The Royal Veterinary and Agricultural University
Food and Resource Economic Institute

KVL A UNIVERSITY OF LIFE SCIENCES

Unit of Economics Working Papers 2001/1

Information Based Estimators for the Non-Stationary Transition Probability Matrix: An Application to the Danish Pork Industry.

Kostas Karantininis

Information Based Estimators for the Non–stationary Transition Probability Matrix:
An Application to the Danish Pork Industry

Kostas Karantininis*

Department of Economics and Natural Resources

The Royal Agricultural University of Copenhagen (KVL)

Abstract

A generalised cross entropy instrumental variables estimator is used to recover the non–stationary transition probability matrix for the Danish pork industry. A technique is also developed to recover missing data points due to category re–specifications. The impact of a set of exogenous variables (prices of pork meat, inputs, and pork substitutes), are evaluated in the form of elasticities. An overall assessment of entry exit and growth of farms is performed.

Keywords: Cross Entropy Estimator, Markov Chain, Pork Industry, Denmark

JEL Classification: C340, C130, Q120

Kostas Karantininis, Associate Professor
Department of Economics & Natural Resources
The Royal Veterinary & Agricultural University (KVL)
Rolighedsvej 23, DK-1958 Frederiksberg C
Copenhagen C, DANMARK
Tel: +45-35282276, Fax: +45-35282295, e-mail: kok@kvl.dk

* E–Mail: kok@kvl.dk. I wish to thank Amos Golan for his encouragement and comments. Financial and other support by the Department of Agricultural Economics and Policy, of the Agricultural University of Wageningen, during the early stages of this research, is gratefully acknowledged. K. Katrakylidis and the Artist provided fruitful discussions and quiet working environments.

Information Based Estimators for the Non–stationary Transition Probability Matrix: An Application to the Danish Pork Industry

Kostas Karantininis
Department of Economics and Natural Resources
The Royal Agricultural University of Copenhagen (KVL)

1. Introduction

The decline in the number of firms in industries and the consequent increase in the average firm size has occupied the economics literature for a very long time. Also the question of what determines the size distribution of firms is a key feature in industrial economic analysis and policy. This becomes increasingly important in the agro–food industry, where farm numbers have been declining drastically over the past decades, whereas farm size increases[1]. Many of the studies on growth and size distribution of firms rely on a very simple stochastic model, which is usually a variant of the well known Gibrat's Law, i.e. firm growth is independent of firm size[2] – although others develop more deterministic models of firm growth (such as Lucas(1978) and Jovanovic 1987, for example). The Markov process in discrete time has been used often as an appropriate tool to describe the movement of economic variables over time[3]. This model has been proven particularly useful given that researchers have very rarely the luxury of longitudinal time–ordered micro data describing movement of individuals between different states. Instead, aggregated data of finite size categories (Markov states) for given time periods are available.

In most of the early applications a purely stochastic Markov process is assumed, and the transition probability matrix (TPM) is assumed to be constant over time – usually referred to as "Stationary Markov Model"[4]. The transition probabilities may vary over time, however, resulting in so–called non–stationary transition probability matrix (NSTPM), adding in this way a deterministic element in the Markov process in the form of a systematic relation with a set of exogenous variables (Telser (1963); and Hallberg (1969) were the first such attempts[5]).

Most studies that use a NSTPM use very strong parametric distributional assumptions, and other restrictions. Traditional estimation techniques fail, or require strong restrictions, because the estimated

---

[1] Evans, 1987b, and the accompanying papers at the symposium on "Empirical analysis of size distribution of farms"
[2] See for example Evans, 1987a for a review
[3] Lee et al. (1977) provide a partial literature list for general Markov studies; Zepeda (1995a&b) also provide a more recent list of mainly agriculture related Markov studies.
[4] See for example Adelman 1958; Padberg 1962; Lee et al., 1977, Oustapassidis, 1986.
[5] See for example Lee, et al., 1977; Disney, et al., 1988; Massow, et al., 1992; Zepeda, 1996a,b, and their cited references.

parameters must satisfy probability assumptions (non–negative, adding up to one)[6]. Most approaches, suffer from the dimensionality problem, and the researcher is restricted in their choice of covariates to be used. A lot of missing data points are also very common due to re–definitions of size categories by the data collectors.

In this paper we use the generalised maximum entropy (ME) formalism, which stems originally out of the Shanon's (1948) information theory and Jaynes (1957a,b). We employ the generalised cross entropy (GCE) formalism by Golan et al. (1996), and we further extend the applications of Golan and Vogel (2000); and Courchane et al., (2000). We use GCE formalism to recover coefficients of the effects of exogenous variables on individual transition probabilities, when a specific functional form (linear) of the relationship is imposed, and decompose the two error terms in this relationship. The developments in this paper are used to recover an NSTPM for the pork industry in Denmark. First, the missing data points that result by re–definitions of size categories are recovered. We use GCE to recover an instrumental variables estimator which is a less restrictive in recovering the NSTPM, and evaluating the impacts of various covariates on transition probabilities. This method allows the use of an extensive set of covariates, their significance is tested with an asymptotic test, and the impact of each covariate on the individual probabilities and size categories is evaluated in the form of elasticities. Prior information on the TPM is introduced using the GCE formalism. The recovered NSTPM is further used to assess Gibrat's law.

In the next Section, we develop the GCE estimator for the stationary Markov model and introduce non-stationarity in Section 2.2. In Section 3 the instrumental variables GCE estimator is derived. In Section 4 we show how to handle problems of re-definition of categories. The application on the Danish pork industry is presented in Section 5. Some concluding remarks are discussed in Section 6.

2. Recovering Markov Transition Probabilities

2.1 An Ill–Posed Problem of Industry Structure: The Stationary Markov Model

The attempt to recover the transition probability matrix, from a set of limited macro data on firm size distribution, is a classic case of an ill-posed problem. As an illustration, the case examined in this study

[6] Telser (1963); and Hallberg (1969)), use OLS estimators, which require strong equality restrictions on the parameters; Lee et al., (1977) impose inequality restrictions; MacRae (1977) suggested a Logit transformation, which automatically satisfies the probabilistic constraints (see Zepeda, 1995a,b for applications).

consists of fifteen census years (i.e. fourteen transitions) for eighteen size groups of pork farms in Denmark – a total of 252 data points. We are attempting to estimate (recover) a transition probability matrix of 18×18 = 324 elements (or 361 elements if we consider entry and exit with an artificial 19th category in a 19x19 transition probability matrix). The ill–posed problem becomes even more pronounced if one considers that a lot of these data points are missing due to re–definition of data groups. The ME formalism is valuable in tackling this problem. In the next Sections we are presenting a non–stationary Markov model. However, it is useful to proceed with a stationary model before.

