



AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
<http://ageconsearch.umn.edu>
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

Hagemann, M., Francksen, T., Latacz-Lohmann, U.: Determinants of internal farm growth in milk production: An investigation using event history analysis. In: Balmann, A., Glauben, T., Graubner, M., Grings, M., Hirschauer, N., Schaft, F., Wagner, P.: Unternehmerische Landwirtschaft zwischen Marktanforderungen und gesellschaftlichen Erwartungen. Schriften der Gesellschaft für Wirtschafts- und Sozialwissenschaften des Landbaues e.V., Band 47, Münster-Hiltrup: Landwirtschaftsverlag (2012), S. 405-419.

DETERMINANTS OF INTERNAL FARM GROWTH IN MILK PRODUCTION: AN INVESTIGATION USING EVENT HISTORY ANALYSIS

Martin Hagemann, Tammo Francksen, Uwe Latacz-Lohmann¹

Abstract

This paper investigates determinants of growth of milk production in German dairy farms. Event history analysis is applied to estimate the likelihood of a farm's moving from a non-growing episode into a growing episode and to assess the impact of various covariates on that likelihood. The analysis is based upon a balanced panel of annual farm accounts from 616 specialised dairy farms from Germany, covering the financial years 1995/96 to 2008/09.

Keywords

Event history analysis; dairy farms; internal growth; determinants of growth

1 Introduction

Motives for farm growth are diverse. They are to some extent inherent to the economic system because, for example, other firms in the supply chain are growing as well and abet bigger batch sizes. Motives are also strongly related to the farm owner's attitude which might be driven by the farm's succession. Moreover, the political framework condition may determine farm growth. However, from an economic perspective the main objective of growth is supposed to be the improvement of the competitiveness of a farm. Hence, it should be a continuous process in order to bring forward technical effort on the farm, making the best use of production factors and thus gain economies of scale.

The political framework of the German milk market is determined by a common market organization for milk of the European Union, which was introduced in 1968. Due to huge overproduction a milk quota was implemented in 1984. Since then the rules for trading the quota changed three times. From 1984 to 1990 the milk quota was linked to land and thus could only be transferred with land. Furthermore, a couple of additional requires like a maximum production per hectare needed to be full filled. The second 'epoch' was from 1990 to 2000 when renting, leasing and sale of milk quota decoupled from farmland was allowed. Since 2000 a milk quota auction is implemented in Germany. There are three trading events per year. In the beginning Germany was divided into 21 trading regions and thus only regional trade of quota was possible (BRÜMMER ET AL., 2003). In 2007 the number of regions was reduced and there are now two trading regions: one in eastern and another one in western Germany. Consequently, milk production capacity can now move from north to south but not from east to west. Nevertheless, the European milk market is in a liberalization process which has a highlight in 2015 when the quota shall be abolished (European Commission, 2009).

Especially the latter change, the introduction of the quota auction, improved transparency of the quota market and enabled larger regional shifts in production (KLEINHANß ET AL., 2010). Hence, this structural change is supposed to be an important event affecting growth of dairy farms within the observed period of this study. Furthermore, it can be assumed that time plays an important role in explaining farming decisions, as it captures unobservable influences such as increasing competition. In the present study a balanced panel of 616 specialized dairy farms from Germany is analyzed over the period from 1996 to 2009 with event history analy-

¹ Department of Agricultural Economics, Christian-Albrechts-Universität zu Kiel, Olshausenstraße 40, D-24098 Kiel, Germany, martin.hagemann@ae.uni-kiel.d

sis². Using this dynamic econometric framework enables us to investigate the impact of farm and farmer characteristics as well as unobservable and not measurable effects on growth. The latter are referred to as the impact of time and can be characterized as forces on farming decisions which are inherent to the economic system: farmers will continually strive for cost reduction, especially in the face of declining milk prices. Besides management optimization this can best be done by capturing economies of scale through expansion.

2 Literature review

Event history analysis was originally developed and thus has been used frequently in the medical science, empirical social research and labor economics.

In the field of agricultural economics applications of event history analysis are rare so far. BURTON ET AL. (2003) point out that the dearth of applications to agricultural adoption is rather surprising as the great advantage of duration analysis is that it deals with both cross-section and time series data. There are a few studies which investigate adoption of political schemes or techniques. Some of them which have been done in the recent past shall be presented here. WYNN ET AL. (2001) investigated the probability and rate of farmer entry into environmental sensitive area schemes in Scotland using multinomial logit and duration analysis, respectively. They found a number of generic factors as important in explaining the entry decision. The duration analysis suggested several factors accelerating scheme entry: an interest in conservation, more adequate information and more extensive systems. BURTON ET AL. (2003) used duration analysis to model the adoption of organic horticultural technology in the UK. They used discrete time models to explore the influence of a range of economic and non-economic determinants. Their results highlight the importance of gender, attitudes to the environment and information networks, as well as systematic effects that influence the adoption decision over the lifetime of the producer and over the survey period. In another study DADI ET AL. (2004) examined the technological adoption behavior over time of smallholder farmers in the East and West Shewa zones of the central highlands of Ethiopia using duration analysis. They focused on fertilizer and herbicide use in the cultivation of tef and wheat which are the major crops in terms of both area and total production of that region. They found out that economic incentives were the most important determinants of the time farmers waited before adopting new technologies; traction power in the form of oxen and infrastructural factors (in particular proximity to markets) also appear to have influences, but less than prices. LÄPPLE (2010) investigated determinants that affect adoption and abandonment of organic farming of the Irish drystock sector in the period from 1981 to 2008. The results highlight that where no attempt is made to account for exit decisions and time effects, important information about sustainable farmer decisions may not be taken into consideration. Growth of dairy farms has not been analyzed with event history analyses so far. However, there are some studies in which farm size, entering and exiting farming as well as farm diversification is investigated. FOLTZ (2004) used a model of sunk costs and farm capital investment to specify two econometric estimations: a random effects Probit model of farm entry and exit and an autocorrelated generalized least squares panel data model of farm size. He found that the price strategy of the New England Dairy Compact, i.e. a formal agreement between the six New England states, which allowed for the establishment of a regional pricing mechanism for fluid milk, reduced farm exits and moderately increased cow numbers. WEISS and BRIGLAUER (2000) analyzed determinants and dynamics of farm diversification using linked census data for Upper-Austria from 1980, 1985 and 1990. With regard to farm size they found out that smaller farms are specialized and also tend to increase the degree of specialization over time more quickly than larger farms. GLAUBEN ET AL. (2003) explored regional differences in farm exit rates using

² In the literature the terminologies event history analysis, analysis of failure times and survival analysis are used synonymously.

county-level data from 326 regions in Western Germany. Econometric cross section estimates indicate that exit rates are higher in regions with smaller farms. HUETTEL and JONGENEEL (2008) analyzed structural change in the dairy sectors of Germany and the Netherlands using the Markov Chain approach. They found out that structural change in the dairy sector is faster in West Germany than in the Netherlands and fastest in the East of Germany.

