenheim

Survival trees - a new method in innovation theory

A successful introduction of a method commonly used in survival analysis into the field of innovation diffusion theory

Anchor Academic Publishing

disseminate knowledge

Burkhard Freiherr von Wangenheim
Survival trees - a new method in innovation theory: A successful introduction of a method commonly used in survival analysis into the field of innovation diffusion theory

ISBN: 978-3-95489-504-5
Fabrication: Anchor Academic Publishing, an Imprint of Diplomica® Verlag GmbH, Hamburg, 2013

Table of contents

1 Introduction

1.1 Context of Thesis

"It is almost universally accepted that technological change and other kinds of innovations are the most important sources of productivity and increased material welfare – and that this has been so for centuries".[1]

On the corporate level, the recognition has succeeded that the implementation and maintenance of a successful innovation management is the key contribution to competitiveness and future growth. For this reason, there is great interest in understanding the processes of innovation and its subsequent diffusion to formulate appropriate policies.

Within the last decades, researchers in management and marketing science have greatly contributed to the development adoption- and diffusion theory by suggesting analytical models for describing and forecasting the diffusion of an innovation in a social system. The main reason for this has been the perceived high failure rate of new products and the consequent needs to improve the related management and marketing decisions.

The explanation why firms do not instantaneously adopt new technology immediately after its commercialisation (i.e. diffusion is a time-intensive process) can be traced to different theories of innovation diffusion advocated in literature. According to early epidemic theories of inter-firm diffusion,[2] diffusion is a disequilibrium process resulting from information symmetries between potential adopters.[3] In contrast to epidemic models, contemporary approaches to technology diffusion are characterised by the dismissal of information spreading as the key explanatory variable of innovation diffusion.[4] Rather, models in general assume that firms behave optimally (i.e. are profit maximizers) and that information pertaining to the technological and economic characteristics of the information is perfect. Within this equilibrium approach there are three categories of models that have been developed in the literature: the rank or

---

[1] Charles Enquist (1997).

[2] The economy wide-degree of diffusion can be decomposed into two elements: Inter-firm diffusion and intra-firm diffusion. Inter-firm diffusion describes a firm's first use of a new technology. Intra-firm diffusion, on the other hand, has not been researched much so far and describes the increasing intensity of technology diffusion. See for literature on inter- and intra diffusion Griliches (1957), Mansfield (1968), Bass (1969) and Hollenstein, Wörter (2004), respectively.

[3] Baptista (2000).

[4] Gourlat, Pentecost (2000).

probit; stock or game theoretic, and order effects models.

In rank or probit models[5] potential adopters of technology have different inherent characteristics and as a result obtain different gross returns from its use.

The essence of stock effect models is that benefits to the marginal adopter from acquisition decreases as the number of previous adopter's increases.[6]

Order effect models are similar to the rank effect models in that the gross returns of a firm adopting a new technology depends upon its position in the order of adoption, with higher order adopting firms achieving a greater return than low-adopters.[7]

Despite the continuing progress of contemporary approaches, the main impetus underlying diffusion research is still the epidemic Bass model[8]. Subsuming the majority of other models derived from that model or independently, this model addresses the market in the aggregate. The typical variable measured is the number of adopters who purchase the product by a certain time t. The emphasis is on the total market response rather than on the individual adopter. Here, the individual characteristics of potential adopters and their impact on the decision-process remain wholly uncovered. Not the individual who decides, whether to adopt or reject an innovation is central to the analysis, but the time-related distribution of the adoption decision dependent on marketing variables.[9] These models cannot explain why a particular individual adopts or rejects an innovation at a specific point in time. Consequently, these models achieve no adequate aggregation of individual adoption decisions. Although the specific managerial implications that these models give should not be questioned in general, they remain limited by the aggregate perspective which they take.

In fact, diffusion theory faces a constant dilemma between disaggregate and aggregate diffusion modelling. Although it is unquestionable that the diffusion process is built upon individual adoption decisions, the persuasion that diffusion models should thus be built upon individual decisions has not yet fully materialized. One reason lies undoubtedly in the substantial modelling obstacles that theory has faced so far in trying to pursue this.

---

5 Ireland, Stoneman (1986).
6 Reingannum (1981a, 1981b, 1989), Quirmbach (1986).
7 Gourlay, Pentecost (2000), p. 3.
8 Bass (1969).
9 Albers (1998), p. 13, Kühnapfel (1995), p. 121.