A stationary TPM using GCE is developed by Lee and Judge (1996), and Golan, et al., (1996). The transition between time t and $t+1$ in the stationary model can be formulated as follows:

$$\mathbf{y}(t+1) = \mathbf{x}'(t)\mathbf{P} + \mathbf{u}(t) \tag{1}$$

where $\mathbf{y}(t+1)$ is a $K\times1$ vector of proportions falling in each of the K Markov states at time $t+1$, and $\mathbf{x}(t)$ are the sample proportions at time t. The TPM is $\mathbf{P}=(\mathbf{p}_1\ \ \mathbf{p}_2\ \dots\ \mathbf{p}_K)$ with each vector $\mathbf{p}'_{\mathbf{k}} = (p_{1k}, p_{2k}, \dots, p_{Kk})$. Finally, $\mathbf{u}(t)$ is a vector of disturbances with zero mean bounded within a specified support vector $\mathbf{v}$. For T transitions, the model can be written more compactly:

$$\underset{(TK\times1)}{\mathbf{y}_T} = \underset{(TK\times K^2)}{(I_K \otimes X_T)}\ \underset{(K^2\times1)}{\mathbf{p}} + \underset{(TK\times1)}{\mathbf{u}_T} \tag{2}$$

where the TPM is now written as a vector: $\mathbf{p} = (\mathbf{p}'_1, \mathbf{p}'_2, \dots, \mathbf{p}'_K)$, I_K is a $K\times K$ identity matrix and $\otimes$ denotes Kronecker product. Each element of the $\mathbf{u}_T$ is parameterised as $u_{it} = \sum_m^M v_m w_{itm}$, where $\mathbf{w}$ is an M–dimensional vector of weights (in the form of probabilities) for each u_{it}, $\mathbf{v}$ is an M–dimensional vector of supports. The support vector can be set to: $\mathbf{v} = \left[-1/K\sqrt{T}, \dots, 0, \dots, 1/K\sqrt{T}\right]'$ (Courchane et. al., 2000). By using GCE, any prior information about $\mathbf{P}$ can be incorporated in the form of a matrix of priors $\mathbf{Q}$. Prior information about the disturbance $\mathbf{u}_T$, call it w^o_{itm}, can be incorporated as well and are assumed to be uniformly symmetric about zero. Let $H(\cdot)$ be the measure of cross entropy, then the GCE is:

$$\min_{\mathbf{p},\mathbf{w}}\left\{H(\mathbf{P},\mathbf{W},\mathbf{Q},\mathbf{W}^o)=\sum_i\sum_j p_{ij}\ln(p_{ij}/q_{ij})+\sum_i\sum_t\sum_m w_{itm}\ln(w_{itm}/w^o_{itm})\right\} \quad (3)$$

subject to the following three sets of constraints:

(a) The $K\times T$ data consistency constraints (Equations (2)); (b) The normalization constraints for both the transition probabilities (K constraints) and the error weights ($K\times T$ constraints): $\sum_j^K p_{ij}=1$, $\sum_m^M w_{itm}=1$; and (c) the K^2 non–negativity constraints for $\mathbf{P}$ and the $K\times T\times M$ constraints for $\mathbf{w}$: $\mathbf{P}\geq 0$, and $\mathbf{w}\geq 0$. The solution to the above system of equations is derived elsewhere and we see no need to repeat the process here (see for example, Golan, et. al., 1996).

2.2 Non–Stationarity

The simplest form of non–stationarity is to assume that $\mathbf{P}$ is varying over time, without any further assumptions on functional relationships with other variables. In other words, the objective is to estimate a different TPM for each transition[7]. It is more interesting, however, to examine what actually makes the TPM vary over time. In this case, the NSTPM can be expressed as:

$$p_{ij}(t)=f_{ij}(\mathbf{z}_{ij}(t),\ \beta_{ij})+e_{ij}(t) \quad (4)$$

where $f_{ij}(\bullet)$ is a function relating each element $p_{ij}(t)$ of the NSTPM to a vector of explanatory variables $z_{ij}(t)$. The β_{ij} are parameters of the $f_{ij}(\bullet)$, and $e_{ij}(t)$ is the disturbance term. The Markov process can now be expressed as:

$$\mathbf{y}(t+1)=\mathbf{x}'(t)[\boldsymbol{\beta}\ \mathbf{z}(t)+\mathbf{e}(t)]+\mathbf{u}(t) \quad (5)$$

MacRae (1977) points out that in most estimation methods, each row of transition probabilities must be formulated to depend on the exact same set of exogenous variables. Furthermore, Lee, et al., 1977; and MacRae, 1977, develop the statistical properties of the disturbance terms $\mathbf{e}$ and $\mathbf{u}$ in (5). We show bellow, that with the use of GCE formalism the disturbances $\mathbf{e}$ and $\mathbf{u}$ can be recovered separately.

[7] Lee et. al. (1977) illustrate such a model and develop a weighted least squares estimator with non–equality restrictions. Lee and Judge (1996) develop a GCE estimator

We can parameterise each β_{ijn} and each e_{ijt} over a discrete finite support space: $\beta_{ijn} = \sum_s^S d_{ijns}\theta_s$, and $e_{ijt} = \sum_h^H g_{ijth}\varphi_h$, where $\boldsymbol{\varphi}, \boldsymbol{\theta}$ are support vectors of size S and H respectively, and $\mathbf{d}$ and $\mathbf{g}$ are the corresponding probabilities to be recovered. The Markov process in (5) now becomes:

$$y_{jt} = \sum_i x_{it}\left[\sum_n^{N_{ij}}\left(\sum_s d_{ijns}\theta_s\right)z_{nt} + \sum_h^H g_{ijth}\varphi_h\right] + \sum_m v_m w_{jtm} \tag{6}$$

where N_{ij} is the number of covariates in the (ij)th cell. Applying GCE we can recover the $\boldsymbol{\beta}$, $\mathbf{e}$, and $\mathbf{u}$ through the recovered values of $\mathbf{d}$, $\mathbf{g}$, and $\mathbf{w}$ respectively. There are alternative ways to impose the standard normalisation and non–negativity constraints on transition probabilities. One can impose additional constraints on the $\mathbf{d}$ (either in the form suggested by Lee, et al. 1977; or by Halberg, 1969). Alternatively, one can assume a multinomial Logit transformation, which satisfies both the normalisation and the non–negativity constraints automatically (MacRae, 1977; Golan et al., 1997).