3 Hypotheses on growth of farms milk production

As stated in the introduction motives of growth in milk production are diverse. In this study growth is defined as increasing in milk quota. Even if motives of growth are diverse, we consider only economic rationales in the following. There is evidence in the literature of farm economics that growth into certain farm sizes will gain economies of scale, decrease production costs and thus improve competitiveness (GOERTZ, 1999; LASSEN ET AL., 2008).

In milk production there are two ways to increase the milk output: first via increasing the milk yield per cow and secondly via herd growth. The actual average herd size in German dairy farms is 46 cows per farm with an average milk yield of roughly 6,800 kg per cow and year (DESTATIS, 2010). However, farm structures are different in Germany. In the northern part farms are bigger than in the southern part of Germany. Hence, there might be regional differences in the motives for farm growth, the way how to grow and in the growth rates.

Based on this background and with regard to the available information on the underlying farm data set the following hypotheses (H) on the development of growth over time and the analyzed covariates for growth of dairy farms were made. The event of growth is defined as the time when a farm leaves the “non-growing episode” (NGE). In terms of event history analysis exiting the NGE (origin state) is equivalent to entering the “growing episode” (GE) (destination state).

H1. The probability of entering the GE increases over time: This hypothesis is based on the assumption that growth is up to a certain degree inherent to the economic system and, for example, other firms in the supply chain abet bigger batch sizes over time. Additionally, it is supposed that a farm needs to improve its competitiveness over time in order to compete on scarce production factors, e.g. land.

H2. Northern farms grow more frequent than farms in the center and south of Germany: The analyzed farms are clustered into three regions: north, center and south. Due to more favorable farm structures for milk production in the north of Germany, especially with regard to land, farms tend to grow more frequent and/or with higher grow rates.

H3. The greater farms milk production in terms of quota the shorter the stagnation episode: Increasing farm sizes abet exploitation of economies of scale going along with realizing cost saving potential. As a result a greater farm tends to feature a higher competitiveness and thus is more able to defray internal farm growth.

H4. The larger the share of own land on total land the higher the probability to enter a GE: A high share of own land on total farmland indicates that the farm has collaterals and thus is supposed to take out a loan at more favorable conditions than farms with less own land. Moreover, the farm has to pay less land rents and therefore faces a better liquidity in times of low milk prices.

H5. The higher the share of grassland on total farmland the longer lasts a stagnation episode: It is assumed that farms in grassland regions feature lower quality of forage in terms of energy content than farms in areas with arable fodder production. Farms in grassland regions tend to have lower milk yields from forage. This makes a cost-efficient milk yield increase difficult. Hence, the probability to enter a growing episode is assumed to be lower.

H6. Higher increase of average milk yield per cow and/or number of cows per farm induces growth in terms of milk quota: Typically, small growing steps were done without purchasing

quota (or until the year 2000 leasing and renting were also possible). Hence up to a certain degree, farmers take the risk to pay a fee for over quota produced milk. However, reaching a certain level of over quota production via milk yield increase and/or herd growth, the risk of paying the fee accelerates a growing step by purchasing milk quota.

H7. A high share of subsidies on total returns decelerates growth: Incentives to improve competitiveness via cost decreasing by exploitation of economies of scale dwindle as a result of higher direct income support.

H8. Younger farmers have a higher incentive to grow: Older farmers tend to have a lower readiness to assume risks. Consequently, in opposite to future-oriented younger farmers, older farmers have a lower motivation to withstand competition pressure by internal growing and its associated risk.

H9. A higher level of agricultural education advances growth: Economic success is, inter alia, based on managerial skills likely going hand in hand with a high education level. Consequently, a solid agricultural education is an important fundament for sustainable growth.

H10. A milk price increase encourages investments in milk production: A considerable milk price increase during a NGE accelerates investments due to a better liquidity. It is further expected that episodes of high milk prices are bringing confidence with regard to the attitude to future perspectives of farmers what probably abets decisions for growth.

H11. Fee for over-quota production reinforces incentives for growing in milk production: A high expenditure for over-quota production indicates that some small growing steps, e.g. via increasing milk yield, have been done without extending quota. In order to avoid or reduce the fee farmers have an increasing incentive to expand their milk quota.

H12. Farms with a higher income per kg milk are more likely to grow: Farms with a relatively high farm income per kg milk are utilizing their capacities more efficient. Due to a higher profitability these farms are more likely to defray growing steps.

H13. The introduction of the quota auction in Germany in the year 2000 supports internal growth: The quota auction brought the possibility to demand three times per year batches of milk quota. Purchasing milk quota has become easier after the introduction of the quota auction in 2000.

4 Methodology

The set of statistical models used to implement event history analysis can be differentiated in three approaches: nonparametric estimation methods, parametric models, and semi-parametric transition rate models. Nonparametric estimation methods are more likely of a descriptive nature and can be used to describe the process under study. These methods are especially helpful for graphical presentation of the survivor function as well as the transition rate.

Parametric models are a general statistical technique through which one can analyze how the transition rate is dependent on a set of covariates. Furthermore, parametric models permit the analysis of the impact of duration dependence and the influence of one or more parallel processes by the use of time-dependent covariates (BLOSSFELD and ROHWER, 2002). In order to illustrate the impact of different parametric model specifications on the estimation results of our empirical study and to compare their outcomes, we select four different methods out of the group of various parametric approaches: the *Exponential Transition Rate Model*, the *Piecewise Constant Exponential Model*, the *Gompertz Model*, and the *Log-Normal Model*.