Most models that allow for illumination of individual adoption behaviour are static in nature, hereby failing to capture the inherent dynamics of the diffusion process which makes plausible aggregation nearly impossible. *This dilemma has forced an explicit distinction between adoption- and diffusion theory. Although this distinction is often taken to frame the sort of analysis that is performed, it is forced by the disability of most diffusion models to persuasively incorporate the naturally inherent individual perspective.*

By recognizing that the diffusion process is built upon individual adoption decisions, the adoption theory should be recognized and modelled much more as the key basement of diffusion theory rather than a theory that is conceptionally and in content different to the diffusion theory. The implication of this is that diffusion models that take the individual perspective simultaneously perform an adoption analysis.

Moreover, diffusion models based on individual adoption decisions offer an opportunity to study the actual pattern of social communication and its impact on product perceptions, preferences and ultimate adoption. Nonetheless, first attempts to establish the diffusion process on the basis of individual adoption decisions faced severe problems in realizing ultimate aggregation.[10] Merely the study by Chatterjee and Eliashberg (1989) provided encouraging empirical evidence for a useful aggregation of individual adoption decisions.[11] Indeed, it has been recognized only recently that the above described dilemma can be solved.

So-called event history data is able to capture the dynamics of the diffusion process while, simultaneously, the individual perspective (micro level) can be preserved.

Eventually, with the introduction of hazard models[12] into diffusion theory, various micro models were found that could effectively deal with event history data and thus allowed for consideration of individual heterogeneity among adopters by incorporating covariate effects into diffusion models. Up to now, most models that have come up in the widely applied field of event history analysis have been applied to diffusion theory, too.[13] It should be said, however, that these applications have taken place only recently making the use of event history data still a novel thought to diffusion theory.

---

[10] Hiebert (1974), Stoneman (1981), Feder, O`Mara (1982), Jensen (1982).
[11] Mahajan, Muller, Bass (1990).
[12] Kalbfleisch, Prentice (1980), Cox, Oakes (1984).
[13] Reingannum (1982), Hannan, Mc Dowell (1984, 1987, 19990), Sinha, Chandrashekaran (1992), Gönül, Srinivasan (1993), Caudil et al. (1995), Gourlat, Pentecost (2000), Litfin (2000).

The main reason for this may lie in the extent of data collection necessary to perform an analysis. Especially, in economic theory, where the necessity for event history data is not obvious, this may prove a vital obstacle; keeping track of each individual and his adoption decision is undoubtedly a more challenging task than simply taking the aggregate approach. Fortunately, with rising technological possibilities, the applicability of event history models has risen, too.

1.2 Contribution of Thesis

With the extension of the non-parametric classification and regression tree method (CART™)[14] to the analysis of censored event data, we are now given the opportunity to move research forward by examining usefulness and applicability of that method for the analysis and forecast of innovation diffusion. The development of the so-called "survival trees" was highly motivated by the need to develop meaningful prognosis rules in medical science.[15] As will be shown later, there are a number of essential parallels between survival analysis in medical science and diffusion analysis in economics. Emergences of new methods in that field are therefore likely to prove applicable in adoption- and diffusion theory (ADT), too.

As the CART™ method itself is still new to economic theory, it should not surprise that no known application of survival trees has taken place in an economic context so far. Indeed, even for the CART™ method only two applications in an economic context are known.[16] Both methods, CART™ and survival trees, have been developed in the area of medical science and seem to spread only slowly to other scientific areas. Economists and other non-medical scientists alike will have to be persuaded of the new insights that these methods offer. As for survival trees, this thesis is the first attempt to do this.

The method offers additional insights into causal relations that traditional methods fail to give and can therefore resemble a powerful contribution to modern diffusion theory. Its interpretational power makes it likely that this method will meet widespread acceptance.

---

[14] Breiman et al. (1984).
[15] Gordon, Olshen (1985).
[16] Haughton, Oulabi (1993), Köllinger, Schade (2004).

1.3 Structure & Internal Pattern of Thesis

I want to briefly put into words the structure of the thesis that is already summarized in the table of contents. I believe this will make it more easily understandable and more coherent. Additionally, I find it important that the reader is aware of the internal pattern underlying this thesis. With this, I mean simple formatting or used terminology decisions.

Let us start with the **structure**: In the course of the thesis, the survival tree method will be introduced within the context of ADT. For this reason, I will provide arguments in favour of dynamic micro models as a means to analyse and forecast innovation diffusion (section 2.1).

As event history data enables us to do this, I will set up the common concepts and ideas of event history data modelling just as the classical methods from this area, all within the context of ADT (2.2, 2.3, and 2.4). This will be done to grasp an understanding of the interpretation and functionality of the event history patterns within the ADT context and is considered essential for understanding the survival tree method and its usefulness in forecasting innovation diffusion.