3. Instrumental Variables Generalised Cross–Entropy Estimator

We show here the generalised cross–entropy (GCE) estimator for the NSTPM similar to (Golan and Vogel, 2000). Let Z_{tn} be a TxN matrix of N covariates in the T time periods. We can incorporate this information into the GCE model by multiplying with Z_{tn} both sides of the data consistency constraint (2):

$$\sum_t z_{tn}y_{tj} = \sum_t\sum_i z_{tn}x_{it}p_{ij} - \sum_t\sum_m z_{tn}v_m w_{jtm}\ , \quad \forall j = 1,\ldots K, \quad n = 1,\ldots,N \tag{7}$$

Priors are introduced in the objective function (3) in the form of matrices $\mathbf{Q}$ (corresponding to the transition probabilities $\mathbf{P}$) and $\mathbf{W}^o$ (for the disturbance probabilities $\mathbf{W}$). The objective is therefore to recover the p_{ij} that are closest as possible to to q_{ij} and satisfy the data. The priors w_{itm}^o are assumed uniformly distributed around zero, hence they add no additional information to the model. The solution to this problem is:

$$\tilde{p}_{ij} = \frac{q_{ij}\exp\left[\sum_t \sum_n x_{it} z_{tn} \tilde{\lambda}_{nj}\right]}{\sum_j q_{ij}\exp\left[\sum_t \sum_n x_{it} z_{tn} \tilde{\lambda}_{nj}\right]} \tag{8}$$

where $\tilde{p}_{ij}$ and $\tilde{\lambda}_{nj}$ are the recovered probabilities and Lagrange multipliers respectively. The following "probability elasticities" (Zepeda, 1995b) show the effect of each z_{tn} on p_{ij}:

$$E^{P}_{ijtn} = \frac{\partial \tilde{p}_{ij}}{\partial z_{tn}} \frac{z_{tn}}{\tilde{p}_{ij}} = x_{it} z_{tn}\left[\tilde{\lambda}_{nj} - \sum_{j} \tilde{p}_{ij}\tilde{\lambda}_{nj}\right] \tag{9}$$

Where E^{P}_{ijtn} measures the percentage change of the n^{th} covariate on the transition probability between states i and j at time t.

Similarly, the following elasticity measures the effect of each exogenous variable on the number of farms:

$$E^{y(t+1)}_{jn} = \frac{\partial y_{j(t+1)}}{\partial z_{tn}} \frac{z_{tn}}{y_{j(t+1)}} = \frac{z_{tn}}{y_{j(t+1)}} \sum_i \left[\tilde{p}_{ij} x^2_{it}\left(\tilde{\lambda}_{nj} - \sum_j \tilde{p}_{ij}\tilde{\lambda}_{nj}\right)\right] \tag{10}$$

evaluated at the means.

4. Re–Definition of Categories and Missing Data Points

It is very common that statistical authorities change the definition of categories when they aggregate industry data. For example, categories 13–18 in Table 1, are given as aggregates in category 13 (1000+) for the years 1984–1994; the definition changes in 1995 where categories 13 and 14 are aggregated as category 13 (1000–1999), whereas categories 15 to 18 are aggregated as category 15 (2000+); the next year and the years since 1996 the farms over 1000 pigs are disaggregated in the six categories shown in Table 1. An opposite aggregation occurs in the smaller size categories: the smaller size categories are collected into 11 categories until 1994 and they are aggregated into four categories since 1996. Given these circumstances, the researcher can either aggregate further in order to have a consistent set throughout, i.e. aggregate categories 1–11 for all years until 1995 into the four current categories, and aggregate categories 13–18 for all years 1995–1998 into one category (category 13) that was in place until 1994. This is however a very inefficient approach because it ignores a lot of important information. Alternatively, one can recover the missing data by treating them as unknown parameters in the ME framework.

Define the missing data points (categories) as $x_{it} = \sum_{m=1}^{M} r_{itm}\eta_m$, where r_{itm} are probabilities to be recovered, and η_m are corresponding supports. Consider that categories ε to ζ ($1 \le \varepsilon < \zeta \le K$) are given as aggregates in category ε for the periods γ to δ ($1 \le \gamma \le \delta \le T$). The data consistency constraint (2) can be modified as follows[8]:

$$\left.\begin{array}{ll} \sum_m r_{ij(t+1)m}\eta_m & \text{if } \gamma \le t+1 \le \delta \\ & \text{and } \varepsilon \le j \le \zeta \\ y_{j(t+1)} & \text{otherwise} \end{array}\right\} = \begin{cases} \sum_{i=1}^{\varepsilon-1} x_{it}p_{ij} + \sum_{i=\varepsilon}^{\zeta}\left(\sum_m r_{ijtm}\eta_m\right)p_{ij} + \sum_{i=\zeta+1}^{K} x_{it}p_{ij} & \text{if } \gamma \le t \le \delta \\ \sum_i x_{it}p_{ij} & \text{otherwise} \end{cases} \tag{11}$$

In addition, the information about the sums of the aggregated categories $x_\varepsilon(t)$ must also be satisfied:

$$\sum_{i=\varepsilon}^{\zeta} x_{it} = \sum_{i=\varepsilon}^{\zeta}\sum_{m=1}^{M} r_{itm}\eta_m = x_\varepsilon(t) \quad \forall \gamma \le t \le \delta \tag{12}$$

The usual normalisation and non–negativity constraints for the probabilities r_{itm} are added. Natural bound for the support vector is the largest of the aggregated categories $x_\varepsilon(t)$:

$$\eta = \left[-\max_t\{x_\varepsilon(t)\},\ldots,0,\ldots,+\max_t\{x_\varepsilon(t)\}\right] \quad \forall \gamma \le t \le \delta \tag{13}$$

The recovered missing data are then given by:

$$\hat{x}_i(t) = \sum_{m=1}^{M} \hat{r}_{itm}\eta_m \quad \forall \gamma \le t \le \delta, \quad \varepsilon \le i \le \zeta \tag{14}$$

5. Structural Changes in the Danish Pork Industry

We use aggregate data on size distribution of pork farms in Denmark between 1984 – 1998 (Table 1)[9]. The pork farm distribution typifies similar trends in pork industry elsewhere (Massow, et al., 1992; Disney, et

[8] The parameterised errors $v_m w_{jtm}$ are omitted for notational simplicity

[9] Before 1984 category definitions were even more severely different and the data base would have been simply unmanageable unless they were aggregated to an extend that all detailed information would have been lost.

al., 1988), as well as the general trend in the farm sector (Evans, 1988b). The total number of farms has decreased from 46,094 to 17,689 between 1984-1998 (a 61.2% decrease). While small farms (less than 49 animals) decreased from 19,483 to 3,693, large farms (more than 1000 animals) increased from 1,207 to 3,956, during the study period. Detailed data for the larger categories do not exist for the entire period, it is however evident that these farms increase constantly at least for the three years that data are available on farms larger than 1,500 animals.