Parametric models are based on specific parametric assumptions about the distribution of episodes durations. Time in these models normally serves as a proxy variable for a latent causal factor that is difficult to measure directly. Applying a specific parametric model in empirical event history analysis often causes the problem that the implicitly made parametric assumptions cannot be adequately justified by strong theoretical arguments. BLOSSFELD and ROHWER

(2002: 176ff.) therefore suggest to use these models with extreme caution. In their opinion, estimating a variety of model specifications and comparing the results seems to be an appropriate strategy. An interesting alternative to the parametric model is the use of semi-parametric models. These models only require the specification of a functional form for the influence of covariates, but leave the shape of the transition rate as unspecified as possible. The most widely applied semi-parametric model is the proportional hazard model by Cox (1972), also referred to as *Cox Model*. As an alternative to the above mentioned parametric methods we also apply the Cox model in this study. Leaving the shape of the transition rate mostly unspecified enables us to proof its effect on the estimation outcomes compared to the results of the parametric models.

Exponential Transition Rate Model: The exponential transition rate model assumes that the residence time (t) in an origin state, i.e. the duration of an episode, can be described by an exponential distribution. It is taken into account that the risk, that an event happens, depends on a set of covariates. This enables the analyst to investigate how the transition rate describing the process under study (in our case the internal growth of a farm's milk production) depends on observable characteristics of the individuals (here dairy farms) and their environment. As BLOSSFELD and ROHWER (2002: 99) mention, the effect of a covariate on the transition rate reflects both: (1) its impact on the speed of the dependent process and (2) its impact on the proportion of individuals who have experienced an event after a certain time.

In the basic exponential model, the transition rate can vary with different constellations of covariates, but is time-constant. Thus, using the exponential model implicitly assumes that the process under consideration is not time-dependent.

A general definition of the time-constant transition rate from origin state j to destination state k is (see BLOSSFELD and ROHWER, 2002: 87)

$$r_{jk}(t) \equiv r_{jk} = \exp\{\alpha_{jk0} + A_{jk1}\alpha_{jk1} + \dots\} = \exp\{A_{jk}\alpha_{jk}\}, \quad (4.1)$$

where A defines the (row) vector of observed covariates and α describes the (column) vector of associated coefficients, both specific with regard to the origin state and the destination state; α_0 is a constant term.

This exponential model, like all parametric transition rate models, can be estimated by using the maximum likelihood method (see, for example, ALLISON, 1984).

Piecewise Constant Exponential Model: The piecewise constant exponential model is only a simple extension of the above described standard exponential model. As the term of the model already refers to, the hazard rates are assumed to be piecewise constant. So the basic idea is to split the whole observation period into several time periods and to assume that transition rates are constant in each of these intervals but can change between them (BLOSSFELD and ROHWER, 2002: 120). By splitting the time axis into intervals I_l ($l=1, \dots, L$), the transition rate r_{jk} from origin state j to destination state k is given as

$$r_{jk}(t) = \exp\{\bar{\alpha}_l^{(jk)} + A^{(jk)}\alpha^{(jk)}\} \quad \text{if } t \in I_l, \quad (4.2)$$

with $A^{(jk)}$ the vector of covariates, and $\alpha^{(jk)}$ the associated vector of coefficients. It is assumed that the effects of the covariates on the process under study are constant across the time periods. In contrast, estimated period-specific constants $\bar{\alpha}_l^{(jk)}$ are allowed to vary across the time intervals I_l . The baseline rate, given by these period-specific constants, serves to examine the force of change over time. Increasing competition pressure in global markets, for instance, could lead to the fact that dairy farmers are increasingly forced to realise cost saving potentials by means of internal growing along with exploitation of economies of scale. Thus,

the longer dairy farmers remain inactive (i.e. with increasing duration in a stagnation episode), the more the competitive conditions increase the pressure to grow in milk production and to improve the competitiveness by exploiting cost saving potentials, for example.

Gompertz Model: Parametric techniques are based on assumptions about the functional form of the transition (hazard) rate and the way the explanatory variables, i.e. the covariates, influence the probability of changing a specific state over time. The functional form of the hazard rate reflects how the probability of the corresponding event changes over time and consequently depicts the parametric model's underlying assumption about time-dependence. Time-dependence in parametric models might be seen as the result of a diffusion process, reflecting the changing relationship of a set of interdependent individual units in a dynamic system over time (BLOSSFELD and ROHWER, 2002: 178; DIEKMANN, 1989: 32ff.). The Gompertz model is an appropriate parametric approach, if substantive theory in the run-up to the empirical event history analysis suggests the modelling of a time path with a monotonically decreasing or increasing transition rate. Referring to this, we hypothesize in our empirical study a causal mechanism that with increasing duration in a non-growing episode of milk production, the hazard rate rises monotonically. Increasing competition pressure over time, for example, leads to the fact that ever more dairy farmers have to consider their own competitive power. Exploitation of economies of scale by growing in milk production could be one way to withstand the increasing pressures of competition. The transition rate r_{jk} of the Gompertz model used in this study is given by

$$r_{jk}(t) = b_{jk} \exp\{c_{jk}t\}, \quad b_{jk} = \exp\{B^{(jk)}\beta^{(jk)}\}, \quad c_{jk} = \gamma_0^{(jk)} \quad (4.3)$$

$B^{(jk)}$ denotes the (row) vector of time-constant covariates, with the first component of the vector assumed to be equal to one. $\beta^{(jk)}$ and $\gamma_0^{(jk)}$ are the coefficients of the model to be estimated. In the Gompertz model, the transition rate monotonically decreases if the parameter c_{jk} is negative and it monotonically increases, if the parameter c_{jk} is positive.