Survival trees have been derived from CART™ and consequently both methods share essential conceptual features. After a general introduction into the CART™ methodology (3.1) and a first introduction in the area of survival trees (3.2), I will attempt to classify the proposals that have come for the construction of survival trees into three building blocks that are commonly used in the construction of CART™ (3.3).

Subsequently, the various proposals that have come up in the construction of survival trees will be evaluated and the merits just as the deficiencies of the method will be discussed (3.4).

I will describe in detail the software applications available for survival tree calculations to facilitate future work on them (4.1). The data that the method will be applied on is presented and the way data was handled is documented (4.2) before I state which of the various options was taken (4.3).

Analysing the results, we will see whether the method can offer new insights into ADT and whether the previously discussed merits & deficiencies of the method hold true or might have to be reconsidered in the discussed context (4.4).

Eventually, I will discuss the central question about the usefulness of the method to forecast innovation diffusion. I will try to relate the method's results and their implications to economic practice. Other related issues and thoughts will be discussed, as well (4.5). Conclusively, main patterns and findings of the thesis will be summarized (4.6).

Let me now explain the **internal pattern** of the thesis relating to measures that were taken to ease functionality and readability of the thesis.

The problem of inconsistent terminology is particularly apparent in event history analyses. If we take, for instance, the denomination "event history data", we can easily find at least five other denominations, all used interchangeably, which may sometimes hamper understanding substantially. I will thus name these cases when they appear and say explicitly which of the various denominations I will use. Additionally, I have developed an index of synonyms in Appendix 7.3 to prevent any confusion.

Other confusion is likely to be caused by the various denominations in ADT. No definite rule can be established as to whether one should use adoption theory or diffusion theory for a specific field under investigation. In this thesis, I claim that these two areas belong essentially together. I will therefore make no distinction between these two areas using the single denomination adoption- and diffusion theory (ADT) throughout this thesis.

Besides, there is no generally agreed structure in the area as to what model belongs to what class of models and so on. The classification of models into micro and macro, static and dynamic models is by no means generally agreed and was adapted from Litfin (2000).

For easier readability and in order to put emphasize on sentences that I consider vital, I will format respective text **bold** or *italic*. In this way, words representing important issues are formatted **bold** to enable easier localization.[17] Italic formatting is used for sentences *that I considered vital for overall understanding*.

I have noticed that the literature on survival trees has picked up momentum within the year 2003 and 2004, especially. This made it difficult to incorporate all new literature in the thesis as it was published while this thesis was written. Yet, I think I

---

[17] Bold was used for the authors of the various proposals in 3.3. because their names stand exemplarily for the method they developed.

have successfully attempted to include all literature until the end of November 2004 in the thesis.

Sometimes, I will sum up findings or provide a brief outlook at the very beginning of a section. I do this to make sure one does not lose track of the findings and is always aware of why a certain section was written.

2 Modelling Censored Event Data in the Context of Innovation Adoption- and Diffusion Theory

In virtually every area of the social sciences, there is great interest in events and their causes. Criminologists study crimes, arrests, convictions, and incarcerations. Medical sociologists are concerned with hospitalizations, visits to a physician and psychotic episodes.[18] As a field of economics, innovation theory investigates and tries to predict the effects of innovations on society. Hereby, the adoption decisions of the members of society play the decisive role.

In each of the above mentioned examples, an event consists of some qualitative change that occurs at a specific point in time. Because events are defined in terms of change over time, it has become increasingly acknowledged that the best way to study events and their causes is to collect event history data.[19] In its simplest form, event history is a "longitudinal record of when events happen to a sample of individuals or collectivities"[20].

In this chapter, I will provide reasons why innovation diffusion analysis and forecast should be performed on the basis of dynamic micro models. These models can be established only on the basis of event history data. As all models from the area of event history analysis are either directly or indirectly based on the hazard rate framework, I will establish this framework to ease understanding of the upcoming presentation of the various parametric, semi-parametric and non-parametric models.

For the upcoming introduction of the survival trees, *it is important to understand the conceptionel parallels between diffusion theory and survival analysis*. These parallels allow us to use models coming from the area of survival analysis for ADT.

---

[18] Allison (1984) p. 9.
[19] Alternatively, data is collected as cross-sectional or panel data. For a comparison of these approaches with event history data collection see Blossfeld, Rohwer (2002), pp. 4-6.
[20] Allison (1984) p.9.