<TABLE 1 about here>

It is hypothesised that pork prices affect pork supply through entry, exit and expansion of pork farms. Similarly, pork feed prices are expected to have a negative effect on pork supply by affecting the restructuring of pork farms, as captured by the NSTPM. Input and output prices of other livestock are expected to affect structural changes: (a) prices of milk, beef, eggs and poultry meat; and (b) input prices: pig composite feeds, poultry and cattle feeds, fertilizer prices (as a proxy of energy costs) and interest rate. These are also expected to affect the decision to expand or contract pork production, and even entry or exit, because they constitute alternative sources of income for pork farmers[10].

The data are available from series of government publications (Landbrugs–Statistik, various issues), and are all converted to 1980–based indices. Data on the distribution of pork farms were also taken from the same series. All numbers for farms were normalised by division with the maximum number in the series.

Several variations of the methods presented above have been tried here[11]. First, the missing data were recovered, and the models that follow were using the completed data set as they were recovered by the two–step procedure described in Section 3 above. In all cases, an artificial 19th category was introduced to account for entry and exit.

Using the recovered complete data, the simple stationary transition probabilities were recovered, as a benchmark situation. The recovered probabilities were very uniform – an indication that the data did not contain sufficient information to pull the probabilities away from the “prior” uniform distribution. The overall

[10] This is only a subset of exogenous variables that might affect the transition probabilities. There exists a plethora of hypotheses concerning factors affecting farm size growth and distribution (see Zepeda (1995 for a review and refernces).
[11] Detailed results of this study can be found in the working paper with the same title at the author’s Institute. Data and GAMS code are also available upon request.

performance of the model was very low[12]: A very high normalised entropy, $S(P)$=0.93; a pseudo–R^2 = 0.07; and X^2 = 7.7.

The Non–Stationary model was tried next. With all the covariates discussed previously, namely input and output prices of pork and other related livestock products. All the exogenous variables are lagged one period. The instrumental variables model and the ME formalism was applied to recover the NSTPM. The $S(P)$ for this matrix is 0.73, the pseudo–R^2 = 0.26; and X^2 = 30.13. The test for non–stationarity gives an ER=22.4, which is significant at the 99% level.

Transition probabilities at the lower left off–diagonal were mostly non–zero, a fact showing that large farms are likely to reduce in size. Instead of forcing these values to be zero, this information is introduced as priors[13]. By using the Cross–Entropy formalism, the model is allowed to select how close the recovered probabilities are to the prior knowledge (Table 2). We construct the matrix of priors using the following process: First construct a matrix of uniform probabilities equal to 1/19. Then, set to zero the elements p_{ij} for any $j \geq i+5$ or $\leq i-5$ for $i,j \neq 19$. Increase the value of the diagonal elements to: $1-\sum_j p_{ij} \ \forall i \neq j$. This process reflects the belief that farms do not grow more than certain rate (maximum five size categories) each time, and secondly, that it is more likely that a farm will remain in the same category than otherwise.

<TABLE 2 about here>

The recovered NSTPM has an $S(P)$=0.505, a corresponding pseudo–R^2 = 0.49 and a X^2=55.4. This is a remarkable improvement compared to both its stationary and non–stationary counterparts. Notice also that the recovered NSTPM in Table 2 preserves much of the structure of the matrix of the priors implying that these prior beliefs are supported by the data (Golan, et al. 1996).

How does the farms' birth, grow and death, relate to this analysis? Consider entry first. Note that as indicated by the last row of the NSTPM entry is most likely to occur in medium size categories (10–12, or 400–999 animals), and somewhat less into the smallest category, and even lesser into the large size

[12] The measures of performance used here are the "normalised cross entropy" (Golan, et al., 1996); McFadens' pseudo–R^2 (McFaden, 1974); and the "Entropy Ratio" (ER) which is distributed as X^2_{K-1} (Courchane et al., 2000).
[13] Restricting lower and upper off–diagonal elements to zero is common practice in many similar studies (Disney, et al., 1988; Zepeda, 1996a)

categories[14]. We must note, however, that although the recovered probability of entry into the large categories is very small (0.001 to 0.003, for the categories 14 to 17, i.e. for farms having more than 1500 animals) the impact is more significant given that the number of farms in these categories is very small compared to the rest[15].

At a first glance, there is no apparent relationship between firm exit and size (Figure 1). Some conclusions however can still be drawn. Take for instance the size categories up to 1–1500 animals (categories 1–12). During the last five–six years their absolute size is relatively uniform among these categories, roughly 1000–1500 farms (Table 1). As we see from the last column, the probability of exit among these categories is larger for categories 7–12 (ranging from 0.12–0.22), whereas it is below 0.1 for all size categories with less than 150 animals (categories 1–6). This probably reflects the fact that the smaller farms are more likely to sustain losses during harsh times than larger ones, and is important, for policy makers, to take notice that although small farms decrease in numbers, they do not disappear[16].

<FIGURE 1 about here>

Farm growth is clearly related to farm size, as shown in Figure 2, however not proportionally. Given that size categories are quite arbitrary, one can not make an immediate connection of this result to Gibrat's law. To do so, we need to consider proportional changes, given the limitations of the aggregation of categories. We calculate the medians of each size category and take the transition probabilities (from Table 2) to the category that has approximately double or triple median size (i.e. a proportional growth). These probabilities are plotted in Figure 3. No significant correlation is fount between median farm size and the transition probabilities to double or triple in size. This is in accordance to Gibrat's Law of proportional growth, although this can not be viewed as a formal test of the Gibrat hypothesis.

<FIGURE 2 about here>
<FIGURE 3 about here>

The impact of the covariates on the transition probabilities and the distribution of farms is given by

---

[14] Probabilities on entry and exit must be interpreted with caution, because the 19th category is really artificial, and the actual probabilities do not need to mean anything, except for their relative size which allows comparison between different categories.

[15] To see this consider that the artificial 19th category was normalised to 1, which when brought back to actual numbers is 8729. A probability 0.001 means that 8.7 farms are entering category 17, which is 6.4% of the existing farms in 1998, and is approximately one third of the farms (27 farms) entering category 17 that year (Table 1).