Log-Normal Model: The log-normal model is suitable if it is theoretically well-substantiated that the transition rate is somehow bell-shaped. Thus, in contrast to the Gompertz model, the log-normal model implies a non-monotonic relationship between transition rate and duration. This means that the transition rate initially increases to a maximum and afterwards decreases in the course of time. Such a bell-shaped transition rate is plausible in theoretical terms if two explicitly not measurable contradictory causal forces influence the process under study. Due to the non-measurability of the causal factors, the duration of an episode has to serve as a proxy variable for them. It is conceivable, for instance, that in a stagnation episode an increasing competitive pressure and the resulting necessity for farms to realize cost saving potentials by exploitation of economies of scale due to internal growth increases the transition rate up to a certain point. On the other hand, if farmers haven't been growing in a long-term view, it can be expected that the willingness of realizing growth steps in the future will continue to be reduced as time passes. Hence, it is possible that the transition rate strongly increases for some years and then slightly decreases. In the standard log-normal model the transition rate $r_{jk}(t)$ can be estimated as (BLOSSFELD and ROHWER, 2002: 207)

$$r_{jk}(t) = \frac{1}{b_{jk} t} \frac{\phi(z_t^{(jk)})}{1 - \Phi(z_t^{(jk)})}, \quad z_t^{(jk)} = \frac{\log(t) - a_{jk}}{b_{jk}}, \quad a_{jk} = A^{(jk)}\alpha^{(jk)}, \quad b_{jk} = \exp(\beta_0^{(jk)}) \quad (4.4)$$

$A^{(jk)}$ is the covariate vector with the first component as a constant equal to one. The associated coefficients vectors $\alpha^{(jk)}$ and $\beta_0^{(jk)}$ are the model parameter to be estimated. The estimated parameter b_{jk} gives further information about the non-monotonic pattern (skew and

kurtosis) of the transition rate. ϕ and Φ denote the standard normal density function and the standard normal distribution function.

Cox Model: Time in parametric models with an assumed time-dependent transition rate normally serves as a proxy variable for a latent causal factor that is difficult to measure directly (cf., BLOSSFELD and ROHWER, 2002: 228). The main weakness of such models is that their results can be influenced by the made parametric assumptions about both: (i) the time-dependence of the process under study, and (ii) the way that explanatory variables influence the risk of having an event. The Cox model however, only requires the specification of a functional form for the influence of covariates, but leaves the shape of the transition rate as unspecified as possible. Thus, results of event history analyses based on the Cox model are no longer influenced by more or less arbitrary assumptions about the time-dependence. The Cox model used in this paper specifies the transition rate as (see, e.g., BLOSSFELD and ROHWER, 2002: 229; LIN, 2003: 821; JENKINS, 2005: 28; BUIS, 2006: 22)

$$r_{jk}(t) = h_{jk}(t) \exp\left\{A^{(jk)}\alpha^{(jk)}\right\} \quad (4.5)$$

Thus, the hazard rate, $r_{jk}(t)$, at time t for a transition from origin state j to destination state k is defined as the product of two factors: (i) an unspecified hazard function for the transition, $h_{jk}(t)$, also called baseline rate, and (ii) an exponentiated linear function of a set of time-constant covariates $A^{(jk)}$ and their associated coefficients $\alpha^{(jk)}$. The baseline rate, which depends on the time t captures the shape of the transition rate and hence summarizes the pattern of “duration dependence”. The model estimation is based on the method of partial likelihood, developed by COX (1975). It should be noted that the partial likelihood method gives estimates of the coefficients of the Cox model, but no direct estimates of the underlying baseline rate.³

5 Data and explanatory variables

The underlying dataset of the present study consist of a balanced panel containing 616 dairy farms in Germany, each farm observed over the period from 1996 to 2009. The data were provided by LAND-DATA GmbH, a leading software house and service provider on the field of agricultural accounting in Germany. The database contains the annual accounts for each farm over the observation period.⁴ In addition to these economic information, the data set encloses a series of socio-economic and other farm-specific characteristics, for instance, such as age and education level of the famer, the farm’s geographical location and a variety of production-related parameters, e.g. milk yield, herd size, amount of farmland, stocking rate.

In order to illustrate the farms under study and their development over time some descriptive statistics are given in Table 1. At this, the descriptive analysis concentrates on the first and last accounting year of the observation period and provides the mean and the standard deviation (in parentheses) of some economic, socio-economic and production-related variables.

As can be seen from Table 1, the analyzed farms increased their herd size during the observation period by an average of about 30%. During the same period the average milk yield per cow and year has been increased by approximately 18%. The amount of a farm’s milk production in 2008/09 was on average 63% higher compared to the financial year 1996/97. At the

³ For further information about the partial likelihood estimation we refer, for instance, to Cox (1975), Blossfeld and Hamerle (1989), Blossfeld and Rohwer (2002), Fan and Jiang (2009).

⁴ In Germany an accounting year in agriculture respectively covers the period from the 1st of July to Jun 30th of the following calendar year, defined as financial year. Thus, the first financial year observed in our dataset is the period from 1st of July 1996 to Jun 30th 1997 while the last one is the period from 1st of July 2008 to Jun 30th 2009.

same time, the nominal value of the whole farm profit rose by 74%. The mainly increased standard deviations of the parameters give an indication of a higher heterogeneity of the observed dairy farms in 2008/09 compared to 1996/97.

Analyzing the farms by their geographical location and, therefore, grouping them into three regions of Germany, in the following called “North”⁵, “Center”⁶, and “South”⁷, shows that the highest proportion of the observed farms are located in the south of Germany (67%). In contrast, 20% of the observed farms are located in the North and only 13% in the Center.

Table 1: Descriptive statistics of the analyzed data set (n=616 farms)

Indicator	Unit	Mean (SD.)	Mean (SD.)
		1996/1997	2008/2009
Herd size	Cows / farm	48 (19)	62 (33)
Milk production	kg / year	256,243 (122,461)	419,803 (286,836)
Milk yield	kg / cow / year	5,794 (1,075)	6,811 (1,435)
Total farmland	ha	48 (25)	61 (35)
Stocking rate	cows / ha	1.11 (0.45)	1.09 (0.36)
Whole farm profit	€ / farm	23,205 (21,159)	40,353 (34,016)
Age of farmer	years	53 (11)	50 (9)
Index of agricultural education	1 = no education 5 = agr. engineer	2.9 (1.0)	3.2 (0.9)

Source: own description

The selection of covariates was done based on data availability. In advance a correlation matrix was compiled in order to test on multi-co-linearity between the variables. Thus the covariables represent a sound set of information on the farms ensuring correlation coefficients of < 0.3 among all selected variables.⁸ In the following we describe the selected covariates out of the data set available, ensured that no multi-co-linearity is given. In the further course of this paper, the impact of these covariates on the speed of the dependent process, i.e. the growth of a farm’s milk production, will be tested and discussed. Accordingly, the corresponding descriptions of the covariates can be seen in the presented result tables later on.