[16] Disney et al. (1988), came to similar conclusion for the pork farms in Southern U.S., whereas Massow, et al. (1992), came to different conclusion for the Ontario pork industry.

calculating the elasticities in 9 and 10 above. The elasticities of transition probabilities for pig prices evaluated at the means are shown in Table 3 (negative values are highlighted)[17]. The pig prices have a negative effect on all transition probabilities to categories 13 and 14, a negative effect on most of transition probabilities to categories 7, 8 and 9, and 1,2,3, and 4. Most of the elasticities for pig prices are positive in most of the upper off–diagonals (except for categories 3, 7, 8, 13 and 14) and negative in most of the lower off–diagonal elements. This is an expected result, indicating that increases in pig prices reduce the probability of firms downsizing, and increase the probability of them increasing in size. Interestingly enough, most of the elements at the left in the last row are also negative indicating that as pork prices increase, entry to the small categories decreases. However, the elasticities for entry in the large categories (15, 16, and 17) are positive and large, indicating a strong influence of pork prices on entry in large categories. Increases in pork prices have a negative effect on exit from large categories, whereas they increase exit for most of the smaller categories, as indicated by the numbers of the last column of Table 3. Notice however that most of the elasticities for the exit category are small.

<TABLE 3 about here>

The cumulated effects of the covariates on the number of farms in each category are given by the category elasticities (Equation 10). These elasticities evaluated at the means are shown in Table 4. As one would expect these numbers are simply a composite of those shown by the probability elasticities, since the category elasticities show the accumulated effect of the covariates over time. Pig prices have a positive elasticity with respect to the size of the largest categories (15, 16, 17, 18). The elasticities for pork prices are negative for categories 12–14, which is in accordance to the transition probability elasticities (Table 3) which are mostly negative for these categories. Similar interpretation can be made for the rest of the size elasticities. The interest rate size elasticities are positive for categories 15–17 and negative for the largest category 18. One possible explanation that Massow et al. (1992) give for a similar result, is that the largest farms may downsize due to high interest rates, which increases the sizes of the immediately lower categories 15–17.

<TABLE 4 about here>

[17] Note that although transition probabilities are zero, transition elasticities do not need to be so. As we see in (9), if $p_{ij}=0$, the elasticity is then simply : $E^{P}_{ijtn} = x_{it} z_{tn} \tilde{\lambda}_{nj}$, which is non–zero as long as the lambdas are non–zero.

6. Concluding Remarks

In this article we have shown several estimators for the non–stationary transition probabilities of Markov process, using generalised cross entropy formalism. An instrumental variables approach was developed and generalised cross entropy estimators for the non–stationary transition probability matrix were derived. The GCE estimator is more efficient and overcomes many of the problems of traditional techniques, such as the OLS and multinomial Logit, especially the problem of dimensionality.

We used the instrumental variables estimator to recover a NSTPM for the Danish pork industry, considering the transitions of 18 size categories of farms over 1984–1998. With GCE formalism we were able to recover data points, missing due to redefinitions of the size categories over time, and thus avoided aggregation of the data. Along the NSTPM we were able to calculate in the form of elasticities, the effects of a number of covariates on transition probabilities and numbers of farms in each category

Overall, this technique is useful when one is faced with such ill–posed problems, with large TPM, and missing data points. It is also a natural tool for allowing the researcher to incorporate efficiently any prior knowledge, in a non–restrictive way.

A major limitation of the instrumental variable technique presented here is that it does not allow for the application of different covariates for each transition probability, which is certainly permissible in the linear model of Section 2.2. The instrumental variable model permits however the application of many covariates, and the calculation of the effect of each of these covariates on each element of the TPM.

References

Adelman, I.G., 1958, A stochastic analysis of the size distribution of firms, Journal of the American Statistical Association, 893-904

Buckwell, A.E. and D.M. Sucksmith, 1979, Projecting farm structural change, Journal of Agricultural Economics, 131-143

Courchane, M., A. Golan, and D. Nickerson, 2000, Estimation and evaluation of loan discrimination – An information approach, Journal of Housing Research, 11

Disney, W.T., P.A. Duffy, and W.E. Hardy Jr, 1988, A Markov Chain analysis of pork farm size distributions in the South, Southern Journal of Agricultural Economics, 57-64

Evans, D.S., 1987a, Tests of alternative theories of firm growth, Journal of Political Economics 95, 657–674

Evans, D.S., 1987b, Empirical analysis of the size distribution of farms: Discussion, American Journal of Agricultural Economics, 484–485

Golan, A., and S.J. Vogel, 2000, Estimation of non–stationary social accounting matrix coefficients with supply side information, Economic Systems Research, 12

Golan, A., G. Judge, and D. Miller, 1996, Maximum entropy econometrics, (John Willey and Sons, Chichester).

Golan, Amos; G. Judge; and J. M. Perloff, 1997, Estimation and inference with censored and ordered multinomial data, Journal of Econometrics, 1997

Hallberg, M.C., 1969, Projecting the size distribution of agricultural firms - an application of a Markov process with non-stationary transition probabilities, American Journal of Agricultural Economics, 289-302

Jaynes, E.T., 1957a, Information theory and statistical mechanics, Physics Review 106, 620-630

Jaynes, E.T., 1957b, Information theory and statistical mechanics II, Physics Review 108, 171-190

Jovanovic, B., 1982, Selection and the evolution of industry, Econometrica, 50(3), 649–670

Landbrugs–Statistik, (Danish agricultural statistics), various issues

Lee, T.C.; G. Judge and A. Zellner, 1977, Estimating the parameters of the Markov probability model from aggregate time series data, 2nd edition (North Holland, Amsterdam)

Lee, T.C. and G. Judge, 1996, Entropy and Cross–Entropy Procedures for Estimating Transition Probabilities from Aggregate Data, in Berry, D.A., C.M. Chaloner, and J.K. Geweke, eds, Bayesian Analysis in Statistics and Econometrics, (John Wiley and Sons) 153–164

Lucas, R.E.Jr., 1978, On the size distribution of firms, Bell Journal of Economics 9: 508-203

MacRae, E.C., 1977, Estimation of time–varying Markov processes with aggregate data, Econometrica 45, 183–198

Massow von, M., A. Weersink, and C.G. Turvey, 1992, Dynamics of structural change in the Ontario hog industry, Canadian Journal of Agricultural Economics, 93-107