(1) *Region*: Two covariates, referred to as “Region South” and “Region Centre”, were programmed as binary coded variables based on a third variable, named “Region North”. These variables indicate the location of the farms in the three observed regions in Germany.

(2) *Milk Quota*: The impact of farm size on growth was measured per 100,000 kg milk quota at the end of a non-growing episode.

(3; 4) *Share of own land and share of grass land*: These variables show the proportion (in %) of own land and grass land on total farmland, respectively.

(5; 6) *Average milk yield increase and average herd growth*: Internal growth in farm’s milk production is possible via milk yield increase calculated in 100 kg per cow and year and / or via herd growth which was considered in number of cows per farm and year. Both parameters were measured on average during a non-growing episode.

(7) *Share of subsidies on returns*: The share of subsidies is measured in percentage of a farm’s returns. Subsidies include in this case all governmental payments. However, the main part is related to a payment per hectare which is paid decoupled from production.

⁵ Containing the federal states: Lower Saxony, Mecklenburg-Western Pomerania, North Rhine Westphalia, and Saxony-Anhalt.

⁶ Containing the federal states: Hesse, Rhineland-Palatinate, Saarland, Saxony, and Thuringia.

⁷ Containing the federal states: Baden-Wuerttemberg and Bavaria.

⁸ According to Nachtigall and Wirtz (2002 91) as well as Bühl and Zöfel (2002 318) correlation coefficients of < 0.5 indicate a “minor correlation”.

(8; 9) *Age of the farmer and level of farmer's agricultural education*: The farmer's age is considered in years. The level of the farmer's agricultural education is measured as an index (1 to 5). Level 1 stands for the lowest education level (no agricultural education), 2 is equal to "still in education", 3 indicates "skilled worker", 4 means "master craftsman" and 5 is the highest level "agricultural engineer".

(10) *Milk price difference*: This is covariate measures the development of a farm's milk price (in € Cent per kg milk) in the last year of a non-growing episode.

(11) *Fee for over quota production*: A farm's fee for over quota production is measured in relation (%) to the milk returns at the end of a non-growing episode. The fee for over-quota production is a prohibitive instrument of the European Union's common milk market organization. The level of the prohibition is determined by the degree of over production per country. As balancing between federal states is allowed the fee had often been low.

(12) *Farm income*: The impact of profitability on growth is measured in € Cent farm income per kg produced milk (quota and overproduction) at the end of a non-growing episode.

(13) *Quota Auction*: In order to investigate the impact of the 'new' trading system for quota, the quota auction, a dummy variable for the introduction in the year 2000 was included into the event history analysis.

6 Results and Discussion

In this empirical study determinants of internal growth in dairy farms are identified using different models of event history analysis. The event under study is the internal growth in farm's milk production, which is measured as an increase of the farm's own milk quota, annually recorded in the underlying data set. Over quota production is not considered as growth because this must not necessarily be an increase of production volume for longevity. Estimating the effects of several covariates on internal growth in milk production by application of (semi-)parametric models requires the consideration of superposition-effects of covariates with regard to the degree of growth i.e. in this study the percentage increase of own quota per growing episode. For example, the farm's size (here measured in kg milk quota) not necessarily abets both small and high internal growth. With regard to consequential costs bigger farms may prefer smaller growing steps (in relative terms) while smaller farms may tend to do high growth when they, for example, change their production system from a 30 cow place stanchion barn into a 60 cow place free stall barn with milking parlor. For this reason the present study will focus on analyzing the impact of the above mentioned explanatory variables (section 5) on big growth steps in milk production. The definition of big growth steps was done by selecting the upper quartile of episodes ($n = 304$) out the data sample with the highest growth rates after leaving a non-growing episode. As a result, an increase of a farm's own milk quota by more than 21.75% of the origin milk quota is defined as a big growth step.

This section continues as follows: First general notes on the different model outputs will be made in order to explain how to interpret the results. Thereafter determinates of internal growth will be present and by discussing them we refer to the hypotheses which were made in section 3.

6.1 General notes on the different model outputs

The Exponential model: A simple model without covariates would treat the data as a sample of homogeneous individual growth episodes. Hence, in estimating a model without covariates we would abstract from all sources of heterogeneity among farms and their growth episodes. As we are interested in analyzing how a farm's growth in milk production depends on observable farm-specific characteristics and their environment, we strive to estimate and proof the impact of the covariates mentioned in chapter 5 on a farmer's decision to grow in high growth steps. With a view to Table 2 (on page 11) and under consideration of the estimation results of

the Exponential model, we get, first of all, a value of the log-likelihood function: -1024.60. Using the likelihood ratio test we can compare this model with the Exponential model without covariates and test, if the null hypothesis that the additionally included covariates do not significantly improve the model fit holds. Here, with a given significance level of 0.01 we conclude that the null hypothesis should be rejected. Thus, at least one of the included covariates significantly improves the model fit. More precisely, there are eleven significant explanatory variables for high growth in milk production: the covariates “milk quota”, “share of grassland”, “average milk yield increase”, “average herd growth”, “share of subsidies on returns”, “age of the farmer”, and “milk price difference” are significantly different from zero at the 0.01 level; the covariates “fee for over quota milk production” and the dummy for the year of the implementation of “milk quota auction” are both significantly different from zero at the level 0.05; and the “share of own land” and “farm income” are significant at the level 0.1 (see Table 2).

The Piecewise constant model: A comparison of this model with the Exponential model based on a likelihood ratio (LR) test shows that the Piecewise constant model provides a highly significant improvement (LR = 171.14). The estimated parameters for the baseline transition rate are significant in four periods of a non-growing episode. Since we define a period as a two year lasting time span, the first two periods, for example, contain the first four financial years of a non-growing episode. During these four periods the baseline transition rate increases from 1.13 to 2.02. This means that with an increasing duration in a non-growing episode the need to grow in milk production (i.e. the non-growing exit rate) increases. With regard to the hypotheses (H) which were made in section 3 the output of this model supports H1. There it was supposed that the probability of entering the growing episode increases overtime as a result of an increasing need to improve competitiveness and other reasons being inherent to the economic system.

Defining the periods, we choose a two year lasting time interval for our estimation. Generally, time periods can be arbitrarily defined, but there is some trade-off. Choosing a large number of time periods provides a better approximation of the baseline rates. However, this implies a large number of coefficients to be estimated. On the other hand, choosing a small number of periods will cause less estimation problems, but there is probably a poorer approximation of the baseline transition rate. Therefore, as a compromise, we define a period as a two year lasting time span.

The Gompertz model: The output of this model provides as well as the Piecewise constant model information on the impact of time on the dependent process under study. Time serves as a proxy variable in event history analysis models (see section 4). As hypothesized in section 3 it is supposed that over time the probability of entering the growing episode increases. This hypothesis is supported by the parameter \hat{c}_{jk} which is positive and significant different from zero at the 0.01 level ($\hat{c}_{jk} = 0.1946$). We therefore conclude that with an increasing duration in a non-growing episode the transition rate increases monotonically due to some unobservable factors (especially an increasing pressure caused by competition).

The Log-normal model: In terms of statistical significance and effect-direction of the analyzed covariates, the results of the Log-normal model are basically the same as the results of the other models. However, one must note that the signs of the estimated parameters must be multiplied by -1 to get coefficients comparable to the other models. The estimated coefficient $\hat{b}_{jk} = \exp(-0.133) = 0.875$ is highly significant. Thus, with regard to the skew and kurtosis of the transition rate function, the included time-constant covariates in the $A^{(jk)}$ -vector of the Log-normal model make the non-monotonic pattern of the transition rate function significant different from a reference function with $b_{jk} = 1$. The reference function describes the skew

and kurtosis of the non-monotonic transition rate in an initial situation. In our case, as b_{jk} is lower than 1, we conclude that the covariates make the non-monotonic pattern of the transition rate function flatter and less skewed to the left than it is the case for the reference function.

The Cox Model: The vector of associated coefficients α has the same interpretation as the corresponding vector in the parametric transition rate models. The estimated coefficients resemble the results for the other models. Thus, with regard to the covariates and in terms of effect-direction and significance of the effects, the same substantive conclusions must be drawn. However, it should be noted, that there is no estimate for the constant anymore. The linear combination of the row vector of covariates and the column vector of associated coefficients cannot contain an intercept because this is absorbed in the baseline rate which doesn't have a specified functional form. Thus, the constant becomes part of the baseline hazard rate in the Cox model. As can be seen from the results, controlling for an unspecified baseline hazard rate in the Cox model does not affect the coefficient in terms of the effect-direction and significance level, compared to the results of the other models estimated.

6.2 Determinants of internal growth

In section 3 hypotheses were made on all analyzed covariates. With regard to time, which serves in event history analysis models as a proxy variable for unobservable or non-measurable time-dependent influences on the process under study, it was assumed that “the probability of entering a growing episode increases over time”. The estimated coefficients for the period-specific constants of the Piecewise constant model as well as the \hat{c}_{jk} -constant of the Gompertz model support this hypothesis for high quota growth. Hence the results document that there is a need for farms to grow which is somehow inherent to the economic system and targets the improvement of competitiveness.

With regard to the farms under study, a dairy farm's amount of ‘*milk quota*’ was treated as a measure for the farm size. The hypothesis “the greater farms milk production in terms of quota the shorter the stagnation episode” is not supported by the results. Indeed, the estimated coefficient for the corresponding covariate is significant but has a negative sign. Hence, growth rate of more than 21.75% per growing episode, which was the criteria for high growth in our study, is more likely for smaller farms than for bigger ones. This could result of a technology change of small farms from a stanchion barn farming system to a free stall barn system with milking parlor. This is a typical step in farm development especially in the southern part of Germany where still about 50% of the cows are tethered. In the range of such a technology change a doubling of the cow herd is distinctive. Once farms changed into a free stall barn farming system they tend to carry out smaller growing steps (in relative terms) in order to defray absolute costs of growth.

The coefficients of the two explanatory variables ‘*Share of own land*’ and ‘*Share of grass land*’ are significant. The share of own land, however, did not accelerate growth as it was assumed in hypothesis 4. There it was hypothesized that a high share of own land might allow for taking out loans at more favorable conditions than farms with less own land could do. The results, however, suggest that farms with high quota growth tend to have less own land. This is a result of nutrient regulations which govern in Germany a maximum application of nitrogen from manure per hectare. Consequently, high growth in milk production demands for appropriate growth of farmland. Individually, purchase of farmland is limited due to potential capital restrictions and regional scarce supply of land.

Table 2:

Estimation results for farms with observed high quota growth after entering a growing episode (GE)