McFadden, D., 1974, The Measurement of Urban Travel Demand, Journal of Public Economics, 3, 303–328

Oustapassidis, K., 1986, Greek Agricultural Cooperatives: radical change of trend recognition? Oxford Agrarian Studies XV, I, 216-39

Padberg, D.I., 1962, The use of Markov processes in measuring changes in market structure, American Journal of Agricultural Economics, 189-199

Shannon, C.E., 1948, A mathematical theory of communication, Bell System Technical Journal, 379-423

Telser, L.G., 1963, Least-squares estimates of transition probabilities, in: Christ, C.F., et. al., Editors, Measurement in economics – Studies in mathematical economics and econometrics in memory of Yehunda Grunfeld, , Stanford University Press (Stanford), 270-292

Zepeda, L., 1995a, Asymmetry and non–stationarity in the farm size distribution of Wisconsin milk producers: An aggregate analysis, American Journal of Agricultural Economics, 837–852

Zepeda, L., 1995b, Technical change and the structure of production: A non–stationary Markov analysis, European Review of Agricultural Economics, 41–60

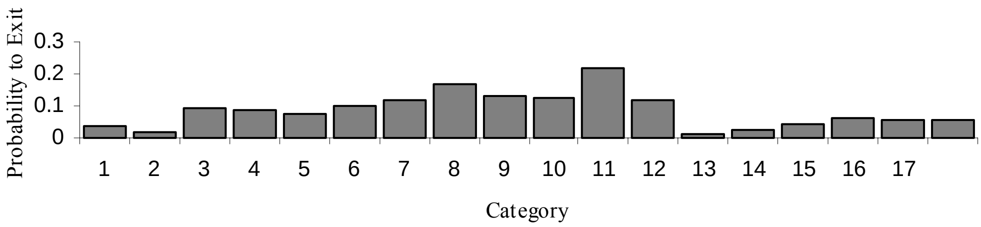

FIgure 1. Probability to Exit by Size Category

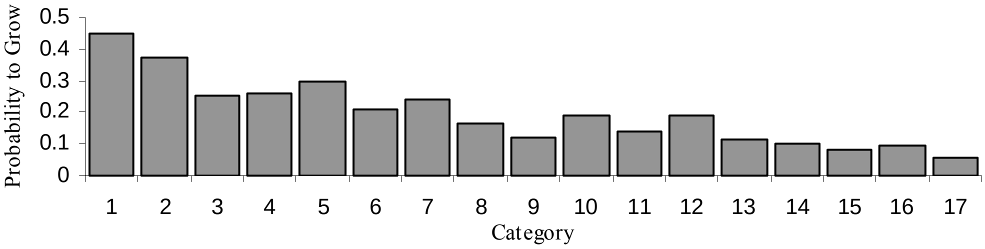

Figure 2. Probability to Grow to Any Size

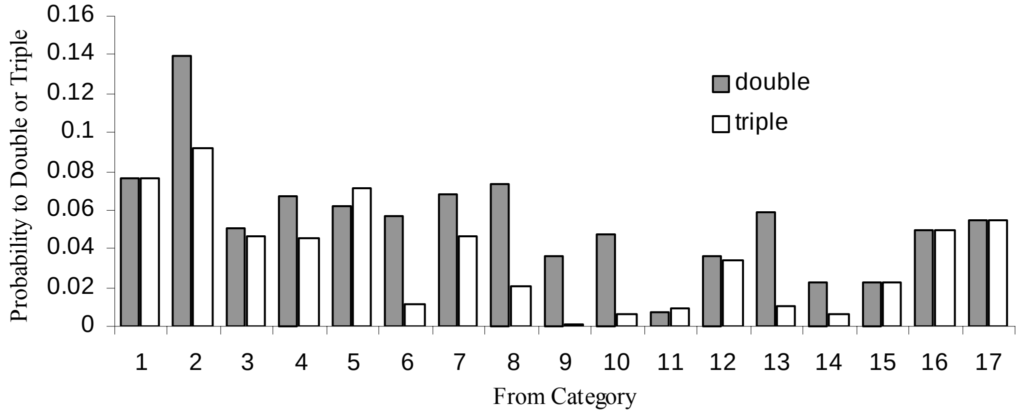

Figure 3. Probability to Double and Probability to Triple in Size

Table 1. Number of Pigs in Denmark by Size Category

	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
1 1-9	5720	4527	4012	3417	3088	2772	2546	2147	1958	1899	1631	**5540**[2]	**5138**[2]	**4417**[2]	**3693**[2]
2 10-29	8729	7508	6267	5699	4993	4372	3783	3558	3279	3167	2570				
3 30-49	5034	5089	4761	4008	3654	2967	2809	2503	2179	2169	1582				
4 50-74	4329	4328	4191	3509	2992	2663	2608	2266	2344	1996	1596	**2393**[3]	**4483**[6]	**4132**[6]	**3527**[6]
5 75-99	3246	3137	2980	2575	2191	1927	1887	1754	1476	1527	1310				
6 100-149	4102	4127	3914	3471	2876	2718	2511	2387	2345	2338	1657	1556			
7 150-199	2599	2690	2561	2372	2306	1891	1884	1818	1728	1535	1378	1194			
8 200-299	3518	3556	3465	3145	2874	2674	2559	2604	2517	2240	1816	1664	**2759**[7]	**2564**[7]	**1507**[7]
9 300-399	2293	2429	2344	2313	2083	2049	2036	1905	1826	1680	1411	1264			
10 400-499	1601	1688	1739	1666	1533	1442	1408	1300	1286	1316	1059	1044	**2428**[8]	**2347**[8]	**3578**[8]
11 500-699	2155	2196	2224	2154	2173	2005	1973	1925	1959	1901	1795	1674			
12 700-999	1561	1552	1648	1655	1694	1701	1645	1760	1682	1774	1685	1679	1595	1605	1428
13 1000-1499	**1207**[1]	**1395**[1]	**1520**[1]	**1706**[1]	**1865**[1]	**2024**[1]	**2253**[1]	**2413**[1]	**2813**[1]	**3317**[1]	**3226**[1]	**2360**[4]	1524	1719	1616
14 1500-1999													817	841	892
15 2000-2999												**1050**[5]	627	719	844
16 3000-4999													328	362	452
17 5000-9999													110	107	134
18 10000+													12	15	18
TOTAL	**46094**	**44222**	**41626**	**37690**	**34322**	**31205**	**29902**	**28340**	**27392**	**26859**	**22716**	**21418**	**19821**	**18828**	**17689**