	Exponential model (Model 1)		Piecewise constant (Model 2)		Gompertz model (Model 3)		Log-normal model (Model 4)		Cox model (Model 5)	
	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE
Period-1			-0,3754	0,5140						
Period-2			1,1330 **	0,5175						
Period-3			1,6116 ***	0,5346						
Period-4			1,6959 ***	0,5610						
Period-5			2,0208 ***	0,5930						
Period-6			0,7178	0,8948						
Constant-A	-0,1177	0,4974					0,1895	0,2992		
Constant-B					-0,1548	0,5024	-0,1330 ***	0,0410		
Constant-C					0,1946 ***	0,0258				
Region South	-0,0704	0,1648	0,2474	0,1688	0,1003	0,1676	-0,0454	0,1038	0,0187	0,1672
Region Center	-0,0691	0,2141	0,1423	0,2203	0,0484	0,2171	-0,0029	0,1314	-0,0097	0,2177
Milk quota at the end of the NGE ¹ (in 100,000 kg)	-0,1738 ***	0,0488	-0,1461 ***	0,0487	-0,1595 ***	0,0486	0,0827 ***	0,0258	-0,1731 ***	0,0490
Share own land (in % of farmland)	-0,0041 *	0,0022	-0,0051 **	0,0022	-0,0045 **	0,0022	0,0026 **	0,0014	-0,0044 **	0,0022
Share of grassland (in % of farmland)	-0,0076 ***	0,0024	-0,0094 ***	0,0024	-0,0083 ***	0,0024	0,0056 ***	0,0015	-0,0084 ***	0,0024
Average milk yield increase per cow and year during the NGE ¹ (in 100 kg)	0,0763 ***	0,0244	0,1124 ***	0,0285	0,0925 ***	0,0260	-0,0423 ***	0,0136	0,0834 ***	0,0258
Average herd growth (in cows per year) during the NGE ¹	0,0248 ***	0,0026	0,0273 ***	0,0026	0,0249 ***	0,0026	-0,0226 ***	0,0021	0,0255 ***	0,0026
Share of subsidies on returns (in % of farm returns)	-0,0760 ***	0,0124	-0,1395 ***	0,0152	-0,1173 ***	0,0148	0,0536 ***	0,0065	-0,0964 ***	0,0141
Age of the farmer (in years)	-0,0231 ***	0,0064	-0,0328 ***	0,0069	-0,0297 ***	0,0067	0,0204 ***	0,0037	-0,0257 ***	0,0066
Level of farmer's agricultural education (index)	0,0816	0,0717	0,1047	0,0727	0,0935	0,0726	-0,0462	0,0414	0,0868	0,0720
Milk price difference TEnd ² to previous year (in € Cent per kg milk)	0,0495 ***	0,0112	0,0472 ***	0,0105	0,0549 ***	0,0105	-0,0403 ***	0,0073	0,0484 ***	0,0110
Fee for over quota milk production TEnd ² (in % of milk returns)	0,0256 **	0,0107	0,0371 ***	0,0107	0,0307 ***	0,0106	-0,0232 **	0,0093	0,0300 ***	0,0107
Farm income TEnd ² (in € Cent per kg milk; quota + over production)	0,0079 *	0,0042	0,0014	0,0045	0,0041	0,0042	-0,0059 *	0,0034	0,0059	0,0044
Dummy for the first year with quota auction (year 2000)	-0,3181 **	0,1521	0,0722	0,1564	0,0007	0,1590	-0,0438	0,0885	-0,1694	0,1568
Log likelihood (starting values)			-1,169,213		-1,180,229		-1,301,806		-2,006,096	
Log likelihood (final estimates)			-1,024,597		-998,058		-950,652		-1,855,429	

Notes: Own calculations based on the use of TDA, a computer program for statistical calculations with special emphasis on methods to analyze panel and event history data. TDA is distributed as freeware under the terms of the GNU General Public License.
 *, **, *** significantly different from zero at the 0.1, 0.05 and 0.01 level, respectively. The signs of the estimated parameters in Model 4 (Log-normal model) must be multiplied by -1 to get coefficients comparable to the other model's estimates.

Thus, renting farmland is more likely and, consequently, the share of own land decreases with high growth in milk production and increasing farm size. The results further show that the share of grass land has a negative impact on growth in the sense that a higher share of grass land reduces the probability to leave the NGE and to carry out a big growing step. Consequently, hypothesis 5 is supported. There it was supposed that the specific characteristics in terms of forage quality of dairy farms in grassland regions decrease the probability to enter a growing episode. Moreover, the infrastructure for grazing might become a limiting factor for high growth as bigger cow herds would demand for larger grazing areas near by the farm. A big growth step could then require a system change from a seasonal grazing to keeping the cows the whole year in the barn.

With regard to the covariates '*Average milk yield increase*' and '*Average herd growth*' the results show that the hypothesis "higher increase of average milk yield per cow and/or number of cows per farm induces growth in terms of milk quota" is supported by the results. In view of the covariate '*Share of subsidies on returns*' we examine the hypothesis if "a high share of subsidies on total returns decelerates growth". The results support this hypothesis, too. Hence, the assumption that a higher direct income support reduces incentives to improve competitiveness via cost decreasing by exploitation of economies of scale is confirmed.

The estimated coefficients for the variable '*Age of the farmer*' show that the stated hypothesis "younger farmers have a higher incentive to grow" is confirmed. However, the '*level of farmer's agricultural education*' does not appear to have a significant influence on growth of a farm's milk production.

An '*increase of the milk price*' within the last year of a stagnation episode has, as assumed in H10 in section 3, a measurable impact on high internal growth. Thus, the estimates suggest that an increased liquidity as a result of higher milk revenue has a significant impact on a farm's ability to grow. Meanwhile the impact of profitability in milk production, measured by the variable '*farm income per kg milk*', does not appear to influence the growth rate.¹ Consequently, the related hypothesis "farms with a higher income per kg milk are more likely to grow" cannot be supported by the estimates. This result is hardly traceable from an economic point of view as we must conclude that farms seem to decide to grow without optimizing the management and consequently the profitability beforehand.

The coefficient of the covariate '*fee for over quota milk production*' is positive in sign and significant. Hence, as stated in H11, in order to avoid or reduce the risk of paying a fee for over quota production farmers have an increasing incentive to expand their milk quota.

It was hypothesized in section 3 that the implementation of the '*quota auction*' in Germany in the year 2000 had a positive impact on growth. However, only the coefficient of the corresponding covariate in the Exponential model is significant positive and, thus, would support this hypothesis. The respective coefficients of all other applied models are insignificant. We therefore assume that the implementation of the milk quota auction rather supports low internal growth than bigger growing steps. This can be explained by the fact that quota prices revealed a high variation from one trading event to the next and a decreasing trend over the observed time period. Hence splitting high growth into smaller steps could be a strategy of farmers to reduce quota-price risk. Another argument for smaller growing steps are follow-up costs of growth which are easier to defray in a small-growth-step strategy.

¹ The coefficients of the tested covariate "Farm income" are only significant different from zero at the level 0.1 in the Exponential model and the Log-normal model, whereas none of the other models attest the significance of this variable.