1. 1000+ for years 1984-1994
2. 1-49 for years 1995-1998
3. 50-99 for year 1995
4. 1000-1999 for year 1995
5. 2000+ for year 1995
6. 50-199 for years 1996-1998
7. 200-399 for years 1996-1998
8. 400-699 for years 1996-1998

Source: Landbrugs-Statistik, various issues

Table 2. Transition Probability Matrix: Non-Stationary - Cross-Entropy

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	0.511	0.076	0.135	0.112	0.061	0.067	0	0	0	0	0	0	0	0	0	0	0	0	0.037
	(0.086)	(0.083)	(0.085)	(0.084)	(0.082)	(0.082)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(0.079)
2	0.037	0.572	0.139	0.092	0.050	0.068	0.023	0	0	0	0	0	0	0	0	0	0	0	0.017
	(0.079)	(0.086)	(0.084)	(0.083)	(0.081)	(0.082)	(0.075)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(0.072)
3	0.080	0.104	0.464	0.051	0.050	0.047	0.043	0.064	0	0	0	0	0	0	0	0	0	0	0.097
	(0.075)	(0.075)	(0.077)	(0.073)	(0.073)	(0.073)	(0.073)	(0.074)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(0.075)
4	0.063	0.113	0.045	0.428	0.054	0.067	0.046	0.049	0.044	0	0	0	0	0	0	0	0	0	0.090
	(0.080)	(0.082)	(0.078)	(0.083)	(0.079)	(0.080)	(0.078)	(0.079)	(0.078)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(0.081)
5	0.059	0.083	0.048	0.050	0.386	0.061	0.062	0.086	0.071	0.017	0	0	0	0	0	0	0	0	0.077
	(0.081)	(0.083)	(0.080)	(0.080)	(0.085)	(0.081)	(0.081)	(0.083)	(0.082)	(0.072)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(0.082)
6	0.069	0.083	0.022	0.032	0.049	0.437	0.047	0.057	0.057	0.012	0.036	0	0	0	0	0	0	0	0.098
	(0.080)	(0.081)	(0.073)	(0.076)	(0.079)	(0.083)	(0.078)	(0.079)	(0.079)	(0.066)	(0.077)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(0.081)
7	0	0.075	0.026	0.045	0.032	0.044	0.418	0.079	0.069	0.022	0.047	0.023	0	0	0	0	0	0	0.120
	(--)	(0.083)	(0.077)	(0.081)	(0.079)	(0.081)	(0.087)	(0.084)	(0.083)	(0.075)	(0.081)	(0.076)	(--)	(--)	(--)	(--)	(--)	(--)	(0.085)
8	0	0	0.016	0.041	0.030	0.043	0.046	0.491	0.044	0.027	0.073	0.021	0	0	0	0	0	0	0.167
	(--)	(--)	(0.068)	(0.075)	(0.073)	(0.076)	(0.076)	(0.081)	(0.076)	(0.073)	(0.078)	(0.071)	(--)	(--)	(--)	(--)	(--)	(--)	(0.080)
9	0	0	0	0.039	0.027	0.045	0.053	0.079	0.503	0.017	0.067	0.036	0	0.002	0	0	0	0	0.130
	(--)	(--)	(--)	(0.078)	(0.075)	(0.079)	(0.080)	(0.081)	(0.084)	(0.071)	(0.081)	(0.077)	(--)	(0.035)	(--)	(--)	(--)	(--)	(0.082)
10	0	0	0	0	0.020	0.020	0.032	0.024	0.023	0.572	0.103	0.048	0.006	0.008	0.022	0	0	0	0.122
	(--)	(--)	(--)	(--)	(0.067)	(0.066)	(0.069)	(0.067)	(0.067)	(0.075)	(0.073)	(0.071)	(0.055)	(0.058)	(0.067)	(--)	(--)	(--)	(0.073)
11	0	0	0	0	0	0.024	0.048	0.077	0.063	0.015	0.415	0.076	0.008	0.009	0.011	0.033	0	0	0.221
	(--)	(--)	(--)	(--)	(--)	(0.072)	(0.077)	(0.078)	(0.078)	(0.068)	(0.081)	(0.078)	(0.060)	(0.062)	(0.064)	(0.074)	(--)	(--)	(0.080)
12	0	0	0	0	0	0	0.015	0.014	0.010	0.033	0.054	0.566	0.046	0.036	0.034	0.036	0.036	0	0.122
	(--)	(--)	(--)	(--)	(--)	(--)	(0.072)	(0.071)	(0.066)	(0.080)	(0.083)	(0.088)	(0.082)	(0.080)	(0.080)	(0.080)	(0.080)	(--)	(0.086)
13	0	0	0	0	0	0	0	0	0	0.085	0.010	0.010	0.768	0.046	0.059	0.010	0.001	0.001	0.010
	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(0.079)	(0.064)	(0.064)	(0.082)	(0.077)	(0.078)	(0.064)	(0.033)	(0.029)	(0.063)
14	0	0	0	0	0	0	0	0	0.004	0.043	0.012	0.028	0.154	0.637	0.067	0.022	0.006	0.005	0.023
	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(0.049)	(0.077)	(0.066)	(0.074)	(0.081)	(0.082)	(0.079)	(0.072)	(0.057)	(0.053)	(0.073)
15	0	0	0	0	0	0	0	0	0	0.079	0.080	0.039	0.069	0.074	0.537	0.037	0.023	0.022	0.041
	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(0.083)	(0.083)	(0.079)	(0.082)	(0.083)	(0.086)	(0.079)	(0.075)	(0.075)	(0.080)
16	0	0	0	0	0	0	0	0	0	0	0.072	0.062	0.070	0.078	0.044	0.515	0.047	0.049	0.062
	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(0.083)	(0.083)	(0.083)	(0.084)	(0.081)	(0.087)	(0.081)	(0.082)	(0.083)
17	0	0	0	0	0	0	0	0	0	0	0	0.063	0.059	0.059	0.031	0.041	0.633	0.055	0.059
	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(0.083)	(0.083)	(0.083)	(0.079)	(0.081)	(0.087)	(0.082)	(0.083)
18	0	0	0	0	0	0	0	0	0	0	0	0	0.061	0.061	0.027	0.038	0.052	0.703	0.057
	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(0.082)	(0.082)	(0.077)	(0.080)	(0.081)	(0.087)	(0.082)
19	0.018	0	0	0	0	0	0	0	0	0.025	0.035	0.031	0.003	0.001	0.002	0.001	0.001	0.003	0.880
	(0.074)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(--)	(0.078)	(0.080)	(0.079)	(0.046)	(0.033)	(0.038)	(0.035)	(0.033)	(0.046)	(0.089)