7 Conclusion

In this empirical study event history analysis is for the first time applied to investigate determinants of internal growth of German dairy farms. Using this dynamic econometric framework enables us to consider explanatory variables that change over time such as milk price or age of the farmer. Additionally, parametric event history analysis models allow considering non-measurable effects on growth of milk production because time serves as a proxy variable in such models. In this case “an increasing competition-pressure” may be an explanation for non-measurable effects on internal growth.

A variety of farm and farmer characteristics, as well as market effects over the study period from 1995 to 2009 have been considered. In conclusion, different types of (semi-)parametric models were applied in this study and provided almost the same estimates of the impact of covariates on the transition rate. This proves the validity and robustness of the estimated results with regard to the dependency of the results on the model-specific assumptions.

The fact that time in event history analysis serves as a proxy variable counting for non-measurable effects on the event under study is a strength of the applied method compared to alternative approaches such as Probit, Tobit and Heckman models. Although, for example, a Heckman model enables one to explain why, at a particular moment in time, some farmers grow in milk production and others do not, these models don't allow explaining the impact of time on internal growth.

In general, the impact of covariates on the ‘event’ under study, i.e. the effect-direction of explanatory variables, depends considerably on the definition of the event. Thus in our case, it is important for further research to investigate whether alternative definitions of growth in a farm's agricultural production (the ‘event’) influence the estimated impact of explanatory variables. Consequently, contradictory causal factors on growth could better be analyzed.

References

- ALLISON, P. (1984): *Event History Analysis, Regression for Longitudinal Event Data* (Beverly Hills, CA, USA: Sage Publications).
- BLOSSFELD, H.-P. and A. HAMERLE (1989): Using Cox model to study multiepisodic processes. In: *Sociological Methods and Research* 17: 432-448.
- BLOSSFELD, H.-P. and G. ROHWER (2002): *Techniques of Event History Modeling, New Approaches to Causal Analysis*. New York, USA and London, UK: Lawrence Erlbaum Associates.
- BRÜMMER, B., J.P. LOY and C. STRUVE (2003): Analysis of milk quota trading in Germany [in German]. In: *German Journal of Agricultural Economics* 52: 275-288.
- BÜHL, A. and P. ZÖFEL (2002): *SPSS 11, Introduction into modern data analysis using windows* [in German]. Munich, DE: Pearson Studium.
- BUIS, M.L. (2006): *An introduction to Survival Analysis*. Department of Social Research Methodology, Vrije Universiteit Amsterdam, Amsterdam, NL.
- BURTON, M., D. RIGBY, and T. YOUNG (2003): Modeling the adoption of organic horticultural technology in the UK using Duration Analysis. In: *The Australian Journal of Agricultural and Resource Economics* 47: 29-54.
- COX, D.R. (1975): Partial likelihood. In: *Biometrika* 62: 269-276.
- COX, D.R. (1972): Regression models and life-tables. In: *Journal of the Royal Statistical Society* 34: 187-220.
- DADI, L., M. BURTON, and A. OZANNE (2004): Duration analysis of technological adoption in Ethiopian agriculture. In: *Journal of Agricultural Economics* 55: 613-631.
- DESTATIS (2010): *Statistical Federal Office* [in German]. <http://www.destatis.de>.
- DIEKMANN, A. (1989): Diffusion and survival models for the process of entry into marriage. In: *Journal of Mathematical Sociology* 14: 31-44.

- EUROPEAN COMMISSION (2009): Regulation (EG) No. 72/2009 of the European Council from January 19th 2009 [in German]. http://ec.europa.eu/index_en.htm.
- FAN, J. and J. Jiang (2009): Non- and Semi-Parametric Modeling in Survival Analysis. In: Fan J., X. Lin and J.S. Lin (Ed.). *New Developments in Biostatistics and Bioinformatics*. Beijing, CN: Higher Education Press: 3-33.
- FOLTZ, J.D. (2004): Entry, exit, and farm size: Assessing an experiment in dairy price policy. In: *American Journal of Agricultural Economics* 86: 594-604.
- GLAUBEN, T., H. TIETJE and C.R. WEISS (2003): Agriculture on the move: Exploring regional differences in farm exit rates. Working Paper EWP 0308, Department of Food Economics and Consumption Studies, University of Kiel.
- GOERTZ, D. (1999): Production costs of milk production in Germany [in German]. Johann Heinrich von Thünen-Institut, Braunschweig, Germany.
- HUETTEL, S. and JONGENEEL, R. (2008): Structural Change in the Dairy Sectors of Germany and the Netherlands – A Markov Chain Analysis. Paper presented as “Contributed paper” at the 12th Congress of the European Association of Agricultural Economists (Ghent, Belgium, August 26-29, 2008).
- JENKINS, S.P. (2005) Survival Analysis. Unpublished manuscript, Institute for Social and Economic Research, University of Essex, Colchester, UK. Downloadable from <http://www.iser.essex.ac.uk/teaching/degree/stephenj/ec968/pdfs/ec968lnotesv6.pdf>.
- KLEINHANß, W., F. OFFERMANN and M. EHRMANN (2010): Evaluation of the Impact of milk quota – Case study Germany. Johann Heinrich von Thünen-Institut, Braunschweig, Germany.
- LÄPPLE, D. (2010): Adoption and abandonment of organic farming: An empirical investigation of the Irish drystock sector. In: *Journal of Agricultural Economics* 61: 697-714.
- LASSEN, B., F. ISERMEYER and C. FRIEDRICH (2008): Milk production in transition –An investigation of regional potentials and leeway in Germany [in German]. Johann Heinrich von Thünen-Institut, Braunschweig, Germany.
- Lin, D.Y. (2003): Survival Analysis. In: Lu Y. and J.-Q. Fang (Ed.): *Advanced Medical Statistics*: Singapore: World Scientific Publishing: 815-835.
- NACHTIGALL, C. and M. WIRTZ (2002): Probability calculation and inferential statistics [in German]. Weinheim und Munich, DE: Juventa Verlag.
- WEISS C.R. and W. BRIGLAUER (2000): Determinants and Dynamics of Farm Diversification. Working Paper EWP 0002, Department of Food Economics and Consumption Studies, University of Kiel.
- WYNN, G., B. CRABTREE and J. POTTS (2001): Modeling farmer entry into the environmentally sensitive area schemes in Scotland In: *Journal of Agricultural Economics* 52: 65-82.