(Values in parentheses are asymptotic standard deviations calculated as the negative inverse of the Hessian)
Source: Estimated

Table 3. Mean Probability Elasticities for Pig Prices

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	0.8	3.2	-19.6	2.5	18.3	6.7	-0.4	-2.5	-0.3	16.8	13.5	2.2	-32.1	-33	19.8	24.5	18.2	20	3.8
2	0	3.8	-32.1	2.6	27.4	9.3	-1.9	-5.3	-1.7	25	19.9	2.1	-51.7	-53	29.8	37.1	27.2	30.1	4.7
3	8.6	11.3	-13.6	10.4	27.6	15	7.3	5	7.4	25.9	22.4	10.1	-27.1	-28	29.2	34.3	27.4	29.5	11.9
4	-1.4	1	-21.2	0.2	15.5	4.3	-2.5	-4.6	-2.4	14.1	10.9	-0.1	-33.2	-34	17	21.5	15.4	17.2	1.6
5	-4.5	-2.8	-18.8	-3.4	7.7	-0.4	-5.3	-6.8	-5.3	6.6	4.4	-3.6	-27.5	-28.1	8.7	12	7.6	8.9	-2.4
6	-3.7	-1.4	-23.1	-2.1	12.9	1.9	-4.8	-6.9	-4.7	11.4	8.4	-2.4	-35	-35.8	14.3	18.8	12.8	14.5	-0.8
7	-0.7	0.9	-14.6	0.4	11.2	3.3	-1.5	-3	-1.4	10.1	7.9	0.2	-23.1	-23.7	12.2	15.4	11.1	12.3	1.3
8	-0.6	1.7	-19.8	0.9	15.7	4.9	-1.7	-3.8	-1.6	14.3	11.3	0.7	-31.5	-32.3	17.1	21.5	15.6	17.4	2.2
9	-1	0.7	-14.8	0.1	10.8	3	-1.8	-3.2	-1.7	9.8	7.6	0	-23.2	-23.8	11.9	15	10.8	12	1.1
10	-5.6	-4.4	-16.1	-4.8	3.2	-2.6	-6.3	-7.4	-6.2	2.5	0.8	-5	-22.4	-22.8	4	6.4	3.2	4.1	-4.1
11	-4.5	-2.8	-18.8	-3.4	7.7	-0.4	-5.3	-6.8	-5.3	6.6	4.3	-3.6	-27.5	-28	8.7	12	7.6	8.9	-2.4
12	-1	0.5	-13.5	0	9.6	2.6	-1.7	-3	-1.7	8.7	6.7	-0.2	-21.1	-21.6	10.5	13.4	9.6	10.7	0.8
13	6.8	7.5	1	7.3	11.8	8.5	6.5	5.9	6.5	11.4	10.4	7.2	-2.6	-2.8	12.2	13.6	11.8	12.3	7.7
14	3.7	4	0.5	3.9	6.4	4.6	3.5	3.1	3.5	6.1	5.6	3.9	-1.4	-1.6	6.6	7.3	6.3	6.6	4.1
15	-1.3	-1	-4.1	-1.1	1.1	-0.5	-1.5	-1.8	-1.5	0.9	0.4	-1.1	-5.8	-5.9	1.3	1.9	1.1	1.3	-0.9
16	-1.3	-1.1	-3.8	-1.1	0.8	-0.6	-1.5	-1.8	-1.5	0.6	0.2	-1.2	-5.4	-5.5	1	1.5	0.8	1	-1
17	-1.1	-0.8	-3.3	-0.9	0.8	-0.4	-1.2	-1.4	-1.2	0.7	0.3	-0.9	-4.6	-4.7	1	1.5	0.8	1	-0.7
18	-1.1	-0.9	-3	-1	0.5	-0.6	-1.2	-1.4	-1.2	0.4	0.1	-1	-4.2	-4.3	0.7	1.1	0.5	0.7	-0.8
19	-11.3	-3.6	-76.8	-6.1	44.5	7.6	-15.2	-22.1	-14.8	39.7	29.3	-7	-117	-120	49.4	64.4	44.2	50.1	-1.7

Source: Estimated

Table 4. Size Elasticities (Cross-Entropy)

PRICES	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Pig Meat	0	3.1	-17.6	1.2	13.1	3.8	-1.7	-3.4	-2.4	10.5	9.3	-1.7	-9.4	-10.5	7.2	8.6	4.5	6.4	-1
Milk	-8.6	-0.5	-4.9	1.6	-1.1	-1.1	6.2	-5.3	-3.8	-11.4	17.8	1.3	3.8	8	-2.1	-4.2	-2.5	-2.1	3.1
Egg	5.7	-3.1	5.4	3.2	-0.7	1.2	-6.2	-6.7	-4.3	28.6	15.8	-3.9	-1.7	-1.6	-2.1	-0.4	0.6	0.4	-6
Cattle Meat	3.2	7.6	-9.7	-0.1	-18.4	-7.8	2.6	23.1	12.1	-17.9	-34.2	6.4	0.9	-6.6	5.4	3.2	-0.3	-0.3	4
Poultry	10.4	-3.6	-36.8	-0.6	-3	-31.2	20.1	4.4	8.8	-76.3	-19.1	6.2	16.1	21.4	-13.3	-15.4	-7.9	-11.4	28.1
Pig feed	3.1	-30.1	115.1	2.2	-19	72.4	-31.3	-35	-24.9	84.3	110.6	-9.6	-7.5	0.9	-35.9	-15	13.6	22.4	-55.2
Cattle Feeds	-7.5	-6.5	9.3	15.6	-3.3	3.7	-2.6	11.2	-5.5	10.9	-13.3	2.1	1.6	-1.5	4.4	1.7	-2.3	-3.4	-3.6
Poultry Feeds	-18	34.9	-82.3	-21.4	29.9	-49.9	25.6	8.8	20.7	-53.5	-59	-2.6	1.8	4.2	23.4	8.8	-9.1	-15.6	39.7
Fertilizers	1.4	0.6	29.7	-5.9	-4.1	13.1	-10.9	-7.2	-4.8	32.7	3.8	-0.1	-2.6	-6.3	2.2	4.2	2.6	3.6	-12.5
Interest Rate	9.9	-1.1	-6.5	2.5	8.5	-3	-3.3	9.7	4.2	-4.2	-33.6	1.6	-3.3	-8.8	10.9	9.1	1	0.1	2.4

Source: Estimated